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

**The SMEFiT Collaboration:**

Jacob J. Ethier,[1,2] Fabio Maltoni,[3,4] Luca Mantani,[3] Emanuele R. Nocera,[2,5] Juan Rojo,[1,2] Emma Slade,[6] Eleni Vryonidou,[7] and Cen Zhang[8,9]

[1] *Department of Physics and Astronomy, Vrije Universiteit Amsterdam, NL-1081 HV Amsterdam, The Netherlands*

[2] *Nikhef Theory Group, Science Park 105, 1098 XG Amsterdam, The Netherlands*

[3] *Centre for Cosmology, Particle Physics and Phenomenology (CP3), Université Catholique de Louvain, B-1348 Louvain-la-Neuve, Belgium*

[4] *Dipartimento di Fisica e Astronomia, Università di Bologna and INFN, Sezione di Bologna, via Irnerio 46, 40126 Bologna, Italy*

[5] *The Higgs Centre for Theoretical Physics, The University of Edinburgh, JCMB, KB, Mayfield Rd, Edinburgh EH9 3JZ, Scotland*

[6] *Rudolf Peierls Centre for Theoretical Physics, University of Oxford, Clarendon Laboratory, Parks Road, Oxford OX1 3PU, United Kingdom*

[7] *Department of Physics and Astronomy, University of Manchester, Oxford Road, Manchester M13 9PL, United Kingdom*

[8] *Institute of High Energy Physics, and School of Physical Sciences, University of Chinese Academy of Sciences, Beijing 100049, China*

[9] *Center for High Energy Physics, Peking University, Beijing 100871, China*

## Abstract

We present a global interpretation of Higgs, diboson, and top quark production and decay measurements from the LHC in the framework of the Standard Model Effective Field Theory (SMEFT) at dimension six. We constrain simultaneously 36 independent directions in its parameter space, and compare the outcome of the global analysis with that from individual and two-parameter fits. Our results are obtained by means of state-of-the-art theoretical calculations for the SM and the EFT cross-sections, and account for both linear and quadratic corrections in the $1/\Lambda^2$ expansion. We demonstrate how the inclusion of NLO QCD and $\mathcal{O}(\Lambda^{-4})$ effects is instrumental to accurately map the posterior distributions associated to the fitted Wilson coefficients. We assess the interplay and complementarity between the top quark, Higgs, and diboson measurements, deploy a variety of statistical estimators to quantify the impact of each dataset in the parameter space, and carry out fits in BSM-inspired scenarios such as the top-philic model. Our results represent a stepping stone in the ongoing program of model-independent searches at the LHC from precision measurements, and pave the way towards yet more global SMEFT interpretations extended to other high-$p_T$ processes as well as to low-energy observables.

arXiv:2105.00006v1 [hep-ph] 30 Apr 2021

# Contents

# 1 Introduction

A powerful, model-independent framework to constrain, identify, and parametrise potential deviations with respect to the predictions of the Standard Model (SM) is provided by the Standard Model Effective Field Theory (SMEFT) [1–3], see also [4] for a review. A particularly attractive feature of the SMEFT is its capability to systematically correlate deviations from the SM between different processes, for example between Higgs and top quark cross-sections, or between high-$p_T$ and flavor observables.

A direct consequence of this model independence is the high dimensionality of the parameter space spanned by the relevant higher-dimensional EFT operators. Indeed, the number of Wilson coefficients constrained in typical SMEFT analyses can vary between just a few up to the several tens or even hundreds, depending on the specific assumptions adopted concerning the flavour, family (non-)universality of the couplings, and CP-symmetry structure (among others) of the UV-complete theory. For this reason, 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.

The phenomenology of the SMEFT has attracted significant attention, with most analyses focusing on specific sectors of the parameter space and groups of processes. Some of these recent studies have targeted the top quark properties [5–8], the Higgs and electroweak gauge sector [9–11], single and double gauge boson production [12–15], vector-boson scattering [14, 16, 17], and flavour and low-energy observables [18–20], among several others. Furthermore, analyses that combine the constraints of different groups of processes in the EFT parameter space, such as the Higgs and electroweak sector with the top quark one [21] or top quark data with $B$-meson observables [22, 23], have also been presented. These and related studies demonstrate that a global interpretation of the SMEFT is unavoidable and makes possible benefiting from hitherto unexpected connections, such as the correlation of the LHCb flavour anomalies [24, 25] at the $B$-meson scale with the high-$p_T$ tails at the LHC [15, 26].

With the ultimate motivation of performing a truly global EFT interpretation of particle physics data, the `SMEFiT` fitting framework was developed in [7] and applied to the analysis of the top quark properties at the LHC as a proof-of-concept. This novel EFT fitting methodology, inspired by techniques deployed by the NNPDF Collaboration to determine the proton's parton distribution functions (PDFs) [27–31], made possible constraining the Wilson coefficients associated to 34 independent dimension-six operators that modify the production cross-sections of top quarks. Our results improved over existing bounds [32] for the wide majority of directions in the SMEFT parameter space and in several cases the associated Wilson coefficients were constrained for the first time. Subsequently, `SMEFiT` was extended with the Bayesian reweighting method [33] developed for PDFs [34, 35] which allows one constraining the EFT parameter space *a posteriori* with novel measurements without requiring a dedicated fit. `SMEFiT` has also been recently applied for the first SMEFT interpretation of vector boson scattering data [14] from the full Run II dataset.

In this work, we complement and extend the `SMEFiT` analysis framework of [7] in several directions. First and foremost, we extend the dimension-six EFT operator basis in order to simultaneously describe top-quark measurements together with Higgs boson production and decay cross-sections, as well as with weak gauge boson pair production from LEP and the LHC. Specifically, we consider Higgs signal strengths, differential distributions, and simplified template cross-section (STXS) measurements from ATLAS and CMS taken at Runs I and II. Furthermore, we account for the most recent top-quark observables from the Run II dataset, such as updated measurements of four-top, top quark pair in association with a $Z$ boson, and differential single-top and top quark pair production. We also include the differential distributions in gauge boson pair production from LEP and the LHC, which constrain complementary directions in the EFT space. In addition, we account in an indirect manner for the information provided by electroweak precision observables (EWPO) from LEP [36] by means of imposing restrictions on specific combinations of the EFT coefficients.

A second improvement as compared to [7] concerns the fitting methodology. On the one hand, the Monte Carlo replica fitting method has been upgraded by means of more efficient

optimizers and the imposition of post-fit quality selection criteria for the replicas. On the other hand, we have implemented a novel, independent approach to constrain the parameter space based on Nested Sampling (NS) by means of the MultiNest algorithm [37]. As opposed to the replica fitting method, which is an optimisation problem, NS aims to reconstruct the posterior probability distribution given the model and the data by means of Bayesian inference. We have cross-validated the performance of the two methods and demonstrated that they lead to equivalent results. The availability of two orthogonal fitting strategies strengthens the robustness of SMEFiT and facilitates the combined interpretation of data from different processes.

From the combination of the improved fitting framework and the extensive input dataset, we derive individual, two-dimensional, and global (marginalised) bounds for 36 independent directions (and 14 dependent ones) in the EFT parameter space. The EFT cross-sections used in this analysis account for either only the linear or for both linear and quadratic effects, $\mathcal{O}(\Lambda^{-2})$ and $\mathcal{O}(\Lambda^{-4})$ respectively, and include NLO QCD corrections whenever available. We demonstrate in detail how the inclusion of NLO QCD and $\mathcal{O}(\Lambda^{-4})$ corrections in the EFT calculations is instrumental in order to accurately pin down the posterior distributions associated to the fitted Wilson coefficients.

By means of information geometry and principal component analysis techniques, we quantify the sensitivity of each of the input datasets to the various Wilson coefficients. We validate these statistical diagnosis tools by means of a series of fits restricted to subsets of processes, such as Higgs-only and top-only EFT analyses. Specifically, we quantify the interplay between the top-quark and Higgs measurements in the determination of EFT degrees of freedom sensitive to both processes, such as the modifications of the top Yukawa coupling. Furthermore, we explore how the EFT fit results are modified when additional, UV-inspired theory restrictions are imposed in the parameter space, and present results for the case of a top-philic model.

The paper is organised as follows. First of all, Sect. 2 discusses the operator basis, flavour assumptions, the fitted degrees of freedom, and the top-philic scenario. Then Sect. 3 describes the top-quark, Higgs, and diboson datasets that are used as input to the analysis together with the corresponding SM and EFT calculations. The methodological improvements in SMEFiT, together with the description of the fit settings, are presented in Sect. 4. The main results of this work, namely the combined SMEFT interpretation of top-quark, Higgs, and diboson measurements at the LHC, are presented and discussed in Sect. 5. Finally, in Sect. 6 we summarise and discuss future steps in this project.

Supplementary information is provided in three appendices. In App. A we present the comparison between the SM and SMEFT theory predictions with the experimental datasets used as input to the fit; in App. B we describe the implementation of the Higgs signal strength measurements; then in App. C we discuss how the results of this work are rendered publicly available and provide usage instruction.

## 2 EFT description of the top, Higgs, and electroweak sectors

In this section we collect the definitions and conventions that will be used to construct the dimension-six operators and the associated degrees of freedom (DoFs) relevant for the theoretical description of the processes considered in this analysis. These are operators that modify the production and decay of Higgs bosons and top quarks at hadron colliders, precision electroweak measurements from LEP/SLC, and gauge-boson pair production cross-sections both

at LEP2 and at the LHC.

First of all, we provide explicit definitions for the operators and for the physical EFT coefficients adopted in this work, as well as the corresponding notational conventions. Following the recommendation of the LHC Top Quark Working Group [32] as well as the strategy of our previous work [7], in the top-quark sector we fit specific degrees of freedom closely related to the experimental measurements, instead of directly using the Warsaw-basis operator coefficients. Our degrees of freedom are therefore linear combinations of the Warsaw-basis operator coefficients, which appear in the interference with SM amplitudes, and represent interactions of physical fields after electroweak symmetry breaking. These combinations are then aligned with physically relevant directions of the parameter space, and thus have a more transparent physical interpretation. They also represent the maximal information that can be extracted from measuring a certain process.

We will then discuss how the constraints provided by the electroweak precision observables (EWPOs) from LEP/SLC can be approximately accounted for by means of a series of restrictions on the EFT parameter space. We also discuss theoretical constraints on the operator coefficients following a more restrictive assumption about the UV-complete theory, namely the so-called top-philic scenario. Finally, we discuss several theoretical relations that must be satisfied by the EFT cross-sections following the requirement that physical cross-sections are positive-definite quantities.

## 2.1 Operator basis and degrees of freedom

**Conventions.**  Let us start by summarizing the notation and conventions that are adopted in this work concerning the relevant dimension-six SMEFT operators. Here we follow the notation of the Warsaw basis presented in [3]. In this notation, flavour indices are labelled by $i, j, k$ and $l$; left-handed quark and lepton fermion $SU(2)_L$ doublets are denoted by $q_i$, $\ell_i$; the right-handed quark singlets by $u_i$, $d_i$, while the right-handed lepton singlets are denoted by $e$, $\mu$, $\tau$ without using flavor index. Given the special role of the top-quark in this work, we use $Q$ and $t$ to denote the left-handed top-bottom doublet and the right-handed top singlet, instead of using $q_3$ and $u_3$. The Higgs doublet is denoted by $\varphi$; the antisymmetric $SU(2)$ tensor by $\varepsilon \equiv i\tau^2$; $\tilde{\varphi} = \varepsilon\varphi^*$; and we define

$$(\varphi^\dagger i\overleftrightarrow{D}_\mu \varphi) \equiv \varphi^\dagger(iD_\mu\varphi) - (iD_\mu\varphi^\dagger)\varphi\,, \qquad (\varphi^\dagger i\overleftrightarrow{D}^I_\mu \varphi) \equiv \varphi^\dagger\tau^I(iD_\mu\varphi) - (iD_\mu\varphi^\dagger)\tau^I\varphi\,, \qquad (2.1)$$

where $\tau^I$ are the Pauli matrices. In the following, $G^A_{\mu\nu}$, $W^I_{\mu\nu}$, and $B_{\mu\nu}$ stand for the $SU(3)$ strong and $SU(2)_L$ and $U(1)_Y$ weak gauge field strengths respectively, and the covariant derivatives include all the relevant interaction terms. For instance, the gluon field strength tensor is given by

$$G^A_{\mu\nu} = \partial_\mu G^A_\nu - \partial_\nu G^A_\mu + g_s f^{ABC} G^B_\mu G^C_\nu\,, \qquad (2.2)$$

where $G^A_\mu$ is the gluon field, $A, B, C$ are color indices in the adjoint representation, $g_s$ is the strong coupling and $f^{ABC}$ are the structure constants of $SU(3)$. Similar definitions hold for the electroweak $W^{\mu\nu}_I$ and $B^{\mu\nu}$ field strength tensors, for instance one has

$$W^I_{\mu\nu} = \partial_\mu W^I_\nu - \partial_\nu W^I_\mu + g_w \epsilon^I_{JK} W^J_\mu W^K_\mu\,, \qquad (2.3)$$

where $g_w$ is the $SU(2)_L$ coupling constant.

**Flavour assumptions.** The number of independent dimension-six operators can be unfeasibly large, if all three generations of the SM fermions are taken into account: there are 2499 in total [38], with 572 four-fermion operators that are in principle relevant for top-quark physics [39]. In this analysis, we follow closely the strategy which we adopted in our previous top-quark sector study [7] and that has been documented in the LHC Top Quark Working Group note [32]: we implement the Minimal Flavour Violation (MFV) hypothesis [40] in the quark sector as the baseline scenario. A slight difference is that instead of a $U(2)_q \times U(2)_u \times U(2)_d$ flavour symmetry among the first two generations, we now impose the $U(2)_q \times U(2)_u \times U(3)_d$ symmetry, under the assumption that the Yukawa couplings are nonzero only for the top quark. This flavour assumption is consistent with the `SMEFT@NLO` model [41], the implementation of automated one-loop calculation in the SMEFT which we will use to the provide theoretical inputs for our global fit, as discussed in the next section.

As a result of the different flavour assumption, the EFT parameter space is further reduced compared to [7]. In particular, the coefficients of operators with right-handed bottom quarks are either set to zero or set equal to the corresponding down-quark ones. Furthermore, we then slightly relax our assumptions by keeping the bottom and charm quark Yukawa operators in our fit, to account for the current LHC sensitivity to these parameters. All other light quark Yukawa operators are set to zero, since we do not expect to have any sensitivity on their coefficients.

Concerning the leptonic sector, the adopted flavour symmetry is $(U(1)_\ell \times U(1)_e)^3$, also following [32]. This assumption sets all the lepton masses as well as their Yukawa couplings to zero in the SM, while leaving independent parameters for each lepton-antilepton pair of a given generation. This is then relaxed by including the $\tau$ Yukawa operator, to account for the expected LHC sensitivity arising from dedicated measurements. In practice, the lepton flavour assumptions do not have implications for the EFT fit given the constraints from $Z$-pole measurements at LEP and SLC, see the discussion below.

**Purely bosonic operators.** Table 2.1 reports the purely bosonic dimension-six operators that modify the production and decay of Higgs bosons as well as the interactions of the electroweak gauge bosons. For each operator, we indicate its definition in terms of the SM fields and the notation that we will use both for the operators and for the Wilson coefficients. These operators modify several important Higgs boson production and decay processes that are (or will become) accessible at the LHC, as well as the production of gauge boson pairs both in electron-positron and in proton-proton collisions.

One can comment on some interesting features of the operators defined in Table 2.1. To begin with, the operators $O_{\varphi WB}$ and $O_{\varphi D}$ are the ones often identified as the $S$ and $T$ oblique parameters, though this identification is basis-dependent and is not strictly correct in the Warsaw basis. Together with several of the two-fermion operators listed in Table 2.2, they are severely constrained by the $Z$-pole and $W$-pole measurements available from LEP and SLC, but with 2 linear combinations left unconstrained. These two combinations in turn modify the electroweak triple gauge boson (TGC) couplings and the Higgs-electroweak interactions. They are thus constrained mainly by the diboson measurements at the LEP2 and the LHC, as well as the Higgs measurements at the LHC. We will discuss this property in more detail in the following section. The operator $O_W$ generates a TGC coupling modification which is purely transversal and is hence constrained only by diboson data.

The rest of the bosonic operators listed in Table 2.1 modify only the Higgs boson couplings,

| Operator | Coefficient | Definition |
|----------|-------------|------------|
| $\mathcal{O}_{\varphi G}$ | $c_{\varphi G}$ | $\left(\varphi^\dagger \varphi\right) G^{\mu\nu}_A G^A_{\mu\nu}$ |
| $\mathcal{O}_{\varphi B}$ | $c_{\varphi B}$ | $\left(\varphi^\dagger \varphi\right) B^{\mu\nu} B_{\mu\nu}$ |
| $\mathcal{O}_{\varphi W}$ | $c_{\varphi W}$ | $\left(\varphi^\dagger \varphi\right) W^{\mu\nu}_I W^I_{\mu\nu}$ |
| $\mathcal{O}_{\varphi WB}$ | $c_{\varphi WB}$ | $(\varphi^\dagger \tau_I \varphi) B^{\mu\nu} W^I_{\mu\nu}$ |
| $\mathcal{O}_{\varphi d}$ | $c_{\varphi d}$ | $\partial_\mu(\varphi^\dagger \varphi)\partial^\mu(\varphi^\dagger \varphi)$ |
| $\mathcal{O}_{\varphi D}$ | $c_{\varphi D}$ | $(\varphi^\dagger D^\mu \varphi)^\dagger(\varphi^\dagger D_\mu \varphi)$ |
| $\mathcal{O}_W$ | $c_{WWW}$ | $\epsilon_{IJK} W^I_{\mu\nu} W^{J,\nu\rho} W^{K,\mu}_\rho$ |

**Table 2.1.** Purely bosonic dimension-six operators that modify the production and decay of Higgs bosons and the interactions of the electroweak gauge bosons. For each operator, we indicate its definition in terms of the SM fields, and the notational conventions that will be used both for the operator and for the Wilson coefficient. The operators $O_{\varphi WB}$ and $O_{\varphi D}$ are severely constrained by the EWPOs together with several of the two-fermion operators from Table 2.2.

and represent degrees of freedom that are accessible only with Higgs data. First, the operators $\mathcal{O}_{\varphi W}$ and $\mathcal{O}_{\varphi B}$ modify the interaction between Higgs bosons and electroweak gauge bosons. At the LHC, they can be probed for example by means of the Higgs decays into weak vector bosons, $h \to ZZ^*$ and $h \to W^+W^-$, as well as in the vector-boson-fusion (VBF) process and in associated production with vector bosons, $hW$ and $hZ$. In addition, the $\mathcal{O}_{\varphi G}$ operator is similar but introduces a direct coupling between the Higgs boson and gluons. It therefore enters the Higgs total width and branching ratios, the production cross section in gluon fusion channel, as well as the associated production channel $t\bar{t}h$. Finally, the $O_{\varphi d}$ operator generates a wavefunction correction to the Higgs boson, which rescales all the Higgs boson couplings in a universal manner.

**Two-fermion operators.** Table 2.2 collects, using the same format as in Table 2.1, the relevant Warsaw-basis operators that contain two fermion fields, either quarks or leptons, plus a single four-lepton operator. From top to bottom, we list the two-fermion operators involving 3rd generation quarks, those involving 1st and 2nd generation quarks, and operators containing two leptonic fields (of any generation). We also include in this list the four-lepton operator $\mathcal{O}_{\ell\ell}$.

The operators that involve a top-quark field, either $Q$ (left-handed doublet) or $t$ (right-handed singlet), are crucial for the interpretation of LHC top-quark measurements. Interestingly, all of them involve at least one Higgs-boson field, which introduces an interplay between the top and Higgs sectors of the SMEFT. For example, the chromo-magnetic dipole operator $O_{tG}$ and the dimension-six Yukawa operator $O_{t\varphi}$ are constrained by both top quark measurements, such as $t\bar{t}h$ associated production, as well as Higgs measurements, such as

| Operator | Coefficient | Definition |
|---|---|---|
| | **3rd generation quarks** | |
| $\mathcal{O}_{\varphi Q}^{(1)}$ | $c_{\varphi Q}^{(1)}$ (*) | $i(\varphi^\dagger \overset{\leftrightarrow}{D}_\mu \varphi)(\bar{Q}\,\gamma^\mu Q)$ |
| $\mathcal{O}_{\varphi Q}^{(3)}$ | $c_{\varphi Q}^{(3)}$ | $i(\varphi^\dagger \overset{\leftrightarrow}{D}_\mu \tau_I \varphi)(\bar{Q}\,\gamma^\mu \tau^I Q)$ |
| $\mathcal{O}_{\varphi t}$ | $c_{\varphi t}$ | $i(\varphi^\dagger \overset{\leftrightarrow}{D}_\mu \varphi)(\bar{t}\,\gamma^\mu t)$ |
| $\mathcal{O}_{tW}$ | $c_{tW}$ | $i(\bar{Q}\tau^{\mu\nu}\tau_I t)\,\tilde{\varphi}\,W_{\mu\nu}^I + \text{h.c.}$ |
| $\mathcal{O}_{tB}$ | $c_{tB}$ (*) | $i(\bar{Q}\tau^{\mu\nu} t)\,\tilde{\varphi}\,B_{\mu\nu} + \text{h.c.}$ |
| $\mathcal{O}_{tG}$ | $c_{tG}$ | $ig_s\,(\bar{Q}\tau^{\mu\nu}T_A t)\,\tilde{\varphi}\,G_{\mu\nu}^A + \text{h.c.}$ |
| $\mathcal{O}_{t\varphi}$ | $c_{t\varphi}$ | $\left(\varphi^\dagger \varphi\right)\bar{Q}\,t\,\tilde{\varphi} + \text{h.c.}$ |
| $\mathcal{O}_{b\varphi}$ | $c_{b\varphi}$ | $\left(\varphi^\dagger \varphi\right)\bar{Q}\,b\,\varphi + \text{h.c.}$ |
| | **1st, 2nd generation quarks** | |
| $\mathcal{O}_{\varphi q}^{(1)}$ | $c_{\varphi q}^{(1)}$ (*) | $\sum_{i=1,2} i(\varphi^\dagger \overset{\leftrightarrow}{D}_\mu \varphi)(\bar{q}_i\,\gamma^\mu q_i)$ |
| $\mathcal{O}_{\varphi q}^{(3)}$ | $c_{\varphi q}^{(3)}$ | $\sum_{i=1,2} i(\varphi^\dagger \overset{\leftrightarrow}{D}_\mu \tau_I \varphi)(\bar{q}_i\,\gamma^\mu \tau^I q_i)$ |
| $\mathcal{O}_{\varphi u}$ | $c_{\varphi u}$ | $\sum_{i=1,2} i(\varphi^\dagger \overset{\leftrightarrow}{D}_\mu \varphi)(\bar{u}_i\,\gamma^\mu u_i)$ |
| $\mathcal{O}_{\varphi d}$ | $c_{\varphi d}$ | $\sum_{i=1,2,3} i(\varphi^\dagger \overset{\leftrightarrow}{D}_\mu \varphi)(\bar{d}_i\,\gamma^\mu d_i)$ |
| $\mathcal{O}_{c\varphi}$ | $c_{c\varphi}$ | $\left(\varphi^\dagger \varphi\right)\bar{q}_2\,c\,\tilde{\varphi} + \text{h.c.}$ |
| | **two-leptons** | |
| $\mathcal{O}_{\varphi \ell_i}^{(1)}$ | $c_{\varphi \ell_i}^{(1)}$ | $i(\varphi^\dagger \overset{\leftrightarrow}{D}_\mu \varphi)(\bar{\ell}_i\,\gamma^\mu \ell_i)$ |
| $\mathcal{O}_{\varphi \ell_i}^{(3)}$ | $c_{\varphi \ell_i}^{(3)}$ | $i(\varphi^\dagger \overset{\leftrightarrow}{D}_\mu \tau_I \varphi)(\bar{\ell}_i\,\gamma^\mu \tau^I \ell_i)$ |
| $\mathcal{O}_{\varphi e}$ | $c_{\varphi e}$ | $i(\varphi^\dagger \overset{\leftrightarrow}{D}_\mu \varphi)(\bar{e}\,\gamma^\mu e)$ |
| $\mathcal{O}_{\varphi \mu}$ | $c_{\varphi \mu}$ | $i(\varphi^\dagger \overset{\leftrightarrow}{D}_\mu \varphi)(\bar{\mu}\,\gamma^\mu \mu)$ |
| $\mathcal{O}_{\varphi \tau}$ | $c_{\varphi \tau}$ | $i(\varphi^\dagger \overset{\leftrightarrow}{D}_\mu \varphi)(\bar{\tau}\,\gamma^\mu \tau)$ |
| $\mathcal{O}_{\tau \varphi}$ | $c_{\tau \varphi}$ | $\left(\varphi^\dagger \varphi\right)\bar{\ell}_3\,\tau\,\varphi + \text{h.c.}$ |
| | **four-lepton** | |
| $\mathcal{O}_{\ell\ell}$ | $c_{\ell\ell}$ | $\left(\bar{\ell}_1 \gamma_\mu \ell_2\right)\left(\bar{\ell}_2 \gamma^\mu \ell_1\right)$ |

**Table 2.2.** Same as Table 2.1 for the operators containing two fermion fields, either quarks or leptons, as well as the four-lepton operator $\mathcal{O}_{\ell\ell}$. The flavor index $i$ runs from 1 to 3. The coefficients indicated with (*) in the second column do not correspond to physical degrees of freedom in the fit, but are rather replaced by $c_{\varphi q_i}^{(-)}$, $c_{\varphi Q_i}^{(-)}$, and $c_{tZ}$ defined in Table 2.3.

Higgs production through gluon fusion. Furthermore, the electroweak-dipole operators, $O_{tW}$ and $O_{tB}$, as well as the current operators, $O_{\varphi Q}^{(3)}$ and $O_{\varphi t}$, can be constrained by the associated production of single top-quarks and Higgs bosons, as well as by the loop-induced Higgs decays into a $Z\gamma$ final state.

In Table 2.2 we also list operators that contain light quark (1st and 2nd generation) and leptonic fields (of any generation). The light quark operators enter the Higgs production through the $Vh$ and VBF channels, as well as the diboson processes. These operators also modify the Higgs boson width and branching ratios. For example, the Higgs decay width to $q\bar{q}\ell^+\ell^-$ becomes modified by operators that induce an effective $Zhq\bar{q}$ vertex, such as $\mathcal{O}_{\varphi u}$. The leptonic operators are relevant for the same reason, once we account for the leptonic decays of the Higgs and gauge bosons. In addition, indirect contributions arise from the $\mathcal{O}_{\varphi\ell_1}^{(3)}$, $\mathcal{O}_{\varphi\ell_2}^{(3)}$, and $\mathcal{O}_{\ell\ell}$ operators, which modify the measurement of the Fermi constant, $G_F$, and this affects the extracted SM parameters. They therefore introduce a universal contribution to all electroweak interactions, and are relevant for $Vh$, VBF, and for the diboson channels.

| DoF | Definition |
|-----|-----------|
| $c_{\varphi Q}^{(-)}$ | $c_{\varphi Q}^{(1)} - c_{\varphi Q}^{(3)}$ |
| $c_{tZ}$ | $-\sin\theta_W c_{tB} + \cos\theta_W c_{tW}$ |
| $c_{\varphi q}^{(-)}$ | $c_{\varphi q}^{(1)} - c_{\varphi q}^{(3)}$ |

**Table 2.3.** Additional degrees of freedom defined from linear combinations of the two-fermion operators listed in Table 2.2. The first two DoFs modify the $t\bar{t}Z$ couplings, while the third combination is introduced for consistency with the first one. These are the DoFs that enter at the fit level, replacing those marked with (*) in Table 2.2.

We point out that most of the operator coefficients defined in Table 2.2 correspond directly to degrees of freedom used in the fit, except for three of them, which are indicated with a (*) in the second column. Instead, following Ref. [32], three additional degrees of freedom are defined from the linear combinations indicated in Table 2.3. These are the DoFs that enter at the fit level, replacing those marked with a (*) in Table 2.2.

Finally, we note that, as mentioned above, here flavour universality in the leptonic sector is not imposed, and thus the coefficients of the operators involving bilinears in the electron, muon, and tau lepton fields are in principle independent. In total we have 23 independent fit parameters, defined from two-fermion operators, plus in addition the four-lepton operator $c_{\ell\ell}$. However, in practice, this flexibility will not be relevant for the present fit due to the constraints from the EWPOs, to be discussed next.

**The role of electroweak precision observables.** At this point, one should note that a subset of the dimension-six operators defined in Tables 2.1 and 2.2 are already well constrained by the electroweak precision observables (EWPO) [42] measured at the $Z$-pole [36] and the $W$-pole at the LEP and SLC electron-position colliders. Given in particular the high accuracy of these LEP measurements, these constraints are known to dominate in many cases when compared to those provided by the LHC cross-sections. Specifically, the operators sensitive

to the EWPO are the following (with definitions presented in Tables 2.1 and 2.2)

$$\mathcal{O}_{\varphi WB}, \mathcal{O}_{\varphi D}, \mathcal{O}_{\varphi q}^{(1)}, \mathcal{O}_{\varphi q}^{(3)}, \mathcal{O}_{\varphi u}, \mathcal{O}_{\varphi d}, \mathcal{O}_{\varphi \ell_i}^{(3)}, \mathcal{O}_{\varphi \ell_i}^{(1)}, \mathcal{O}_{\varphi e/\mu/\tau}, \mathcal{O}_{\ell\ell}. \tag{2.4}$$

Note that, with $i = 1, 2, 3$, these add up to 16 operators, rather than the 10 which would correspond to the flavour universal configuration in the leptonic sector.

Fourteen linear combinations of the coefficients associated to these 16 operators are constrained by the LEP EWPOs [43], leaving therefore only two linear combinations unconstrained. These two remaining unconstrained directions can be determined from the information contained in diboson production cross-sections [38, 44, 45] as well as by the Higgs production and decay measurements.

For completeness, the 14 linear combinations of bosonic and two-fermion Wilson coefficients which are constrained by the EWPOs measured at LEP are the following:

$$\begin{aligned}
&\frac{1}{4} g_1^2 \left( -2c_{\varphi\ell_1}^{(3)} - 2c_{\varphi\ell_2}^{(3)} + c_{\ell\ell} \right) - \frac{c_{\varphi D} g_w^2}{4} - g_1 g_w c_{\varphi WB}, \\
&c_{\varphi\ell_i}^{(3)} - f\left(-\frac{1}{2}, -1\right) + f\left(\frac{1}{2}, 0\right), \quad i = 1, 2, 3, \\
&f\left(-\frac{1}{2}, -1\right) - \frac{c_{\varphi\ell_i}^{(3)}}{2} - \frac{c_{\varphi\ell_i}^{(1)}}{2}, \quad i = 1, 2, 3, \\
&f(0, -1) - \frac{c_{\varphi e}}{2}, \quad f(0, -1) - \frac{c_{\varphi\mu}}{2}, \quad f(0, -1) - \frac{c_{\varphi\tau}}{2}, \\
&f\left(\frac{1}{2}, \frac{2}{3}\right) - \frac{c_{\varphi q}^{(-)}}{2}, \quad f\left(-\frac{1}{2}, -\frac{1}{3}\right) - \frac{c_{\varphi q}^{(-)}}{2} - c_{\varphi q}^{(3)} \\
&f\left(0, \frac{2}{3}\right) - \frac{c_{\varphi u}}{2}, \quad f\left(0, -\frac{1}{3}\right) - \frac{c_{\varphi d}}{2},
\end{aligned} \tag{2.5}$$

where the function $f$ is defined by:

$$f(T_3, Q) = \left( -\frac{c_{\varphi\ell_1}^{(3)}}{2} - \frac{c_{\varphi\ell_2}^{(3)}}{2} + \frac{c_{\ell\ell}}{4} - \frac{c_{\varphi D}}{4} \right) \left( \frac{g_1^2 Q}{g_w^2 - g_1^2} + T_3 \right) - c_{\varphi WB} \frac{Q g_1 g_w}{g_w^2 - g_1^2}, \tag{2.6}$$

where $g_1$ and $g_w$ are the corresponding electroweak couplings.

While in this work we do not explicitly include any EWPO data in the present fit, we still need to account for the information that they provide on the SMEFT parameter space. This is achieved by assuming that the EWPOs are precise enough to allow us to set the 14 linear combinations of Eq. (2.5) to zero in our fit. The remaining two degrees of freedom can be parametrized by, say, $c_{\varphi WB}$ and $c_{\varphi D}$, if the following replacements are made

$$\begin{pmatrix} c_{\varphi\ell_i}^{(3)} \\ c_{\varphi\ell_i}^{(1)} \\ c_{\varphi e/\mu/\tau} \\ c_{\varphi q}^{(-)} \\ c_{\varphi q}^{(3)} \\ c_{\varphi u} \\ c_{\varphi d} \\ c_{\ell\ell} \end{pmatrix} = \begin{pmatrix} -\frac{1}{t_W} & -\frac{1}{4t_W^2} \\ 0 & -\frac{1}{4} \\ 0 & -\frac{1}{2} \\ \frac{1}{t_W} & \frac{1}{4s_W^2} - \frac{1}{6} \\ -\frac{1}{t_W} & -\frac{1}{4t_W^2} \\ 0 & \frac{1}{3} \\ 0 & -\frac{1}{6} \\ 0 & 0 \end{pmatrix} \begin{pmatrix} c_{\varphi WB} \\ c_{\varphi D} \end{pmatrix}. \tag{2.7}$$

These relations will emulate the impact of LEP EWPOs in the fit, and allow us to produce a consistent fit without explicitly including the EWPOs.

Thanks to these 14 constraints, the 7 and 24 operators listed in Tables 2.1 and 2.2 respectively are then reduced to 17 independent degrees of freedom to be constrained by the LHC experimental data and the LEP diboson cross-sections. This allows us to set bounds on all operator coefficients listed in Tables 2.1 and 2.2. Of course, the bounds on the 16 operators of Eq. (2.4) will be highly correlated as indicated by Eq. (2.7). When presenting results for the independent DoFs, for example when evaluating the Fisher Information matrix or the principal components, we will select $c_{\varphi WB}$ and $c_{\varphi D}$, with the understanding that the replacements of Eq. (2.7) have been made. Note that it has been argued that the diboson channels at the LHC can in principle compete with EWPO [46, 47], which indicates that in an accurate fit one should always include the full set of EWPO constraints explicitly, as has been done, for example, in the combined Higgs/electroweak fits of [10, 21]. We however leave this option to future work.

**Four-fermion top quark operators.** We finally discuss the four-quark operators which involve the top quark fields and thus modify the production of top quarks at hadron colliders. The dimension-six four-fermion operators sensitive to top quarks can be classified into two categories: operators composed by four heavy quark fields (top and/or bottom quarks) and operators composed by two light and two heavy quark fields. The physical degrees of freedom corresponding to four-heavy and two-light-two-heavy interactions that we use in the present analysis are constructed in terms of suitable linear combinations of the four fermion coefficients in the Warsaw basis, whose corresponding operators are defined as

$$
\begin{aligned}
\mathcal{O}_{qq}^{1(ijkl)} &= (\bar{q}_i \gamma^\mu q_j)(\bar{q}_k \gamma_\mu q_l), \\
\mathcal{O}_{qq}^{3(ijkl)} &= (\bar{q}_i \gamma^\mu \tau^I q_j)(\bar{q}_k \gamma_\mu \tau^I q_l), \\
\mathcal{O}_{qu}^{1(ijkl)} &= (\bar{q}_i \gamma^\mu q_j)(\bar{u}_k \gamma_\mu u_l), \\
\mathcal{O}_{qu}^{8(ijkl)} &= (\bar{q}_i \gamma^\mu T^A q_j)(\bar{u}_k \gamma_\mu T^A u_l), \\
\mathcal{O}_{qd}^{1(ijkl)} &= (\bar{q}_i \gamma^\mu q_j)(\bar{d}_k \gamma_\mu d_l), \\
\mathcal{O}_{qd}^{8(ijkl)} &= (\bar{q}_i \gamma^\mu T^A q_j)(\bar{d}_k \gamma_\mu T^A d_l), \\
\mathcal{O}_{uu}^{(ijkl)} &= (\bar{u}_i \gamma^\mu u_j)(\bar{u}_k \gamma_\mu u_l), \\
\mathcal{O}_{ud}^{1(ijkl)} &= (\bar{u}_i \gamma^\mu u_j)(\bar{d}_k \gamma_\mu d_l), \\
\mathcal{O}_{ud}^{8(ijkl)} &= (\bar{u}_i \gamma^\mu T^A u_j)(\bar{d}_k \gamma_\mu T^A d_l),
\end{aligned}
\tag{2.8}
$$

where recall that $i, j, k, l$ are fermion generation indices. In Table 2.4 we provide the definition of all degrees of freedom that enter the fit in terms of the coefficients of Warsaw basis operators of Eq. (2.8). Within our flavour assumptions, the coefficients associated to different values of the generation indices $i$ ($i = 1, 2$) or $j$ ($j = 1, 2, 3$) will be the same.

