PRELIMINARY

LHC EFT WG Area 4 - Fits and related systematics

Editors: N. Berger, J. de Blas, F. Canelli, N. Castro,
P. Govoni and G. Petrucciani

April 30, 2021

So far, the activities of this Area 4 have focused on:

1. Identifying different computational fitting frameworks publicy available with capabilities for performing EFT studies.
2. Reviewing the fitting experience in the SMEFT framework by the LHC ATLAS and CMS collaborations, and discussing/exploring the possibility of combined analyses.
3. Discussing the usability and possible improvements in the presentation of public experimental results, so they can be efficiently used for global studies combining information from different processes.

These topics were covered in two meetings: the first meeting on January 27, 2021, dedicated exclusively to the presentation of fitting frameworks, and a second meeting, together with the Area 3 of the LHC EFT WG –dedicated to *Experimental measurements and observables*– on February 22, 2021. In this document we briefly summarize some of the points discussed in these meetings.

1 Fitting frameworks

The following, publicy available, codes were discussed during the meeting on January 27, 2021. In alphabetic order:

1.1 EFTFitter

EFTfitter [1] provides a computational framework to study constraints on any model. It uses a Bayesian statistical approach for parameter inference and relies on the assumption of Gaussian measurements. There is however no restriction on the dependence of the observables on the model parameters. It must be noted that the code per-se only provides the statistical tool. Any set of observables, together with their predictions and experimental values must be provided externally. See Ref. [2] for a dedicated SMEFT study in the Top/Bottom sector making use of EFTfitter.

1.2 Fitmaker

The current implementation of Fitmaker is focused on SMEFT analyses of Electroweak (EW), Top, diboson and Higgs observables (see [3] for a recent global analysis of the SMEFT constraints from all these processes). Dimension-6 SMEFT contributions are considered at the linear level (interference with the SM). The code allows both frequentist and Bayesian statistical analyses via χ^2 minimization and Nested sampling algorithms, respectively. Some of the limitations of the code is that theory uncertainties (both in the SM and beyond) are not directly implemented for all observables (included in others, in the uncertainty of the experimental inputs).

1.3 HEPfit

HEPfit [4] is designed as a flexible tool to perform fits to any type of experimental contraints and in any particular model. Both model and observables can be implemented externally by the user, or chosen among a library of models/observables available. The SM and the dimension-6 SMEFT are currently implemented for the study or EW, Higgs, diboson and Flavor observables. Dimension-6 SMEFT corrections are currently included at the linear level. Including quadratic contributions as well as other types of observables, e.g. Top, is work in progress. The default statistical framework is Bayesian but the code can be used as library to compute the likelihood and be used in any framework. Theory uncertainties can be included via nuisance parameters and experimental observables added in basically any numerical form. In particular, current work focuses on the implementation of Deep Neural Networks as input/output of the code [5].

1.4 Sfitter

Sfitter [6] is another general purpose tool for likelihood calculation and parameter inference, capable but not restricted to SMEFT studies. The current implementation of the SMEFT can be used for studies of EW, Higgs, diboson and Top observables. As some of the other codes it is capable of studying different types of observables, including kinematical distributions. Dimension-6 SMEFT contributions to physical processes are implemented with the possibility of including quadratic corrections in the new physics.

1.5 SMEFit

As in the case of some of the previous codes, the scope of the SMEfit [7] goes beyond EFT analyses. The code itself was built upon experience in global PDF determinations and can be also used for that purpose, as well as for a simulatenous determination of PDFs with EFT coefficients. The SMEFT implementation is available for various processes in the Top, Higgs and diboson sector. The SMEFit methodology is based in the generations of large samples of MC replicas to reconstruct the probability distribution in the space of experimental data, obtaining the SMEFT coefficients via frequentist parameter inference (cost function minimization). It also allows Bayesian inference studies via Nested sampling methods, to produce robust statistical results when non-linearities are present (though at the price of reduced speed when the number of parameters is large).

1.6 Further considerations

During the discussion time, it was emphasized the need to have the possibility to do cross-checks between the results obtained by the different codes, so one can test the different SMEFT implementations or the compatibility of the different statistical engines/algorithms used by each tool. For this purpose it was suggested to:

1. Define a series of *benchmark fitting scenarios* to perform the comparison.

These benchmarks could be selected simply by identifying common subsets of observables and operators available in two or more codes. Given the overlap and complementarities in terms of physics processes implemented in each of these codes, by comparing the results for several of these benchmarks it should be possible to validate the implementation in a given tool against at least one other code, for most of the processes.

A proper comparison, of course, also requires selecting common SMEFT assumptions [1], some of which, e.g. flavor symmetries, may depend on the types of processes considered in each benchmark. In particular, for a first comparison and validation of the basic SMEFT implementations, the suggestion is to perform the comparisons including only leading-order EFT effects of $O(\Lambda^{-2})$, using trivial (flat) priors for the SMEFT Wilson coefficients and neglecting theory uncertainties. Some of these assumptions could be lifted in future cross checks. For the actual comparison of results, and given the previous assumptions, it would be enough to compare the positions of the mode and covariance matrices for the Wilson coefficients. For the purpose of identifying the source of possible discrepancies, however, it is also recommended to provide directly the results for 1D fits assuming only one operator at a time. A more in-depth cross check could include comparing predictions for different observables as well as other relevant statistical information. For that purpose, it was also suggested to:

2. Agree on a common output format so that one can easily compare the statistical results for the model parameters to be inferred from the data as well as the predictions for different observables from the fits.

As second point discussed at the end of the meeting was related to usability of these codes by the experimental collaborations as part of the analyses, e.g. to account for existing bound in operators entering directly in the processes they study. [2] For this purpose, it would be good to know if there is any particular requirement from ATLAS/CMS/LHCb so that these tools can be easily implemented in the experimental analysis workflow.

[1] Relating results in different operator bases could be done using publicy available tools, e.g. [8].

[2] In this regard, however, it is also