Comparing with our previous EFT analysis of the top quark sector [7], in this work due to the different flavor assumptions several degrees of freedom that were used there as independent fit parameters are now absent. In particular, the coefficients $c_{QtQb}^1$ and $c_{QtQb}^8$ are set to zero. In addition, four-heavy operators that involve right-handed bottom quarks are not free parameters anymore. The correspondence between these four-heavy degrees of

| DoF | Definition (in Warsaw basis notation) |
|-----|----------------------------------------|
| $c_{QQ}^{1}$ | $2c_{qq}^{1(3333)} - \frac{2}{3}c_{qq}^{3(3333)}$ |
| $c_{QQ}^{8}$ | $8c_{qq}^{3(3333)}$ |
| $c_{Qt}^{1}$ | $c_{qu}^{1(3333)}$ |
| $c_{Qt}^{8}$ | $c_{qu}^{8(3333)}$ |
| $c_{tt}^{1}$ | $c_{uu}^{(3333)}$ |
| $c_{Qq}^{1,8}$ | $c_{qq}^{1(i33i)} + 3c_{qq}^{3(i33i)}$ |
| $c_{Qq}^{1,1}$ | $c_{qq}^{1(ii33)} + \frac{1}{6}c_{qq}^{1(i33i)} + \frac{1}{2}c_{qq}^{3(i33i)}$ |
| $c_{Qq}^{3,8}$ | $c_{qq}^{1(i33i)} - c_{qq}^{3(i33i)}$ |
| $c_{Qq}^{3,1}$ | $c_{qq}^{3(ii33)} + \frac{1}{6}(c_{qq}^{1(i33i)} - c_{qq}^{3(i33i)})$ |
| $c_{tq}^{8}$ | $c_{qu}^{8(ii33)}$ |
| $c_{tq}^{1}$ | $c_{qu}^{1(ii33)}$ |
| $c_{tu}^{8}$ | $2c_{uu}^{(i33i)}$ |
| $c_{tu}^{1}$ | $c_{uu}^{(ii33)} + \frac{1}{3}c_{uu}^{(i33i)}$ |
| $c_{Qu}^{8}$ | $c_{qu}^{8(33ii)}$ |
| $c_{Qu}^{1}$ | $c_{qu}^{1(33ii)}$ |
| $c_{td}^{8}$ | $c_{ud}^{8(33jj)}$ |
| $c_{td}^{1}$ | $c_{ud}^{1(33jj)}$ |
| $c_{Qd}^{8}$ | $c_{qd}^{8(33jj)}$ |
| $c_{Qd}^{1}$ | $c_{qd}^{1(33jj)}$ |

**Table 2.4.** Definition of the four-fermion degrees of freedom that enter into the fit in terms of the coefficients of Warsaw basis operators of Eq. (2.8). These DoFs are classified into four-heavy (upper) and two-light-two-heavy (bottom part) operators. The flavor index $i$ is either 1 or 2, and $j$ is either 1, 2 or 3: with our flavor assumptions, these coefficients will be the same regardless of the specific values that $i$ and $j$ take.

freedom from [7] and those of the present work is

$$c^1_{Qb} = c^1_{Qd}, \quad c^8_{Qb} = c^8_{Qd}, \quad c^1_{tb} = c^1_{td}, \quad c^8_{tb} = c^8_{td}. \tag{2.9}$$

Furthermore, we do not have $c^{1,8}_{Qb,tb}$ in the present fit anymore. These considerations explain why the 11 four-heavy operators of our previous study are now reduced to the 5 listed in Table 2.4.

All in all, in total we end up with 5 degrees of freedom involving four heavy quark fields and 14 involving two light and two heavy quark fields, for a total of 19 independent parameters at the fit level associated to four-quark operators. The more stringent flavour assumptions restricting the four-heavy operators imply that the constraints that we will obtain in the present fit for the four-fermion operators will be superior, thanks to these new constraints as well as the addition of the latest top production measurements from Run II of the LHC.

**Overview of the degrees of freedom.** We summarise in Table 2.5 the degrees of freedom considered in the present work. These are associated either to the Wilson coefficients of Warsaw-basis operators or to linear combinations of those. We categorize the DoFs into five disjoint classes, from top to bottom: four-quark (two-light-two-heavy), four-quark (four-heavy), four-lepton, two-fermion, and purely bosonic DoFs. We end up with 50 EFT coefficients that enter the theory predictions associated to the processes input to the fit, of which 36 are independent. The 16 DoFs displayed in the last columns are subject to the 14 constraints from the EWPOs listed in Eq. (2.7), leaving only 2 independent combinations to be constrained by the fit. When presenting results for the independent DoFs, for example when evaluating the Fisher Information matrix, we will select $c_{\varphi WB}$ and $c_{\varphi D}$, for illustration purposes. Then in Table 2.6 we indicate the notation that will be used to indicate the EFT coefficients listed in Table 2.5 in the subsequent sections, as well as in the released output files with the results of the global analysis, where again only two of the 16 EFT coefficients labelled in blue are independent fit parameters.

## 2.2  The top-philic scenario

The four-fermion operators defined in the previous section and listed in Table 2.4 correspond to a specific set of assumptions concerning the flavour structure of the UV-completion of the Standard Model. However, there exist well-motivated BSM scenarios that suggest further restrictions in the SMEFT parameter space spanned by these four-fermion operators. Therefore, phenomenological explorations of the SMEFT would benefit from comparing results obtained in different scenarios concerning the possible UV completion, from more restrictive to more general.

With this motivation, we have implemented a new feature in the `SMEFiT` analysis framework which allows one to implement arbitrary restrictions in the EFT parameter space, for example those motivated by specific BSM scenarios or existing constraints such as those from EWPO, as discussed in the previous section. As a proof of concept, here we will present results for the top-philic scenario introduced in [32]. This scenario is not constructed by imposing a specific flavour symmetry, but rather by assuming that new physics couples predominantly to the third-generation left-handed doublet, the third-generation right-handed up-type quark singlet, the gauge bosons, and the Higgs boson. In other words, that new physics interacts mostly with the top and bottom quarks as well as with the bosonic sector. The top-philic

| Class | $N_{\text{dof}}$ | 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) |

**Table 2.5.** Summary of the degrees of freedom considered in the present work. We categorize these DoFs into five disjoint classes: four-quark (two-light-two-heavy), four-quark (four-heavy), four-lepton, two-fermion, and purely bosonic DoFs. The 16 DoFs displayed in the last columns are subject to 14 constraints from the EWPOs, leaving only 2 independent combinations to be constrained by the fit.

| Class | DoF | Notation |
|-------|-----|----------|
| four-quark (two-light-two-heavy) | $c_{Qq}^{1,8},\ c_{Qq}^{1,1},\ c_{Qq}^{3,8},$ | c81qq, c11qq, c83qq, |
|  | $c_{Qq}^{3,1},\ c_{tq}^{8},\ c_{tq}^{1},$ | c13qq, c8qt, c1qt, |
|  | $c_{tu}^{8},\ c_{tu}^{1},\ c_{Qu}^{8},$ | c8ut, c1ut, c8qu, |
|  | $c_{Qu}^{1},\ c_{td}^{8},\ c_{td}^{1},$ | c1qu, c8dt, c1dt, |
|  | $c_{Qd}^{8},\ c_{Qd}^{1}$ | c8qd, c1qd |
| four-quark (four-heavy) | $c_{QQ}^{1},\ c_{QQ}^{8},\ c_{Qt}^{1},$ | cQQ1, cQQ8, cQt1, |
|  | $c_{Qt}^{8},\ c_{tt}^{1}$ | cQt8, ctt1 |
| four-lepton | $c_{\ell\ell}$ | cll |
| two-fermion (+ bosonic fields) | $c_{t\varphi},\ c_{tG},\ c_{b\varphi},$ | ctp, ctG, cbp, |
|  | $c_{c\varphi},\ c_{\tau\varphi},\ c_{tW},$ | ccp, ctap, ctW, |
|  | $c_{tZ},\ c_{\varphi Q}^{(3)},\ c_{\varphi Q}^{(-)},$ | ctZ, c3pQ3, cpQM, |
|  | $c_{\varphi t},\ c_{\varphi\ell_1}^{(1)},\ c_{\varphi\ell_1}^{(3)},$ | cpt, cpl1, c3pl1, |
|  | $c_{\varphi\ell_2}^{(1)},\ c_{\varphi\ell_2}^{(3)},\ c_{\varphi\ell_3}^{(1)},$ | cpl2, c3pl2, cpl3, |
|  | $c_{\varphi\ell_3}^{(3)},\ c_{\varphi e},\ c_{\varphi\mu},$ | c3pl3, cpe, cpmu, |
|  | $c_{\varphi\tau},\ c_{\varphi q}^{(3)},\ c_{\varphi q}^{(-)},$ | cpta, c3pq, cpqMi, |
|  | $c_{\varphi u},\ c_{\varphi d}$ | cpui, cpdi |
| Purely bosonic | $c_{\varphi G},\ c_{\varphi B},\ c_{\varphi W},$ | cpG, cpB, cpW, |
|  | $c_{\varphi d},\ c_{\varphi WB},\ c_{\varphi D},$ | cpd, cpWB, cpD, |
|  | $c_{WWW}$ | cWWW |

**Table 2.6.** The notation that will be used to indicate the EFT coefficients listed in Table 2.5 in the subsequent sections, as well as in the released output files with the results of the global analysis. Only two of the 16 EFT coefficients labelled in blue are independent fit parameters.

scenario satisfies the flavour assumptions that we are imposing in this work, but is based on a more restrictive theoretical assumption.

The restrictions in the EFT parameter space introduced by the top-philic assumption lead to a number of relations between the DoFs listed in Table 2.5. These relations are the following:

$$
\begin{aligned}
c_{QDW} &= c_{Qq}^{3,1}, \\
c_{QDB} &= 6c_{Qq}^{1,1} = \frac{3}{2}c_{Qu}^{1} = -3c_{Qd}^{1}, \\
c_{tDB} &= 6c_{tq}^{1} = \frac{3}{2}c_{tu}^{1} = -3c_{td}^{1}, \\
c_{QDG} &= c_{Qq}^{1,8} = c_{Qu}^{8} = c_{Qd}^{8}, \\
c_{tDG} &= c_{tq}^{8} = c_{tu}^{8} = c_{td}^{8}, \\
c_{Qq}^{3,8} &= 0,
\end{aligned}
\tag{2.10}
$$

which can be implemented as an additional restriction at the fitting level. Therefore, we now have 9 equations that relate a subset of the 14 two-heavy-two-light degrees of freedom listed in Table 2.5 among them, which leave 5 independent two-heavy-two-light degrees of freedom. The number of operators coupling the top quark with gauge bosons, as well as that of the four-heavy operators, is not modified. By comparing with Table 2.5, we see that in the top-philic scenario the EFT fit will constrain 41 DoFs, of which 27 are independent.

In principle, the top-philic assumption also implies non-trivial correlations between the light-fermion couplings to the gauge and Higgs bosons. However, following our strategy to include the EWPO, most of them are already set to zero, while the two remaining degrees of freedom are not affected. The same assumptions also imply that the light fermion Yukawa operator coefficients are proportional to the SM Yukawa couplings. As will be shown in Sect. 5, imposing the additional relations of the top-philic scenario leads to more stringent bounds on all the relevant Wilson coefficients, due to the fact that the same amount of experimental information is now used to constrain a significantly more limited parameter space.

## 2.3   Cross-section positivity

The constraint that physical cross-sections are (semi-)positive definite quantities can also be accounted for in global SMEFT analyses. This positivity requirement has different implications depending on whether the EFT expansion is considered up to either the linear or quadratic level.

The expansion up to linear terms, $\mathcal{O}(\Lambda^{-2})$, does not automatically lead to positive-definite cross sections, as in this case the new physics terms are generated by interference with the SM amplitudes, and their sign and size directly depend on the Wilson coefficients $c_i$. Imposing the positivity of the cross sections will therefore set (possibly one-sided) bounds on the Wilson coefficients. This can be easily implemented in the fitting procedure if helpful. In fact, we do not find the need to do so, since the fitted experimental data already leads to positive-definite cross-sections.

The expansion up to quadratic terms, $\mathcal{O}(\Lambda^{-4})$, i.e. specifically those coming from squaring the linearly expanded amplitudes, obviously automatically leads to positive definite cross sections. No constraints on the Wilson coefficients can therefore be obtained or need to be imposed. On the other hand, verifying the positivity of the cross section at the quadratic level

provides a sanity check that the theoretical calculation of the various contributions is correctly performed, also taking into account the MC generation uncertainties. The conditions that have to be met are simple to obtain. Consider the SMEFT Lagrangian

$$\mathcal{L} = \mathcal{L}_{\text{SM}} + \sum_{i=1}^{n_{\text{op}}} \frac{c_i}{\Lambda^2} \mathcal{O}_i \,, \tag{2.11}$$

where $\mathcal{O}_i$ stand for dimension-six operators and $c_i$ are the corresponding Wilson coefficients, which we assume to be real. Any observable calculated using this Lagrangian can be written as a quadratic form

$$
\begin{aligned}
\Sigma \;=\;& c_0^2 \Sigma_{00} \\
+\;& c_0 c_1 \Sigma_{01} + c_1 c_0 \Sigma_{10} + c_0 c_2 \Sigma_{02} + \ldots \\
+\;& c_1^2 \Sigma_{11} + c_1 c_2 \Sigma_{12} + c_1 c_3 \Sigma_{13} + \ldots \\
=\;& \boldsymbol{c}^T \cdot \boldsymbol{\Sigma} \cdot \boldsymbol{c}.
\end{aligned}
\tag{2.12}
$$

The first line corresponds to the SM contribution, where $c_0$ is an auxiliary coefficient that can be set to unity at the end. The second line corresponds to the linear $\mathcal{O}(\Lambda^{-2})$ EFT contributions, while the third line to the $\mathcal{O}(\Lambda^{-4})$ contributions. $\boldsymbol{\Sigma}$ is by construction a symmetric matrix.[1]

Given that a physical cross-section must be either positive or null, the matrix $\boldsymbol{\Sigma}$ must be semi-positive-definite. We can therefore use the Sylvester criterion that states that a symmetric matrix is semi-positive-definite if and only if all principal minors are greater or equal to zero. As a simple example, the constraints coming from the $2 \times 2$ minors are:

$$\left( \Sigma_{ii} \Sigma_{jj} - \Sigma_{ij}^2 \right) \geq 0 \,, \qquad i,j = 0, \ldots n_{\text{op}} \,. \tag{2.13}$$

We have verified that the Sylvester criterion, and Eq. (2.13) in particular, are satisfied by the EFT calculations used as input to the present analysis.

# 3 Experimental data and theoretical calculations

In this section we present the experimental measurements and the theoretical computations used to constrain the SMEFT operators introduced in Sect. 2. We focus in turn on each of the three groups of LHC processes that we consider in the current analysis: top quark, Higgs boson, and gauge boson pair production.

## 3.1 Top-quark production data

The top-quark production measurements included in this analysis belong to four different categories: inclusive top-quark pair production, top-quark pair production in association with vector bosons or heavy quarks, inclusive single top-quark production, and single top-quark production in association with vector bosons. In the following we present the datasets that belong to each of these categories. Top-quark pair production in association with a Higgs boson is discussed in Sect. 3.2.

---

[1]Note that with respect to the convention where $\sigma = \sigma_{\text{SM}} + \sum_{i=1}^{n_{\text{op}}} c_i \sigma_i + \sum_{i<j}^{n_{\text{op}}} c_i c_j \sigma_{ij}$, one has to account for factors of 2, *e.g.*, $\sigma_i = 2\Sigma_{i0}$ and $\sigma_{ij} = 2\Sigma_{ij}$ for $i \neq j$.

**Inclusive top-quark pair production.** The experimental measurements of inclusive top-quark pair production included in this analysis are summarised in Table 3.1. For each of them, we indicate the dataset label, the center of mass energy $\sqrt{s}$, the integrated luminosity $\mathcal{L}$, the final state or the specific production mechanism, the physical observable, the number of data points $n_{\text{dat}}$, and the publication reference. Measurements indicated with a (*) were not included in our earlier analysis [7].

The bulk of the measurements correspond to datasets already included in [7]: at 8 TeV, the ATLAS top-quark pair invariant mass distribution [48] and the CMS top-quark pair normalized invariant rapidity distribution [49], both in the lepton+jets final state, the CMS top-quark pair normalized invariant mass and rapidity two-dimensional distribution in the dilepton final state [50], and the ATLAS and CMS $W$ helicity fractions [51, 52]; at 13 TeV, the CMS top-quark pair invariant mass distributions in the lepton+jets and dilepton final states based on integrated luminosities of up to $\mathcal{L} = 35.8$ fb$^{-1}$ [53–55]. In addition to these, we now consider further top-quark pair invariant mass distributions: at 8 TeV, the ATLAS measurement in the dilepton final state [56]; at 13 TeV, and the ATLAS and CMS measurements, respectively in the lepton+jets and dilepton final states, corresponding to an integrated luminosity of $\mathcal{L} = 35.8$ fb$^{-1}$ [57, 58]. We also include top-quark pair charge asymmetry measurements: the ATLAS and CMS combined dataset at 8 TeV [59], and the ATLAS dataset at 13 TeV [60].

Although several distributions differential in various kinematic variables are available for the measurements presented in [49, 50, 53–58], only one of them can typically be included in the fit at a time. The reason is that experimental correlations between pairs of distributions are unknown: including more than one distribution at a time will therefore result in a double counting. An exception to this state of affairs is represented by the ATLAS measurement of [48], which is provided with the correlations among differential distributions. Unfortunately, they significantly deteriorate the fit quality when an analysis of all the available distributions is attempted, a fact that questions their reliability (see also [61, 62]). We therefore include only one distribution also in this case. In general, we include the invariant mass distribution $m_{t\bar{t}}$, whose high-energy tail is known to be particularly sensitive to deviations from the SM expectations. For [49] we include instead the invariant rapidity distribution as in our earlier analysis [7], due to difficulties in achieving an acceptable fit quality to $m_{t\bar{t}}$.

The additional top-quark pair measurements considered in this work do not expand the kinematic coverage in the EFT parameter space in comparison to those already included in [7]. Nevertheless, they provide additional weight for the inclusive top-quark pair differential distributions in the global fit, which are known to provide the dominant constraints on several of the EFT coefficients. All in all, we end up with $n_{\text{dat}} = 94$ data points in this category.

Additional sensitivity to EFT effects could be achieved by means of LHC Run-II measurements with an extended coverage in the invariant mass or transverse momentum. However, differential distributions based on luminosities larger than $\mathcal{L} \simeq 36$ fb$^{-1}$ are not available yet: the statistical precision of the data, and consequently their constraining power, remain therefore limited. For instance, the ATLAS fully hadronic final state measurement [63] is available, but it exhibits larger uncertainties than in the cleaner lepton+jets and dilepton final states. Furthermore, some measurements are not reconstructed at the parton level, as required in our analysis. This is the case of the ATLAS and CMS measurements at high top-quark transverse momentum [63, 64], that are based on reconstructing boosted topologies, and of the dilepton distributions from ATLAS [65], that are restricted to the particle level.

Concerning theoretical calculations, the SM cross-sections are evaluated at NLO using

| Dataset | $\sqrt{s}$, $\mathcal{L}$ | Info | Observables | $n_{\mathrm{dat}}$ | Ref |
|---|---|---|---|---|---|
| ATLAS_tt_8TeV_ljets | 8 TeV, 20.3 fb$^{-1}$ | lepton+jets | $d\sigma/dm_{t\bar{t}}$ | 7 | [48] |
| CMS_tt_8TeV_ljets | 8 TeV, 20.3 fb$^{-1}$ | lepton+jets | $1/\sigma d\sigma/dy_{t\bar{t}}$ | 10 | [49] |
| CMS_tt2D_8TeV_dilep | 8 TeV, 20.3 fb$^{-1}$ | dileptons | $1/\sigma d^2\sigma/dy_{t\bar{t}}dm_{t\bar{t}}$ | 16 | [50] |
| ATLAS_tt_8TeV_dilep (*) | 8 TeV, 20.3 fb$^{-1}$ | dileptons | $d\sigma/dm_{t\bar{t}}$ | 6 | [56] |
| CMS_tt_13TeV_ljets_2015 | 13 TeV, 2.3 fb$^{-1}$ | lepton+jets | $d\sigma/dm_{t\bar{t}}$ | 8 | [53] |
| CMS_tt_13TeV_dilep_2015 | 13 TeV, 2.1 fb$^{-1}$ | dileptons | $d\sigma/dm_{t\bar{t}}$ | 6 | [55] |
| CMS_tt_13TeV_ljets_2016 | 13 TeV, 35.8 fb$^{-1}$ | lepton+jets | $d\sigma/dm_{t\bar{t}}$ | 10 | [54] |
| CMS_tt_13TeV_dilep_2016 (*) | 13 TeV, 35.8 fb$^{-1}$ | dileptons | $d\sigma/dm_{t\bar{t}}$ | 7 | [58] |
| ATLAS_tt_13TeV_ljets_2016 (*) | 13 TeV, 35.8 fb$^{-1}$ | lepton+jets | $d\sigma/dm_{t\bar{t}}$ | 9 | [57] |
| ATLAS_WhelF_8TeV | 8 TeV, 20.3 fb$^{-1}$ | $W$ hel. fract | $F_0, F_L, F_R$ | 3 | [51] |
| CMS_WhelF_8TeV | 8 TeV, 20.3 fb$^{-1}$ | $W$ hel. fract | $F_0, F_L, F_R$ | 3 | [52] |
| ATLAS_CMS_tt_AC_8TeV (*) | 8 TeV, 20.3 fb$^{-1}$ | charge asymmetry | $A_C$ | 6 | [59] |
| ATLAS_tt_AC_13TeV (*) | 13 TeV, 139 fb$^{-1}$ | charge asymmetry | $A_C$ | 5 | [60] |

**Table 3.1.** The experimental measurements of inclusive top-quark pair production at the LHC considered in the present analysis. For each dataset we indicate the label, the center of mass energy $\sqrt{s}$, the integrated luminosity $\mathcal{L}$, the final state or the specific production mechanism, the physical observable, the number of data points $n_{\mathrm{dat}}$, and the publication reference. Measurements indicated with (*) were not included in [7]. We also include in this category the $W$ helicity fractions from top quark decay and the charge asymmetries.

MadGraph5_aMC@NLO [66] and supplemented with NNLO $K$-factors [67, 68]. The input PDF set is NNPDF3.1NNLO no-top [69], to avoid possible contamination between PDF and EFT effects.[2] The EFT cross-sections are evaluated with MadGraph5_aMC@NLO [66] combined with the SMEFT@NLO model [41]. Unless otherwise specified, the same EFT settings will be used also for the other processes considered in this analysis. Specifically, NLO QCD effects to the EFT corrections are accounted systematically whenever available.

**Associated top-quark pair production.** Table 3.2 lists, in the same format as Table 3.1, the experimental measurements for top quark pair production in association with heavy quarks or weak vector bosons. The dataset considered in [7] consisted of the CMS measurements of total cross-sections for $t\bar{t}t\bar{t}$ and $b\bar{b}b\bar{b}$ at 13 TeV [72, 73], and in the ATLAS and CMS measurements of inclusive $tW$ and $tZ$ production at 8 TeV and 13 TeV [74–77]. In the present analysis, we augment this dataset with the most updated measurements of total cross-sections for $t\bar{t}t\bar{t}$ and $t\bar{t}b\bar{b}$ production at 13 TeV: for $t\bar{t}b\bar{b}$, with the ATLAS and CMS

| Dataset | $\sqrt{s}, \mathcal{L}$ | Info | Observables | $N_{\text{dat}}$ | Ref |
|---|---|---|---|---|---|
| CMS_ttbb_13TeV | 13 TeV, 2.3 fb$^{-1}$ | total xsec | $\sigma_{\text{tot}}(t\bar{t}b\bar{b})$ | 1 | [72] |
| CMS_ttbb_13TeV_2016 (*) | 13 TeV, 35.9 fb$^{-1}$ | total xsec | $\sigma_{\text{tot}}(t\bar{t}b\bar{b})$ | 1 | [81] |
| ATLAS_ttbb_13TeV_2016 (*) | 13 TeV, 35.9 fb$^{-1}$ | total xsec | $\sigma_{\text{tot}}(t\bar{t}b\bar{b})$ | 1 | [80] |
| CMS_tttt_13TeV | 13 TeV, 35.9 fb$^{-1}$ | total xsec | $\sigma_{\text{tot}}(t\bar{t}t\bar{t})$ | 1 | [73] |
| CMS_tttt_13TeV_run2 (*) | 13 TeV, 137 fb$^{-1}$ | total xsec | $\sigma_{\text{tot}}(t\bar{t}t\bar{t})$ | 1 | [78] |
| ATLAS_tttt_13TeV_run2 (*) | 13 TeV, 137 fb$^{-1}$ | total xsec | $\sigma_{\text{tot}}(t\bar{t}t\bar{t})$ | 1 | [79] |
| CMS_ttZ_8TeV | 8 TeV, 19.5 fb$^{-1}$ | total xsec | $\sigma_{\text{tot}}(t\bar{t}Z)$ | 1 | [74] |
| CMS_ttZ_13TeV | 13 TeV, 35.9 fb$^{-1}$ | total xsec | $\sigma_{\text{tot}}(t\bar{t}Z)$ | 1 | [75] |
| CMS_ttZ_ptZ_13TeV (*) | 13 TeV, 77.5 fb$^{-1}$ | total xsec | $d\sigma(t\bar{t}Z)/dp_T^Z$ | 4 | [83] |
| ATLAS_ttZ_8TeV | 8 TeV, 20.3 fb$^{-1}$ | total xsec | $\sigma_{\text{tot}}(t\bar{t}Z)$ | 1 | [76] |
| ATLAS_ttZ_13TeV | 13 TeV, 3.2 fb$^{-1}$ | total xsec | $\sigma_{\text{tot}}(t\bar{t}Z)$ | 1 | [77] |
| ATLAS_ttZ_13TeV_2016 (*) | 13 TeV, 36 fb$^{-1}$ | total xsec | $\sigma_{\text{tot}}(t\bar{t}Z)$ | 1 | [82] |
| CMS_ttW_8_TeV | 8 TeV, 19.5 fb$^{-1}$ | total xsec | $\sigma_{\text{tot}}(t\bar{t}W)$ | 1 | [74] |
| CMS_ttW_13TeV | 13 TeV, 35.9 fb$^{-1}$ | total xsec | $\sigma_{\text{tot}}(t\bar{t}W)$ | 1 | [75] |
| ATLAS_ttW_8TeV | 8 TeV, 20.3 fb$^{-1}$ | total xsec | $\sigma_{\text{tot}}(t\bar{t}W)$ | 1 | [76] |
| ATLAS_ttW_13TeV | 13 TeV, 3.2 fb$^{-1}$ | total xsec | $\sigma_{\text{tot}}(t\bar{t}W)$ | 1 | [77] |
| ATLAS_ttW_13TeV_2016 (*) | 13 TeV, 36 fb$^{-1}$ | total xsec | $\sigma_{\text{tot}}(t\bar{t}W)$ | 1 | [82] |

**Table 3.2.** Same as Table 3.1, now for the production of top quark pairs in association with heavy quarks and with weak vector bosons.

measurements based on $\mathcal{L} = 137$ fb$^{-1}$ [78, 79]; for $\sigma_{\text{tot}}(t\bar{t}b\bar{b})$, with the ATLAS and CMS measurements based on $\mathcal{L} = 36$ fb$^{-1}$ [80, 81]. These measurements are comparatively more precise than the measurements already included in [7] thanks to the increased luminosity.

Concerning top-quark pair production in association with an electroweak gauge boson, we include here the ATLAS total cross-section measurements of $t\bar{t}W$ and $t\bar{t}Z$ based on $\mathcal{L} = 36$ fb$^{-1}$ [82], as well as the CMS differential measurements of $d\sigma(t\bar{t}Z)/dp_T^Z$ based on $\mathcal{L} = 78$ fb$^{-1}$ [83], which is the first differential measurement of $t\bar{t}V$ associated production presented at the LHC. We do not include the still preliminary ATLAS measurement of $\sigma_{\text{tot}}(t\bar{t}Z)$ based on $\mathcal{L} = 139$ fb$^{-1}$ [84]. The $t\bar{t}V$ measurements are especially useful to constrain EFT effects that modify the electroweak couplings of the top-quark. In total, we include $n_{\text{dat}} = 20$ data points in the category of $t\bar{t}$ associated production with heavy quark pairs or weak vector bosons.

Theoretical predictions are computed at NLO both in the SM and in the EFT. We use MCFM for the SM cross-sections and SMEFT@NLO for the EFT corrections, with NLO QCD effects accounted for exactly for the 2-fermion operators. The exception is the $p_T^Z$ distribution in $t\bar{t}Z$ events, for which MadGraph5_aMC@NLO is used instead to evaluate the SM cross-section at NLO.

| Dataset | $\sqrt{s}, \mathcal{L}$ | Info | Observables | $N_{\mathrm{dat}}$ | Ref |
|---|---|---|---|---|---|
| CMS_t_tch_8TeV_inc | **8 TeV, 19.7 fb$^{-1}$** | $t$-channel | $\sigma_{\mathrm{tot}}(t), \sigma_{\mathrm{tot}}(\bar{t})$ | 2 | [85] |
| ATLAS_t_tch_8TeV | **8 TeV, 20.2 fb$^{-1}$** | $t$-channel | $d\sigma(tq)/dy_t$ | 4 | [87] |
| CMS_t_tch_8TeV_dif | **8 TeV, 19.7 fb$^{-1}$** | $t$-channel | $d\sigma/d|y^{(t+\bar{t})}|$ | 6 | [86] |
| CMS_t_sch_8TeV | **8 TeV, 19.7 fb$^{-1}$** | $s$-channel | $\sigma_{\mathrm{tot}}(t+\bar{t})$ | 1 | [89] |
| ATLAS_t_sch_8TeV | **8 TeV, 20.3 fb$^{-1}$** | $s$-channel | $\sigma_{\mathrm{tot}}(t+\bar{t})$ | 1 | [88] |
| ATLAS_t_tch_13TeV | **13 TeV, 3.2 fb$^{-1}$** | $t$-channel | $\sigma_{\mathrm{tot}}(t), \sigma_{\mathrm{tot}}(\bar{t})$ | 2 | [90] |
| CMS_t_tch_13TeV_inc | **13 TeV, 2.2 fb$^{-1}$** | $t$-channel | $\sigma_{\mathrm{tot}}(t), \sigma_{\mathrm{tot}}(\bar{t})$ | 2 | [92] |
| CMS_t_tch_13TeV_dif | **13 TeV, 2.3 fb$^{-1}$** | $t$-channel | $d\sigma/d|y^{(t+\bar{t})}|$ | 4 | [91] |
| CMS_t_tch_13TeV_2016 (*) | **13 TeV, 35.9 fb$^{-1}$** | $t$-channel | $d\sigma/d|y^{(t)}|$ | 5 | [93] |

**Table 3.3.** Same as Table 3.1, now for inclusive single $t$ production both in the $t$- and the $s$-channels.

**Inclusive single top-quark production.** We now move to consider the inclusive production of single top-quarks, both in the $t$-channel and in the $s$-channel ($tW$ associated production is discussed separately below). Table 3.3 displays the information on the experimental data for these processes that is being considered in the present analysis. The dataset in this category that was already included in our previous analysis [7] consisted, at 8 TeV, of the $t$-channel total cross-sections and in the top-quark rapidity differential distributions from CMS [85, 86] and from ATLAS [87], and in the $s$-channel total cross-sections from ATLAS [88] and CMS [89]; at 13 TeV, in the $t$-channel total cross-sections and top-quark rapidity differential distributions from ATLAS [90] and CMS [91, 92].

Here we augment this dataset with one additional measurement, namely the CMS top-quark rapidity differential cross-section for $t$-channel single top-quark production at 13 TeV based on $\mathcal{L} = 35.9$ fb$^{-1}$ [93]. As customary, we consider the distribution reconstructed at parton level for consistency with the theoretical predictions. No differential measurements of single top-quark production based on the Run II dataset have been presented by ATLAS so far. Furthermore, while the ATLAS and CMS combination of total cross-sections for single top-quark production at 7 TeV and 8 TeV has been presented in [94], here we include instead the original individual measurements. We end up with $n_{\mathrm{dat}} = 27$ data points in this category.

The calculation of the SM and EFT cross-sections has been carried out with the same settings as for inclusive $t\bar{t}$ production. Note that for single top we work with a 5-flavour number scheme (5FNS) where the bottom quark is considered as massless, and thus enters the initial state of the reaction, see [95] for details. The NNLO QCD $K$-factors in the 5FNS are obtained from the calculation of [96].

**Associated single top-quark production with weak bosons.** Finally, in Table 3.4 we consider the experimental measurements on the associated production of single top-quarks together with a weak gauge boson $V$. The dataset in this category that was already part of our original analysis [7] consisted of the total inclusive cross-sections for $tW$ production by

| Dataset | $\sqrt{s}, \mathcal{L}$ | Info | Observables | $N_{\mathrm{dat}}$ | Ref |
|---|---|---|---|---|---|
| `ATLAS_tW_8TeV_inc` | **8 TeV, 20.2 fb$^{-1}$** | inclusive (dilepton) | $\sigma_{\mathrm{tot}}(tW)$ | 1 | [97] |
| `ATLAS_tW_inc_slep_8TeV (*)` | **8 TeV, 20.2 fb$^{-1}$** | inclusive (single lepton) | $\sigma_{\mathrm{tot}}(tW)$ | 1 | [103] |
| `CMS_tW_8TeV_inc` | **8 TeV, 19.7 fb$^{-1}$** | inclusive | $\sigma_{\mathrm{tot}}(tW)$ | 1 | [98] |
| `ATLAS_tW_inc_13TeV` | **13 TeV, 3.2 fb$^{-1}$** | inclusive | $\sigma_{\mathrm{tot}}(tW)$ | 1 | [99] |
| `CMS_tW_13TeV_inc` | **13 TeV, 35.9 fb$^{-1}$** | inclusive | $\sigma_{\mathrm{tot}}(tW)$ | 1 | [100] |
| `ATLAS_tZ_13TeV_inc` | **13 TeV, 36.1 fb$^{-1}$** | inclusive | $\sigma_{\mathrm{tot}}(tZq)$ | 1 | [102] |
| `ATLAS_tZ_13TeV_run2_inc (*)` | **13 TeV, 139.1 fb$^{-1}$** | inclusive | $\sigma_{\mathrm{fid}}(t\ell^+\ell^-q)$ | 1 | [104] |
| `CMS_tZ_13TeV_inc` | **13 TeV, 35.9 fb$^{-1}$** | inclusive | $\sigma_{\mathrm{fid}}(Wb\ell^+\ell^-q)$ | 1 | [101] |
| `CMS_tZ_13TeV_2016_inc (*)` | **13 TeV, 77.4 fb$^{-1}$** | inclusive | $\sigma_{\mathrm{fid}}(t\ell^+\ell^-q)$ | 1 | [105] |

**Table 3.4.** Same as Table 3.1, now for single top quark production in association with electroweak gauge bosons.

ATLAS and CMS at 8 TeV [97, 98] and at 13 TeV [99, 100], as well as in the ATLAS and CMS measurements of the $tZ$ total cross-sections at 13 TeV [101, 102], in the latter case restricted to the fiducial region in the $Wb\ell^+\ell^-q$ final state.

In addition to these datasets, we include here several new measurements of $tW$ and $tZ$ production. First of all, we include a new total cross-section measurement of $tW$ production by ATLAS at 8 TeV [103]. This measurement is carried out in the single lepton channel, and thus does not overlap with [97], which instead was obtained in the two leptons with one central $b$-jet channel. Then we include the ATLAS measurement of the fiducial cross-section for $tZ$ production [104] using the $t\ell^+\ell^-q$ final state (in the tri-lepton channel) based on the full Run II luminosity of $\mathcal{L} = 139$ fb$^{-1}$. In this analysis, the cross-section measurement differs from the background-only hypothesis (dominated by $t\bar{t}Z$ and dibosons) by more than five sigma and thus corresponds to an observation of this process. We also consider the corresponding measurement from CMS, where the observation of $tZ$ associated production is reported by reconstructing the $t\ell^+\ell^-q$ final state [105] based on a luminosity of $\mathcal{L} = 77.4$ fb$^{-1}$. No differential distributions for $tZ$ have been reported so far. The settings of the theoretical calculations for these $n_{\mathrm{dat}} = 9$ data points are the same as of [7].

In addition to these measurements, both ATLAS and CMS have measured differential distributions in $tW$ production at 13 TeV based on a luminosity of $\mathcal{L} = 35.9$ fb$^{-1}$ [106, 107]. However, these measurements are reported at the particle rather than at the parton level, and therefore they are not suitable for inclusion in the present analysis, which is restricted to top-quark level observables. We also note that CMS has reported on the EFT interpretation of the associated production of top-quarks, including with vector bosons, in an analysis based on a luminosity of $\mathcal{L} = 41.5$ fb$^{-1}$ [108].

Combining the four categories discussed above, the present analysis contains $n_{\mathrm{dat}} = 150$ top-quark cross-sections, to be compared with $n_{\mathrm{dat}} = 103$ in [7]. In Sect. 5.3 we will quantify

| Dataset | $\sqrt{s}$, $\mathcal{L}$ | Info | Observables | $n_{\text{dat}}$ | Ref. |
|---------|----------------|------|-------------|------|------|
| ATLAS_CMS_SSinc_RunI (*) | 7+8 TeV, 20 fb$^{-1}$ | Incl. $\mu_i^f$ | $gg$F, VBF, $Vh$, $t\bar{t}h$ $h \to \gamma\gamma, VV, \tau\tau, b\bar{b}$ | 20 | [109] |
| ATLAS_SSinc_RunI (*) | 8 TeV, 20 fb$^{-1}$ | Incl. $\mu_i^f$ | $h \to Z\gamma, \mu\mu$ | 2 | [110] |
| ATLAS_SSinc_RunII (*) | 13 TeV, 80 fb$^{-1}$ | Incl. $\mu_i^f$ | $gg$F, VBF, $Vh$, $t\bar{t}h$ $h \to \gamma\gamma, WW, ZZ, \tau\tau, b\bar{b}$ | 16 | [111] |
| CMS_SSinc_RunII (*) | 13 TeV, 36.9 fb$^{-1}$ | Incl. $\mu_i^f$ | $gg$F, VBF, $Wh$, $Zh$ $t\bar{t}h$ $h \to \gamma\gamma, WW, ZZ, \tau\tau, b\bar{b}$ | 24 | [112] |

**Table 3.5.** Same as Table 3.1 now for the measurements of the inclusive signal strenghts, Eq. (3.2), in Higgs production and decay from the LHC Run I and Run II.

the impact of the new top-quark measurements by comparing two fits, one based on the dataset of [7] and one based on the extended top-quark dataset included here.

## 3.2 Higgs production and decay

We now turn to the Higgs boson production and decay measurements. We consider first inclusive cross-section measurements, presented as signal strengths normalised to the SM predictions, and then differential distributions and STXS measurements.

**Signal strengths.** First of all, we consider the inclusive Higgs boson production signal strengths $\mu_i^f$ measured by ATLAS and CMS from LHC Run I and Run II. These signal strengths are defined for each combination of production and decay channels in terms of cross-section $\sigma_i$ and the branching fraction $B_f$ as

$$\mu_i^f \equiv \frac{\sigma_i \times B_f}{(\sigma_i)_{\text{SM}} \times (B_f)_{\text{SM}}} = \mu_i \cdot \mu^f = \left(\frac{\sigma_i}{(\sigma_i)_{\text{SM}}}\right)\left(\frac{B_f}{(B_f)_{\text{SM}}}\right), \qquad (3.1)$$

that is, as the ratio of the experimentally measured production cross-sections in specific decay channels to the corresponding (state-of-the-art) SM predictions. These inclusive signal strengths can also be expressed as

$$\mu_i^f = \left(\frac{\sigma_i}{(\sigma_i)_{\text{SM}}}\right)\left(\frac{\Gamma(h \to f)}{\Gamma(h \to f)|_{\text{SM}}}\right)\left(\frac{\Gamma(h \to \text{all})}{\Gamma(h \to \text{all})|_{\text{SM}}}\right)^{-1}, \qquad (3.2)$$

in terms of the partial and total decay widths. The measurements of signal strengths that we consider in the present analysis are collected in Table 3.5. In contrast to the differential distributions and STXS discussed below, these signal strengths are typically extrapolated to the full phase space and do not include selection or acceptance cuts.

For the LHC Run I, we take into account the inclusive measurements of Higgs boson production and decay rates from the ATLAS and CMS combination based on the full 7 and 8 TeV datasets [109]. Specifically, we include the 20 measurements presented in Table 8 of [109]. These measurements correspond to five different production channels ($gg$F, VBF, $Wh$, $Zh$, $tth$) for five final states ($\gamma\gamma$, $ZZ$, $WW$, $\tau\tau$, $b\bar{b}$), excluding those combinations that

are either not measured with a meaningful precision or not measured at all. We account for the experimental correlations between the measured signal strengths using the information provided in [109]. In addition to these ATLAS+CMS combination results from Run I, we also include two more signal strengths measurements from Run I, namely the ATLAS constraints on the $Z\gamma$ and $\mu\mu$ decays from [110].

For the LHC Run II, we consider the ATLAS measurement of signal strengths corresponding to an integrated luminosity of $\mathcal{L} = 80$ fb$^{-1}$ [111], and the CMS measurement corresponding to an integrated luminosity of $\mathcal{L} = 35.9$ fb$^{-1}$ [112]. As in the case of the Run I signal strengths, we keep into account correlations between the various production and final state combinations. The ATLAS combination contains 16 signal strengths for the $gg$F, VBF, $Vh$ and $t\bar{t}h$ production channels and the $\gamma\gamma$, $ZZ$, $WW$, $\tau\tau$ and $b\bar{b}$ final states. As in the case of Run I, measurements are sometimes not available for all final states for a given production channel, for example the $h \to b\bar{b}$ decay is not available for $gg$F while $\tau\tau$ is not provided in the case of $Vh$ associate production. The CMS analysis contains 24 signal strengths measurements in the $gg$F, VBF, $Wh$, $Zh$, and $t\bar{t}h$ production channels for the same final states as in the ATLAS case. Results for the $WW$, $ZZ$ and $\gamma\gamma$ final states are available for all production channels, while for the other final states, $\mu\mu$, $\tau\tau$, and $b\bar{b}$, signal strength measurements are only available for specific production channels. In total, we have $n_{\mathrm{dat}} = 62$ measurements of Higgs inclusive signal strengths from Runs I and II.

Concerning the theoretical calculations corresponding to these measurements, the SM production cross-sections and decay branching fractions are obtained from the associated experimental publications. In turn, these represent the most updated available predictions, and are provided in the LHC Higgs Cross-Section Working Group (HXSWG) reports [113–115]. As in the case of top-quark production processes, EFT calculations are obtained at NLO QCD using `MadGraph5_aMC@NLO` [66] with the `SMEFT@NLO` model. Additional details about the implementation of EFT corrections to the Higgs signal strengths are provided in App. B.

**Differential distributions and STXS.** Table 3.6 summarizes the experimental measurements of differential distributions and STXS for Higgs boson production and decay at the LHC considered in the present analysis. Whenever one has a potential double counting between a signal strength measurement and the corresponding differential distribution or STXS measurement, we always select the latter, which provides more information on the EFT parameter space due to its enhanced kinematical sensitivity.

To being with, we consider the ATLAS and CMS differential distributions in the Higgs boson kinematic variables obtained from the combination of the $h \to \gamma\gamma$, $h \to ZZ$, and (in the CMS case) $h \to b\bar{b}$ final states at 13 TeV based on $\mathcal{L} = 36$ fb$^{-1}$ [116, 117]. Specifically, we consider the differential distributions in the Higgs boson transverse momentum $p_T^h$. These distributions are particularly sensitive probes of new physics, for instance through new particles circulating in the gluon-fusion loop.

We also include the ATLAS measurement of the associated production of Higgs bosons, $Vh$, in the $h \to b\bar{b}$ final state at 13 TeV [118]. These measurements, performed in kinematic fiducial volumes defined in the simplified template cross-section framework, correspond to an integrated luminosity of $\mathcal{L} = 79.8$ fb$^{-1}$. Specifically, here we include the data corresponding to the 5-POI (parameters of interest) category, which refers to three cross-sections for $Zh$ production in the bins $75 < p_T^Z < 150$ GeV, $150 < p_T^Z < 250$ GeV, and $p_T^Z > 250$ GeV, and two cross-sections for $Wh$ production, one for $150 < p_T^W < 250$ GeV and the other for

| Dataset | $\sqrt{s}, \mathcal{L}$ | Info | Observables | $N_{\text{dat}}$ | Ref |
|---|---|---|---|---|---|
| CMS_H_13TeV_2015 (*) | 13 TeV, 35.9 fb$^{-1}$ | $gg$F, VBF, $Vh$, $t\bar{t}h$ <br> $h \to ZZ, \gamma\gamma, b\bar{b}$ | $d\sigma/dp_T^h$ | 9 | [116] |
| ATLAS_ggF_13TeV_2015 (*) | 13 TeV, 36.1 fb$^{-1}$ | $gg$F, VBF, $Vh$, $t\bar{t}h$ <br> $h \to ZZ(\to 4l)$ | $d\sigma/dp_T^h$ | 9 | [117] |
| ATLAS_Vh_hbb_13TeV (*) | 13 TeV, 79.8 fb$^{-1}$ | $Wh$, $Zh$ | $d\sigma^{(\text{fid})}/dp_T^W$ <br> $d\sigma^{(\text{fid})}/dp_T^Z$ | 2 <br> 3 | [118] |
| ATLAS_ggF_ZZ_13TeV (*) | 13 TeV, 79.8 fb$^{-1}$ | $gg$F, $h \to ZZ$ | $\sigma_{\text{ggF}}(p_T^h, N_{\text{jets}})$ | 6 | [111] |
| CMS_ggF_aa_13TeV (*) | 13 TeV, 77.4 fb$^{-1}$ | $gg$F, $h \to \gamma\gamma$ | $\sigma_{\text{ggF}}(p_T^h, N_{\text{jets}})$ | 6 | [119] |

**Table 3.6.** Same as Table 3.1 for differential distributions and STXS for Higgs production and decay.

$p_T^W > 250$ GeV. Gauge bosons are reconstructed by means of their leptonic decays.

Then we also include selected differential measurements presented in the ATLAS Run II Higgs combination paper [111]. Specifically, we include the measurements of Higgs production in gluon fusion, $gg \to h$, in different bins of $p_T^h$ and in the number of jets in the event. These measurements are presented as $\sigma_i \times B_{ZZ}/B_{ZZ}^{(\text{SM})}$, since the $ZZ$ branching fraction is used to normalise the data. We include the 0-jet cross-section, the 1-jet cross-section for $p_T^h < 60$ GeV, $60 \leq p_T^h \leq 120$ GeV, and $120 \leq p_T^h \leq 200$ GeV, and the $\leq 1$ jet and $\leq 2$ jet cross-sections for $p_T^h \geq 200$ GeV and $p_T^h < 200$ GeV respectively.

Furthermore, we consider the differential Higgs boson production measurements presented by CMS at 13 TeV based on an integrated luminosity of $\mathcal{L} = 77.4$ fb$^{-1}$ and corresponding to the final state $\gamma\gamma$ [119]. The STXS measurements associated to different final-state topologies and kinematic values such as $p_T^h$ are presented. These inclusive measurements are dominated by the gluon-fusion production channel. Note that the CMS diphoton measurement of [119] supersedes [120], which was based on the 2016 dataset only.

Whenever available, the information on the experimental correlated systematic uncertainties is included. As mentioned above, the SM theoretical predictions are taken from the HXSWG reports [113–115]. In total, we include $n_{\text{dat}} = 35$ measurements of differential cross-sections and STXS on Higgs production and decay from the LHC Run II.

We note that additional Higgs production and decay measurements have been recently presented by ATLAS and CMS based on the full Run II luminosity of $\mathcal{L} = 139$ fb$^{-1}$. Two examples of these are the CMS measurement of the $p_T^h$ distribution in the $h \to WW$ fully leptonic final state [121] and the updated ATLAS measurement of $Vh$ associated production in the $b\bar{b}$ final state [122]. These measurements are however not expected to modify significantly the results of the present analysis, since the constraints they provide on the EFT parameter space are already covered by other measurements, and their inclusion is left for future work.

| Dataset | $\sqrt{s}$, $\mathcal{L}$ | Info | Observables | $N_{\mathrm{dat}}$ | Ref |
|---|---|---|---|---|---|
| LEP2_WW_diff (*) | [182, 296] GeV | LEP-2 comb | $d^2\sigma(WW)/dE_{\mathrm{cm}}d\cos\theta_W$ | 40 | [123] |
| ATLAS_WZ_13TeV_2016 (*) | 13 TeV, 36.1 fb$^{-1}$ | fully leptonic | $d\sigma^{(\mathrm{fid})}/dm_T^{WZ}$ | 6 | [124] |
| ATLAS_WW_13TeV_2016 (*) | 13 TeV, 36.1 fb$^{-1}$ | fully leptonic | $d\sigma^{(\mathrm{fid})}/dm_{e\mu}$ | 13 | [125] |
| CMS_WZ_13TeV_2016 (*) | 13 TeV, 35.9 fb$^{-1}$ | fully leptonic | $d\sigma^{(\mathrm{fid})}/dp_T^Z$ | 11 | [126] |

**Table 3.7.** Same as Table 3.1 for the differential distributions of gauge boson pair production from LEP-2 and the LHC.

## 3.3 Diboson production from LEP and the LHC

We complement the constraints provided by the Higgs data with those provided by diboson production cross-sections measured by LEP and the LHC. The dataset is summarised in Table 3.7. To begin with, we consider the LEP-2 legacy measurements of $WW$ production [123]. Specifically, we include the cross-sections differential in $\cos\theta_W$ in four different bins in the center of mass energy, from $\sqrt{s} = 182$ GeV up to $\sqrt{s} = 206$ GeV. Here $\theta_W$ is defined as the polar angle of the produced $W^-$ boson with respect to the incoming electron beam direction. Each set of bins with a different center-of-mass energy correspond to a different integrated luminosity, ranging between $\mathcal{L} = 163.9$ pb$^{-1}$ and 630.5 pb$^{-1}$. For each value of $\sqrt{s}$, there are 10 bins in $\cos\theta_W$, adding up to a total of $n_{\mathrm{dat}} = 40$ data points. The theoretical calculations of the SM predictions, which include higher-order electroweak but not NLO QCD corrections, are also taken from [123]. For this process, the squared terms in the EFT proportional to $c_i c_j/\Lambda^{-4}$ are small and will be neglected.

Concerning the LHC datasets, we include measurements of the differential distributions for $W^\pm Z$ production at 13 TeV from ATLAS [127] and CMS [126] based on a luminosity of $\mathcal{L} = 36.1$ fb$^{-1}$. In both cases, the two gauge bosons are reconstructed by means of the fully leptonic final state, whereby events of the type $WZ \to \ell^+\ell^-\ell^{(\prime)\pm}$ are selected. The different leptonic final states are then combined into an inclusive measurement. For the ATLAS measurement three fiducial distributions are presented, respectively differential in $p_T^W$, $p_T^Z$ and $m_T^{WZ}$. As indicated in Table 3.7, in this analysis, our baseline choice will be to include the $m_T^{WZ}$ distribution, which extends up to transverse masses of $m_T^{WZ} = 600$ GeV. In the case of the corresponding CMS measurement, normalised differential distributions in $p_T^Z$, $m_{WZ}$, $p_T^W$, and $p_T^{\mathrm{jet,lead}}$ are available. Here the baseline will be the $p_T^Z$ distribution.

In addition to these measurements, we also consider the differential distributions for $WW$ production from ATLAS at 13 TeV based on a luminosity of $\mathcal{L} = 36.1$ fb$^{-1}$ [125]. Events are selected by requiring one electron and one muon in the final state, corresponding to the decay mode $WW \to e^\pm\nu\mu^\pm\nu$. Several differential distributions in the fiducial region are provided, including $m_{e\mu}$, $p_T^{e\mu}$ and $|y_{e\mu}|$. Here our baseline choice will be the $m_{e\mu}$ distribution, the invariant mass of the dilepton system, which reaches values of up to $m_{e\mu} \simeq 1$ TeV. The total number of data points in the LHC diboson category is $n_{\mathrm{dat}} = 30$.

Other diboson measurements from the LHC have been presented but their EFT interpretation is left for future work. For instance, the data for the CMS differential distributions of $WW$ production at 13 TeV based on $\mathcal{L} = 36.1$ fb$^{-1}$ [128] is still preliminary. ATLAS has presented recent measurements of the differential cross-sections in four-lepton events in 13 TeV

| Category | Processes | $n_{\mathrm{dat}}$ |
|---|---|---|
| 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** |

**Table 3.8.** The number of data points $n_{\mathrm{dat}}$ in our baseline dataset for each of the categories of processes considered here.

based on $\mathcal{L} = 139$ fb$^{-1}$ [129], though here the measured distributions receive contributions from single $Z$ and Higgs boson production, in addition to those from $ZZ$ production.

The theoretical predictions for the SM cross-sections of these LHC diboson processes are accurate to NNLO QCD and were computed with `MATRIX` [130]. The EFT contributions for this process include NLO QCD corrections and take into account the constraints from Eq. (2.5) to express the calculation in terms of only three Wilson coefficients, one being the triple-gauge operator $c_{WWW}$ and the other two the purely bosonic coefficients $c_{\varphi D}$ and $c_{\varphi WB}$. This choice is ultimately arbitrary and has no physical implications; any other two coefficients out of Eq. (2.4) would lead to the same results. Its only motivation is to facilitate the event generation of the diboson processes.

## 3.4 Dataset and theory overview and EFT sensitivity

We conclude this section by presenting an overview of the datasets considered (and of the corresponding theoretical calculations), summarizing their dependence on the EFT coefficients defined in Sect. 2, and quantifying the sensitivity that each process has on these coefficients by means of information geometry.

**Dataset overview.** In Table 3.8 we summarise the number of data points in our baseline dataset for each of the data categories and processes considered in this analysis, as well as the total per category and the overall total. We include 150, 97, and 70 cross-sections from top-quark production, Higgs boson production and decay, and diboson production cross-sections from LEP and the LHC respectively in the baseline dataset, for a total of 317 cross-section measurements.

**Overview of theoretical calculations.** Table 3.9 displays a summary of the theoretical calculations used for the description various datasets included in the present analysis. We indicate, for both the SM and the SMEFT contributions to the cross-sections, the perturbative accuracy and the codes used to produce the corresponding theoretical predictions.

**Dependence on the EFT coefficients.** In order to interpret the results of the global EFT analyses which will be presented in Sect. 5, it is useful to collect the dependence of the various datasets described in this section with respect to the degrees of freedom defined in Sect. 2. Table 3.10 indicates which EFT coefficients contribute to the theoretical description of each of the processes considered in this analysis. Recall that the 16 coefficients listed in Eq. (2.4) are related among them by the EWPO relations, and that only two of them are independent.

In Table 3.10 we display from top to bottom the coefficients associated to the two-light-two-heavy, four-heavy, four-lepton, two-fermion plus bosonic, and purely bosonic dimension-six operators. The Higgs measurements are separated between the Run I and Run II datasets, and in the latter case also between signal strengths and differential distributions and STXS. A check mark outside (inside) brackets indicates that a given process constrains the corresponding coefficients starting at $\mathcal{O}(\Lambda^{-2})$ ($\mathcal{O}(\Lambda^{-4})$) at LO. Entries labelled with (b) indicate that the sensitivity to the associated coefficients enters via bottom-initiated processes, which arise due to contributions from the $b$-PDF in the 5FNS adopted here.

Several observations can be drawn from this table. First of all, we observe that the four-heavy coefficients are constrained only by the $t\bar{t}Q\bar{Q}$ production data, either $t\bar{t}t\bar{t}$ or $t\bar{t}b\bar{b}$. Such measurements also depend on the 2-light-2-heavy operators, as well as on $c_{tG}$, although in practice this correlation is small. Furthermore, the four-heavy coefficients are essentially left undetermined at $\mathcal{O}(\Lambda^{-2})$, and can only be meaningfully constrained only the quadratic corrections are accounted for. One can also note how the two-light-two-heavy operators are constrained by top-quark pair production (inclusive and in association with vector bosons) as well as by the $t\bar{t}h$ production measurements. As will be shown below, by far the dominant constraints on these coefficients arise from the differential distributions in inclusive top quark pair production.

Concerning the two-fermion operators, most of them are constrained both by top and by Higgs production process. Recall that the top and Higgs sectors are connected, among others, by means of the gluon-fusion production process (with its virtual top-quark loop) as well as by $t\bar{t}h$ associated production. In particular, we note that $c_{t\varphi}$, which modifies the top Yukawa coupling, is constrained by these Higgs production measurements. The purely bosonic operators exhibit sensitivity only to Higgs and diboson processes, since these do not affect the properties of top quarks. The diboson data is uniquely sensitive to the triple-gauge coefficient $c_{WWW}$, which modifies the triple (and quartic) electroweak gauge couplings, as well as to $c_{\varphi D}$ and $c_{\varphi WB}$, which are also constrained by Higgs data.

**The Fisher matrix and information geometry.** The information presented in Table 3.10 does not allow one to compare the sensitivity brought in by different datasets on a given EFT coefficient. To achieve this, here we adopt the ideas underlying information geometry [131] and define the Fisher information matrix $I_{ij}$ as

$$I_{ij}(\boldsymbol{c}) = -\mathrm{E}\left[\frac{\partial^2 \ln f(\boldsymbol{\sigma}_{\mathrm{exp}}|\boldsymbol{c})}{\partial c_i \partial c_j}\right], \qquad i,j = 1,\ldots,n_{\mathrm{op}}, \tag{3.3}$$

| Category | Process | SM | Code/Ref | SMEFT |
|---|---|---|---|---|
| Top quark production | $t\bar{t}$ (incl) | NNLO QCD | `MG5_aMC` NLO + NNLO $K$-fact | NLO QCD |
| | $t\bar{t} + V$ | NLO QCD | `MG5_aMC` NLO | LO QCD + NLO SM $K$-fact |
| | single-$t$ (incl) | NNLO QCD | `MG5_aMC` NLO + NNLO $K$-fact | NLO QCD |
| | $t + V$ | NLO QCD | `MG5_aMC` NLO | LO QCD + NLO SM $K$-fact |
| | $t\bar{t}t\bar{t}$, $t\bar{b}t\bar{b}$ | NLO QCD | `MG5_aMC` NLO | LO QCD + NLO SM $K$-fact |
| Higgs production and decay | $gg \to h$ | NNLO QCD + NLO EW | HXSWG | NLO QCD |
| | VBF | NNLO QCD + NLO EW | HXSWG | LO QCD |
| | $h + V$ | NNLO QCD + NLO EW | HXSWG | NLO QCD |
| | $ht\bar{t}$ | NNLO QCD + NLO EW | HXSWG | NLO QCD |
| | $h \to X$ | NNLO QCD + NLO EW | HXSWG | NLO QCD ($X = b\bar{b}$) LO QCD ($X \neq b\bar{b}$) |
| Diboson production | $e^+e^- \to W^+W^-$ | NNLO QCD + NLO EW | LEP EWWG | LO QCD |
| | $pp \to VV'$ | NNLO QCD | `MG5_aMC` NLO + NNLO $K$-fact | NLO QCD |

**Table 3.9.** Summary of the theoretical calculations used for the description various datasets included in the present analysis. We indicate, for both the SM and the SMEFT contributions to the cross-sections, the perturbative accuracy and the codes used to produce the corresponding theoretical predictions. In all cases, the EFT cross-sections are evaluated with `MG5_aMC` interfaced to `SMEFT@NLO`. See the text for more details and the corresponding references.

| Class | DoF | $t\bar{t}$ | $t\bar{t}V$ | $t$ | $tV$ | $t\bar{t}Q\bar{Q}$ | $h$ ($\mu_i^f$, Run-I) | $h$ ($\mu_i^f$, Run-II) | $h$ (STXS, Run-II) | $VV$ |
|---|---|---|---|---|---|---|---|---|---|---|
| 2-heavy-2-light | $c_{Qq}^{1,8}$ | ✓ | ✓ | | | ✓ | ✓ | ✓ | ✓ | |
| | $c_{Qq}^{1,1}$ | (✓) | (✓) | | | ✓ | (✓) | (✓) | (✓) | |
| | $c_{Qq}^{3,8}$ | ✓ | ✓ | (✓) | (✓) | ✓ | ✓ | ✓ | ✓ | |
| | $c_{Qq}^{3,1}$ | (✓) | (✓) | ✓ | ✓ | ✓ | (✓) | (✓) | (✓) | |
| | $c_{tq}^{8}$ | ✓ | ✓ | | | ✓ | ✓ | ✓ | ✓ | |
| | $c_{tq}^{1}$ | (✓) | (✓) | | | ✓ | (✓) | (✓) | (✓) | |
| | $c_{tu}^{8}$ | ✓ | ✓ | | | ✓ | ✓ | ✓ | ✓ | |
| | $c_{tu}^{1}$ | (✓) | (✓) | | | ✓ | (✓) | (✓) | (✓) | |
| | $c_{Qu}^{8}$ | ✓ | ✓ | | | ✓ | ✓ | ✓ | ✓ | |
| | $c_{Qu}^{1}$ | (✓) | (✓) | | | ✓ | (✓) | (✓) | (✓) | |
| | $c_{td}^{8}$ | ✓ | ✓ | | | ✓ | ✓ | ✓ | ✓ | |
| | $c_{td}^{1}$ | (✓) | (✓) | | | ✓ | (✓) | (✓) | (✓) | |
| | $c_{Qd}^{8}$ | ✓ | ✓ | | | ✓ | ✓ | ✓ | ✓ | |
| | $c_{Qd}^{1}$ | (✓) | (✓) | | | ✓ | (✓) | (✓) | (✓) | |
| 4-heavy | $c_{QQ}^{1}$ | | | | | ✓ | | | | |
| | $c_{QQ}^{8}$ | | | | | ✓ | | | | |
| | $c_{Qt}^{1}$ | | | | | ✓ | | | | |
| | $c_{Qt}^{8}$ | | | | | ✓ | | | | |
| | $c_{tt}^{1}$ | | | | | ✓ | | | | |
| 4-lepton | $c_{ll}$ | | | ✓ | ✓ | | ✓ | ✓ | ✓ | ✓ |
| 2-fermion +bosonic | $c_{t\varphi}$ | | | | | | ✓ | ✓ | ✓ | |
| | $c_{tG}$ | ✓ | ✓ | | | ✓ | ✓ | ✓ | ✓ | |
| | $c_{b\varphi}$ | | | | | | ✓ | ✓ | ✓(b) | |
| | $c_{c\varphi}$ | | | | | | ✓ | ✓ | | |
| | $c_{\tau\varphi}$ | | | | | | ✓ | ✓ | | |
| | $c_{tW}$ | ✓ | | ✓ | ✓ | | ✓ | ✓ | | |
| | $c_{tZ}$ | | ✓ | | ✓ | | ✓ | ✓ | | |
| | $c_{\varphi Q}^{(3)}$ | | ✓(b) | ✓ | ✓ | | ✓(b) | ✓(b) | ✓(b) | |
| | $c_{\varphi Q}^{(-)}$ | | ✓ | | ✓ | | ✓ | ✓ | ✓(b) | |
| | $c_{\varphi t}$ | | ✓ | | ✓ | | ✓ | ✓ | | |
| | $c_{\varphi l_i}^{(1)}$ | | | | | | ✓ | ✓ | | ✓ |
| | $c_{\varphi l_i}^{(3)}$ | | | ✓ | ✓ | | ✓ | ✓ | ✓ | ✓ |
| | $c_{\varphi e}$ | | | | | | ✓ | ✓ | | ✓ |
| | $c_{\varphi \mu}$ | | | | | | ✓ | ✓ | | |
| | $c_{\varphi \tau}$ | | | | | | ✓ | ✓ | | |
| | $c_{\varphi q}^{(3)}$ | | ✓ | ✓ | ✓ | | ✓ | ✓ | ✓ | ✓ |
| | $c_{\varphi q}^{(-)}$ | | ✓ | | ✓ | | ✓ | ✓ | ✓ | ✓ |
| | $c_{\varphi u}$ | | ✓ | | ✓ | | ✓ | ✓ | ✓ | ✓ |
| | $c_{\varphi d}$ | | ✓ | | ✓ | | ✓ | ✓ | ✓ | ✓ |
| purely bosonic | $c_{\varphi G}$ | | | | | | ✓ | ✓ | ✓ | |
| | $c_{\varphi B}$ | | | | | | ✓ | ✓ | ✓ | |
| | $c_{\varphi W}$ | | | | | | ✓ | ✓ | ✓ | |
| | $c_{\varphi d}$ | | | | | | ✓ | ✓ | ✓ | |
| | $c_{\varphi D}$ | | | ✓ | ✓ | | ✓ | ✓ | ✓ | ✓ |
| | $c_{\varphi WB}$ | | | ✓ | ✓ | | ✓ | ✓ | ✓ | ✓ |
| | $c_{WWW}$ | | | | | | | | | ✓ |

**Table 3.10.** Overview indicating which EFT coefficients contribute to the theoretical description of each of the processes considered in this global analysis.

where E[ ] indicates the expectation value and $f(\boldsymbol{\sigma}_{\mathrm{exp}}|\boldsymbol{c})$ indicates the relation between a set of experimental measurements and the assumed true values of the EFT coefficients $\boldsymbol{c}$. The covariance matrix in the EFT parameter space, $C_{ij}(\boldsymbol{c})$, is then bounded by the Fisher information matrix:

$$C_{ij} \geq \left(I^{-1}\right)_{ij}, \tag{3.4}$$

which is known as the Cramer-Rao bound. The diagonal entries of the Cramer-Rao bound are $C_{ii} = (\delta c_i)^2 \geq (I^{-1})_{ii}$ and indicate that the smallest possible uncertainty achievable on the coefficient $c_i$ given the input data is $\delta c_i|^{(\mathrm{best})} = \sqrt{(I^{-1})_{ii}}$.

For a given set of EFT coefficients, comparing the values of $I_{ij}$ between different datasets highlights those which provide the highest information. The larger the entries of the Fisher matrix, the better (smaller uncertainty) that these coefficients can be constrained by the data considered. The Fisher information matrix can also be understood as a metric in model space. If one has two sets of coefficients $\boldsymbol{c}_a$ and $\boldsymbol{c}_b$, corresponding to two different points in the EFT parameter space, then the local distance between them is defined as

$$d_{\mathrm{loc}}(\boldsymbol{c}_a, \boldsymbol{c}_b) = \left[\sum_{i,j}(\boldsymbol{c}_a - \boldsymbol{c}_b)_i I_{ij}(\boldsymbol{c}_a)(\boldsymbol{c}_a - \boldsymbol{c}_b)_j\right]^{1/2}, \tag{3.5}$$

a feature which provides a robust method to quantify how (di)similar are two points in this model space.

If one has $n_{\mathrm{dat}}$ experimental measurements $\sigma_m^{(\mathrm{exp})}$ whose theoretical predictions depend on $n_{\mathrm{op}}$ coefficients $\boldsymbol{c}$, assuming that these measurements are Gaussianly distributed, one has

$$f(\boldsymbol{\sigma}_{\mathrm{exp}})|\boldsymbol{c}) = \prod_{m=1}^{n_{\mathrm{dat}}} \frac{1}{\sqrt{2\pi\delta_{\mathrm{exp},m}^2}} \exp\left(-\frac{\left(\sigma_m^{(\mathrm{exp})} - \sigma_m^{(\mathrm{th})}(\boldsymbol{c})\right)^2}{2\delta_{\mathrm{exp},m}^2}\right), \tag{3.6}$$

where $\delta_{\mathrm{exp},m}$ stands for the total experimental uncertainty associated to this cross-section measurement. Here we neglect for simplicity the point-by-point correlations, the extension to the full correlation covariance matrix is straightforward. The theoretical predictions that enter Eq. (3.6) include the SM contribution as well as the terms linear and quadratic on the Wilson coefficients,

$$\sigma_m^{(\mathrm{th})}(\boldsymbol{c}) = \sigma_m^{(\mathrm{sm})} + \sum_{i=1}^{n_{\mathrm{op}}} c_i \sigma_{m,i}^{(\mathrm{eft})} + \sum_{i<j}^{n_{\mathrm{op}}} c_i c_j \sigma_{m,ij}^{(\mathrm{eft})}, \tag{3.7}$$

where we assume $\Lambda = 1$ TeV, so that one can write

$$-\ln f(\boldsymbol{\sigma}_{\mathrm{exp}}|\boldsymbol{c}) = \sum_{m=1}^{n_{\mathrm{dat}}} \frac{1}{2\delta_{\mathrm{exp},m}^2}\left(\left(\sigma_m^{(\mathrm{exp})} - \sigma_m^{(\mathrm{sm})}\right) - \sum_{i=1}^{n_{\mathrm{op}}} c_i \sigma_{m,i}^{(\mathrm{eft})} - \sum_{i<j}^{n_{\mathrm{op}}} c_i c_j \sigma_{m,ij}^{(\mathrm{eft})}\right)^2 + A, \tag{3.8}$$

where $A$ is a constant that does not depend on the value of the coefficients, and thus the Fisher information matrix can be evaluated using Eq. (3.3) to yield

$$I_{ij} = \mathrm{E}\left[\sum_{m=1}^{n_{\mathrm{dat}}} \frac{1}{\delta_{\mathrm{exp},m}^2}\left(\sigma_{m,ij}\left(\sigma_m^{(\mathrm{th})} - \sigma_m^{(\mathrm{exp})}\right) + \left(\sigma_{m,i}^{(\mathrm{eft})} + \sum_{l=1}^{n_{\mathrm{op}}} c_l \sigma_{m,il}^{(\mathrm{eft})}\right)\left(\sigma_{m,j}^{(\mathrm{eft})} + \sum_{l'=1}^{n_{\mathrm{op}}} c_{l'} \sigma_{m,jl'}\right)\right)\right], \tag{3.9}$$

where this expectation value can be evaluated by averaging over the $N_\text{rep}$ replicas (or $N_\text{spl}$ samples) that provide a sampling of the probability density in the space of coefficients within our approach, see also Sect. 4.

The Fisher information matrix becomes specially simple if we restrict ourselves to the linear approximation [10], i.e. $\sigma_{m,ij} = 0$, since in this case

$$I_{ij} = \sum_{m=1}^{n_\text{dat}} \frac{\sigma_{m,i}^{(\text{eft})} \sigma_{m,j}^{(\text{eft})}}{\delta_{\text{exp},m}^2} \,, \tag{3.10}$$

which is independent of the values of the coefficients $\boldsymbol{c}$ and therefore of the actual fit results. The diagonal entries of the Fisher matrix $I_{ii}$ are then given by the sum over a given dataset (or group of processes) of the square of the linear EFT cross-sections over the total experimental uncertainty. In the specific case of one-parameter fits, the Cramer-Rao bound reads

$$\delta c_i \geq \left( \sum_{m=1}^{n_\text{dat}} \frac{\sigma_{m,i}^{(\text{eft})} \sigma_{m,j}^{(\text{eft})}}{\delta_{\text{exp},m}^2} \right)^{-1/2} \,, \qquad i = 1, \ldots, n_\text{op} \,, \tag{3.11}$$

which, provided that the sum is over the global dataset, can be used to cross-check the results of individual (one-parameter) fits.

One should emphasize that the absolute size of the entries of the Fisher matrix does not contain physical information: one is always allowed to redefine the overall normalisation of an operator such that $c_i \sigma_{m,i}^{(\text{eft})} = c_i' \sigma_{m,i}'^{(\text{eft})}$, with $c_i' = B_i c_i$ and $\sigma_{m,i}^{(\text{eft})} = \sigma_{m,i}'^{(\text{eft})}/B_i$ with $B_i$ being arbitrary constants. However, the relative value of $I_{ii}$ between two groups of processes is independent of the choice of operator normalisation and thus conveys meaningful information. For this reason, in the following we present results for the Fisher information matrix such that the sum of the diagonal entries associated to a given EFT coefficient adds up to a fixed reference value.

Fig. 3.1 displays the values of the diagonal entries of the Fisher information matrix, Eq. (3.9), evaluated for the same groups of processes as in Table 3.10. The normalisation is such that the sum of the entries associated to each coefficients adds up to 100. We show results for the Fisher information both at the linear level, Eq. (3.10), and with the quadratic corrections included, Eq. (3.9), in the left and right panels respectively. The entries in blue indicate those groups of processes which provide more than 75% of the information on the corresponding EFT coefficient. Entries in grey indicate relative contributions of less than 10%.

The information contained in Fig. 3.1 is consistent with that of Table 3.10, but now we can identify, for each coefficients, which datasets provide the dominant constraints. For instance, one observes that the two-light-two-heavy operators are overwhelmingly constrained by inclusive top quark pair production data, except for $c_{Qq}^{3,1}$ for which single top is the most important set of processes. In the case of the two-fermion operators, the leading constraints typically arise from Higgs data, in particular from the Run II signal strengths measurements. Two exceptions are $c_{\varphi t}$, which at the linear level (but not at the quadratic one) is dominated by $t\bar{t}V$, and the chromo-magnetic operator $c_{tG}$, for which inclusive $t\bar{t}$ production is most important. Also for the purely bosonic operators the Higgs data provides most of the information, except for $c_{WWW}$, as expected since this operator is only accessible in diboson processes. Furthermore, one observes that the $\mathcal{O}(\Lambda^{-4})$ corrections induce in most cases a moderate change in the Fisher information matrix, but in others they can significantly alter the balance between

31
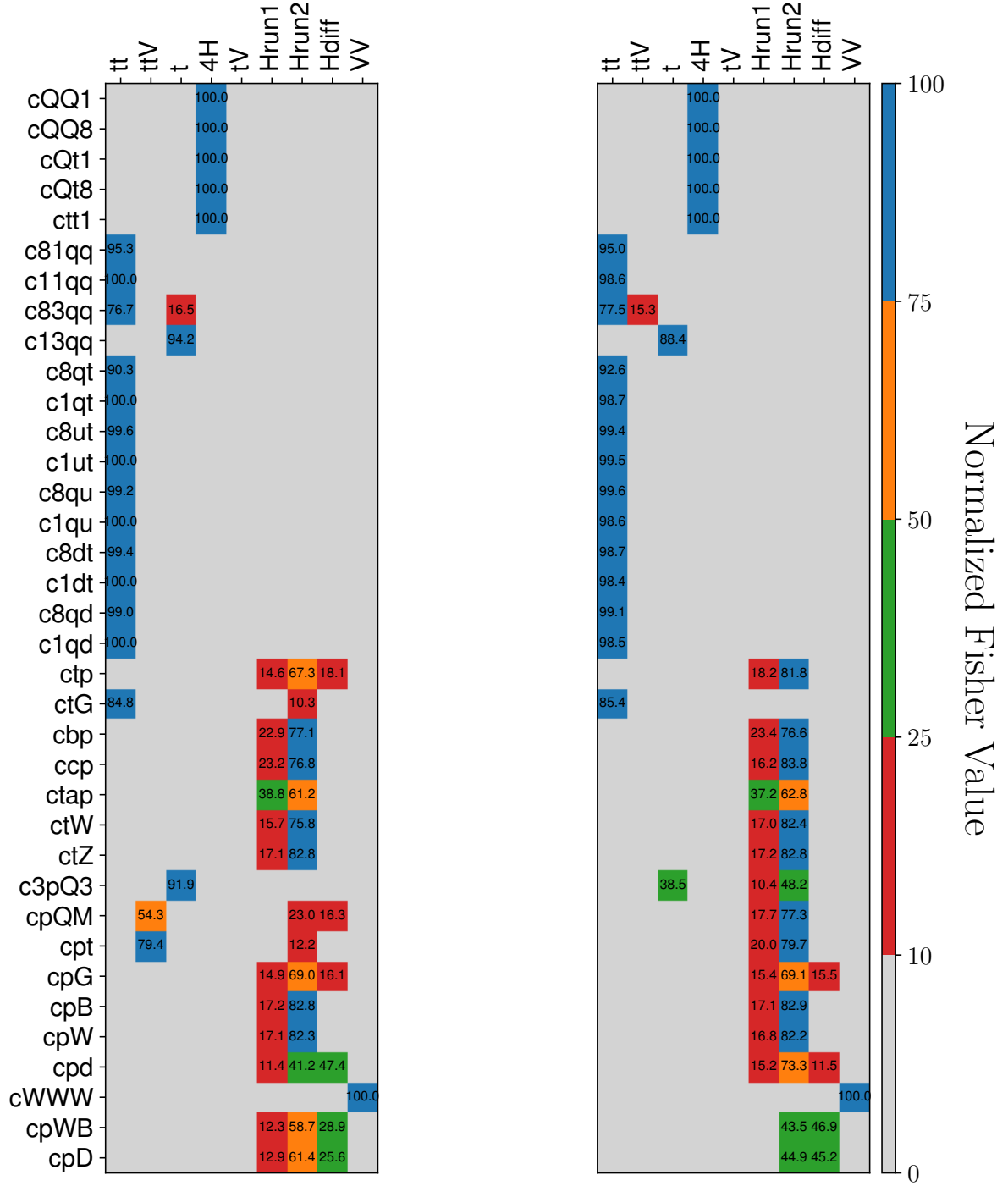
**Figure 3.1.** The values of the diagonal entries of the Fisher information matrix, Eq. (3.9), evaluated for the same groups of processes as in Table 3.10. The normalisation here is such that the sum of the entries associated to each EFT coefficient adds up to 100. We show results for the Fisher information matrix both at the linear level, Eq. (3.10), and with the quadratic corrections included, left and right panels respectively. For entries in the heat map larger than 10, we also indicate the corresponding numerical values. Note that the sum of the entries over columns does not have a physical interpretation.

processes. As a representative example, some of the two-fermion operators such as $c_{\varphi t}$ and $c_{\varphi Q}^{(3)}$ become dominated by the Higgs data only once quadratic corrections are accounted for.

Another relevant application of the Fisher information matrix is the determination of optimal directions in the EFT parameter space by means of principal component analysis (PCA), and in particular the assessment of whether or not the coefficients basis adopted for the fit contains flat directions. We will discuss this related application in the Sect. 4.5.

# 4 Fitting methodology

In this section we describe the fitting methodology that is used in this work to map the EFT parameter space spanned by the Higgs, diboson, and top quark data. In addition to results obtained with the Monte Carlo replica fitting (MCfit) method presented in Ref. [7], now we also determine the posterior probability distributions in the parameter space using the `MultiNest` Nested Sampling (NS) algorithm [37, 132], a robust sampling procedure that is completely orthogonal to the MCfit method and that is based on Bayesian inference.

We begin with a brief discussion of the log-likelihood function and the treatment of uncertainties that is adopted in the fit. We then present the individual $\chi^2$ profiles associated to each EFT coefficient in the quadratic fits and discuss the eventual presence of (quasi-)degenerate minima. Subsequently, the main features of the NS and MCFit strategies used in the global fit are summarized, including several improvements that have been implemented in the latter technique. The results obtained with the two methods are also benchmarked. Finally, we carry out a principal component analysis (PCA) to determine the linear combinations of parameters that have the highest and lowest variabilities given our global dataset and assess the possible presence of flat directions.

## 4.1 Log-likelihood

The overall fit quality is quantified by the log-likelihood, or $\chi^2$ function, defined as

$$\chi^2(\boldsymbol{c}) \equiv \frac{1}{n_{\text{dat}}} \sum_{i,j=1}^{n_{\text{dat}}} \left( \sigma_i^{(\text{th})}(\boldsymbol{c}) - \sigma_i^{(\text{exp})} \right) (\text{cov}^{-1})_{ij} \left( \sigma_j^{(\text{th})}(\boldsymbol{c}) - \sigma_j^{(\text{exp})} \right) , \qquad (4.1)$$

where $\sigma_i^{(\text{exp})}$ and $\sigma_i^{(\text{th})}(\boldsymbol{c})$ are the central experimental data and corresponding theoretical predictions for the $i$-th cross-section, respectively. The total covariance matrix, $\text{cov}_{ij}$, should contain all relevant sources of experimental and theoretical uncertainties. Assuming the latter are normally distributed, and that they are uncorrelated with the experimental uncertainties, this total covariance matrix can be expressed as a sum of the separate experimental and theoretical covariance matrices [133, 134],

$$\text{cov}_{ij} = \text{cov}_{ij}^{(\text{exp})} + \text{cov}_{ij}^{(\text{th})} . \qquad (4.2)$$

As usual, the experimental covariance matrix is constructed from all sources of statistical and systematic uncertainties that are made available by the experiments (as discussed in Sect. 3). Moreover, the correlated multiplicative uncertainties are treated via the '$t_0$' prescription [135] in the fit, while the standard experimental definition is used to quote the resulting $\chi^2$ values.

Concerning the theoretical covariance matrix, $\text{cov}^{(\text{th})}$, its contributions depend on the specific type of processes considered. In the case of the top quark and LHC diboson production

cross-sections, we compute the SM predictions using the best possible theoretical accuracy. In doing so, we also evaluate the uncertainty associated to the input PDFs and their correlations, as discussed in Ref. [7]. These computations are based on the NNPDF3.1 no-top fit [69], a global PDF determination based on a dataset that excludes all measurements that are used in the present SMEFT analysis.

For the Higgs production and decay measurements, we take instead the SM predictions from the experimental publications, which in turn are collected from the HXSWG reports. In such a case, the total theory uncertainty is available, which includes both PDF errors and missing higher order uncertainties (MHOUs). The total theory uncertainty for Higgs measurements is therefore included in the fit covariance matrix by means of the correlation prescription provided in the corresponding ATLAS and CMS publications.

## 4.2 Individual fits from the $\chi^2$ profiles

Individual (one-parameter) fits correspond to varying a single EFT coefficient while keeping the rest fixed to their SM values. While such fits neglect the correlations between the different coefficients, they provide a useful baseline for the global analysis, since there the CL intervals will be by construction looser (or at best, similar) as compared to those of the one-parameters fits. They are also computationally inexpensive, as they can be carried out analytically from a scan of the $\chi^2$ profile without resorting to numerical methods. Another benefit is that they facilitate the comparison between different EFT analyses, which may adopt different fitting bases but whose individual bounds should be similar provided they are based on comparable data sets and theoretical calculations.

In the scenario where a single EFT coefficient, $c_j$, is allowed to vary while the rest are set to zero, the theoretical cross-section (for $\Lambda = 1$ TeV) given by Eq. (3.7) simplifies to

$$\sigma_m^{(\text{th})}(c_j) = \sigma_m^{(\text{sm})} + c_j \sigma_{m,j}^{(\text{eft})} + c_j^2 \sigma_{m,jj}^{(\text{eft})} , \tag{4.3}$$

which results in a quartic polynomial form for the $\chi^2$ when inserted into Eq. (4.1), namely

$$\chi^2(c_j) = \sum_{k=0}^{4} a_k (c_j)^k . \tag{4.4}$$

Restricting the analysis to the linear order in the EFT expansion further simplifies Eq. (4.4) to a parabolic form,

$$\chi^2(c_j) = \sum_{k=0}^{2} a_k (c_j)^k = \chi_0^2 + b(c_j - c_{j,0})^2 , \tag{4.5}$$

where $c_{j,0}$ is the value of $c_j$ at the minimum of the parabola, and in this case linear error propagation (Gaussian statistics) is applicable.

To determine the values of the quartic polynomial coefficients $a_k$ in Eq. (4.4), it is sufficient to fit this functional form to a scan of the $\chi^2$ profile obtained by varying the EFT coefficient $c_j$ when all other coefficients are set to their SM value. The associated 95% CL interval to the coefficient $c_j$ can then be determined by imposing the condition

$$\chi^2(c_j) - \chi^2(c_{j,0}) \equiv \Delta\chi^2 \leq 5.91 . \tag{4.6}$$

We note that if the size of the quadratic $\mathcal{O}(\Lambda^{-4})$ corrections is sizable, there will be more than one solution for $c_{j,0}$ and one might end up with pairwise disjoint CL intervals.
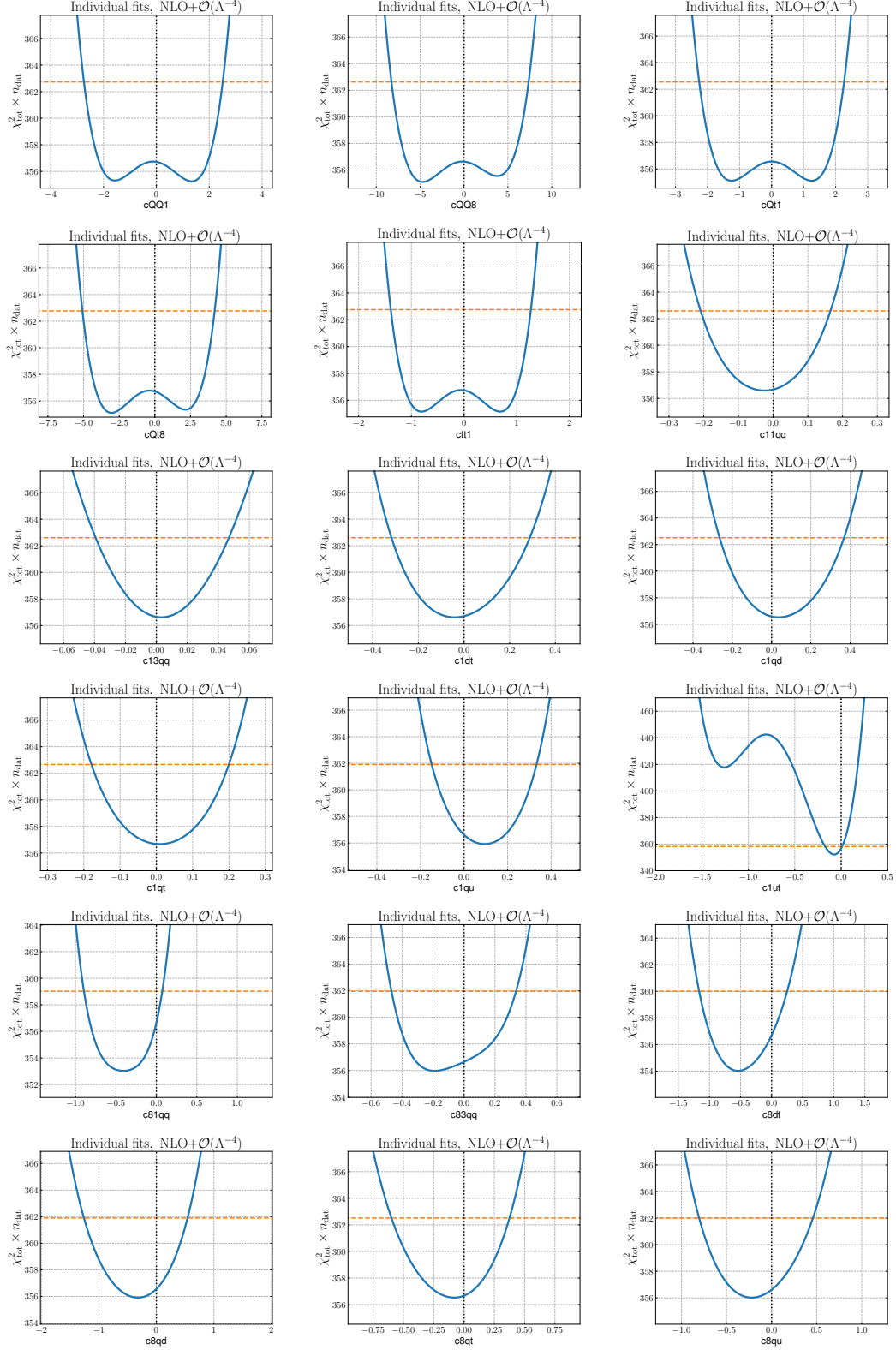
**Figure 4.1.** Results of quartic polynomial fits to the $\chi^2$ profiles obtained in one-parameter scans for each EFT coefficient, with all others set to their SM values. We show the absolute $\chi^2$ for the $n_{\mathrm{dat}} = 317$ data points of the global dataset calculated with the $t_0$ prescription, with the horizontal (vertical) line indicating the corresponding 95% CL ranges (the SM prediction).

35

**Figure 4.2.** Fig. 4.1 continued.
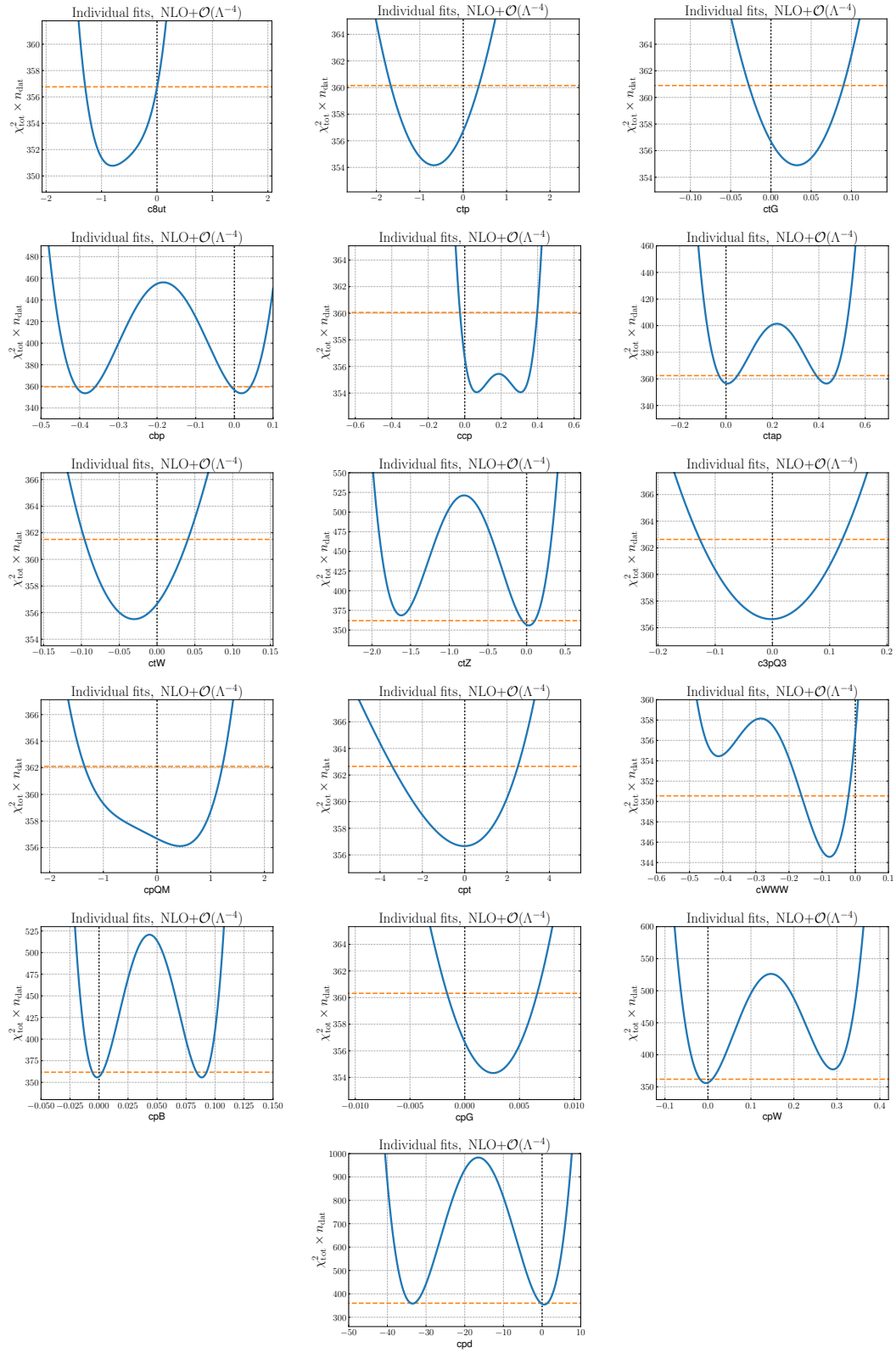
Figs. 4.1 and 4.2 display the results of quartic polynomial fits to the $\chi^2$ profiles obtained in the one-parameter scans of each EFT coefficient, based on the $n_{\text{dat}} = 317$ data points of the global dataset and the baseline theory settings (where higher-order QCD and EFT corrections are accounted for). Here the absolute $\chi^2$ is evaluated with the $t_0$ prescription, and we also display the corresponding 95% CL ranges (vertical line) and the SM expectation (horizontal line). We show the 34 profiles associated to the independent EFT coefficients in Table 2.5 which are not constrained by the EWPOs. These profiles are shown in the following order: four-heavy, two-light-two-heavy, two-fermion, and purely bosonic coefficients.

From the $\chi^2$ profiles displayed in Figs. 4.1 and 4.2 one can observe, on the one hand, how for several of the coefficients the parabolic approximation performs reasonably well, indicating the dominance of the linear EFT corrections. On the other hand, other coefficients deviate from the parabolic behaviour in a striking manner, including several degrees of freedom that exhibit two quasi-degenerate solutions, one being "SM-like" and the other distinctly non-zero. It is important to identify in particular which coefficients display such degenerate solutions in the one-parameter fits, since these might lead to a multi-modal posterior distributions in the case of the global analysis.

From the inspection of these $\chi^2$ profiles one can identify three categories of EFT coefficients whose individual profiles are poorly described by the parabolic approximation. First of all, one has the case of coefficients such as the four-heavy operators, for which a quartic profile with two quasi-degenerate solutions distributed symmetrically around the SM value is observed. Secondly, there are coefficients such as $c_{WWW}$, $c_{tZ}$, and $c_{tu}^1$ which display a second solution far from the SM-like one but which corresponds to higher values of the $\chi^2$, and hence does not modify the calculation of the CL intervals. We note that close to these absolute minima, the parabolic approximation can be reliably adopted. In both cases, the resulting CL intervals remain non-disjoint. Thirdly, one finds coefficients that exhibit quasi-degenerate solutions leading to disjoint CL intervals, where again one solution is SM-like and the other is far from the SM value. Examples of this category are the operators that modify the bottom, charm, and tau lepton Yukawa interactions, $c_{\varphi b}$, $c_{\varphi c}$, and $c_{\varphi \tau}$, and the purely bosonic operator $c_{\varphi d}$. Such degenerate solutions are likely to propagate to the global fit where all operators are simultaneously varied, and indeed as will be discussed in Sect. 5 the presence of these quasi-degenerate minima on the one-parameter fits has consequences at the level of posterior probability distributions in the global case.

Another useful application of the parameter bounds obtained from these individual fits is to help defining in an automated manner the suitable initial sampling ranges for each EFT coefficient in the global fits based on the MCfit and NS approaches. With this motivation, the individual bounds corresponding to a given input dataset and settings of the theoretical calculations are evaluated by default before each fitting run. Let us also mention that the bounds on the EFT coefficients obtained with the method presented here for the one-parameter analyses (quartic fits to the $\chi^2$ profiles) are found to be in agreement with the corresponding results obtained the NS and MCfit approaches.

## 4.3 Nested Sampling

The main approach that is adopted in this work to constrain the EFT parameter space is Nested Sampling (NS), specifically the version implemented in the `MultiNest` algorithm [37]. In comparison to MCfit, which is an optimisation problem aimed to determine the best-fit values for each of the replicas, NS is based on sampling the figure of merit $\chi^2$ to determine its

dependence on the Wilson coefficients and locate the region of maximum likelihood. Since NS is completely independent from the MCfit procedure, its availability makes possible validating the robustness of the resulting bounds in EFT parameter space via two orthogonal methods.

The starting point of NS is Bayes' theorem, which allows one to evaluate the probability distribution of a set of parameters $\boldsymbol{c}$ associated to a model $\mathcal{M}(\boldsymbol{c})$ given a set of experimental measurements $\mathcal{D}$,

$$P(\boldsymbol{c}|\mathcal{D}, \mathcal{M}) = \frac{P(\mathcal{D}|\mathcal{M}, \boldsymbol{c})P(\boldsymbol{c}|\mathcal{M})}{P(\mathcal{D}|\mathcal{M})} \,. \tag{4.7}$$

Here $P(\boldsymbol{c}|\mathcal{D}, \mathcal{M})$ represents the posterior probability of the model parameters given the assumed model and the observed experimental data, $P(\mathcal{D}|\mathcal{M}, \boldsymbol{c}) = \mathcal{L}(\boldsymbol{c})$ is the likelihood (conditional probability) of the experimental measurements given the model and a specific choice of parameters, and $P(\boldsymbol{c}|\mathcal{M}) = \pi(\boldsymbol{c})$ is the prior distribution for the model parameters. The denominator in Eq. (4.7), $P(\mathcal{D}|\mathcal{M}) = \mathcal{Z}$, is known as the Bayesian evidence and ensures the normalisation of the posterior distribution,

$$\mathcal{Z} = \int \mathcal{L}(\boldsymbol{c})\pi(\boldsymbol{c})d\boldsymbol{c} \,, \tag{4.8}$$

where the integration is carried out over the domain of the model parameters $\boldsymbol{c}$.

The key ingredient of Nested Sampling is to utilise the ideas underlying Bayesian inference to map the $n$-dimensional integral over the prior density in model parameter space $\pi(\boldsymbol{c})d\boldsymbol{c}$, where $n$ represents the dimensionality of $\boldsymbol{c}$, into a one-dimensional function of the form

$$X(\lambda) = \int_{\{\boldsymbol{c}:\mathcal{L}(\boldsymbol{c})>\lambda\}} \pi(\boldsymbol{c})d\boldsymbol{c} \,. \tag{4.9}$$

In this expression, the prior mass $X(\lambda)$ corresponds to the (normalised) volume of the prior density $\pi(\boldsymbol{c})d\boldsymbol{c}$ associated with values of the model parameters that lead to a likelihood $\mathcal{L}(\boldsymbol{c})$ greater than the parameter $\lambda$. Note that by construction, the prior mass $X$ decreases monotonically from the limiting value $X = 1$ to $X = 0$ as $\lambda$ is increased. The integration of $X(\lambda)$ extends over the regions in the model parameter space contained within the fixed-likelihood contour defined by the condition $\mathcal{L}(\boldsymbol{c}) = \lambda$. This property allows the evidence to be expressed as,

$$\mathcal{Z} = \int_0^1 \mathcal{L}(X)dX \,, \tag{4.10}$$

where $\mathcal{L}(X)$ is defined as the inverse function of $X(\lambda)$, which always exists provided the likelihood is a continuous and smooth function of the model parameters. Therefore, the transformation from $\boldsymbol{c}$ to $X$ in Eq. (4.9) achieves a mapping of the prior distribution into infinitesimal elements, sorted by their associated likelihood $\mathcal{L}(\boldsymbol{c})$.

The next step of the NS algorithm is to define a decreasing sequence of values in the prior volume, that is now parameterised by the prior mass $X$. In other words, one slices the prior volume into a large number of small regions

$$1 = X_0 > X_1 > \ldots X_\infty = 0 \,, \tag{4.11}$$

and then evaluates the likelihood at each of these values, $\mathcal{L} = \mathcal{L}(X_i)$. This way, all of the $\mathcal{L}_i$ values can be summed in order to evaluate the integral for the Bayesian evidence, Eq. (4.10).

Since in general the likelihood $\mathcal{L}(\boldsymbol{c})$ exhibits a complex dependence on the model parameters $\boldsymbol{c}$, the summation in Eq. (4.10) must be evaluated numerically using *e.g.* Monte Carlo

integration methods. In practice, one draws $N_{\text{live}}$ points from the parameter prior volume $\pi(\boldsymbol{c})$, known as *live points*, and orders the likelihood values from smallest to largest, including the starting value of the prior mass at $X_0 = 1$. As samples are drawn from the prior volume, the live point with the lowest likelihood $\mathcal{L}_i$ is removed from the set and replaced by another live point drawn from the same prior distribution but now under the constraint that its likelihood is larger than $\mathcal{L}_i$. This sampling process is repeated until the entire hyper-volume $\pi(\boldsymbol{c})$ of the prior parameter space has been covered, with ellipsoids of constrained likelihood being assigned to the live-points as the prior volume is scanned.

While the end result of the NS procedure is the estimation of the Bayesian evidence $\mathcal{Z}$, as a byproduct one also obtains a sampling of the posterior distribution associated to the EFT coefficients expressed as

$$\{\boldsymbol{c}^{(k)}\}\,, \qquad k = 1, \ldots, N_{\text{spl}}\,, \tag{4.12}$$

with $N_{\text{spl}}$ indicating the number of samples drawn by the final NS iteration. One can then compute expectation values, variances, and correlations of the model parameters by evaluating the MC sum over these posterior samples together with their associated weights, in the same manner as averages are carried out over the $N_{\text{rep}}$ replicas in the MCfit method.

**Prior volume.** An important input for NS is the choice of prior volume $\pi(\boldsymbol{c})$ in the model parameter space. In this analysis, we adopt flat priors defined by ranges in parameter space for the coefficients $\boldsymbol{c}$. A suitable choice of prior volume where the sampling takes place is important to speed up the NS algorithm: a range too wide will make the optimisation less efficient, while a range too narrow might bias the results by cutting specific regions of the parameter space that are relevant. Furthermore, using a common range for all parameters should be avoided, since the range of intrinsic variation will be rather different for each of the EFT coefficients, as illustrated also by the one-parameter fits reported in Figs. 4.1 and 4.2.

Taking these considerations into account, we adopt here the following strategy. First, a single model parameter $c_i$ is allowed to vary while all others are set to their SM value, $c_j = 0$ for $j \neq i$. The $\chi^2(c_i)$ is then scanned in this direction to determine the values $c_i^{(\text{min})}$ and $c_i^{(\text{max})}$ satisfying the condition $\chi^2/n_{\text{dat}} = 4$. We then repeat this procedure for all parameters and end up with a hyper-volume defined by pairs of values

$$\pi(\boldsymbol{c}) = \left[\left(c_i^{(\text{min})}, c_i^{(\text{max})}\right)\,, \quad i = 1, \ldots, n_{\text{op}}\right]\,, \tag{4.13}$$

which then defines the initial prior volume. At this point, one performs an initial exploratory NS global analysis using this volume to study the posterior probability distribution for each EFT coefficient. Our final analysis is then obtained by manually adjusting the initial sampling ranges until the full posterior distributions are captured for the chosen prior volume. For parameters that are essentially unconstrained in the global fit, such as the four-heavy operators in the case of linear EFT calculations, a hard boundary of $(-50, 50)$ is imposed (for $\Lambda = 1$ TeV) .

**Performance.** In order to increase the efficiency of the posterior probability estimation by NS, we enable the "constant efficiency mode" in MultiNest, which adjusts the total volume of ellipsoids spanning the live points so that the sampling efficiency is close to its associated hyperparameter set by the user. With 24 cpu cores, we are able to achieve an accurate posterior for the linear EFT fits in around 30 minutes using 500 live points, a target efficiency

of 0.05, and an evidence tolerance of 0.5, which results in $N_{\rm spl} \simeq 5000$ posterior samples. To ensure the stability of our final results, we chose 1000 live points and a target efficiency of 0.005, which yields $\simeq 1.5 \times 10^4$ samples for the linear analysis and $\simeq 10^4$ samples for an analysis that includes also the quadratic EFT corrections. With these settings, our final global analyses containing the simultaneous determination of $n_{\rm op} \simeq 36$ coefficients take $\sim 3.5$ hours running in 24 cpu cores, with a similar performance for linear and quadratic EFT fits.

The NS method is especially suitable to tackle parameter spaces of moderate dimensionality. Being based purely on sampling, it is not affected by limitations in minimisation methods such as ending up in local minima. It is also more robust upon the presence of fluctuations, and does not require specifying certain hyperparameters such as the learning rates which are used in MCfit. The main limitation of NS is that, as in all sampling methods, the execution times grows exponentially with $n_{\rm op}$, the dimensionality of the model parameter space. For parameter spaces of dimensionality greater than around 50, the current NS implementation that we use becomes unpractically slow and MCfit becomes the most suitable strategy available.

### 4.4   The Monte Carlo replica method revisited

The `SMEFiT` analysis of Ref. [7] was based on the Monte Carlo replica approach (MCfit), which in turn was inspired by the NNPDF analysis of the quark and gluon substructure of protons. The MCfit method aims to construct a sampling of the probability distribution in the space of the experimental data, which then translates into a sampling of the probability distribution in the space of the EFT coefficients through an optimisation procedure where the best-fit values of the coefficients for each replica, $\boldsymbol{c}^{(k)}$, are determined.

Given an experimental measurement of a hard-scattering cross-section, denoted by $\sigma_i^{(\rm exp)}$, with total uncorrelated uncertainty $\delta_i^{(\rm stat)}$ and $n_{\rm sys}$ correlated systematic uncertainties $\delta_{i,\alpha}^{(\rm sys)}$, the $N_{\rm rep}$ artificial MC replicas of the experimental data are generated as

$$\sigma_i^{(\rm art)(k)} = \sigma_i^{(\rm exp)}\left(1 + r_i^{(k)}\delta_i^{(\rm stat)} + \sum_{\alpha=1}^{n_{\rm sys}} r_{i,\alpha}^{(k)}\delta_{i,\alpha}^{(\rm sys)}\right), \quad k = 1,\ldots,N_{\rm rep}, \qquad (4.14)$$

where the index $i$ runs from 1 to $n_{\rm dat}$ and $r_i^{(k)}$, $r_{i,\alpha}^{(k)}$ are univariate Gaussian random numbers. Correlations between data points induced by systematic uncertainties are accounted for by ensuring that $r_{i,\alpha}^{(k)} = r_{i',\alpha}^{(k)}$. It can be show that central values, variances, and covariances evaluated by averaging over the MC replicas reproduce the corresponding experimental values.

A fit to the $n_{\rm op}$ degrees of freedom $\boldsymbol{c}/\Lambda$ is then performed for each of the MC replicas generated by Eq. (4.14). These best-fit values are determined from the minimisation of the cost function

$$E^{(k)}(\boldsymbol{c}) \equiv \frac{1}{n_{\rm dat}}\sum_{i,j=1}^{n_{\rm dat}} \left(\sigma_i^{(\rm th)}\!\left(\boldsymbol{c}^{(k)}\right) - \sigma_i^{(\rm art)(k)}\right)(\mathrm{cov}^{-1})_{ij}\left(\sigma_j^{(\rm th)}\!\left(\boldsymbol{c}^{(k)}\right) - \sigma_j^{(\rm art)(k)}\right), \qquad (4.15)$$

where $\sigma_i^{(\rm th)}(\boldsymbol{c}^{(k)})$ indicates the theoretical prediction for the $i$-th cross-section evaluated with the $k$-th set of EFT coefficients. This process results in a collection of $\boldsymbol{c}^{(k)}$ best-fit coefficient values from which estimators such as expectation values, variances, and correlations are evaluated. The overall fit quality is then evaluated using Eq. (4.1), where the central experimental values are compared to the mean theoretical prediction computed by the resulting fit replicas.

As mentioned in Sect. 4.1, various theoretical uncertainties are also included in the $\chi^2$ definition for some datasets. A consistent treatment of theoretical uncertainties in the fitting procedure means that these are not only included in the fit via the covariance matrix in Eqs. (4.15), but also in the corresponding replica generation. In other words, the replicas are sampled according to a multi-Gaussian distribution defined by the total covariance matrix Eq. (4.2) which receives contributions both of experimental and of theoretical origin. We therefore account for such errors in the generation of Monte Carlo replicas [133] using Eq. (4.14).

There are numerous advantages of using the MCfit method for global EFT analyses. First, it does not require specific assumptions about the underlying probability distribution of the fit parameters, and in particular does not rely on the Gaussian approximation. Secondly, the computational cost scales in a much milder way with the number of operators $n_{\mathrm{op}}$ included in the fit as compared to NS. Thirdly, it can be used to assess the impact of new datasets in the fit *a posteriori* with the Bayesian reweighting formalism.

In comparison with [7], several improvements have been implemented to increase the efficiency and accuracy of the MCfit procedure used in this analysis:

**Optimisation.** In the top quark sector analysis of [7], the minimisation of Eq. (4.15) was achieved by a gradient descent method which relies on local variations of the error function. This choice is advantageous since $E^{(k)}$ is at most a quartic form of the fit parameters, see Eq. (4.4) and its generalisation to multiple operators, and therefore evaluating its gradient is computationally efficient.

Since in the present analysis our parameter space is more complex, the optimiser that we use now to determine the best-fit values of the degrees of freedom $\boldsymbol{c}^{(k)}$ within MCfit is a trust-region algorithm `trust-constr` available in the `SciPy` package. An advantage of using `trust-constr` in this context is that it allows one to provide the optimiser with any combination of constraints on the coefficients, including existing bounds. This is a rather useful feature, since in many cases of interest one would like to restrict the EFT parameter space based on theoretical considerations, such as when accounting for the LEP EWPOs or in the top-philic scenario discussed in Sect. 2.

**Initial sampling range and bounds.** For each MC replica fit, the initial values of the fit coefficients $\boldsymbol{c}^{(k)}$ are initialised at random within a pre-defined range. This sampling range, as well as the boundaries imposed on the minimisation procedure for the poorly constrained parameters, are taken to be the same as those used in the NC procedure. That is, the sampling ranges for the global fits are derived from a one-parameter $\chi^2$ scanning procedure subsequently inflated to cover a sufficiently large parameter hyper-volume.

**Cross-validation.** Given the large dimensionality of the considered EFT parameter space, it is conceivable that the optimiser algorithm ends up fitting the statistical fluctuations of the experimental data rather than the underlying physical law. One way to prevent the minimiser from over-fitting the data is to use look-back cross-validation stopping. In this method, each replica dataset is randomly split with equal probability into two disjoint sets, known as the training and validation sets. Only the data points in the training set are then used to compute the figure of merit being minimised, Eq. (4.15), while the data points in the validation set are monitored alongside the fit. The random assignment of the data points to the training or

**Figure 4.3.** The best-fit values and 95% CL intervals for a global fit based on linear EFT calculations, comparing the outcome of the NS and MCfit methods. We display the results corresponding to the 50 coefficients listed in Table 2.5 (except for $c_{\ell\ell} = 0$), of which 36 are independent fit parameters,

validation sets is different for each MC replica, and the splitting only occurs for experiments that contain more than 5 bins in the distribution. The fit is run for a fixed large number of iterations, and then the optimal stopping point of the fit is then determined as the iteration for which the figure of merit evaluated on the validation set, $E_{\mathrm{val}}^{(k)}$, exhibits a global minimum. All in all, it is found that the risk of over-fitting is small and that MCfit results with and without cross-validation applied are reasonably similar.

**Quality selection criteria.** One disadvantage of optimisation strategies such as MCfit is that as the parameter space space is increased, the minimiser might sometimes converge on a local, rather than on the global, minimum. This is specially problematic in the quadratic EFT fits which often display quasi-degenerate minima, as illustrated by the $\chi^2$ profiles of Figs. 4.1 and 4.2. For this reason, it is important to implement post-fit quality selection criteria that indicate when a fitted replica should be kept and when it should be discarded. Here, a MC replica is kept if the total error function of the replica dataset, $E_{\mathrm{tot}}^{(k)}$, satisfies $E_{\mathrm{tot}}^{(k)} \leq 3$.

**Benchmarking.** Fig. 4.3 compares the outcome of global fits obtained with either the NS or MCfit method, all other settings identical. Specifically, here we show the best-fit values and 95% CL intervals for global fits based on linear EFT calculations. We provide the results corresponding to the 50 coefficients listed in Table 2.5 (except for $c_{\ell\ell}$, which is set to zero by the EWPOs) of which 36 are independent fit parameters. We will further discuss the physical interpretation of these results in Sect. 5, here we only aim to establish that the two methods indeed lead to equivalent results.

The comparison of Fig. 4.3 demonstrates that in general the two methods are in excellent agreement, both in terms of best-fit values and of the corresponding uncertainties. This said, for specific coefficients one observes small differences, with MCfit in general tending to provide

**Figure 4.4.** Comparison of the $\chi^2$ distributions evaluated over replicas and posterior samples in global linear EFT fits based on the MCfit and NC methods, respectively. The corresponding 95% CL intervals on the EFT coefficients are displayed in Fig. 4.3.

somewhat looser bounds. The reason for this behaviour is that optimisation-based methods such as MCfit can be distorted by fitting inefficiencies, such as when the optimiser finds a local, rather than global, minimum. This phenomenon is further illustrated in Fig. 4.4, which compares the $\chi^2$ distributions evaluated over replicas and posterior samples in the MCfit and NC methods respectively. We observe that the MCfit distribution exhibits broader tails, implying that the bounds obtained this way might in some cases be slightly over-conservative.

Fig. 4.3, as well as the corresponding benchmark comparison for fits based on quadratic EFT calculations, demonstrates that results obtained with either NS or MCfit are statistically equivalent. In the rest of this work, we will adopt NS as the baseline method, since its not affected by potential inefficiencies in the minimisation procedure and, as discussed above, can produce global fits within a reasonable execution time.

## 4.5   Principal Component Analysis

Principal Component Analysis (PCA) represents a valuable tool to identify the combinations of degrees of freedom that exhibit the largest and smallest variabilities in a linear algebra problem. This identification has many applications, for instance, a large gap in variability suggests that the effective dimensionality of the problem is smaller than the nominal one, and thus dimensional reduction methods are advantageous to simplify the solution. Furthermore, directions in the parameter space with very small variability are difficult to constrain from data and are identified with flat directions. Such flat directions might compromise the reliability of the obtained results i.e. in Hessian EFT fits.[3]

---

[3]The PCA method can also be exploited to efficiently carry out linear SMEFT fits [136].

Here we apply the PCA technique combined with Singular Value Decomposition (SVD) to global fits based on linear EFT calculations. The goal is to ascertain the presence of possible flat directions, identify large gaps in variability between the principal components, and determine the relation between the physical fitting basis and these principal components. The starting point is the expression for the cross-section as a function of the EFT coefficients, Eq. (3.7), truncated at the linear order,

$$\sigma_m^{(\mathrm{th})}(\boldsymbol{c}) = \sigma_m^{(\mathrm{sm})} + \sum_{i=1}^{n_{\mathrm{op}}} c_i \sigma_{m,i}^{(\mathrm{eft})}, \qquad m = 1 \ldots, n_{\mathrm{dat}}, \tag{4.16}$$

where recall that we have set $\Lambda = 1$ TeV. We then define a matrix $K$ of dimensions $n_{\mathrm{dat}} \times n_{\mathrm{op}}$ and (dimensionless) components $K_{mi} = \sigma_{m,i}^{(\mathrm{eft})}/\delta_{\mathrm{exp},m}$, where $\delta_{\mathrm{exp,m}}$ is the same total experimental error that appears in the evaluation of the Fisher information matrix Eq. (3.9). By means of SVD, we can decompose this matrix $K$ as

$$K = U W V^\dagger, \tag{4.17}$$

where $U$ ($V$) is a $n_{\mathrm{dat}} \times n_{\mathrm{dat}}$ ($n_{\mathrm{op}} \times n_{\mathrm{op}}$) unitary matrix and $W$ is an $n_{\mathrm{dat}} \times n_{\mathrm{op}}$ diagonal matrix with semi-positive real entries, called the singular values, which are ordered by decreasing magnitude. The larger a singular value, the higher the variability of its principal component and the higher the likelihood that this component will be well constrained from the fit.

The elements of the symmetric matrix $V$ in Eq. (4.17) contain the principal components associated to each of the $n_{\mathrm{op}}$ singular values. These correspond to a linear superposition of the original coefficients, that is, we have that

$$\mathrm{PC}_k = \sum_{i=1}^{n_{\mathrm{op}}} a_{ki} c_i, \quad k = 1, \ldots, n_{\mathrm{op}}, \qquad \left( \sum_{i=1}^{n_{\mathrm{op}}} a_{ki}^2 = 1 \ \forall k \right) \tag{4.18}$$

where the larger the value of the squared coefficient $a_{kl}^2$, the larger the relative weight of the associated EFT coefficient in this specific (normalised) principal component. By means of the matrix $V$ (and its inverse), one can rotate between the original fitting basis and the one defined by the principal components.

Fig. 4.5 displays the singular values $\lambda_i$, that is, the elements of the diagonal matrix $W$ in the decomposition of Eq. (4.17), for the $n_{\mathrm{op}} = 36$ principal components associated to the global fit settings summarised in Tables 2.5 and 3.8. From the definition of the matrix $K$, a singular value $\lambda_i \simeq 1$ corresponds to a direction in the parameter space where the magnitude of the (linear) EFT corrections is of the same size as the associated experimental uncertainties. We observe that there are three flat directions (principal components with vanishing singular value), which as shown below can be associated to linear combinations of the four-heavy operators. Except for these three flat directions, there are no large hierarchies in the distribution of singular values, indicating that the physical dimensionality of our problem coincides with that of the chosen fitting basis. The principal component with the highest singular value is dominated by the bosonic operator $c_{\varphi D}$, which modifies the Higgs-gauge interactions and is well constrained by the EWPOs.

Then Fig. 4.6 displays a heat map with the values of the (squared) coefficients $a_{ki}^2$ that relate the original fitting basis to the principal components, Eq. (4.18), and whose associated eigenvalues are displayed in the upper panel. For those entries with $a_{ki}^2 \geq 0.1$ we also indicate the corresponding numerical value. Since the principal components are normalised, the sum
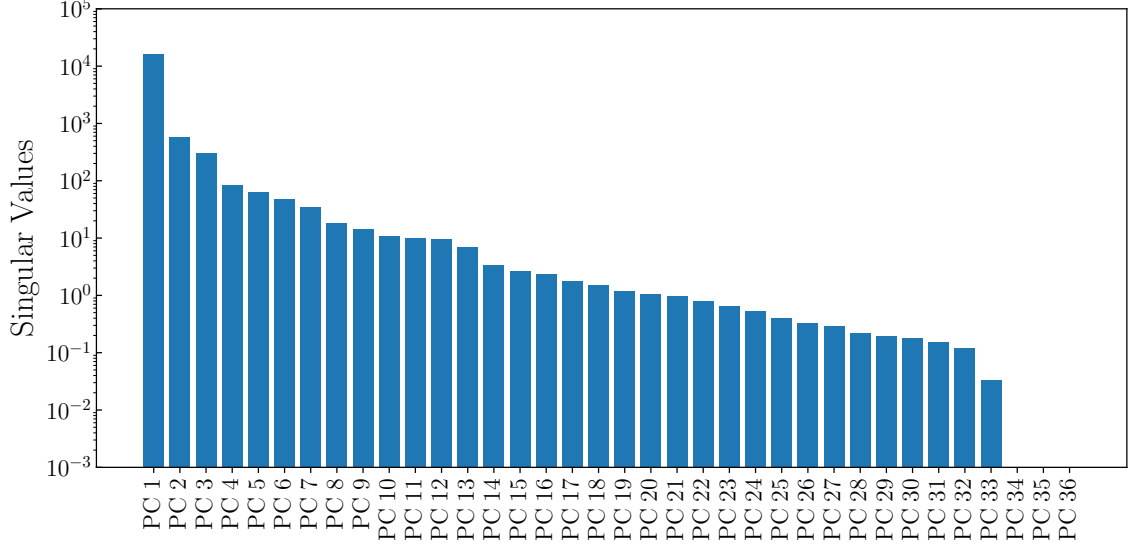
**Figure 4.5.** The distribution of singular values $\lambda_i$ (the elements of the diagonal of the matrix $W$), for the principal components evaluated for the global fit settings summarised in Tables 2.5 and 3.8. In the linear EFT approximation where the PCA analysis is carried out, there exist three flat directions (with vanishing singular values) associated to the four-heavy operators.

of the entries associated to a given row in the heat map adds up to unity. Note also that in this table we have chosen the purely bosonic coefficients $c_{\varphi WB}$ and $c_{\varphi D}$ to represent the two directions that are left unconstrained by the EWPOs, see the discussion in Sect. 2.1.

From Fig. 4.6 one can observe that some principal components ($\mathrm{PC}_k$) are dominated by a single EFT coefficient from the fitting basis. Examples of this are $c_{\varphi D}$ (for $k = 1$), $c_{\varphi G}$ ($k = 2$), $c_{\varphi B}$ ($k = 3$), $c_{tG}$ ($k = 4$), and $c_{Qq}^{3,1}$ ($k = 7$). These PCs have associated reasonably large singular values, $\lambda_k \gtrsim 10$, and therefore one expects that the corresponding coefficients will be well constrained from the fit. Other principal components are instead composed by a superposition of two or at most three coefficients, for instance $c_{\varphi WB}$ and $c_{\varphi Q}^{(3)}$ are combined into $\mathrm{PC}_{14}$ and $\mathrm{PC}_{15}$ with similar weight each. On the other hand, several PCs arise instead from the combination of a large number of EFT coefficients without any of them dominating. This is the case e.g. for PCs associated to combinations involving of the two-light-two-heavy operators, such $k = 22, 23$ and $25$, where no single squared coefficient $a_{ki}^2$ is larger than 0.4.

The three flat directions (vanishing singular values) observed in Fig. 4.5 can be traced back to linear combinations for four-heavy operators, specifically to the following combinations:

$$
\begin{aligned}
\mathrm{PC}_{34} &= 0.91 c_{QQ}^1 - 0.42 c_{tt}^1 \,, \\
\mathrm{PC}_{35} &= 0.62 c_{Qt}^1 - 0.56 c_{Qt}^8 + 0.49 c_{QQ}^8 \,, \\
\mathrm{PC}_{36} &= 0.78 c_{Qt}^1 - 0.50 c_{QQ}^8 + 0.38 c_{tQ}^8 \,,
\end{aligned}
\tag{4.19}
$$

where we don't indicate the contributions with $a_{ki}^2 < 0.1$. This implies that, in a linear EFT fit, one can only constrain two directions out of the five four-heavy operators considered. These flat directions disappear only once we consider the quadratic corrections to the $t\bar{t}t\bar{t}$ and $t\bar{t}b\bar{b}$ cross-sections.

The results of this PCA indicate that our choice of fitting basis, summarised in Table 2.5,
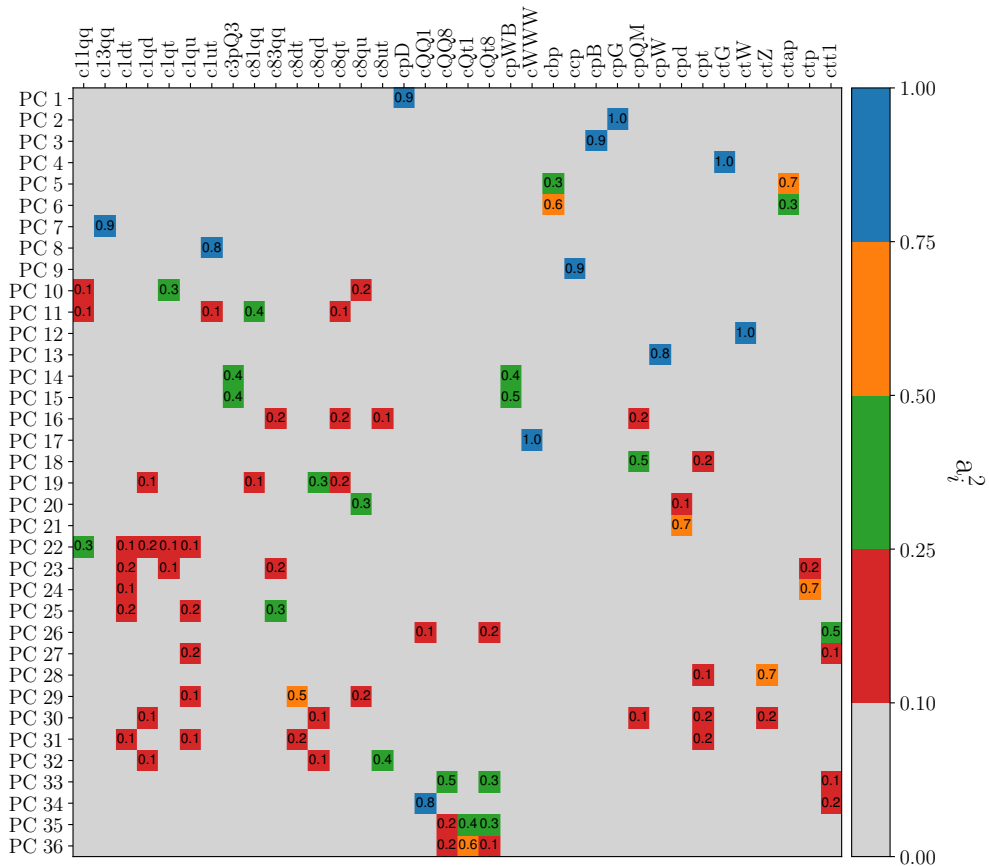
**Figure 4.6.** Heat map displaying the values of the squared coefficients $a_{ki}^2$ that relate the original fitting basis to the principal components, Eq. (4.18), whose associated singular values were reported in Fig. 4.5. For the entries with $a_{ki}^2 \geq 0.1$ we also indicate the corresponding numerical value.

represents a sensible option for which well-defined constraints will be obtained from the fit, up to the previous caveat concerning the four-heavy operators. Therefore, in our case there is no advantage in carrying out the fit in the rotated basis spanned by the principal components Eq. (4.18) rather than in the original one. Furthermore, the lack of large hierarchies in the distribution of singular values reflects the fact that the true dimensionality of the problem coincides with that of the original basis.

While for the global dataset genuine flat directions are either absent or removed by quadratic corrections, this might not be in general the case if we consider fits to reduced datasets. In such scenario, one could consider deploying the PCA method to reduce the dimensionality of the EFT parameter space by removing the directions with singular values below some threshold before inverting back to the physical basis. We note that a similar strategy has been successfully applied to construct compressed Hessian PDF sets in [137, 138]. However, in this work we use PCA as a diagnosis tool to guide the selection of the fitting basis, and postpone to future work its application to carry out EFT fits in the PC rotated basis.

# 5 Results

We now present the main results of this work: the determination of the best-fit values, confidence level intervals, and posterior probability distributions associated to the $n_{\rm op} = 50$ EFT coefficients (of which 36 are independent) listed in Table 2.5 from the global interpretation of Higgs, top quark, and diboson cross-section measurements. As motivated in Sect. 4, the results shown here have been obtained with the NS approach, and we have verified that equivalent results are obtained with the MCfit method.

First of all, we discuss the quality of the fit, both for the total and for individual datasets. Second, we present the bounds and posterior probability distributions for the various EFT coefficients, assess their consistency with the SM hypothesis, and determine their pattern of cross-correlations. Third, we study the dependence of our results on the choice of input dataset, in particular with fits based only on top or Higgs data, as well as on that of the theory settings, where the impact of the NLO QCD corrections to the EFT cross-sections is quantified. Finally, we present EFT fits in the top-philic scenario, where the parameter space is restricted by constraints motivated by specific UV-complete models. The comparison between SM and SMEFT theory predictions with the experimental dataset used as input to the fit is then collected in App. A.

## 5.1 Fit quality

To begin with, we investigate the quality of the fit in terms of the $\chi^2$ values for the individual datasets as well as for the global one. The values that will be provided here correspond to a modified version of Eq. (4.1), specifically

$$\chi^2 \equiv \frac{1}{n_{\rm dat}} \sum_{i,j=1}^{n_{\rm dat}} \left( \left\langle \sigma_i^{\rm (th)}\left(\boldsymbol{c}^{(k)}\right)\right\rangle - \sigma_i^{\rm (exp)}\right)({\rm cov}^{-1})_{ij}\left( \left\langle \sigma_j^{\rm (th)}\left(\boldsymbol{c}^{(k)}\right)\right\rangle - \sigma_j^{\rm (exp)}\right), \qquad (5.1)$$

where the average over the theory predictions is evaluated over the the $N_{\rm spl}$ samples provided by NS, and the covariance matrix is evaluated with the experimental definition [139]. Note that in general the average over theory predictions does not correspond to the theory prediction evaluated using the average value of the Wilson coefficients,

$$\left\langle \sigma_i^{\rm (th)}\left(\boldsymbol{c}^{(k)}\right)\right\rangle \neq \sigma_i^{\rm (th)}\left(\left\langle \boldsymbol{c}^{(k)}\right\rangle\right), \qquad (5.2)$$

due to the presence of the quadratic corrections to the EFT cross-sections.

With the figure of merit defined in Eq. (5.1), we collect in Tables 5.1 and 5.2 the values of the $\chi^2$ per data point corresponding to the baseline settings of our analysis. We display both the values based on the SM theory predictions as well as the best-fit SMEFT results obtained with $\mathcal{O}(\Lambda^{-2})$ and $\mathcal{O}(\Lambda^{-4})$ calculations. Note that, for ease of reference, in these tables each dataset has associated a hyperlink pointing to the original publication. For those datasets for which more than one differential distribution is available, we indicate the specific one used in the fit. Then Table 5.3 presents the summary of these $\chi^2$ values now indicating the total values for each group of processes as well as for the global dataset. Furthermore, the results of Tables 5.1 and 5.2 are graphically represented in Fig. 5.1.

Let us discuss first the $\chi^2$ results evaluated in terms of the groups of processes, listed in Table 5.3. One can observe that the global $\chi^2$ per data point improves from 1.05 when using SM theory to 0.96 (linear) and 1.01 (quadratic) once SMEFT corrections are accounted for.

| Dataset | $n_{\text{dat}}$ | $\chi^2_{\text{SM}}$ | $\chi^2_{\text{EFT}}$ $\mathcal{O}(\Lambda^{-2})$ | $\chi^2_{\text{EFT}}$ $\mathcal{O}(\Lambda^{-4})$ |
|---|---|---|---|---|
| ATLAS_tt_8TeV_ljets_mtt | 7 | 2.95 | 1.51 | 2.76 |
| ATLAS_tt_8TeV_dilep_mtt | 6 | 0.09 | 0.11 | 0.08 |
| CMS_tt_8TeV_ljets_ytt | 10 | 0.91 | 0.96 | 0.88 |
| CMS_tt2D_8TeV_dilep_mttytt | 16 | 1.63 | 1.63 | 1.63 |
| CMS_tt_13TeV_ljets_2015_mtt | 8 | 0.94 | 0.63 | 0.59 |
| CMS_tt_13TeV_dilep_2015_mtt | 6 | 1.30 | 1.16 | 1.24 |
| CMS_tt_13TeV_ljets_2016_mtt | 10 | 1.99 | 1.17 | 1.28 |
| CMS_tt_13TeV_dilep_2016_mtt | 7 | 2.28 | 2.84 | 2.36 |
| ATLAS_tt_13TeV_ljets_2016_mtt | 7 | 0.99 | 0.64 | 0.91 |
| ATLAS_CMS_tt_AC_8TeV | 6 | 0.86 | 0.79 | 0.78 |
| ATLAS_tt_AC_13TeV | 5 | 0.28 | 0.32 | 0.46 |
| ATLAS_WhelF_8TeV | 3 | 1.97 | 1.42 | 1.44 |
| CMS_WhelF_8TeV | 3 | 0.30 | 0.55 | 0.54 |
| ATLAS_ttZ_8TeV | 1 | 1.31 | 0.89 | 1.39 |
| ATLAS_ttZ_13TeV | 1 | 0.01 | 0.12 | 0.05 |
| ATLAS_ttZ_13TeV_2016 | 1 | 0.001 | 0.36 | 0.08 |
| CMS_ttZ_8TeV | 1 | 0.04 | 0.14 | 0.03 |
| CMS_ttZ_13TeV | 1 | 0.90 | 0.17 | 0.47 |
| CMS_ttZ_13TeV_pTZ | 4 | 0.73 | 0.83 | 0.93 |
| ATLAS_ttW_8TeV | 1 | 1.33 | 1.32 | 1.34 |
| ATLAS_ttW_13TeV | 1 | 0.83 | 0.82 | 0.83 |
| ATLAS_ttW_13TeV_2016 | 1 | 0.23 | 0.03 | 0.00 |
| CMS_ttW_8TeV | 1 | 1.54 | 1.21 | 1.57 |
| CMS_ttW_13TeV | 1 | 0.03 | 0.02 | 0.12 |
| CMS_ttbb_13TeV | 1 | 4.96 | 3.02 | 5.77 |
| CMS_ttbb_13TeV_2016 | 1 | 1.75 | 0.52 | 2.37 |
| ATLAS_ttbb_13TeV_2016 | 1 | 0.91 | 0.92 | 0.90 |
| CMS_tttt_13TeV | 1 | 0.05 | 0.02 | 0.08 |
| CMS_tttt_13TeV_run2 | 1 | 0.05 | 1.10 | 2.04 |
| ATLAS_tttt_13TeV_run2 | 1 | 2.35 | 0.74 | 0.30 |

**Table 5.1.** The values of the $\chi^2$ per data point corresponding to the baseline settings of our analysis. We indicate the results for the $t\bar{t}$ datasets, both in inclusive production and in association with vector bosons or heavy quarks. We display the SM values and then the best-fit SMEFT results obtained in analyses based on theory predictions at either $\mathcal{O}(\Lambda^{-2})$ or $\mathcal{O}(\Lambda^{-4})$ accuracy. Each dataset has a hyperlink pointing to the original publication. For those datasets for which more than one differential distribution is available, we indicate the specific ones used in the fit.

| Dataset | $n_{\mathrm{dat}}$ | $\chi^2_{\mathrm{SM}}/n_{\mathrm{dat}}$ | $\chi^2_{\mathrm{EFT}}/n_{\mathrm{dat}}$ $\mathcal{O}\!\left(\Lambda^{-2}\right)$ | $\chi^2_{\mathrm{EFT}}/n_{\mathrm{dat}}$ $\mathcal{O}\!\left(\Lambda^{-4}\right)$ |
|---|---|---|---|---|
| CMS_t_tch_8TeV_inc | 2 | 0.29 | 0.29 | 0.28 |
| ATLAS_t_tch_8TeV | 4 | 0.89 | 0.91 | 0.87 |
| CMS_t_tch_8TeV_diff_Yt | 6 | 0.11 | 0.11 | 0.12 |
| CMS_t_sch_8TeV | 1 | 1.26 | 1.01 | 1.36 |
| ATLAS_t_sch_8TeV | 1 | 0.08 | 0.64 | 0.01 |
| ATLAS_t_tch_13TeV | 2 | 0.01 | 0.01 | 0.01 |
| CMS_t_tch_13TeV_inc | 2 | 0.35 | 0.34 | 0.32 |
| CMS_t_tch_13TeV_diff_Yt | 4 | 0.46 | 0.47 | 0.46 |
| CMS_t_tch_13TeV_2016_diff_Yt | 5 | 0.58 | 0.59 | 0.59 |
| ATLAS_tZ_13TeV_inc | 1 | 0.00 | 0.03 | 0.02 |
| ATLAS_tZ_13TeV_run2_inc | 1 | 0.05 | 0.03 | 0.01 |
| CMS_tZ_13TeV_inc | 1 | 0.66 | 0.35 | 0.58 |
| CMS_tZ_13TeV_2016_inc | 1 | 1.23 | 0.30 | 0.94 |
| ATLAS_tW_8TeV_inc | 1 | 0.02 | 0.09 | 0.00 |
| ATLAS_tW_slep_8TeV_inc | 1 | 0.13 | 0.07 | 0.17 |
| CMS_tW_8TeV_inc | 1 | 0.00 | 0.02 | 0.00 |
| ATLAS_tW_13TeV_inc | 1 | 0.52 | 0.41 | 0.58 |
| CMS_tW_13TeV_inc | 1 | 3.79 | 5.11 | 3.18 |
| ATLAS_CMS_SSinc_RunI | 22 | 0.86 | 0.84 | 0.89 |
| ATLAS_SSinc_RunII | 16 | 0.54 | 0.55 | 0.55 |
| CMS_SSinc_RunII | 24 | 0.77 | 0.71 | 0.67 |
| ATLAS_ggF_ZZ_13TeV | 6 | 0.96 | 0.83 | 0.82 |
| CMS_ggF_aa_13TeV | 6 | 1.05 | 1.03 | 1.04 |
| ATLAS_H_13TeV_2015_pTH | 9 | 1.11 | 1.11 | 1.10 |
| CMS_H_13TeV_2015_pTH | 9 | 0.80 | 0.79 | 0.80 |
| ATLAS_WH_Hbb_13TeV | 2 | 0.10 | 0.03 | 0.15 |
| ATLAS_ZH_Hbb_13TeV | 3 | 0.50 | 0.33 | 0.23 |
| ATLAS_WW_13TeV_2016_memu | 13 | 1.64 | 1.64 | 1.67 |
| ATLAS_WZ_13TeV_2016_mTWZ | 6 | 0.85 | 0.81 | 0.86 |
| CMS_WZ_13TeV_2016_pTZ | 11 | 1.42 | 1.44 | 1.43 |
| LEP_eeWW_182GeV | 10 | 1.38 | 1.38 | 1.38 |
| LEP_eeWW_189GeV | 10 | 0.88 | 0.88 | 0.89 |
| LEP_eeWW_198GeV | 10 | 1.61 | 1.61 | 1.61 |
| LEP_eeWW_206GeV | 10 | 1.09 | 1.09 | 1.08 |

**Table 5.2.** Same as Table 5.1 now for the single top datasets (inclusive and in association with gauge bosons), the Higgs production and decay measurements (signal streghts and differential distributions), and the LEP and LHC diboson cross-sections.

| Dataset | $n_{\text{dat}}$ | $\chi^2_{\text{SM}}$ | $\chi^2_{\text{EFT}}$ $\mathcal{O}(\Lambda^{-2})$ | $\chi^2_{\text{EFT}}$ $\mathcal{O}(\Lambda^{-4})$ |
|---|---|---|---|---|
| $t\bar{t}$ inclusive | 94 | 1.36 | 1.14 | 1.23 |
| $t\bar{t} + V$ | 14 | 0.65 | 0.60 | 0.68 |
| single-top inclusive | 27 | 0.43 | 0.45 | 0.43 |
| single-top $+V$ | 9 | 0.71 | 0.71 | 0.61 |
| $t\bar{t}b\bar{b}$ & $t\bar{t}t\bar{t}$ | 6 | 1.68 | 1.05 | 1.91 |
| Higgs signal strenghts (Run I) | 22 | 0.86 | 0.84 | 0.89 |
| Higgs signal strenghts (Run II) | 40 | 0.68 | 0.65 | 0.62 |
| Higgs differential & STXS | 35 | 0.88 | 0.84 | 0.83 |
| Diboson (LEP+LHC) | 70 | 1.31 | 1.31 | 1.32 |
| **Total** | **317** | **1.05** | **0.96** | **1.01** |

**Table 5.3.** Summary of the $\chi^2$ results listed in Tables 5.1 and 5.2. We indicate the total values for each group of processes as well as for the global dataset.

Considering the fit quality to the various groups of processes, we observe that the description of the inclusive top-quark pair cross-sections is clearly improved once EFT corrections are accounted for, with $\chi^2 = 1.36$ in the SM decreasing to 1.14 and 1.23 in the linear and quadratic EFT fits respectively. For the rest of the datasets, and in particular for the Higgs and diboson measurements, the overall EFT fit quality is similar to that obtained using SM calculations.

Inspection of the $\chi^2$ values associated to individual datasets reported in Tables 5.1 and 5.2, as well as their graphical representation from Fig. 5.1, reveals that in some cases the agreement between the prior SM theoretical calculations and the data is poor. This is the case, in particular, for some of the inclusive $t\bar{t}$ datasets such as `ATLAS_tt_8TeV_ljets_mtt` and `CMS_tt_13TeV_dilep_2016_mtt`, binned in terms of the top-quark pair invariant mass distribution $m_{t\bar{t}}$. Such relatively high values of the $\chi^2$ do not necessarily imply the need for some New Physics effects, but could also be explained by issues with the modelling of the experimental systematic correlations in differential distributions, as discussed for the ATLAS 8 TeV lepton+jets data in [61, 68, 140]. Nevertheless, when all the inclusive $t\bar{t}$ datasets are considered collectively, a reasonable value of $\chi^2_{\text{SM}} = 1.36$ for the $n_{\text{dat}} = 94$ data points in the fit is obtained. In Sect. 5.3 we will assess the stability of the global fit results by presenting fit variants with the individual datasets that lead to a poor $\chi^2$ removed. We will also study variants where distributions sensitive to the high energy behaviour of the EFT, such as $m_{t\bar{t}}$ in top quark pair production, have their bins with $m_{t\bar{t}} \gtrsim 1$ TeV removed from the fit.

Beyond the inclusive $t\bar{t}$ datasets, there are some other instances of a sub-optimal agreement between SM theory and data. In all cases, there exist comparable measurements of the same process, either from the same experiment at a different center-of-mass energy $\sqrt{s}$ or from a different experiment at the same value of $\sqrt{s}$, for which the $\chi^2$ reveals good consistency with the SM. These include `CMS_ttbb_13TeV`, when comparing with the same measurement based

**Figure 5.1.** Graphical representation of the results of Tables 5.1 and 5.2, displaying the values of the $\chi^2$ per data point, Eq. (5.1) for the all datasets used as input in the fit. The $\chi^2$ values are shown for the SM and for the two global SMEFT baseline fits, based on theory calculations at either the linear $\mathcal{O}(\Lambda^{-2})$ or quadratic $\mathcal{O}(\Lambda^{-4})$ order in the EFT expansion.

on the full 2016 luminosity, and `CMS_tW_13TeV_inc`, where again the ATLAS measurement at the same $\sqrt{s}$ exhibit as good $\chi^2$. All in all, one finds a reasonable description of the global input dataset when using SM cross-sections which is further improved in the EFT fit.

For the rest of this section, when presenting the results of fits corresponding to variations of the baseline settings, such as fits based on reduced datasets, we will only indicate the $\chi^2$ values associated to groups of processes using the same format as Table 5.3, rather than to individual datasets, and comment when required on the results for the latter.

## 5.2 Constraints on the EFT parameter space

Following this assessment of the fit quality, we move to present the constraints on the SMEFT parameter space that can be derived from the present global fit. We will present results for the $n_{\text{op}} = 50$ Wilson coefficients listed in Table 2.5, with the understanding that only 36 of them are linearly independent.[4] Specifically, we provide the 95% confidence level intervals for each EFT coefficients, study their posterior probability distributions, evaluate the pattern of their correlations, and compare the marginalised bounds with those obtained in individual fits where only one coefficient is varied at a time. We will also assess the overall consistency of the fit results with respect to the Standard Model hypothesis. The results discussed here correspond to the global dataset with the baseline theory settings for both $\mathcal{O}(\Lambda^{-2})$ and $\mathcal{O}(\Lambda^{-4})$ theory calculations. Fits based on either reduced datasets or different theory settings are then discussed in Sects. 5.3 and 5.4 respectively.

---

[4]We note that the EWPO constraints of Eq. (2.7) set the four-lepton operator to zero, $c_{\ell\ell} = 0$, and hence we exclude this coefficient from the plots and tables of this section.

**Figure 5.2.** The normalised posterior probability distributions associated to each of the $n_{\text{op}} = 50$ fit coefficients considered in the present analysis, for both the linear and quadratic EFT fits. Note that the $x$-axis range is different in each case. From top to bottom and from left to right, we display the four-heavy, two-light-two-heavy, two-fermion, and purely bosonic coefficients. Only 36 of these coefficients are independent as indicated in Table 2.5.

**Posterior distributions.** Fig. 5.2 displays the normalised posterior probability distributions associated to each of the $n_{\text{op}} = 50$ fit coefficients considered in the present analysis, for the linear (blue) and quadratic (orange) EFT fits. As discussed in Sect. 4.3, the NS prior sampling volumes have been optimised to ensure that the posterior distribution associated to each coefficient is fully contained within them. One can observe how in general the $\mathcal{O}(\Lambda^{-4})$ corrections modify significantly the distributions that is obtained from the linear fits, for instance by shifting its median or by decreasing its variance. For several coefficients, the posterior distributions would be poorly described in the Gaussian approximation, and in some cases one finds multi-modal distributions such as for the Yukawa operators $c_{\varphi c}$, $c_{\varphi b}$, and $c_{\varphi \tau}$. Such double-humped distributions can be traced back to the (quasi)-degenerate minima in the individual $\chi^2$ profiles reported in Figs. 4.1 and 4.2. We can also observe how the four-heavy coefficients can only be meaningfully constrained in the quadratic fit. All in all, inclusion of the quadratic EFT corrections modifies in a significant manner the posterior distributions associated to most of the fit coefficients as compared to the linear approximation.

**Figure 5.3.** The best-fit (median) value of the EFT coefficients $c_i/\Lambda^2$ and their associated 95% CL intervals for the global fits based on either linear or quadratic EFT calculations, whose posterior distributions are represented in Fig. 5.2. The dashed horizontal line indicates the SM expectation.

**Confidence level intervals.** From the posterior probability distributions displayed in Fig. 5.2, one can derive the marginalised 95% CL intervals on the EFT coefficients both for the linear and quadratic fits. These results are collected in Table 5.4 (for $\Lambda = 1$ TeV) and represented graphically in Fig. 5.3. In addition, Table 5.4 also includes the corresponding obtained in individual NS fits, where only one operator is varied at a time and the rest are set to their SM values (recall the $\chi^2$ profiles from Figs. 4.1 and 4.2). We will further discuss the outcome of these individual fits below.

From the marginalised bounds displayed in Fig. 5.3, one can observe how the uncertainties associated to the fit coefficients are in general reduced in the quadratic fit in comparison to the linear one. The 95% CL interval is disjoint for the Yukawa coefficients $c_{b\varphi}$ and $c_{\tau\varphi}$ in the quadratic fit, with both a SM-like solution and a second one far from the SM. For the linear fit, we find that all EFT coefficients agree with the SM expectation at the 95% CL level. For the quadratic fit instead, this is not the case for two of the two-light-two-heavy coefficients, specifically for $c_{tu}^1$ and $c_{qu}^8$. These deviations from the SM predictions can be traced back to the quadratic corrections entering the high-energy $t\bar{t}$ differential cross-sections.

The global fit results of Fig. 5.3 are further scrutinized in Fig. 5.4, which displays both the magnitude of the 95% CL intervals and the 68% CL residuals compared to the SM hypothesis associated to the linear and quadratic EFT fits. In the upper panel, the horizontal line indicates the boundaries of the sampling volume used for the poorly-constrained coefficients as explained in Sect. 4.3. From these comparisons, one can observe how the inclusion of quadratic corrections leads to markedly more stringent bounds for most of the fit coefficients, a trend which is specially significant for the four-heavy (unconstrained in the linear fit) and two-light-two-heavy operators which modify the properties of the top quark. The only exception are the Yukawa coefficients, since there the quadratic corrections introduce a second degenerate solution thus enlarging the magnitude of the CL interval.

| Class | DoF | 95% CL bounds, $\mathcal{O}\left(\Lambda^{-2}\right)$ | | 95% CL bounds, $\mathcal{O}\left(\Lambda^{-4}\right)$, | |
| --- | --- | --- | --- | --- | --- |
| | | Individual | Marginalised | Individual | Marginalised |
| 4H | $c_{QQ}^1$ | [-5.906,23.733] | [-46.883,47.593] | [-2.329,2.026] | [-3.169,3.849] |
| | $c_{QQ}^8$ | [-27.763,44.356] | [-48.663,41.092] | [-6.904,5.881] | [-11.660,8.780] |
| | $c_{Qt}^1$ | [-48.237,44.311] | [-47.790,46.974] | [-1.851,1.862] | [-1.436,1.390] |
| | $c_{Qt}^8$ | [-6.151,20.010] | [-48.468,40.316] | [-4.249,3.362] | [-3.277,2.533] |
| | $c_{tt}^1$ | [-2.691,11.480] | [-25.102,46.406] | [-1.158,1.014] | [-0.869,0.785] |
| 2L2H | $c_{Qq}^{1,8}$ | [-0.181,0.134] | [-2.633,3.588] | [-0.238,0.120] | [-0.539,0.056] |
| | $c_{Qq}^{1,1}$ | [-0.318,0.327] | [-1.382,4.046] | [-0.134,0.140] | [-0.078,0.266] |
| | $c_{Qq}^{3,8}$ | [-0.640,0.539] | [-3.803,4.701] | [-0.346,0.286] | [-0.123,0.540] |
| | $c_{Qq}^{3,1}$ | [-0.031,0.038] | [-0.146,0.296] | [-0.033,0.039] | [-0.201,0.097] |
| | $c_{tq}^8$ | [-0.244,0.314] | [-2.713,2.579] | [-0.243,0.233] | [-0.355,0.388] |
| | $c_{tq}^1$ | [-0.258,0.295] | [-2.610,1.358] | [-0.144,0.102] | [-0.171,0.080] |
| | $c_{tu}^8$ | [-0.351,0.172] | [-4.801,6.450] | [-0.542,0.126] | [-0.357,0.430] |
| | $c_{tu}^1$ | [-0.136,-0.000] | [-0.910,0.105] | [-0.161,-0.006] | [-0.295,-0.104] |
| | $c_{Qu}^8$ | [-0.329,0.329] | [-2.801,5.902] | [-0.552,0.221] | [-1.010,-0.091] |
| | $c_{Qu}^1$ | [-0.138,0.750] | [-4.285,7.006] | [-0.117,0.174] | [-0.212,0.081] |
| | $c_{td}^8$ | [-0.696,0.407] | [-7.460,5.118] | [-0.790,0.264] | [-0.868,0.356] |
| | $c_{td}^1$ | [-0.640,0.466] | [-5.973,1.922] | [-0.281,0.201] | [-0.455,0.123] |
| | $c_{Qd}^8$ | [-0.478,0.484] | [-11.524,1.885] | [-0.778,0.325] | [-1.083,0.299] |
| | $c_{Qd}^1$ | [-0.426,0.608] | [-11.695,0.337] | [-0.156,0.285] | [-0.215,0.239] |
| 2F | $c_{t\varphi}$ | [-1.303,0.367] | [-4.821,2.493] | [-1.318,0.385] | [-3.488,1.882] |
| | $c_{tG}$ | [-0.018,0.029] | [-0.072,0.123] | [-0.018,0.030] | [-0.017,0.051] |
| | $c_{b\varphi}$ | [-0.006,0.041] | [-0.034,0.073] | [-0.007,0.036]∪[-0.404,-0.361] | [-0.029,0.052]∪[-0.428,-0.338] |
| | $c_{c\varphi}$ | [-0.022,0.116] | [-0.212,0.144] | [-0.002,0.376] | [-0.094,0.469] |
| | $c_{\tau\varphi}$ | [-0.026,0.035] | [-0.028,0.040] | [-0.024,0.040]∪[0.398,0.463] | [-0.023,0.039]∪[0.382,0.475] |
| | $c_{tW}$ | [-0.091,0.029] | [-0.287,0.142] | [-0.084,0.030] | [-0.243,0.111] |
| | $c_{tZ}$ | [-0.037,0.097] | [-15.088,5.794] | [-0.042,0.091] | [-1.121,0.985] |
| | $c_{\varphi l_1}^{(1)}$ | [-0.002,0.005] | [-0.144,0.159] | [-0.002,0.005] | [-0.051,0.095] |
| | $c_{\varphi l_1}^{(3)}$ | [-0.011,0.008] | [-0.029,0.039] | [-0.013,0.008] | [-0.043,0.020] |
| | $c_{\varphi l_2}^{(1)}$ | [-0.002,0.005] | [-0.144,0.159] | [-0.002,0.005] | [-0.051,0.095] |
| | $c_{\varphi l_2}^{(3)}$ | [-0.011,0.008] | [-0.029,0.039] | [-0.013,0.008] | [-0.043,0.020] |
| | $c_{\varphi l_3}^{(1)}$ | [-0.002,0.005] | [-0.144,0.159] | [-0.002,0.005] | [-0.051,0.095] |
| | $c_{\varphi l_3}^{(3)}$ | [-0.011,0.008] | [-0.029,0.039] | [-0.013,0.008] | [-0.043,0.020] |
| | $c_{\varphi e}$ | [-0.004,0.010] | [-0.288,0.318] | [-0.005,0.010] | [-0.102,0.190] |
| | $c_{\varphi \mu}$ | [-0.004,0.010] | [-0.288,0.318] | [-0.005,0.010] | [-0.102,0.190] |
| | $c_{\varphi \tau}$ | [-0.004,0.010] | [-0.288,0.318] | [-0.005,0.010] | [-0.102,0.190] |
| | $c_{\varphi q}^{(3)}$ | [-0.011,0.008] | [-0.029,0.039] | [-0.013,0.008] | [-0.043,0.020] |
| | $c_{\varphi Q}^{(3)}$ | [-0.104,0.102] | [-0.634,0.985] | [-0.103,0.095] | [-0.675,0.254] |
| | $c_{\varphi q}^{(-)}$ | [-0.069,0.051] | [-0.073,0.057] | [-0.077,0.052] | [-0.033,0.043] |
| | $c_{\varphi Q}^{(-)}$ | [-0.820,1.446] | [-3.253,7.999] | [-1.072,0.980] | [-1.339,1.813] |
| | $c_{\varphi u}$ | [-0.007,0.003] | [-0.212,0.192] | [-0.006,0.003] | [-0.127,0.068] |
| | $c_{\varphi d}$ | [-0.001,0.003] | [-0.096,0.106] | [-0.002,0.003] | [-0.034,0.063] |
| | $c_{\varphi t}$ | [-2.142,2.111] | [-4.611,12.532] | [-2.988,1.979] | [-3.306,3.637] |
| B | $c_{\varphi G}$ | [-0.002,0.005] | [-0.024,0.009] | [-0.002,0.005] | [-0.016,0.007] |
| | $c_{\varphi B}$ | [-0.005,0.002] | [-0.721,0.271] | [-0.005,0.002]∪[0.085,0.092] | [-0.094,0.100] |
| | $c_{\varphi W}$ | [-0.017,0.007] | [-0.430,0.328] | [-0.016,0.007] | [-0.144,0.257] |
| | $c_{\varphi WB}$ | [-0.069,0.051] | [-0.262,0.288] | [-0.077,0.052] | [-0.090,0.188] |
| | $c_{\varphi d}$ | [-0.263,1.128] | [-1.776,2.651] | [0.497,1.216]∪[-0.355,0.471] | [-1.424,1.340] |
| | $c_{\varphi D}$ | [-0.020,0.008] | [-0.636,0.575] | [-0.019,0.010] | [-0.381,0.203] |
| | $c_{WWW}$ | [-1.084,1.289] | [-1.080,1.384] | [-0.196,0.251] | [-0.183,0.233] |

**Table 5.4.** The 95% CL bounds for all the EFT coefficients considered in this analysis, for both individual and global (marginalised) fits obtained using either linear or quadratic EFT calculations.

**Figure 5.4.** The magnitude of the 95% CL intervals (top) and the value of the 68% CL residuals compared to the SM hypothesis (bottom panel) for the fit results displayed in Fig. 5.3.

The 68% CL residuals displayed in the bottom panel of Fig. 5.4 are defined by

$$R(c_i) \equiv \frac{(c_i|_{\mathrm{EFT}} - c_i|_{\mathrm{SM}})}{\delta c_i}, \qquad i = 1, \dots, n_{\mathrm{op}}, \tag{5.3}$$

with $c_i|_{\mathrm{EFT}}$ being the median of the posterior distribution from the EFT fit, $c_i|_{\mathrm{SM}} = 0$, and $\delta c_i$ is the total fit uncertainty for this parameter. We can observe that $|R(c_i)| \lesssim 1$ for most of the fit coefficients, both for the linear and quadratic cases, with exceptions in the case of the two-light-two-heavy coefficients, such as $c_{qd}^1$ in the linear fit and $c_{tu}^1$ in the quadratic one.

Nevertheless, for a large enough number of EFT coefficients one would expect a fraction of these residuals to be larger than unity, even if the SM is the underlying theory. Fig. 5.5 then displays the normalised distribution of these fit residuals. While these coefficients are correlated among them (see the following discussion) and thus cannot be treated as independent variables, the shapes of these distributions are reasonably close to a Gaussian, specially

55

**Figure 5.5.** The (normalised) distribution of the fit residuals shown in the bottom panel of Fig. 5.4.

for the linear fit, highlighting again the overall consistency of the fit results with the SM expectations.

**Correlations.** The correlation coefficient between any two fit coefficients $c_i$ and $c_j$ can be evaluated as follows,

$$\rho(c_i, c_j) = \frac{\left(\frac{1}{N_{\rm spl}} \sum_{k=1}^{N_{\rm spl}} c_i^{(k)} c_j^{(k)}\right) - \langle c_i \rangle \langle c_j \rangle}{\delta c_i \delta c_j}, \qquad i, j = 1, \ldots, n_{\rm op}, \tag{5.4}$$

where $N_{\rm spl}$ denotes the number of samples produced by NS, $\langle c_i \rangle$ indicates the mean value of this coefficient, and, as in Fig. 5.3, $\delta c_i$ is the corresponding uncertainty (standard deviation). The values of Eq. (5.4) are displayed in Fig. 5.6 separately for the linear and quadratic fits. We display only the numerical values for the pair-wise coefficient combinations for which the correlation coefficient is numerically significant, $|\rho(c_i, c_j)| \geq 0.5$. The pairs $(c_i, c_j)$ that do not appear in Fig. 5.6 have a correlation coefficient below this threshold.

We observe how the majority of the fit coefficients are loosely correlated among them, that is, their correlations being $|\rho| \leq 0.5$. One also finds that while several of the two-light-two-heavy coefficients turn out to be strongly correlated at the linear EFT level, this pattern disappears once the quadratic corrections are accounted for. Concerning the two-fermion operators, the correlation patterns present at the linear level are also often reduced in the quadratic fits. For instance, $c_{tZ}$ displays a strong correlation with $c_{\varphi B}$ at the linear level which is then washed out by the quadratic effects. The purely bosonic operators exhibit in general more stable correlations, for example $c_{\varphi WB}$ is strongly anti-correlated with $c_{\varphi D}$ in a manner which is similar in the linear and the quadratic fits. Furthermore, we do not find any pair of fit coefficients where the quadratic corrections flip the sign of their correlation.

In general, from Fig. 5.6 one can conclude that only a moderate subset of Wilson coefficients end up being strongly (anti-)correlated among them after the fit, specially so once
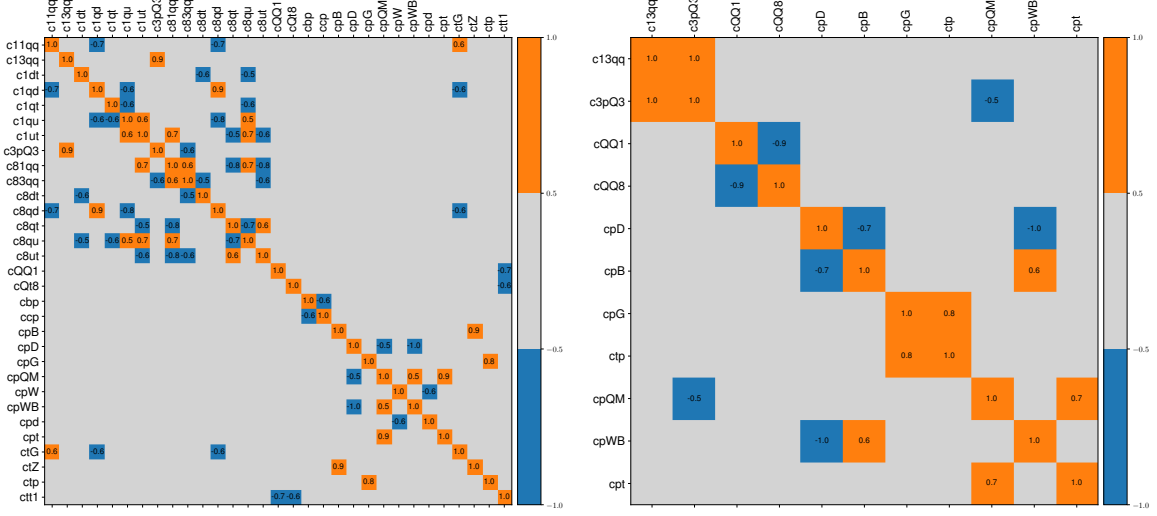
**Figure 5.6.** The correlation coefficients $\rho(c_i, c_j)$ between the EFT coefficients in the linear (left) and quadratic (right panel) fits. We only display the entries with significant (anti-)correlation, $|\rho| \geq 0.5$. Pairs of coefficients $(c_i, c_j)$ that do not displayed here have a correlation coefficient below this threshold.

quadratic EFT corrections are taken into account. This finding is partially explained by our wide input dataset, which makes possible constraining independently most if not all the EFT degrees of freedom.

**Individual fits.** As motivated in Sect. 4.2, individual (one-parameter) fits have several important applications. These include representing a benchmark reference for the global fit results, where the obtained bounds can only loosen as compared to one-parameter fits. The 95% CL bounds associated to the one-parameter linear and quadratic EFT fits were reported in Table 5.4, and the corresponding graphical comparison with the marginalised global fit results is displayed in Fig. 5.7.

Considering first the results of the linear fits, one can observe how in all cases the individual bounds are tighter (or at most comparable) than the marginalised ones by a large amount, around a factor ten or more in most cases. These differences are particularly striking for some of the two-fermion operators, in particular for $c_{tZ}$, as well as for bosonic operators such as $c_{\varphi B}$ and $c_{\varphi W}$, for which the differences between the individual and marginalised results can be as large as two orders of magnitude. Specifically, in the cases of $c_{tZ}$ and $c_{\varphi B}$, the 95% CL intervals are increased as follows when going from the individual to the marginalised fits:

$$c_{tZ}: \quad [-0.04, 0.010] \ \text{(individual)} \quad \text{vs} \quad [-15, 5.8] \ \text{(marginalised)},$$
$$c_{\varphi B}: \quad [-0.005, 0.002] \ \text{(individual)} \quad \text{vs} \quad [-0.7, 0.3] \ \text{(marginalised)}.$$

This effect clearly emphasizes the importance of adopting a fitting basis as wide as possible, in order to avoid obtaining artificially stringent bounds simply because one is being blind to other relevant directions of the parameter space. One important exception of this rule would be those cases where one is guided by specific UV-complete models, which motivate

**Figure 5.7.** Comparison of the magnitude of 95% CL intervals in the global (marginalised) and individual fits at the linear (top) and quadratic (bottom) level, see also Table 5.4.

the reduction in the parameter space to a subset of operators. We also note that the triple gauge operator $c_W$ is one of the few coefficients whose individual and marginalised bounds are identical: this can be traced back to the fact that this operator is very weakly correlated with other coefficients (see also Fig. 5.6), being constrained exclusively by the diboson data.

Inspection of the corresponding results from the quadratic fits, bottom panel of Fig. 5.7, reveals that the differences between individual and marginalised bounds are in general smaller as compared to the linear case. This effect is particularly visible for the two-light-two-heavy and the four-heavy operators, for which one finds that the individual fits underestimate the magnitude of the 95% CL interval by around a factor two on average, rather than by a factor 10 as in the linear case. The situation is instead similar to the linear fits for the two-fermion and the purely bosonic operators, and for example now also for $c_{tZ}$, $c_{\varphi B}$ and $c_{\varphi W}$ one finds large differences between marginalised and individual fits. One should point out, however, that even on those cases where the magnitude of the bound does not vary much, the central

**Figure 5.8.** Representative results for two-parameter fits carried out at linear order in the EFT. We display the 95% CL ellipses obtained for different data subsets and for the complete dataset, labelled as "All Data (2D)". For reference, we also show the marginalised bounds obtained in the global fit. The black square in the center of the plot indicates the SM value.

best-fit values can still shift in a non-negligible manner. This is illustrated by the two-light-two-heavy operators $c_{tu}^1$ and $c_{Qu}^8$, for which one obtains

$$c_{tu}^1 : \qquad [-0.16, -0.01] \text{ (individual)} \qquad \text{vs} \qquad [-0.30, -0.10] \text{ (marginalised)} ,$$
$$c_{Qu}^8 : \qquad [-0.55, 0.22] \text{ (individual)} \qquad \text{vs} \qquad [-1.01, -0.09] \text{ (marginalised)} ,$$

hence for instance while $c_{tu}^1$ agrees with the SM at the 95% CL in the individual fit, this is not the case in the marginalised one. From this comparison, one can conclude that for the two-light-two-heavy and the four-heavy operators most of the sensitivity is already captured by the individual fits once quadratic corrections are included.

**Two-parameter fits.** To complement the insights provided by individual fits, it can also be instructive to carry out two-parameter fits, specially to investigate the relative interplay between specific pairs of EFT coefficients. In such fits, two coefficients are allowed to vary simultaneously while the rest are set to zero. To illustrate the information that can be provided by such two-parameter fits, Fig. 5.8 displays representative results for fits performed at the linear order. We display the 95% CL ellipses obtained when different subsets of data are used as input, as well as for the complete dataset, labelled as "All Data (2D)". For reference, we also show here the marginalised bounds obtained from the global fit.

59

To begin with, the upper panels of Fig. 5.8 display two-parameter fits for the three possible pair-wise combinations of the $c_{t\varphi}$, $c_{tG}$, and $c_{\varphi G}$ coefficients, which connect Higgs production in gluon fusion with top quark pair production, see also the Fisher information table of Fig. 3.1. These comparisons illustrate the relative impact of the various dataset in constraining each coefficient. For example, from the $(c_{t\varphi}, c_{tG})$ fit we see that the sensitivity of $c_{tG}$ is driven by $t\bar{t}$ data, while the Higgs measurements have a flat direction resulting in a elongated ellipse. When top and Higgs data are combined, the resulting 95% CL ellipse is much smaller as compared to the results obtained separately from the two groups of processes. Note that, as in the case of the individual fits reported in Fig. 5.7, also for two-parameter fits the obtained bounds are more stringent as compared to the global marginalised results. Similar considerations apply to the $(c_{\varphi G}, c_{tG})$ fit, while from the $(c_{\varphi G}, c_{t\varphi})$ one learns that the sensitivity is still dominated by the Higgs signal strengths rather than by the differential cross-section measurements.

Then the bottom panels of Fig. 5.8 display two-parameter fits involving the two-light-two-heavy coefficients $c_{Qq}^{1,8}$, $c_{Qq}^{3,8}$, $c_{tu}^{8}$, $c_{td}^{8}$, and $c_{tq}^{8}$, all of which are constrained mostly from top quark pair differential distributions as indicated by the Fisher information matrix. Here the scope is to illustrate the relative sensitivity provided by some of the $t\bar{t}$ datasets that enter the fit: single-inclusive $m_{t\bar{t}}$ distributions, the double-differential $(m_{t\bar{t}}, y_{t\bar{t}})$ distributions, and $t\bar{t}V$ measurements. The results confirm both that the $m_{t\bar{t}}$ distributions completely dominate the fit of these coefficients, and that the marginalised CL ellipses are rather broader than for the two-dimensional fits. The latter is again in agreement with the results of the individual linear fits, reported in the upper panel of Fig. 5.7.

## 5.3 Dataset dependence

The discussion so far has focused on the output of the global fits obtained for the baseline dataset summarised in Tables 3.1 to 3.4. Here we aim to assess the dependence of these results with respect to the choice of input dataset. With this purpose, we consider here fits for the following variations:

- A fit which includes only top quark measurements. This fit makes possible quantifying the interplay between the top and the Higgs data in the global fit.

- A fit which includes only Higgs boson production and decay data, which provides complementary information as compared to the top-only fit.

- A fit which includes only top quark measurements, but now restricted to the same dataset as in our original study from [7]. This comparison allows one to assess the impact in the top-only EFT fit of the new LHC top quark measurements that have become available in the last two years.

- A fit where the diboson data is removed, to determine how much weight the diboson cross-sections carry in the global fit results.

- A fit where all high-energy bins, defined as those bins probing the region $E \gtrsim 1$ TeV, are removed. The motivation for such a fit is to study how important are the constraints provided by the high-energy region in the global fit results, which in turn is an important input to establish the validity of the EFT approximation.

- A fit where those datasets displaying poor agreement with the SM cross-sections are removed. Specifically, here one removes the datasets whose $\chi^2$ differs by more than $3\sigma$

| Dataset | $n_{\mathrm{dat}}$ | $\chi^2_{\mathrm{SM}}$ | $\chi^2_{\mathrm{eft}}$ | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | | baseline | top-only (2021) | top-only (2018) | Higgs-only | diboson excluded | high-$E$ excluded | poor $\chi^2_{\mathrm{sm}}$ excluded |
| $t\bar{t}$ incl. | 94 | 1.36 | 1.23 | 1.24 | 1.23 (63) | | 1.23 | 1.28 (78) | 0.95 (70) |
| $t\bar{t}V$ | 14 | 0.65 | 0.68 | 0.65 | 0.68 (8) | — | 0.68 | 0.68 | 0.67 |
| single-$t$ incl. | 27 | 0.43 | 0.43 | 0.43 | 0.36 (22) | — | 0.43 | 0.43 | 0.43 |
| $tV$ | 9 | 0.71 | 0.61 | 0.58 | 0.76 (6) | — | 0.61 | 0.63 | 0.29 (8) |
| $t\bar{t}Q\bar{Q}$ | 6 | 1.68 | 1.91 | 2.03 | 4.38 (2) | — | 1.93 | 1.94 | 1.26 (4) |
| **Top total** | 150 | 1.10 | 1.02 | 1.03 | 1.01 (101) | — | 1.03 | 1.02 | 0.77 (123) |
| Higgs $\mu_i^f$ (RI) | 22 | 0.86 | 0.89 | — | — | 0.84 | 0.89 | 0.90 | 0.90 |
| Higgs $\mu_i^f$ (RII) | 40 | 0.68 | 0.62 | — | — | 0.68 | 0.62 | 0.62 | 0.62 |
| Higgs STXS | 35 | 0.88 | 0.83 | — | — | 0.82 | 0.83 | 0.84 | 0.83 |
| **Higgs total** | 97 | 0.79 | 0.76 | — | — | 0.75 | 0.76 | 0.76 | 0.76 |
| **Diboson** | 70 | 1.31 | 1.32 | — | — | — | — | 1.32 | 1.32 |
| **Total $n_{\mathrm{dat}}$** | 317 | 317 | 317 | 150 | 101 | 97 | 247 | 301 | 290 |
| **Total $\chi^2$** | — | **1.05** | **1.01** | **1.03** | **1.01** | **0.75** | **0.92** | **1.01** | **0.90** |

**Table 5.5.** Same as Table 5.3 for EFT fits obtained from variations of the baseline dataset. We list the results of the following fits: including only top quark measurements (either for the 2018 or the current dataset); a Higgs-only dataset; without the diboson cross-sections; with the high-energy bins excluded; and with the datasets with a poor $\chi^2_{\mathrm{sm}}$ excluded. In all cases, the quadratic EFT corrections are accounted for. The numbers in parentheses indicate the number of data points, in the case that these are different from those of the baseline settings (listed in the second column).

from their statistical expectation assuming the SM hypothesis. While such disagreements between data and SM theory could very well indicate hints of BSM physics, they can also be explained by for example issues with the experimental correlation models. Hence, this fit allows us to verify to which extent the global fit results are determined from the datasets that disagree the most with the SM predictions.

Note that, as explained in Sect. 3, for the purposes of categorisation into datasets the $t\bar{t}h$ cross-sections are considered part of the Higgs measurements. Furthermore, we note that all these fits are based on quadratic EFT calculations and that the constraints provided by the EWPOs on the EFT parameter space are always accounted for.

To begin with, Table 5.5 collects the values of the $\chi^2$ per data points for EFT fits obtained from variations of the input dataset. We list the results of the various fits described above: including only top quark measurements (either from the current or the 2018 dataset); with a Higgs-only dataset; without the diboson cross-sections; with the high-energy bins excluded; and with the datasets with a poor $\chi^2_{\mathrm{sm}}$ excluded. The numbers in parentheses indicate the number of data points, in the case that these are different from those of the baseline settings listed in the second column. We observe how the description of the Higgs cross-sections is essentially unaffected in these fits with reduced datasets. Concerning the total $\chi^2$ for the top data, we see that it is stable in the fit where the high-energy bins are removed, but that is markedly improved (from 1.02 to 0.77) in the fit where the datasets with poor $\chi^2_{\mathrm{sm}}$ are excluded.

Then in Fig. 5.9 we compare the magnitude of the 95% CL bounds, same as in the upper panel of Fig. 5.4, between the global fit results with those obtained in the top-only and Higgs-

only fits. As mentioned above, these fits allow us to assess the interplay between the top and the Higgs data in the global analysis, in other words, to identify what are the main benefits of the simultaneous mapping of the EFT parameter space as compared to carrying out separate fits to each group of processes. First of all, we note that the global fit bounds are more stringent for all the EFT coefficients than in either the top-only or Higgs-only fit, highlighting the overall consistency of the two datasets. Secondly, the cross-talk of the top and Higgs data is found to be most relevant for the two-fermion coefficients $c_{\varphi t}$ and $c_{\varphi Q}^{(-)}$, whose bounds are improved by around a factor 2 in the global fit as compared to the top-only fit. Another operator that benefits from the global fit is $c_{\varphi G}$, which is unconstrained in the top-only fit but bound in the global fit is clearly improved as compared to the Higgs-only fit. This study shows how by breaking degeneracies one gains information in the global fit as compared to the partial ones, sometimes in unexpected directions in the parameter space such as for $c_{\varphi G}$ in this case.

Next, Fig. 5.10 displays a similar comparison as in Fig. 5.9 now comparing first the outcome of the global fit with that of a fit where the diboson cross-sections have been removed, and second comparing two top-only fits, namely the fit displayed in the upper panel of Fig. 5.9 with a fit based on the same dataset as our previous study from [7]. The fit without diboson data demonstrates that the constraints provided by the diboson cross-sections are negligible in comparison with those provided by the Higgs data (and the EWPOs) for all coefficients considered in the fit, except for the triple gauge operator $c_W$. This result is consistent with the Fisher information analysis of Fig. 3.1, and indicates that, apart from $c_W$, the diboson data does not provide competitive information on the EFT parameter space in the context of a global fit.

The comparison of the two top-only fits in the bottom panel of Fig. 5.10 illustrates how for all coefficients the bounds are improved thanks to the more recent LHC measurements.[5] The improvement is consistent across the board, quantifies the additional information brought in by the new top-quark cross-section measurements (see Table 5.5) and confirms that the broader and more diverse the input dataset is, the more stringent the resulting constraints on the EFT parameter space that will be obtained.

To conclude this discussion of the dataset dependence of our results, we consider the outcome of two more fits: first, one where the datasets exhibiting poor agreement with the SM predictions are excluded, and second, another where all bins sensitive to the high-energy region, defined as $E \gtrsim 1$ TeV, are removed. The best-fit values and 95% CL intervals of these two fits are compared with the baseline results in Fig. 5.11. As indicated in Table 5.5, in the fit where those datasets with poor $\chi^2_{\rm sm}$ have been removed, one is essentially cutting away 27 points from top quark production, mostly from the inclusive $t\bar{t}$ category. The only coefficients that are affected by this reduction in the dataset are some of the two-light-two-heavy operators, whose bounds are somewhat enlarged consistently with the loss of experimental information. Specifically, this has the consequence that the 95% CL intervals for $c_{tu}^1$ and $c_{Qu}^8$ now agree with the SM expectation, as opposed to the global fit configuration.
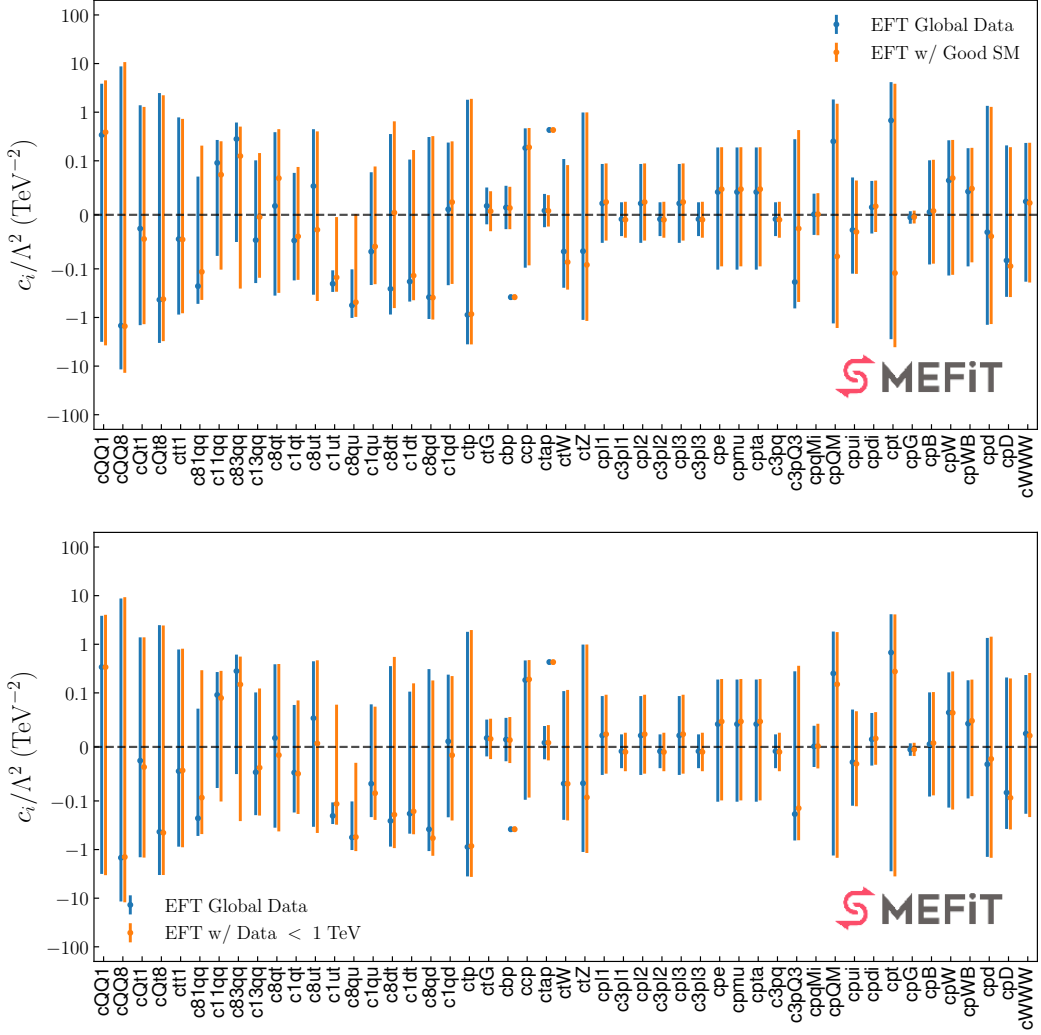
Concerning the outcome of the fit without the high-energy bins, as expected the only differences are observed again for the two-light-two-heavy coefficients. The results are similar as in the previous case: an increase in the coefficient uncertainties and an improved agreement of $c_{tu}^1$ and $c_{Qu}^8$ with the SM predictions. From these comparisons, one can conclude that the global fit is not dominated by the high-energy regions where the EFT validity could be

---

[5]Recall that now we consider the top Yukawa coefficient $c_{t\varphi}$ as part of the Higgs dataset.

**Figure 5.9.** Same as upper panel of Fig. 5.4 now comparing the global fit results with those obtained in a top-only (upper) and Higgs-only (lower panel) fits.

questioned, except for some of the two-light-two-heavy coefficients. Hence, we have established that the large residuals reported in Fig. 5.4 for $c_{tu}^1$ and $c_{Qu}^8$ arise from fitting the large-$m_{t\bar{t}}$ bins of the top quark pair production datasets that, to begin with, are described sub-optimally by the SM predictions.

## 5.4 Impact of NLO QCD corrections in the EFT cross-sections

In addition to the choice of input dataset, another important factor that determines the outcome of a global analysis such as the present one is the accuracy of the EFT theoretical calculations. Here we assess the role played at the level of the fit results by the inclusion of NLO QCD corrections to the EFT cross-sections. As indicated in Table 3.9, our baseline fit includes these NLO corrections to the EFT calculations whenever available, so now we switch

**Figure 5.10.** Same as Fig. 5.9, now comparing the global fit with a no-diboson fit (upper) and the two top-only fits with different datasets (lower panel).

them off deliberately to quantify how much they affect the fit outcome.[6] In the following, the theoretical predictions for the SM cross-sections, based on the state-of-the-art calculations, remain unchanged, and only the EFT ones are modified as compared to the baseline settings.

Table 5.6 compares the values of the $\chi^2$ for the various groups of processes in quadratic fits with and without NLO QCD corrections to the EFT cross-sections, as well as for the associated SM results. One can observe how the fit quality worsens if NLO QCD effects are not accounted for, with the total dataset (top-quark dataset) increasing to $\chi^2 = 1.05$ (1.11) from $\chi^2 = 1.01$ (1.02) in the baseline. On the other hand, the fit quality is essentially unchanged for the Higgs and diboson datasets, irrespective of whether or not NLO QCD corrections are accounted for. One can therefore trace back this increase in the $\chi^2$ values when LO EFT calculations are used to the top-quark pair distributions, which contribute to

---

[6]This study is also motivated by the fact that many EFT fits rely on LO QCD for the EFT cross-sections.

**Figure 5.11.** Same as Fig. 5.3 comparing the global fit results with those of the fit excluding datasets with poor $\chi^2_{\mathrm{sm}}$ (upper) and with the fit where the bins with $E \gtrsim 1$ TeV are removed (bottom panel).

almost one third of the total dataset.

Fig. 5.12 then displays the posterior probability distributions and the corresponding 95% CL intervals for the Wilson coefficients, comparing the results of quadratic fits with and without NLO corrections to the EFT cross-sections. One can observe that these posterior distributions can be severely distorted when LO EFT calculations are used as compared to the baseline, for instance in terms of a shift in the best-fit values or due to an increase in the width of the distributions. As an illustration, the operator that modifies the charm Yukawa interaction, $c_{c\varphi}$, exhibits a bimodal distribution once NLO effects are accounted for, while the dominant solution for the $c_{\varphi t}$ coefficient is far from the SM in the LO fit but SM-like in the NLO case. This comparison highlights how the inclusion of NLO QCD corrections is necessary to achieve robust results for the fitted EFT coefficients. Alternatively, one could account for the missing higher-order uncertainties (MHOUs) in the EFT cross-sections, which

65
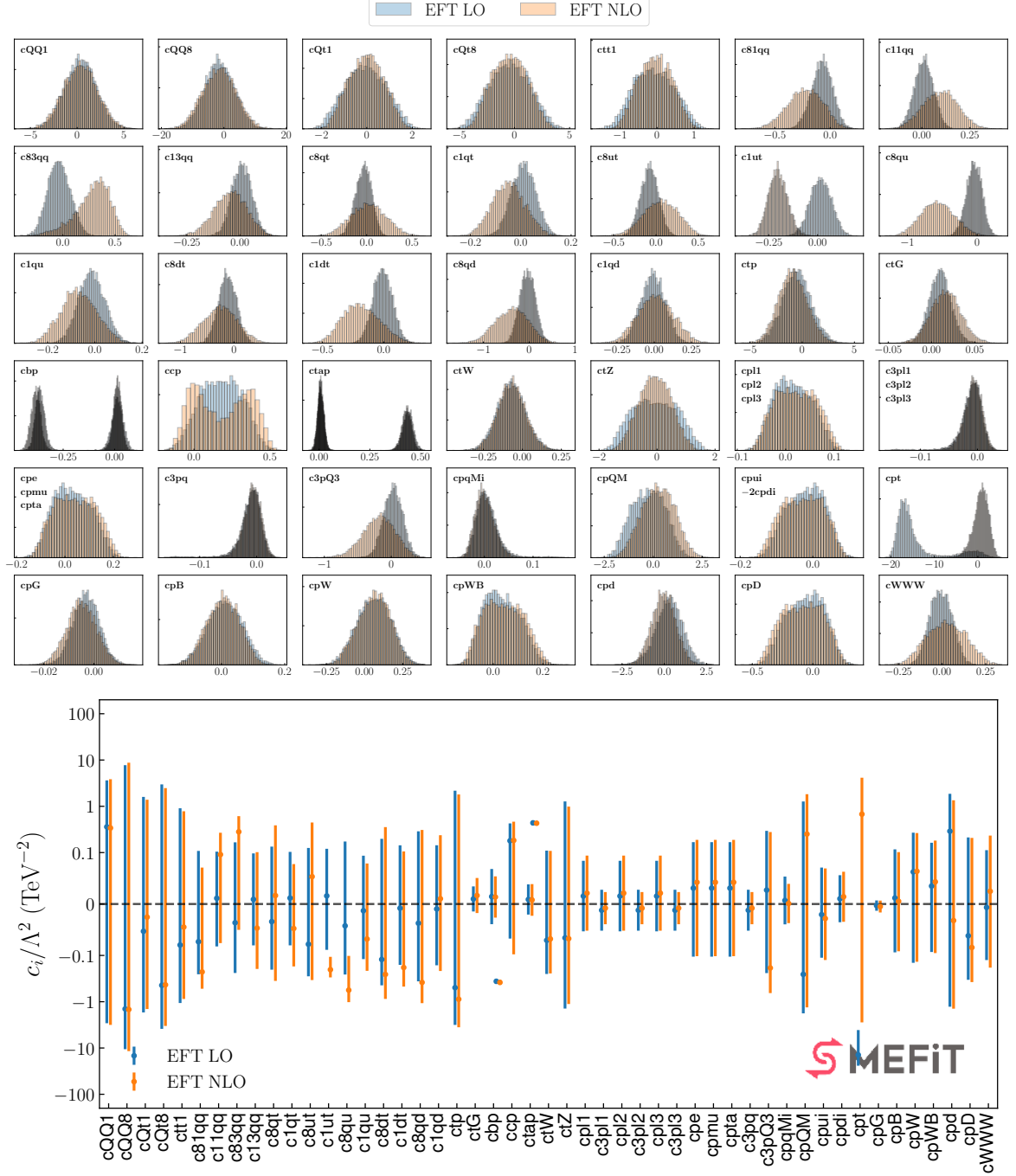
**Figure 5.12.** Top: comparison of the posterior probability distributions of the Wilson coefficients between quadratic fits with and without NLO QCD corrections to the EFT cross-sections. Bottom: the corresponding 95% CL intervals, compared to the SM expectation.

| Dataset | $n_{\mathrm{dat}}$ | $\chi^2_{\mathrm{SM}}$ | $\chi^2_{\mathrm{EFT}}$ (baseline) | $\chi^2_{\mathrm{EFT}}$ (LO QCD in EFT) | $\chi^2_{\mathrm{EFT}}$ (top-philic) |
|---|---|---|---|---|---|
| $t\bar{t}$ inclusive | 94 | 1.36 | 1.23 | 1.37 | 1.29 |
| $t\bar{t}V$ | 14 | 0.65 | 0.68 | 0.58 | 0.68 |
| single top inclusive | 27 | 0.43 | 0.43 | 0.43 | 0.43 |
| $tV$ | 9 | 0.71 | 0.61 | 0.72 | 0.75 |
| $t\bar{t}Q\bar{Q}$ | 6 | 1.68 | 1.91 | 1.90 | 1.88 |
| **Top quark total** | 150 | 1.10 | 1.02 | 1.11 | 1.07 |
| Higgs $\mu_i^f$ (Run I) | 22 | 0.86 | 0.89 | 0.88 | 0.89 |
| Higgs $\mu_i^f$ (Run II) | 40 | 0.68 | 0.62 | 0.64 | 0.62 |
| Higgs differential & STXS | 35 | 0.88 | 0.83 | 0.81 | 0.83 |
| **Higgs total** | 97 | 0.79 | 0.76 | 0.76 | 0.76 |
| **Diboson** | 70 | 1.31 | 1.32 | 1.32 | 1.31 |
| **Global dataset** | **317** | **1.05** | **1.01** | **1.05** | **1.03** |

**Table 5.6.** Same as Table 5.3 now for fits based on variations of the theory settings as compared to the baseline ones. Specifically, we provide the results of a fit where the EFT cross-sections are evaluated at LO in the QCD expansion, as well as those of the top-philic scenario where the parameter space has been restricted as described in Sect. 2.2. In both cases, quadratic EFT corrections are being included. Note that the SM cross-sections are always evaluated using state-of-the-art theory calculations.

are usually neglected, using for instance the approach advocated in [133, 134]. Implementing MHOUs systematically is expected to improve the overall compatibility of EFT fits performed with and without NLO QCD corrections.

Interestingly, for specific operators such as the two-light-two-heavy coefficients and the triple gauge operator $c_W$, neglecting NLO corrections leads to both a shift in the best-fit values but also to somewhat reduced 95% CL intervals. Such behaviour may arise when NLO contributions to the quadratic EFT cross-sections introduce novel new directions in the parameter space, for example with coefficients that are absent at the LO level. Nevertheless, only once all sources of theory uncertainty are included it will be possible to ascertain in an unambiguous manner whether or not NLO QCD corrections lead to more precise (in addition to more accurate) global EFT analyses.

In the specific case of the $c_{\varphi t}$ coefficient, one can verify that the corresponding individual $\chi^2$ profile (analog of Fig. 4.2 for LO fits) does not exhibit the second solution identified in Fig. 5.12, and hence it must be induced by the cross-talk with other coefficients in the fit. To validate this hypothesis, Fig. 5.13 displays the outcome of two-parameter quadratic fits for $(c_{\varphi t}, c_{tZ})$ and $(c_{\varphi t}, c_{\varphi W})$ comparing the results of the LO EFT fit with its NLO counterpart. In both cases, the LO two-parameter fits based on the full dataset favour the solution far from the SM, while the NLO ones instead favour the SM-like one. The explanation for this behaviour can be traced back to the fact that the non-SM solution is disfavored by higher-order EFT corrections to $hZ$ associated production, in particular the loop-induced gluon-gluon contributions.

Another remarkable effect of the NLO QCD corrections to the EFT cross-sections can be
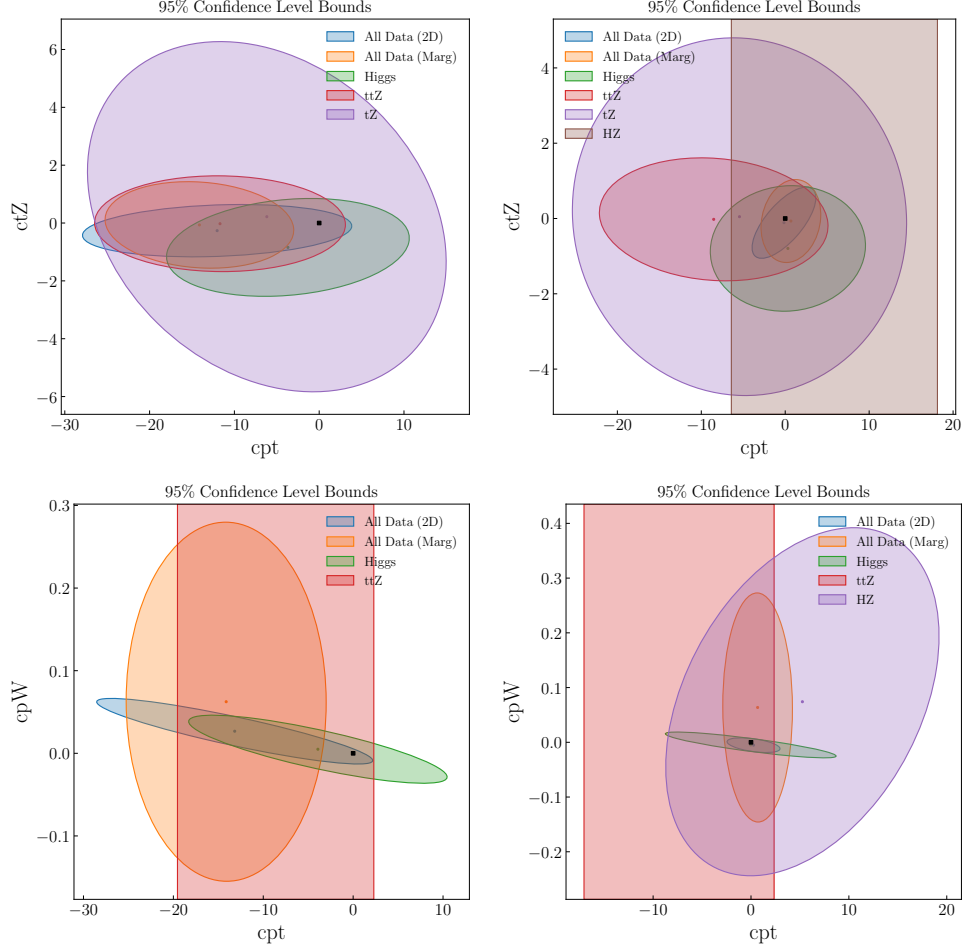
**Figure 5.13.** Same as Fig. 5.8 for the two-parameter quadratic fits of $(c_{\varphi t}, c_{tZ})$ (upper) and $(c_{\varphi t}, c_{\varphi W})$ (lower panels) comparing the results of the LO EFT fit (left) with its NLO counterpart (right panels).

observed as a modification of the correlation patterns. Fig. 5.14 displays the same correlations maps as in Fig. 5.6 now for global fits based on LO EFT calculations at the linear and quadratic level. Specially for the linear fits, we observe that correlations become more sizable in general (higher fraction of colored entries) in the LO case, while these are reduced once NLO corrections are accounted for. This feature demonstrates how NLO QCD effects may reduce parameter correlations by introducing additional sensitivity to the fit coefficients given the same input dataset.

## 5.5 The top-philic scenario

To conclude this section, we present results for a global EFT fit carried out in the top-philic scenario defined in Sect. 2.2. In this scenario, we have the 9 equations of Eq. (2.10) that relate a subset of the 14 two-heavy-two-light coefficients listed in Table 2.5 among them, leaving 5 independent parameters to be constrained in the fit. Given the more constraining
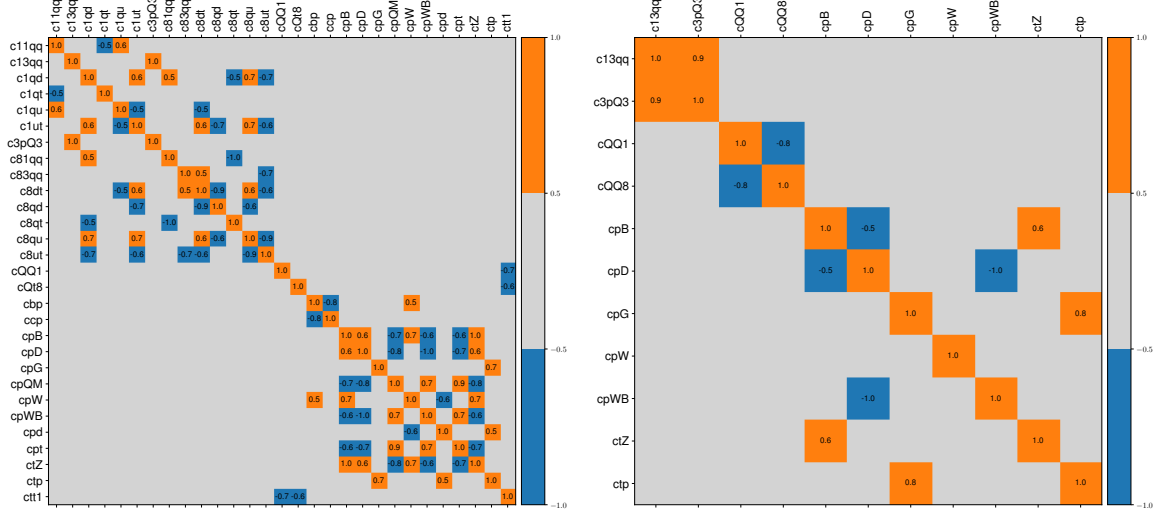
**Figure 5.14.** Same as Fig. 5.6 for LO EFT calculations in linear (left) and quadratic (right) fits.

assumptions associated to the top-philic scenario, one expects to find an improvement in the bounds of the two-light-two-heavy EFT operators due to the fact that the parameter space is being restricted by theoretical considerations, rather than by data in this case.

The values of the $\chi^2$ for each group of datasets in the top-philic scenario were reported in Table 5.6. One finds that the $\chi^2$ to the top quark cross-sections is somewhat worsened with respect to the baseline fit (1.07 versus 1.02), while remaining better than in the SM case. The same considerations apply at the level of the total dataset, driven by the different fit quality to the top quark observables, while the Higgs and diboson data are unaffected. This slightly inferior description of the top quark experimental data in the top-philic scenario can be explained by the reduction in the EFT parameter space, which restricts the number of independent directions that can be explored by the fit. As we show below, the main benefit of the top-philic scenario is instead that, within its assumptions, one achieves more stringent limits for many of the operators that modify top quark properties.

Fig. 5.15 displays the 95% CL intervals for the EFT coefficients comparing the global fit results with those of the top-philic scenario. The only operators that are affected in a significant manner turn out to be the two-light-two-heavy operators, with the bounds in several of them such as $c_{td}^1$, $c_{Qq}^{1,1}$, and $c_{tq}^1$ improving by almost an order of magnitude. The fact that only the bounds on the two-light-two-heavy operators are modified is consistent with the top-philic scenario, given that only this specific group of EFT coefficients is being constrained by its model assumptions.

It is worth emphasizing at this point that, from the technical point of view, carrying out global EFT fits with specific restrictions in the parameter space motivated by UV-completions, such as those arising in the top-philic scenario and leading to Fig. 5.15, is relatively straight-forward. Indeed, the most efficient fitting strategy would be to start from the broadest possible parameter space, and once the corresponding fit has been performed, introduce model assumptions relating EFT coefficients in a systematic manner. This way one can connect with specific models for UV-completions of the SM, which typically result in a rather smaller

**Figure 5.15.** Same as Fig. 5.2 comparing the global fit results with the same fit in the top-philic scenario defined by the relations in Eq. (2.10).

number of EFT coefficients to be constrained from data.

# 6  Summary and outlook

In this work we have presented an extensive interpretation of Higgs, diboson, and top quark measurements from the LHC in the framework of the Standard Model Effective Field Theory. By combining the most updated experimental data with state-of-the-art theory calculations, both in the SM and in the EFT, we have provided bounds on 50 directions in the SMEFT parameter space of which 36 correspond to independent parameters. We have quantified in detail the relative impact that the different types of processes have in the results of this global EFT analysis, both in terms of fits with dataset variations and by means of statistical diagnosis tools such as information geometry techniques and principal component analysis. Our analysis highlights the overall complementarity of the various input processes, further motivating the need for a global interpretation of LHC measurements. We have also demonstrated how, within such a global EFT analysis, genuinely flat directions are essentially absent since each process and each kinematic bin of a distribution constrains separate combinations of the fit parameters. The robustness of our fitting methodology has been cross-validated by deploying two completely independent methods, MCfit and NS, for mapping the EFT parameter space.

We have also extensively quantified the role played in the global analysis by the inclusion of NLO QCD corrections to the EFT cross-sections, whose automation has been recently achieved. We find that the posterior probability distributions of the fit parameters can be modified in non-trivial ways by these NLO QCD effects, shifting the best-fit value, modifying the magnitude of the 95% CL intervals, and even inducing multi-modal distributions. These findings demonstrate that available LHC data is already sensitive to NLO effects in the EFT cross-sections, further highlighting the importance of accounting for them in a systematic

manner to achieve both accurate and precise results.

One could consider several directions in which the present study could be extended. To begin with, one would like to include directly the constraints provided by LEP's electroweak precision observables, rather than in an approximate manner as done in the present work. Furthermore, it would be beneficial to add to the global analysis new high-$p_T$ observables providing complementary information on the Higgs and gauge sectors of the SMEFT, such as for instance vector boson scattering (VBS), $Z$ production in vector boson fusion (VBF), or high-mass Drell-Yan production, all of them constraining the electroweak interactions. Other processes that one might consider in this context are single-inclusive jet, dijet, and multijet production, which are sensitive to several directions in the parameter space not covered by other processes, specifically to a large number of four-fermion operators.

From the point of view of theoretical calculations, it would be important to systematize the study of higher-order terms in the EFT expansion, considering in particular double insertions of dimension-six operators and representative subsets of dimension-eight operators. We point out that, within our fitting methodology, accounting for these higher order terms is technically straightforward. Along the same lines, one could extend the fitting formalism to account for all sources of theoretical uncertainties and their correlations in a systematic manner, both for the SM and for the EFT calculations, something which is done only partially here. In addition, it should be interesting to develop statistically optimal observables for EFT analyses, such as those based on deep learning [141–143], making possible complementing the constraints obtained at the level of unfolded cross-sections with those extracted directly at the detector level.

Another promising research direction is that of combining the global EFT interpretation of high-$p_T$ observables at the LHC presented here with that of flavour data from LHCb and from other $B$-factories such as Belle. The urgency of a simultaneous EFT analyses of high-$p_T$ and flavour data has been further highlighted by the recent evidence reported by the LHCb experiment for the violation of lepton flavour universality (LFU) in $B$-meson decays [25]. These findings demand exploiting the flexibility of the EFT framework in order to comprehensively map the allowed signatures of eventual LFU violation in high-$p_T$ cross-sections at the LHC and elsewhere.

Likewise, it would be important to account for the constraints provided by low energy processes in the SMEFT parameter space, from neutrino data and electric dipole moment measurements to the anomalous magnetic moment of the muon. In the latter case, a $4.1\sigma$ deviation with respect to the SM expectation has recently been reported [144], confirming and strengthening one of the most puzzling anomalies in particle physics. Indeed, the ultimate goal of our program would be a truly global EFT interpretation including all processes sensitive to the sought-for UV completion of the SM, making sure no stone is left unturned in the ongoing quest to unravel the new particles and interactions that lie beyond the Standard Model.

# A    Comparison with experimental data

In this appendix, we present a systematic comparison between the experimental data used as input to the fit with the corresponding theoretical cross-sections based both on the SM and on the best-fit SMEFT results, either at linear $\mathcal{O}(\Lambda^{-2})$ or at quadratic $\mathcal{O}(\Lambda^{-4})$ accuracy. In these comparisons, the experimental measurements will be presented both in terms of unshifted central values, where the error band represents the total uncertainty, and once the best-fit systematic shifts have been subtracted, such that the error band contains only the statistical component. Note that the evaluation of the shifted data is only possible whenever the full breakup of the experimental systematic uncertainties is made available by in `HepData`. If this is not the case, for example when only the full experimental correlation matrix is provided or no information on correlations is released, we will display only the unshifted data.

To begin with, Fig. A.1 displays the comparison between experimental data and the best-fit EFT theory predictions (for linear and quadratic fits) in the case of representative differential top quark pair and single top quark production datasets. Both the data and the EFT fit results are normalised to the central value of the SM theory prediction. This implies that the more the fit results deviate from unity, the larger the best-fit EFT effects are for this specific observable. Furthermore, the error band in the EFT prediction indicates the associated 95% CL interval evaluated over the $N_{\rm spl}$ samples produced by the NS method.

From these comparisons, one can observe how for some datasets the best-fit EFT results move in the direction of the experimental data, for instance in the case of the $m_{t\bar{t}}$ distributions at large invariant masses for inclusive $t\bar{t}$ production. This is an important kinematic region in the fit, since energy-growing effects increase the EFT sensitivity. Interestingly, in the highest $m_{t\bar{t}}$ bins for some of the 13 TeV top datasets the 95% CL interval associated to the EFT prediction does not include the SM expectation. In the case of the single-top $t$-channel differential cross-sections, the EFT fit results are very close to the SM predictions, indicating that EFT effects are well constrained for this process at the scale of the present experimental uncertainties. We also note that the uncertainty band associated to the EFT prediction can turn out to be rather different in the $\mathcal{O}(\Lambda^{-2})$ fits as compared to the $\mathcal{O}(\Lambda^{-4})$ ones, with the latter in general being more precise than the former for the processes considered here.

Then Fig. A.2 displays the same comparison between data and the SM and EFT predictions as in Fig. A.1 now for the $W$ helicity fractions, the single-top $s$-channel and $tV$ total cross-sections, the four-heavy-quark fiducial measurements, the LHC diboson differen-
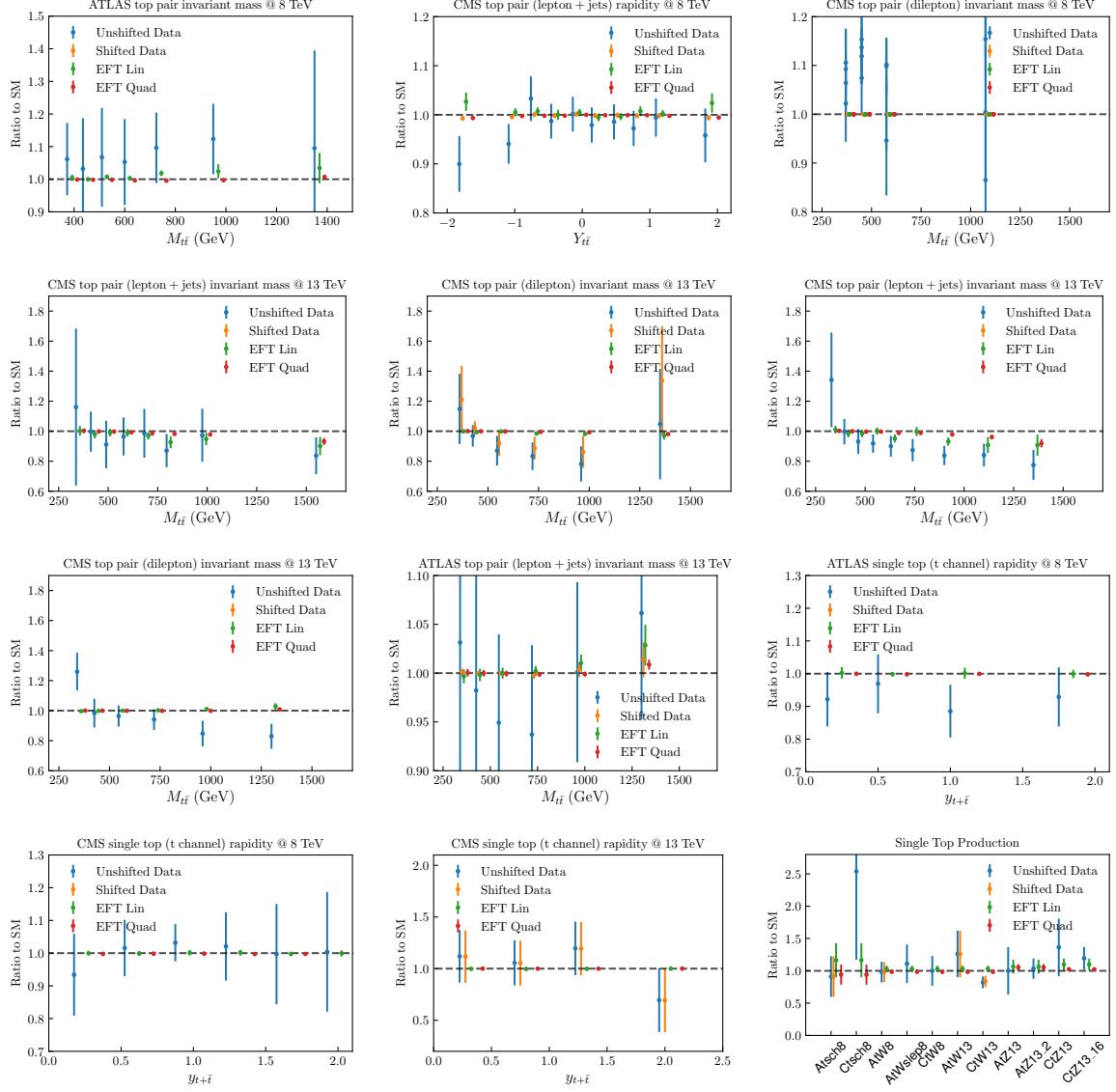
**Figure A.1.** Comparison between experimental data and best-fit EFT theory predictions (both for the linear and the quadratic fits) for representative differential top quark pair and single top quark production datasets. Both the data and the EFT fit results are normalised to the central value of the SM cross-section. The data is presented both with unshifted central values (where the band represents the total experimental error) and once the best-fit systematic shifts have been subtracted (so that the error band contains only the statistical component). The error band in the EFT prediction indicates the 95% CL interval evaluated over the $N_{\mathrm{spl}}$ samples produced by the NS method.
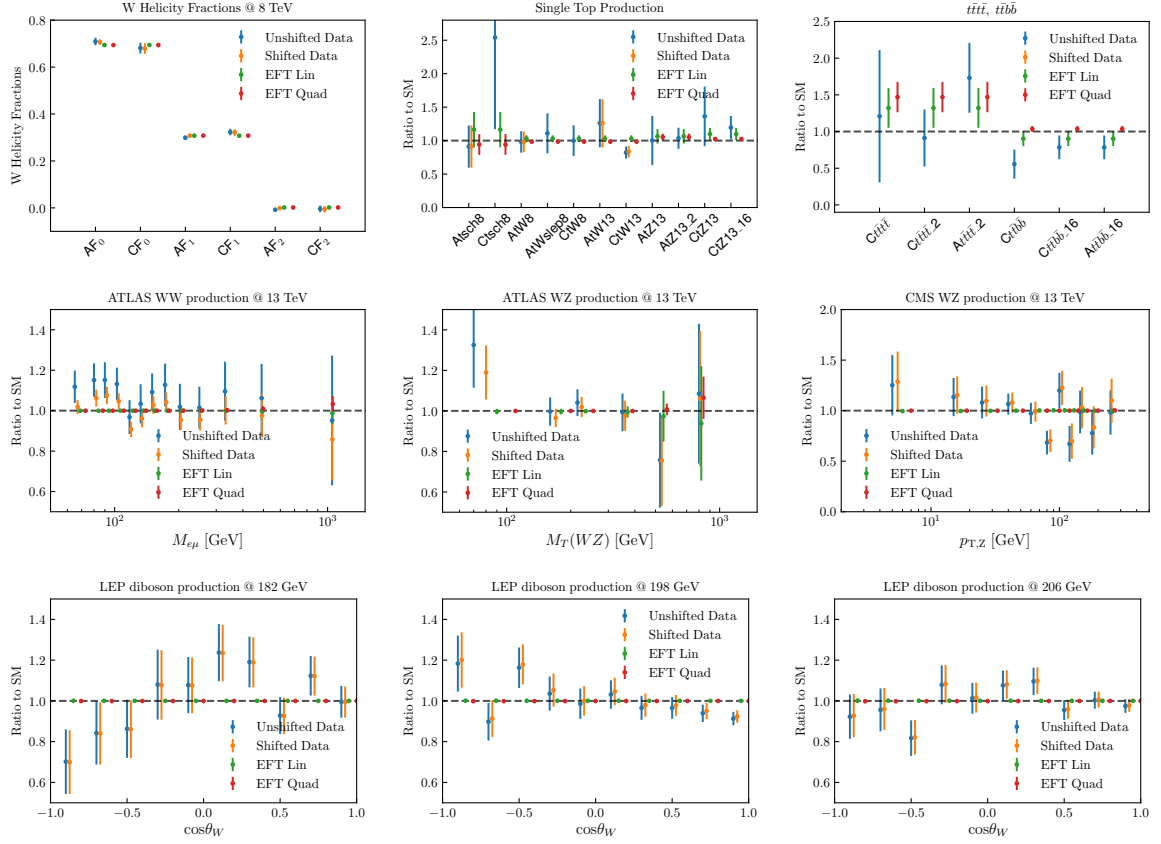
**Figure A.2.** Same as Fig. A.1 now for the $W$ helicity fractions, the single-top $s$-channel and $tV$ total cross-sections, the four-heavy quark fiducial measurements, the LHC diboson differential distributions at 13 TeV, and the LEP diboson cross-sections at different center of mass energies.

tial distributions at 13 TeV, and the LEP diboson cross-sections at different center-of-mass energies. Note that contrary to the rest of the datasets, the comparison for the $W$ helicity fraction is carried out at the absolute rather than at the normalised level. Concerning the single top measurements, the best-fit EFT predictions tend to move towards the experimental data, which in most cases is somewhat higher than the SM prediction. For some datasets, such as single-top $s$-channel cross-section at 8 TeV and the $tW$ cross-sections at 13 TeV, the agreement between ATLAS and CMS is at best marginal and thus the EFT fit interpolates between the two measurements. A similar behaviour is observed for the $t\bar{t}t\bar{t}$ cross-sections at 13 TeV. Furthermore, as was the case for the processes considered in Fig. A.1, the EFT fit uncertainties appear to be reduced in the quadratic case.

Moving to the LEP and LHC diboson datasets, one finds that for electron-positron collisions the EFT fit result is very close to the SM cross-section with a vanishing uncertainty. This result is likely to be related to the constraints imposed by the EWPOs as well as by the LHC diboson data. Nevertheless, the SM predictions are in good agreement with the LEP data for all four center-of-mass energies considered to begin with. In the case of the LHC measurements, for the ATLAS $m_{e\mu}$ and $m_T^{WZ}$ distributions in the $W^+W^-$ and $WZ$ final states, respectively, the data is in good agreement with the SM and the net effect of

74
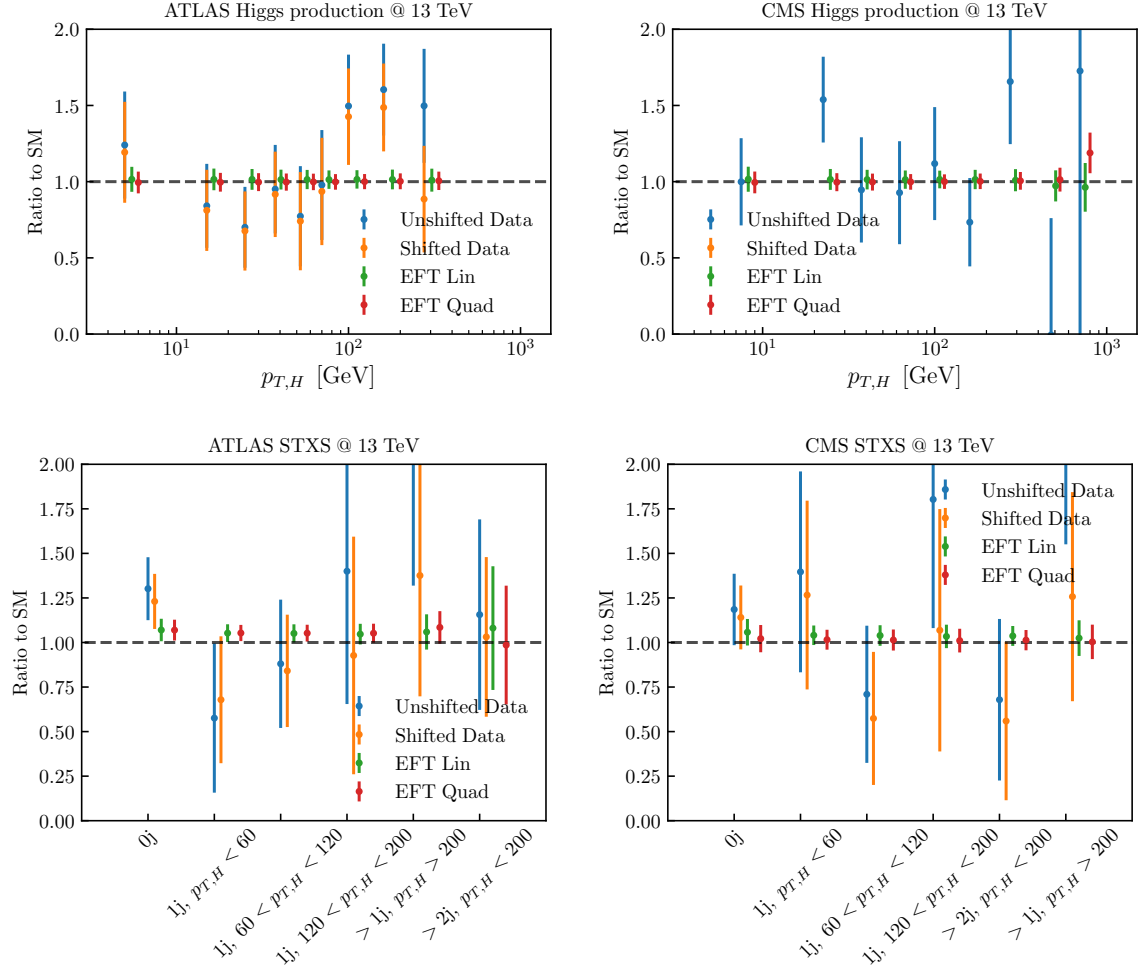
**Figure A.3.** Same as Fig. A.1 for representative Higgs measurements from ATLAS and CMS at $\sqrt{s} = 13$ TeV, namely the $p_{T,H}$ distributions summing over all production modes and final states (upper panels), and the Simplified Template Cross-Section measurements (bottom panels) corresponding to the $ZZ$ (left) and the $\gamma\gamma$ final states (right panel).

the EFT corrections is small, except perhaps for the highest-energy bin of the $m_{e\mu}$ and $m_T^{WZ}$ distribution. Similar considerations apply for the CMS 13 TeV $WZ$ dataset, where we observe good agreement between theory and data also for the high $p_T^Z$ region.

Concerning the comparison between experimental data and theory calculations for the Higgs production and decay measurements, Fig. A.3 displays representative Higgs measurements from ATLAS and CMS at $\sqrt{s} = 13$ TeV, namely the $p_{T,H}$ distributions inclusive over all production modes and final states, and the Simplified Template Cross-Section measurements corresponding to the $ZZ$ and the $\gamma\gamma$ final states for ATLAS and CMS respectively. Then Fig. A.4 summarizes the results corresponding to Higgs boson signal strengths for different production mechanisms and decay channels. Specifically, we show the ATLAS+CMS Run I combination and the ATLAS and CMS Run II measurements at 13 TeV. Note that, by construction, in the signal strengths the SM predictions correspond to $\mu_i^{(f)} = 1$, see the discussion of App. B for more details.
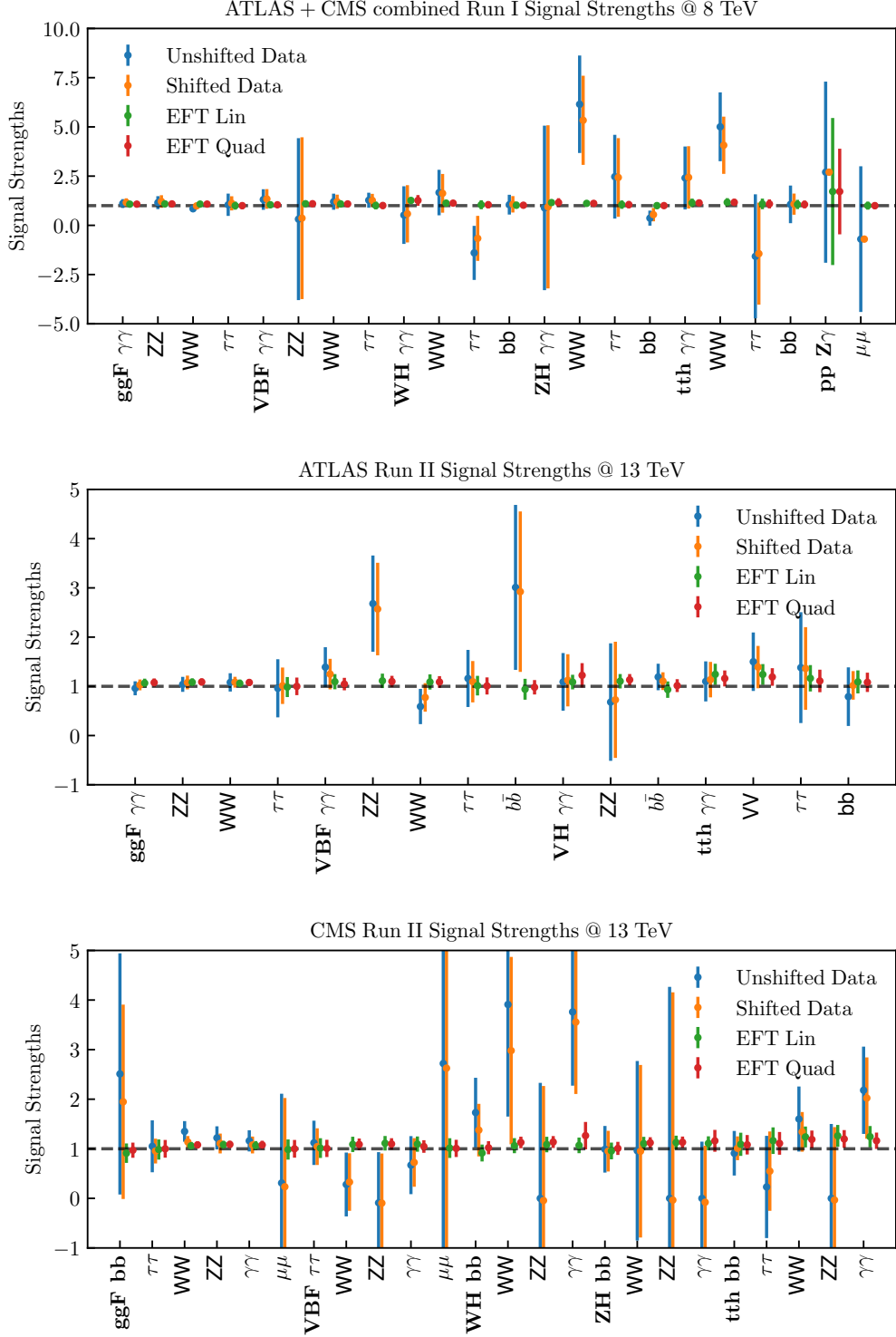
**Figure A.4.** Same as Fig. A.1 for the Higgs boson signal strengths corresponding to different production mechanisms and decay channels. From top to bottom we show the ATLAS+CMS Run I combination and the ATLAS and CMS Run II measurements at 13 TeV. Note that by the definition of the signal strengths the SM predictions correspond to $\mu_i^{(f)} = 1$ in all cases, see also App. B.

In the case of the differential Higgs distributions, we can observe the good agreement both the SM and the EFT predictions within the relatively large experimental uncertainties. For these distributions, the EFT effects can reach a magnitude of up to a few percent in the global fit. For instance, for the CMS $p_{T,H}$ distribution in the quadratic fit, the best-fit results are $\simeq 15\%$ higher than the SM for the $p_{T,H} = 1$ TeV bin. For the case of the signal strengths, also a fair agreement is found, though for some combinations of production channel and decay mode the experimental uncertainties still remain rather large.

## B   Implementation of Higgs signal strengths

In this appendix, we describe how the Higgs signal strengths have been implemented in the present analysis. For a generic Higgs production and decay cross-section, denoted as $\sigma(pp \to h \to X)$, the experimentally measured signal strength is defined as the product of the production cross-section times the corresponding branching ratio, normalised to the Standard Model predictions:

$$\mu_{pp \to h \to X}^{(\mathrm{exp})} = \frac{\sigma^{(\mathrm{exp})}(pp \to h \to X)}{\sigma^{(\mathrm{sm})}(pp \to h \to X)} = \frac{\sigma^{(\mathrm{exp})}(pp \to h \to X)}{\sigma^{(\mathrm{sm})}(pp \to h) \times \mathrm{BR}^{(\mathrm{sm})}(h \to X)} \,, \qquad (\mathrm{B.1})$$

which in turn can be expressed as

$$\mu_{pp \to h \to X}^{(\mathrm{exp})} = \frac{\sigma^{(\mathrm{exp})}(pp \to h \to X)}{\sigma^{(\mathrm{sm})}(pp \to h \to X)} = \frac{\sigma^{(\mathrm{exp})}(pp \to h \to X)}{\sigma^{(\mathrm{sm})}(pp \to h) \times \left(\Gamma_X^{(\mathrm{sm})}/\Gamma_{\mathrm{tot}}^{(\mathrm{sm})}\right)} \,, \qquad (\mathrm{B.2})$$

where $\Gamma_{\mathrm{tot}}$ indicates the total Higgs width and $\Gamma_X$ is the partial width for the decay into the specific final state $X$. In this work, we assume that the Higgs boson decays only to known particles and hence set to zero its branching ratio to invisible final states. The theoretical prediction corresponding to the measurement of the signal strengths Eq. (B.2) in the SMEFT is given by

$$\mu_{pp \to h \to X}^{(\mathrm{th})}(\boldsymbol{c}) = \left( \frac{\sigma^{(\mathrm{EFT})}(pp \to h)(\boldsymbol{c}) \times \left(\Gamma_X^{(\mathrm{EFT})}(\boldsymbol{c})/\Gamma_{\mathrm{tot}}^{(\mathrm{EFT})}(\boldsymbol{c})\right)}{\sigma^{(\mathrm{sm})}(pp \to h) \times \left(\Gamma_X^{(\mathrm{sm})}/\Gamma_{\mathrm{tot}}^{(\mathrm{sm})}\right)} \right) \,, \qquad (\mathrm{B.3})$$

At any order in the EFT expansion, we can express Eq. (B.3) as

$$\mu_{pp \to h \to X}^{(\mathrm{th})}(\boldsymbol{c}) = \left( \frac{\sigma^{(\mathrm{EFT})}(pp \to h)(\boldsymbol{c})}{\sigma^{(\mathrm{sm})}(pp \to h)} \right) \times \left( \frac{\Gamma_X^{(\mathrm{EFT})}(\boldsymbol{c})}{\Gamma_X^{(\mathrm{sm})}} \right) \times \left( \frac{\Gamma_{\mathrm{tot}}^{(\mathrm{EFT})}(\boldsymbol{c})}{\Gamma_{\mathrm{tot}}^{(\mathrm{sm})}} \right)^{-1} \,, \qquad (\mathrm{B.4})$$

where the total width is evaluated as the sum of all relevant partial decay widths,

$$\mu_{pp \to h \to X}^{(\mathrm{th})}(\boldsymbol{c}) = \left( \frac{\sigma^{(\mathrm{EFT})}(pp \to h)(\boldsymbol{c})}{\sigma^{(\mathrm{sm})}(pp \to h)} \right) \times \left( \frac{\Gamma_X^{(\mathrm{EFT})}(\boldsymbol{c})}{\Gamma_X^{(\mathrm{sm})}} \right) \times \left( \frac{\sum_Y \Gamma_Y^{(\mathrm{EFT})}(\boldsymbol{c})}{\sum_Z \Gamma_Z^{(\mathrm{sm})}} \right)^{-1} \,, \qquad (\mathrm{B.5})$$

where $X$, $Y$, and $Z$ indicate possible (SM) final states in which the Higgs boson can decay. Assuming now that $\Lambda = 1$ TeV and working at linear order in the EFT expansion one has

$$\sigma^{(\mathrm{EFT})}(pp \to h)(\boldsymbol{c}) = \sigma^{(\mathrm{sm})}(pp \to h) + \sum_{i=1}^{n} c_i \kappa_{\sigma,i} \,, \qquad (\mathrm{B.6})$$

$$\Gamma_X^{(\text{EFT})}(\boldsymbol{c}) = \Gamma_X^{(\text{sm})} + \sum_{i=1}^{n} c_i \kappa_{\gamma_x,i} \,, \tag{B.7}$$

where $n$ is the number of independent dimension-6 operators in the fitting basis and $\{\kappa_{\sigma,i}\}$ and $\{\kappa_{\gamma_x,i}\}$, are the (absolute) EFT corrections associated to the production cross section and partial width, respectively, corresponding to the operator $c_i$.

Inserting now the linear EFT expansion into Eq. (B.5), one gets

$$
\begin{aligned}
\mu_{pp \to h \to X}^{(\text{th})}(\boldsymbol{c}) &= \left( \frac{\sigma^{(\text{sm})} + \sum_{i=1}^{n} c_i \kappa_{\sigma,i}}{\sigma^{(\text{sm})}} \right) \times \left( \frac{\Gamma_X^{(\text{sm})} + \sum_{j=1}^{n} c_j \kappa_{\gamma_x,i}}{\Gamma_X^{(\text{sm})}} \right) \\
&\quad \times \left( \frac{\sum_Y \left( \Gamma_Y^{(\text{sm})} + \sum_{k=1}^{n} c_k \kappa_{\gamma_y,i} \right)}{\sum_Z \Gamma_Z^{(\text{sm})}} \right)^{-1} \\
&= \left( 1 + \sum_{i=1}^{n} c_i \frac{\kappa_{\sigma,i}}{\sigma^{(\text{sm})}} \right) \times \left( 1 + \sum_{j=1}^{n} c_j \frac{\kappa_{\gamma_x,j}}{\Gamma_X^{(\text{sm})}} \right) \left( \sum_Y \left( \frac{\Gamma_Y^{(\text{sm})}}{\Gamma_{\text{tot}}^{(\text{sm})}} + \sum_{k=1}^{n} c_k \frac{\kappa_{\gamma_y,k}}{\Gamma_{\text{tot}}^{(\text{sm})}} \right) \right)^{-1} \,,
\end{aligned}
\tag{B.8}
$$

which can be further simplified to

$$\mu_{pp \to h \to X}^{(\text{th})}(\boldsymbol{c}) = \left( 1 + \sum_{i=1}^{n} c_i \frac{\kappa_{\sigma,i}}{\sigma^{(\text{sm})}} \right) \times \left( 1 + \sum_{j=1}^{n} c_j \frac{\kappa_{\gamma_x,j}}{\Gamma_X^{(\text{sm})}} \right) \times \left( 1 + \sum_Y \left( \sum_{k=1}^{n} c_k \frac{\kappa_{\gamma_y,k}}{\Gamma_{\text{tot}}^{(\text{sm})}} \right) \right)^{-1} . \tag{B.9}$$

If we Taylor expand the last term at $\mathcal{O}(\Lambda^{-2})$, we obtain the required expression for the theoretical prediction of the Higgs signal strengths at linear order in the EFT,

$$\mu_{pp \to h \to X}^{(\text{th})}(\boldsymbol{c}) = 1 + \sum_{i=1}^{n} c_i \left[ \frac{\kappa_{\sigma,i}}{\sigma^{(\text{sm})}} + \frac{\kappa_{\gamma_x,i}}{\Gamma_X^{(\text{sm})}} - \sum_Y \left( \frac{\kappa_{\gamma_y,i}}{\Gamma_{\text{tot}}^{(\text{sm})}} \right) \right] . \tag{B.10}$$

For simplicity, we can replace the total Higgs decay width in the SM with the corresponding branching fractions,

$$\text{BR}_X^{(\text{sm})} \equiv \frac{\Gamma_X^{(\text{sm})}}{\Gamma_{\text{tot}}^{(\text{sm})}} \quad \to \quad \frac{1}{\Gamma_{\text{tot}}^{(\text{sm})}} = \frac{\text{BR}_X^{(\text{sm})}}{\Gamma_X^{(\text{sm})}} \,, \tag{B.11}$$

which allows us to express Eq. (B.10) as

$$\mu_{pp \to h \to X}^{(\text{th})}(\boldsymbol{c}) = 1 + \sum_{i=1}^{n} c_i \left[ \frac{\kappa_{\sigma,i}}{\sigma^{(\text{sm})}} + \frac{\kappa_{\gamma_x,i}}{\Gamma_X^{(\text{sm})}} - \sum_Y \left( \frac{\kappa_{\gamma_y,i}}{\Gamma_Y^{(\text{sm})}} \text{BR}_Y^{(\text{sm})} \right) \right] . \tag{B.12}$$

Hence we find that we can evaluate the Higgs signal strengths in the EFT as

$$\mu_{pp \to h \to X}^{(\text{th})}(\boldsymbol{c}) = 1 + \sum_{i=1}^{n} c_i \beta_{pp \to h \to X,i} \,, \tag{B.13}$$

where we have defined

$$\beta_{pp \to h \to X,i} \equiv \frac{\kappa_{\sigma,i}}{\sigma^{(\text{sm})}} + \frac{\kappa_{\gamma_x,i}}{\Gamma_X^{(\text{sm})}} - \sum_Y \left( \frac{\kappa_{\gamma_y,i}}{\Gamma_Y^{(\text{sm})}} \text{BR}_Y^{(\text{sm})} \right) . \tag{B.14}$$

Note that in this notation we use $\beta$ to indicate relative EFT corrections while the $\kappa$ always indicate instead absolute corrections. A similar expression, although somewhat more cumbersome, can be derived to account for the quadratic EFT contributions to the Higgs signal strengths.

# C  Usage of SMEFiT results

The results of this work are made available via the website of the `SMEFiT` project:

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

as well as via the corresponding public `GitHub` repository:

https://github.com/LHCfitNikhef/SMEFiT/

Specifically, we provide the full set of $N_{\rm spl}$ samples corresponding to the NS fits presented in this work for the $n_{\rm op} = 50$ Wilson coefficients. We recall that not all of these coefficients are associated to independent degrees of freedom, and that 14 of them are constrained by the EWPO relations as discussed in Sect. 2.

These $N_{\rm spl}$ samples provide a representation of the probability density associated to the EFT coefficients. From these samples, it is easy to evaluate statistical estimators such as means, standard deviations, and correlations, e.g.,

$$\langle c_i \rangle = \frac{1}{N_{\rm spl}} \sum_{k=1}^{N_{\rm spl}} c_i^{(k)}, \quad i = 1, \ldots, n_{\rm op}, \tag{C.1}$$

$$\sigma_{c_i} = \left( \frac{1}{N_{\rm spl} - 1} \sum_{k=1}^{N_{\rm spl}} \left( c_i^{(k)} - \langle c_i \rangle \right)^2 \right)^{1/2}, \quad i = 1, \ldots, n_{\rm op}, \tag{C.2}$$

$$\rho(c_i, c_j) = \left( \frac{1}{N_{\rm spl}} \sum_{k=1}^{N_{\rm spl}} c_i^{(k)} c_j^{(k)} - \langle c_i \rangle \langle c_j \rangle \right) \Big/ \sigma_{c_i} \sigma_{c_k}, \quad i, j = 1, \ldots, n_{\rm op}, \tag{C.3}$$

as well as other estimators such as confidence level intervals and higher moments beyond the quadratic approximation. One should emphasize that, as discussed in Sect. 5, the Gaussian approximation is in general not reliable for most of the EFT parameters in the case that the quadratic $\mathcal{O}(\Lambda^{-4})$ corrections are accounted for.

These samples spanning the probability density in the space of EFT Wilson coefficients could also be used to quantify *a posteriori* the impact in the fit of new measurements by means of the Bayesian reweighting method presented in [33].

To facilitate the usage of the results presented in this work, we also make available via the `GitHub` repository a `Python` analysis code that takes as input the samples corresponding to a specific EFT fit and then evaluates means, correlations, and 95% CL intervals, as well as produces most of the plots and statistical estimators displayed in Sect. 5. In a future installment of our global EFT analysis, we plan to release the full `smefit` fitting code and the associated theoretical and experimental inputs, together with a complete documentation and user-friendly examples.

# References

[1] S. Weinberg, *Baryon and Lepton Nonconserving Processes*, *Phys. Rev. Lett.* **43** (1979) 1566–1570.

[2] W. Buchmuller and D. Wyler, *Effective Lagrangian Analysis of New Interactions and Flavor Conservation*, *Nucl. Phys.* **B268** (1986) 621–653.

[3] B. Grzadkowski, M. Iskrzynski, M. Misiak, and J. Rosiek, *Dimension-Six Terms in the Standard Model Lagrangian*, *JHEP* **10** (2010) 085, [arXiv:1008.4884].

[4] I. Brivio and M. Trott, *The Standard Model as an Effective Field Theory*, *Phys. Rept.* **793** (2019) 1–98, [arXiv:1706.08945].

[5] A. Buckley, C. Englert, J. Ferrando, D. J. Miller, L. Moore, K. Nördstrom, M. Russell, and C. D. White, *Results from TopFitter*, *PoS* **CKM2016** (2016) 127, [arXiv:1612.02294].

[6] A. Buckley, C. Englert, J. Ferrando, D. J. Miller, L. Moore, M. Russell, and C. D. White, *Constraining top quark effective theory in the LHC Run II era*, *JHEP* **04** (2016) 015, [arXiv:1512.03360].

[7] N. P. Hartland, F. Maltoni, E. R. Nocera, J. Rojo, E. Slade, E. Vryonidou, and C. Zhang, *A Monte Carlo global analysis of the Standard Model Effective Field Theory: the top quark sector*, *JHEP* **04** (2019) 100, [arXiv:1901.05965].

[8] I. Brivio, S. Bruggisser, F. Maltoni, R. Moutafis, T. Plehn, E. Vryonidou, S. Westhoff, and C. Zhang, *O new physics, where art thou? A global search in the top sector*, *JHEP* **02** (2020) 131, [arXiv:1910.03606].

[9] A. Biekötter, T. Corbett, and T. Plehn, *The Gauge-Higgs Legacy of the LHC Run II*, *SciPost Phys.* **6** (2019) 064, [arXiv:1812.07587].

[10] J. Ellis, C. W. Murphy, V. Sanz, and T. You, *Updated Global SMEFT Fit to Higgs, Diboson and Electroweak Data*, *JHEP* **06** (2018) 146, [arXiv:1803.03252].

[11] E. da Silva Almeida, A. Alves, N. Rosa Agostinho, O. J. Éboli, and M. Gonzalez-Garcia, *Electroweak Sector Under Scrutiny: A Combined Analysis of LHC and Electroweak Precision Data*, *Phys. Rev. D* **99** (2019), no. 3 033001, [arXiv:1812.01009].

[12] J. Baglio, S. Dawson, S. Homiller, S. D. Lane, and I. M. Lewis, *Validity of standard model EFT studies of VH and VV production at NLO*, *Phys. Rev. D* **101** (2020), no. 11 115004.

[13] S. Alioli, W. Dekens, M. Girard, and E. Mereghetti, *NLO QCD corrections to SM-EFT dilepton and electroweak Higgs boson production, matched to parton shower in POWHEG*, *JHEP* **08** (2018) 205, [arXiv:1804.07407].

[14] J. J. Ethier, R. Gomez-Ambrosio, G. Magni, and J. Rojo, *SMEFT analysis of vector boson scattering and diboson data from the LHC Run II*, arXiv:2101.03180.

[15] A. Greljo and D. Marzocca, *High-$p_T$ dilepton tails and flavor physics*, *Eur. Phys. J.* **C77** (2017), no. 8 548, [arXiv:1704.09015].

[16] R. Gomez-Ambrosio, *Studies of Dimension-Six EFT effects in Vector Boson Scattering*, *Eur. Phys. J. C* **79** (2019), no. 5 389, [arXiv:1809.04189].

[17] A. Dedes, P. Kozów, and M. Szleper, *SM EFT effects in Vector-Boson Scattering at the LHC*, arXiv:2011.07367.

[18] J. Aebischer, J. Kumar, P. Stangl, and D. M. Straub, *A Global Likelihood for Precision Constraints and Flavour Anomalies*, *Eur. Phys. J. C* **79** (2019), no. 6 509, [`arXiv:1810.07698`].

[19] A. Falkowski, M. González-Alonso, and Z. Tabrizi, *Reactor neutrino oscillations as constraints on Effective Field Theory*, *JHEP* **05** (2019) 173, [`arXiv:1901.04553`].

[20] A. Falkowski, M. González-Alonso, and K. Mimouni, *Compilation of low-energy constraints on 4-fermion operators in the SMEFT*, *JHEP* **08** (2017) 123, [`arXiv:1706.03783`].

[21] J. Ellis, M. Madigan, K. Mimasu, V. Sanz, and T. You, *Top, Higgs, Diboson and Electroweak Fit to the Standard Model Effective Field Theory*, `arXiv:2012.02779`.

[22] S. Bißmann, C. Grunwald, G. Hiller, and K. Kröninger, *Top and Beauty synergies in SMEFT-fits at present and future colliders*, `arXiv:2012.10456`.

[23] S. Bruggisser, R. Schäfer, D. van Dyk, and S. Westhoff, *The Flavor of UV Physics*, `arXiv:2101.07273`.

[24] A. Pich, *Flavour Anomalies*, *PoS* **LHCP2019** (2019) 078, [`arXiv:1911.06211`].

[25] **LHCb** Collaboration, R. Aaij et al., *Test of lepton universality in beauty-quark decays*, `arXiv:2103.11769`.

[26] J. Fuentes-Martin, A. Greljo, J. Martin Camalich, and J. D. Ruiz-Alvarez, *Charm physics confronts high-$p_T$ lepton tails*, *JHEP* **11** (2020) 080, [`arXiv:2003.12421`].

[27] **The NNPDF** Collaboration, R. D. Ball et al., *A determination of parton distributions with faithful uncertainty estimation*, *Nucl. Phys.* **B809** (2009) 1–63, [`arXiv:0808.1231`].

[28] **The NNPDF** Collaboration, R. D. Ball et al., *A first unbiased global NLO determination of parton distributions and their uncertainties*, *Nucl. Phys.* **B838** (2010) 136, [`arXiv:1002.4407`].

[29] **NNPDF** Collaboration, R. D. Ball et al., *Parton distributions for the LHC Run II*, *JHEP* **04** (2015) 040, [`arXiv:1410.8849`].

[30] J. Rojo, *Machine Learning tools for global PDF fits*, in *13th Conference on Quark Confinement and the Hadron Spectrum (Confinement XIII) Maynooth, Ireland, July 31-August 6, 2018*, 2018. `arXiv:1809.04392`.

[31] S. Forte and S. Carrazza, *Parton distribution functions*, `arXiv:2008.12305`.

[32] D. Barducci et al., *Interpreting top-quark LHC measurements in the standard-model effective field theory*, `arXiv:1802.07237`.

[33] S. van Beek, E. R. Nocera, J. Rojo, and E. Slade, *Constraining the SMEFT with Bayesian reweighting*, *SciPost Phys.* **7** (2019), no. 5 070, [`arXiv:1906.05296`].

[34] R. D. Ball, V. Bertone, F. Cerutti, L. Del Debbio, S. Forte, et al., *Reweighting and Unweighting of Parton Distributions and the LHC W lepton asymmetry data*, *Nucl.Phys.* **B855** (2012) 608–638, [arXiv:1108.1758].

[35] **The NNPDF** Collaboration, R. D. Ball et al., *Reweighting NNPDFs: the W lepton asymmetry*, *Nucl. Phys.* **B849** (2011) 112–143, [arXiv:1012.0836].

[36] **ALEPH, DELPHI, L3, OPAL, SLD, LEP Electroweak Working Group, SLD Electroweak Group, SLD Heavy Flavour Group** Collaboration, S. Schael et al., *Precision electroweak measurements on the Z resonance*, *Phys. Rept.* **427** (2006) 257–454, [hep-ex/0509008].

[37] F. Feroz, M. P. Hobson, E. Cameron, and A. N. Pettitt, *Importance Nested Sampling and the MultiNest Algorithm*, arXiv:1306.2144.

[38] R. Alonso, E. E. Jenkins, A. V. Manohar, and M. Trott, *Renormalization Group Evolution of the Standard Model Dimension Six Operators III: Gauge Coupling Dependence and Phenomenology*, *JHEP* **04** (2014) 159, [arXiv:1312.2014].

[39] J. A. Aguilar-Saavedra, *Effective four-fermion operators in top physics: A Roadmap*, *Nucl. Phys.* **B843** (2011) 638–672, [arXiv:1008.3562]. [Erratum: Nucl. Phys.B851,443(2011)].

[40] G. D'Ambrosio, G. F. Giudice, G. Isidori, and A. Strumia, *Minimal flavor violation: An Effective field theory approach*, *Nucl. Phys.* **B645** (2002) 155–187, [hep-ph/0207036].

[41] C. Degrande, G. Durieux, F. Maltoni, K. Mimasu, E. Vryonidou, and C. Zhang, *Automated one-loop computations in the SMEFT*, arXiv:2008.11743.

[42] Z. Han and W. Skiba, *Effective theory analysis of precision electroweak data*, *Phys. Rev.* **D71** (2005) 075009, [hep-ph/0412166].

[43] A. Falkowski and F. Riva, *Model-independent precision constraints on dimension-6 operators*, *JHEP* **02** (2015) 039, [arXiv:1411.0669].

[44] C. Grojean, W. Skiba, and J. Terning, *Disguising the oblique parameters*, *Phys. Rev.* **D73** (2006) 075008, [hep-ph/0602154].

[45] I. Brivio and M. Trott, *Scheming in the SMEFT... and a reparameterization invariance!*, *JHEP* **07** (2017) 148, [arXiv:1701.06424]. [Addendum: JHEP05,136(2018)].

[46] Z. Zhang, *Time to Go Beyond Triple-Gauge-Boson-Coupling Interpretation of W Pair Production*, *Phys. Rev. Lett.* **118** (2017), no. 1 011803, [arXiv:1610.01618].

[47] C. Grojean, M. Montull, and M. Riembau, *Diboson at the LHC vs LEP*, *JHEP* **03** (2019) 020, [arXiv:1810.05149].

[48] **ATLAS** Collaboration, G. Aad et al., *Measurements of top-quark pair differential cross-sections in the lepton+jets channel in pp collisions at $\sqrt{s} = 8$ TeV using the ATLAS detector*, *Eur. Phys. J.* **C76** (2016), no. 10 538, [arXiv:1511.04716].

[49] **CMS** Collaboration, V. Khachatryan et al., *Measurement of the differential cross section for top quark pair production in pp collisions at $\sqrt{s} = 8$ TeV*, *Eur. Phys. J.* **C75** (2015), no. 11 542, [arXiv:1505.04480].

[50] **CMS** Collaboration, A. M. Sirunyan et al., *Measurement of double-differential cross sections for top quark pair production in pp collisions at $\sqrt{s} = 8$ TeV and impact on parton distribution functions*, *Eur. Phys. J.* **C77** (2017), no. 7 459, [arXiv:1703.01630].

[51] **ATLAS** Collaboration, M. Aaboud et al., *Measurement of the W boson polarisation in $t\bar{t}$ events from pp collisions at $\sqrt{s} = 8$ TeV in the lepton + jets channel with ATLAS*, *Eur. Phys. J.* **C77** (2017), no. 4 264, [arXiv:1612.02577].

[52] **CMS** Collaboration, V. Khachatryan et al., *Measurement of the W boson helicity fractions in the decays of top quark pairs to lepton + jets final states produced in pp collisions at $\sqrt{s} = 8$ TeV*, *Phys. Lett.* **B762** (2016) 512–534, [arXiv:1605.09047].

[53] **CMS** Collaboration, V. Khachatryan et al., *Measurement of differential cross sections for top quark pair production using the lepton+jets final state in proton-proton collisions at 13 TeV*, *Phys. Rev.* **D95** (2017), no. 9 092001, [arXiv:1610.04191].

[54] **CMS** Collaboration, A. M. Sirunyan et al., *Measurement of differential cross sections for the production of top quark pairs and of additional jets in lepton+jets events from pp collisions at $\sqrt{s} = 13$ TeV*, *Phys. Rev.* **D97** (2018), no. 11 112003, [arXiv:1803.08856].

[55] **CMS** Collaboration, A. M. Sirunyan et al., *Measurement of normalized differential $t\bar{t}$ cross sections in the dilepton channel from pp collisions at $\sqrt{s} = 13$ TeV*, *JHEP* **04** (2018) 060, [arXiv:1708.07638].

[56] **ATLAS** Collaboration, M. Aaboud et al., *Measurement of top quark pair differential cross-sections in the dilepton channel in pp collisions at $\sqrt{s} = 7$ and 8 TeV with ATLAS*, *Phys. Rev. D* **94** (2016), no. 9 092003, [arXiv:1607.07281]. [Addendum: Phys.Rev.D 101, 119901 (2020)].

[57] **ATLAS** Collaboration, G. Aad et al., *Measurements of top-quark pair differential and double-differential cross-sections in the $\ell$+jets channel with pp collisions at $\sqrt{s} = 13$ TeV using the ATLAS detector*, *Eur. Phys. J. C* **79** (2019), no. 12 1028, [arXiv:1908.07305].

[58] **CMS** Collaboration, A. M. Sirunyan et al., *Measurements of $t\bar{t}$ differential cross sections in proton-proton collisions at $\sqrt{s} = 13$ TeV using events containing two leptons*, *JHEP* **02** (2019) 149, [arXiv:1811.06625].

[59] **ATLAS, CMS** Collaboration, M. Aaboud et al., *Combination of inclusive and differential $t\bar{t}$ charge asymmetry measurements using ATLAS and CMS data at $\sqrt{s} = 7$ and 8 TeV*, *JHEP* **04** (2018) 033, [arXiv:1709.05327].

[60] **ATLAS** Collaboration, *Inclusive and differential measurement of the charge asymmetry in $t\bar{t}$ events at 13 TeV with the ATLAS detector*, .

[61] S. Amoroso et al., *Les Houches 2019: Physics at TeV Colliders: Standard Model Working Group Report*, in *11th Les Houches Workshop on Physics at TeV Colliders: PhysTeV Les Houches*, 3, 2020. arXiv:2003.01700.

[62] S. Bailey and L. Harland-Lang, *Differential Top Quark Pair Production at the LHC: Challenges for PDF Fits*, *Eur. Phys. J. C* **80** (2020), no. 1 60, [arXiv:1909.10541].

[63] **ATLAS** Collaboration, G. Aad et al., *Measurements of top-quark pair single- and double-differential cross-sections in the all-hadronic channel in pp collisions at $\sqrt{s} = 13$ TeV using the ATLAS detector*, *JHEP* **01** (2021) 033, [arXiv:2006.09274].

[64] **CMS** Collaboration, A. M. Sirunyan et al., *Measurement of differential $t\bar{t}$ production cross sections using top quarks at large transverse momenta in pp collisions at $\sqrt{s} = 13$ TeV*, *Phys. Rev. D* **103** (2021), no. 5 052008, [arXiv:2008.07860].

[65] **ATLAS** Collaboration, G. Aad et al., *Measurement of the $t\bar{t}$ production cross-section and lepton differential distributions in eμ dilepton events from pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, *Eur. Phys. J. C* **80** (2020), no. 6 528, [arXiv:1910.08819].

[66] J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, et al., *The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations*, *JHEP* **1407** (2014) 079, [arXiv:1405.0301].

[67] M. Czakon, D. Heymes, and A. Mitov, *Dynamical scales for multi-TeV top-pair production at the LHC*, *JHEP* **04** (2017) 071, [arXiv:1606.03350].

[68] M. Czakon, N. P. Hartland, A. Mitov, E. R. Nocera, and J. Rojo, *Pinning down the large-x gluon with NNLO top-quark pair differential distributions*, *JHEP* **04** (2017) 044, [arXiv:1611.08609].

[69] **NNPDF** Collaboration, R. D. Ball et al., *Parton distributions from high-precision collider data*, *Eur. Phys. J.* **C77** (2017), no. 10 663, [arXiv:1706.00428].

[70] A. Greljo, S. Iranipour, Z. Kassabov, M. Madigan, J. Moore, J. Rojo, M. Ubiali, and C. Voisey, *Parton distributions in the SMEFT from high-energy Drell-Yan tails*, arXiv:2104.02723.

[71] S. Carrazza, C. Degrande, S. Iranipour, J. Rojo, and M. Ubiali, *Can New Physics hide inside the proton?*, *Phys. Rev. Lett.* **123** (2019), no. 13 132001, [arXiv:1905.05215].

[72] **CMS** Collaboration, A. M. Sirunyan et al., *Measurements of $t\bar{t}$ cross sections in association with b jets and inclusive jets and their ratio using dilepton final states in pp collisions at $\sqrt{s} = 13$ TeV*, *Phys. Lett.* **B776** (2018) 355–378, [arXiv:1705.10141].

[73] **CMS** Collaboration, A. M. Sirunyan et al., *Search for standard model production of four top quarks with same-sign and multilepton final states in proton–proton collisions at $\sqrt{s} = 13$ TeV*, *Eur. Phys. J.* **C78** (2018), no. 2 140, [arXiv:1710.10614].

[74] **CMS** Collaboration, V. Khachatryan et al., *Observation of top quark pairs produced in association with a vector boson in pp collisions at $\sqrt{s} = 8$ TeV*, *JHEP* **01** (2016) 096, [arXiv:1510.01131].

[75] **CMS** Collaboration, A. M. Sirunyan et al., *Measurement of the cross section for top quark pair production in association with a W or Z boson in proton-proton collisions at $\sqrt{s}$ = 13 TeV*, JHEP **08** (2018) 011, [arXiv:1711.02547].

[76] **ATLAS** Collaboration, G. Aad et al., *Measurement of the $t\bar{t}W$ and $t\bar{t}Z$ production cross sections in pp collisions at $\sqrt{s}$ = 8 TeV with the ATLAS detector*, JHEP **11** (2015) 172, [arXiv:1509.05276].

[77] **ATLAS** Collaboration, M. Aaboud et al., *Measurement of the $t\bar{t}Z$ and $t\bar{t}W$ production cross sections in multilepton final states using 3.2 $fb^{-1}$ of pp collisions at $\sqrt{s}$ = 13 TeV with the ATLAS detector*, Eur. Phys. J. **C77** (2017), no. 1 40, [arXiv:1609.01599].

[78] **CMS** Collaboration, A. M. Sirunyan et al., *Search for production of four top quarks in final states with same-sign or multiple leptons in proton-proton collisions at $\sqrt{s}$ = 13 TeV*, Eur. Phys. J. C **80** (2020), no. 2 75, [arXiv:1908.06463].

[79] **ATLAS** Collaboration, G. Aad et al., *Evidence for $t\bar{t}t\bar{t}$ production in the multilepton final state in proton–proton collisions at $\sqrt{s}$ = 13 TeV with the ATLAS detector*, Eur. Phys. J. C **80** (2020), no. 11 1085, [arXiv:2007.14858].

[80] **ATLAS** Collaboration, M. Aaboud et al., *Measurements of inclusive and differential fiducial cross-sections of $t\bar{t}$ production with additional heavy-flavour jets in proton-proton collisions at $\sqrt{s}$ = 13 TeV with the ATLAS detector*, JHEP **04** (2019) 046, [arXiv:1811.12113].

[81] **CMS** Collaboration, A. M. Sirunyan et al., *Measurement of the $t\bar{t}b\bar{b}$ production cross section in the all-jet final state in pp collisions at $\sqrt{s}$ = 13 TeV*, Phys. Lett. B **803** (2020) 135285, [arXiv:1909.05306].

[82] **ATLAS** Collaboration, M. Aaboud et al., *Measurement of the $t\bar{t}Z$ and $t\bar{t}W$ cross sections in proton-proton collisions at $\sqrt{s}$ = 13 TeV with the ATLAS detector*, Phys. Rev. D **99** (2019), no. 7 072009, [arXiv:1901.03584].

[83] **CMS** Collaboration, A. M. Sirunyan et al., *Measurement of top quark pair production in association with a Z boson in proton-proton collisions at $\sqrt{s}$ = 13 TeV*, JHEP **03** (2020) 056, [arXiv:1907.11270].

[84] **ATLAS Collaboration** Collaboration, *Measurements of the inclusive and differential production cross sections of a top-quark-antiquark pair in association with a Z boson at $\sqrt{s}$ = 13 TeV with the ATLAS detector*, Tech. Rep. ATLAS-CONF-2020-028, CERN, Geneva, Aug, 2020.

[85] **CMS** Collaboration, V. Khachatryan et al., *Measurement of the t-channel single-top-quark production cross section and of the $| V_{tb} |$ CKM matrix element in pp collisions at $\sqrt{s}$ = 8 TeV*, JHEP **06** (2014) 090, [arXiv:1403.7366].

[86] **CMS Collaboration** Collaboration, A. M. Sirunyan et al., *Single top t-channel differential cross section at 8 TeV*, CMS-PAS-TOP-14-004 (2014) https://cds.cern.ch/record/1956681.

[87] **ATLAS** Collaboration, M. Aaboud et al., *Fiducial, total and differential cross-section measurements of t-channel single top-quark production in pp collisions at 8 TeV using data collected by the ATLAS detector*, *Eur. Phys. J.* **C77** (2017), no. 8 531, [arXiv:1702.02859].

[88] **ATLAS** Collaboration, G. Aad et al., *Evidence for single top-quark production in the s-channel in proton-proton collisions at $\sqrt{s}$ =8 TeV with the ATLAS detector using the Matrix Element Method*, *Phys. Lett.* **B756** (2016) 228–246, [arXiv:1511.05980].

[89] **CMS** Collaboration, V. Khachatryan et al., *Search for s channel single top quark production in pp collisions at $\sqrt{s}$ = 7 and 8 TeV*, *JHEP* **09** (2016) 027, [arXiv:1603.02555].

[90] **ATLAS** Collaboration, M. Aaboud et al., *Measurement of the inclusive cross-sections of single top-quark and top-antiquark t-channel production in pp collisions at $\sqrt{s}$ = 13 TeV with the ATLAS detector*, *JHEP* **04** (2017) 086, [arXiv:1609.03920].

[91] **CMS** Collaboration, V. Khachatryan et al., *Measurement of the differential cross section for t-channel single-top-quark production at $\sqrt{s}$ = 13 TeV*, *CMS-PAS-TOP-16-004* (2016) https://cds.cern.ch/record/2151074.

[92] **CMS** Collaboration, A. M. Sirunyan et al., *Cross section measurement of t-channel single top quark production in pp collisions at $\sqrt{s}$ = 13 TeV*, *Phys. Lett.* **B772** (2017) 752–776, [arXiv:1610.00678].

[93] **CMS** Collaboration, A. M. Sirunyan et al., *Measurement of differential cross sections and charge ratios for t-channel single top quark production in proton–proton collisions at $\sqrt{s}$ = 13 TeV*, *Eur. Phys. J. C* **80** (2020), no. 5 370, [arXiv:1907.08330].

[94] **ATLAS, CMS** Collaboration, M. Aaboud et al., *Combinations of single-top-quark production cross-section measurements and $|f_{LV} V_{tb}|$ determinations at $\sqrt{s}$ = 7 and 8 TeV with the ATLAS and CMS experiments*, *JHEP* **05** (2019) 088, [arXiv:1902.07158].

[95] E. R. Nocera, M. Ubiali, and C. Voisey, *Single Top Production in PDF fits*, *JHEP* **05** (2020) 067, [arXiv:1912.09543].

[96] E. L. Berger, J. Gao, C. P. Yuan, and H. X. Zhu, *NNLO QCD Corrections to t-channel Single Top-Quark Production and Decay*, *Phys. Rev.* **D94** (2016), no. 7 071501, [arXiv:1606.08463].

[97] **ATLAS** Collaboration, G. Aad et al., *Measurement of the production cross-section of a single top quark in association with a W boson at 8 TeV with the ATLAS experiment*, *JHEP* **01** (2016) 064, [arXiv:1510.03752].

[98] **CMS** Collaboration, S. Chatrchyan et al., *Observation of the associated production of a single top quark and a W boson in pp collisions at $\sqrt{s}$ =8 TeV*, *Phys. Rev. Lett.* **112** (2014), no. 23 231802, [arXiv:1401.2942].

[99] **ATLAS** Collaboration, M. Aaboud et al., *Measurement of the cross-section for producing a W boson in association with a single top quark in pp collisions at $\sqrt{s}$ = 13 TeV with ATLAS*, *JHEP* **01** (2018) 063, [arXiv:1612.07231].

[100] **CMS** Collaboration, A. M. Sirunyan et al., *Measurement of the production cross section for single top quarks in association with W bosons in proton-proton collisions at $\sqrt{s} = 13$ TeV*, *JHEP* (2018) [arXiv:1805.07399].

[101] **CMS** Collaboration, A. M. Sirunyan et al., *Measurement of the associated production of a single top quark and a Z boson in pp collisions at $\sqrt{s} =$ TeV*, *Phys. Lett.* **B779** (2018) 358–384, [arXiv:1712.02825].

[102] **ATLAS** Collaboration, M. Aaboud et al., *Measurement of the production cross-section of a single top quark in association with a Z boson in proton–proton collisions at 13 TeV with the ATLAS detector*, *Phys. Lett.* **B780** (2018) 557–577, [arXiv:1710.03659].

[103] **ATLAS** Collaboration, G. Aad et al., *Measurement of single top-quark production in association with a W boson in the single-lepton channel at $\sqrt{s} = 8$ TeV with the ATLAS detector*, arXiv:2007.01554.

[104] **ATLAS** Collaboration, G. Aad et al., *Observation of the associated production of a top quark and a Z boson in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, *JHEP* **07** (2020) 124, [arXiv:2002.07546].

[105] **CMS** Collaboration, A. M. Sirunyan et al., *Observation of Single Top Quark Production in Association with a Z Boson in Proton-Proton Collisions at $\sqrt{s} = 13$ TeV*, *Phys. Rev. Lett.* **122** (2019), no. 13 132003, [arXiv:1812.05900].

[106] **ATLAS** Collaboration, M. Aaboud et al., *Measurement of differential cross-sections of a single top quark produced in association with a W boson at $\sqrt{s} = 13$ TeV with ATLAS*, *Eur. Phys. J.* **C78** (2018), no. 3 186, [arXiv:1712.01602].

[107] **CMS Collaboration** Collaboration, *Measurement of differential cross sections for single top quark production in association with a W boson at $\sqrt{s} = 13$ TeV*, Tech. Rep. CMS-PAS-TOP-19-003, CERN, Geneva, 2020.

[108] **CMS Collaboration** Collaboration, *Using associated top quark production to probe for new physics within the framework of effective field theory*, Tech. Rep. CMS-PAS-TOP-19-001, CERN, Geneva, 2020.

[109] **ATLAS, CMS** Collaboration, G. Aad et al., *Measurements of the Higgs boson production and decay rates and constraints on its couplings from a combined ATLAS and CMS analysis of the LHC pp collision data at $\sqrt{s} = 7$ and 8 TeV*, *JHEP* **08** (2016) 045, [arXiv:1606.02266].

[110] **ATLAS** Collaboration, G. Aad et al., *Measurements of the Higgs boson production and decay rates and coupling strengths using pp collision data at $\sqrt{s} = 7$ and 8 TeV in the ATLAS experiment*, *Eur. Phys. J.* **C76** (2016), no. 1 6, [arXiv:1507.04548].

[111] **ATLAS** Collaboration, G. Aad et al., *Combined measurements of Higgs boson production and decay using up to 80 $fb^{-1}$ of proton-proton collision data at $\sqrt{s} = 13$ TeV collected with the ATLAS experiment*, *Phys. Rev. D* **101** (2020), no. 1 012002, [arXiv:1909.02845].

[112] **CMS** Collaboration, A. M. Sirunyan et al., *Combined measurements of Higgs boson couplings in proton–proton collisions at* $\sqrt{s} = 13\,TeV$, *Eur. Phys. J. C* **79** (2019), no. 5 421, [arXiv:1809.10733].

[113] **LHC Higgs Cross Section Working Group** Collaboration, S. Heinemeyer et al., *Handbook of LHC Higgs Cross Sections: 3. Higgs Properties*, arXiv:1307.1347.

[114] **LHC Higgs Cross Section Working Group** Collaboration, D. de Florian et al., *Handbook of LHC Higgs Cross Sections: 4. Deciphering the Nature of the Higgs Sector*, arXiv:1610.07922.

[115] **LHC Higgs Cross Section Working Group** Collaboration, S. Dittmaier et al., *Handbook of LHC Higgs Cross Sections: 2. Differential Distributions*, arXiv:1201.3084.

[116] **CMS** Collaboration, A. M. Sirunyan et al., *Measurement and interpretation of differential cross sections for Higgs boson production at* $\sqrt{s} = 13\,TeV$, *Phys. Lett. B* **792** (2019) 369–396, [arXiv:1812.06504].

[117] **ATLAS** Collaboration, M. Aaboud et al., *Combined measurement of differential and total cross sections in the* $H \to \gamma\gamma$ *and the* $H \to ZZ^* \to 4\ell$ *decay channels at* $\sqrt{s} = 13\,TeV$ *with the ATLAS detector*, *Phys. Lett. B* **786** (2018) 114–133, [arXiv:1805.10197].

[118] **ATLAS** Collaboration, M. Aaboud et al., *Measurement of VH,* $H \to b\bar{b}$ *production as a function of the vector-boson transverse momentum in 13 TeV pp collisions with the ATLAS detector*, *JHEP* **05** (2019) 141, [arXiv:1903.04618].

[119] **CMS** Collaboration, *Measurements of Higgs boson production via gluon fusion and vector boson fusion in the diphoton decay channel at* $\sqrt{s} = 13\,TeV$, .

[120] **ATLAS** Collaboration, M. Aaboud et al., *Measurements of Higgs boson properties in the diphoton decay channel with 36 fb$^{-1}$ of pp collision data at* $\sqrt{s} = 13\,TeV$ *with the ATLAS detector*, *Phys. Rev. D* **98** (2018) 052005, [arXiv:1802.04146].

[121] **CMS** Collaboration, A. M. Sirunyan et al., *Measurement of the inclusive and differential Higgs boson production cross sections in the leptonic WW decay mode at* $\sqrt{s} = 13\,TeV$, *JHEP* **03** (2021) 003, [arXiv:2007.01984].

[122] **ATLAS** Collaboration, G. Aad et al., *Measurements of WH and ZH production in the* $H \to b\bar{b}$ *decay channel in pp collisions at 13 TeV with the ATLAS detector*, *Eur. Phys. J. C* **81** (2021), no. 2 178, [arXiv:2007.02873].

[123] **ALEPH, DELPHI, L3, OPAL, LEP Electroweak** Collaboration, S. Schael et al., *Electroweak Measurements in Electron-Positron Collisions at W-Boson-Pair Energies at LEP*, *Phys. Rept.* **532** (2013) 119–244, [arXiv:1302.3415].

[124] **ATLAS** Collaboration, M. Aaboud et al., *Measurement of* $W^{\pm}Z$ *production cross sections and gauge boson polarisation in pp collisions at* $\sqrt{s} = 13\,TeV$ *with the ATLAS detector*, *Eur. Phys. J. C* **79** (2019), no. 6 535, [arXiv:1902.05759].

[125] **ATLAS** Collaboration, M. Aaboud et al., *Measurement of fiducial and differential $W^+W^-$ production cross-sections at $\sqrt{s} = 13$ TeV with the ATLAS detector, Eur. Phys. J. C* **79** (2019), no. 10 884, [`arXiv:1905.04242`].

[126] **CMS** Collaboration, A. M. Sirunyan et al., *Measurements of the $pp \rightarrow WZ$ inclusive and differential production cross section and constraints on charged anomalous triple gauge couplings at $\sqrt{s} = 13$ TeV, JHEP* **04** (2019) 122, [`arXiv:1901.03428`].

[127] **ATLAS Collaboration** Collaboration, *Measurement of $W^{\pm}Z$ production cross sections and gauge boson polarisation in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, Tech. Rep. ATLAS-CONF-2018-034, CERN, Geneva, Jul, 2018.

[128] **CMS Collaboration** Collaboration, *Studies of* W$^+$W$^-$ *production at $\sqrt{s} = 13$ TeV*, Tech. Rep. CMS-PAS-SMP-18-004, CERN, Geneva, 2020.

[129] **ATLAS** Collaboration, *Measurements of differential cross-sections in four-lepton events in 13 TeV proton-proton collisions with the ATLAS detector*, .

[130] M. Grazzini, S. Kallweit, and M. Wiesemann, *Fully differential NNLO computations with MATRIX, Eur. Phys. J. C* **78** (2018), no. 7 537, [`arXiv:1711.06631`].

[131] J. Brehmer, F. Kling, T. Plehn, and T. M. P. Tait, *Better Higgs-CP Tests Through Information Geometry, Phys. Rev.* **D97** (2018), no. 9 095017, [`arXiv:1712.02350`].

[132] F. Feroz and M. Hobson, *Multimodal nested sampling: an efficient and robust alternative to MCMC methods for astronomical data analysis, Mon. Not. Roy. Astron. Soc.* **384** (2008) 449, [`arXiv:0704.3704`].

[133] **NNPDF** Collaboration, R. Abdul Khalek et al., *Parton Distributions with Theory Uncertainties: General Formalism and First Phenomenological Studies, Eur. Phys. J. C* **79** (2019), no. 11 931, [`arXiv:1906.10698`].

[134] **NNPDF** Collaboration, R. Abdul Khalek et al., *A first determination of parton distributions with theoretical uncertainties, Eur. Phys. J.* **C** (2019) 79:838, [`arXiv:1905.04311`].

[135] **The NNPDF** Collaboration, R. D. Ball et al., *Fitting Parton Distribution Data with Multiplicative Normalization Uncertainties, JHEP* **05** (2010) 075, [`arXiv:0912.2276`].

[136] G. T. Bodwin and H. S. Chung, *New method for fitting coefficients in standard model effective theory, Phys. Rev. D* **101** (2020), no. 11 115039, [`arXiv:1912.09843`].

[137] S. Carrazza, S. Forte, Z. Kassabov, J. I. Latorre, and J. Rojo, *An Unbiased Hessian Representation for Monte Carlo PDFs, Eur. Phys. J.* **C75** (2015), no. 8 369, [`arXiv:1505.06736`].

[138] S. Carrazza, S. Forte, Z. Kassabov, and J. Rojo, *Specialized minimal PDFs for optimized LHC calculations, Eur. Phys. J.* **C76** (2016), no. 4 205, [`arXiv:1602.00005`].

[139] R. D. Ball, S. Carrazza, L. Del Debbio, S. Forte, J. Gao, et al., *Parton Distribution Benchmarking with LHC Data, JHEP* **1304** (2013) 125, [`arXiv:1211.5142`].

[140] **ATLAS Collaboration** Collaboration, *Determination of the parton distribution functions of the proton from ATLAS measurements of differential W and Z/γ* and tt̄ cross sections*, Tech. Rep. ATL-PHYS-PUB-2018-017, CERN, Geneva, Aug, 2018.

[141] J. Brehmer, K. Cranmer, G. Louppe, and J. Pavez, *Constraining Effective Field Theories with Machine Learning*, Phys. Rev. Lett. **121** (2018), no. 11 111801, [arXiv:1805.00013].

[142] R. T. D'Agnolo and A. Wulzer, *Learning New Physics from a Machine*, Phys. Rev. D **99** (2019), no. 1 015014, [arXiv:1806.02350].

[143] S. Chen, A. Glioti, G. Panico, and A. Wulzer, *Parametrized classifiers for optimal EFT sensitivity*, arXiv:2007.10356.

[144] **Muon g-2** Collaboration, B. Abi et al., *Measurement of the Positive Muon Anomalous Magnetic Moment to 0.46 ppm*, Phys. Rev. Lett. **126** (2021), no. 14 141801, [arXiv:2104.03281].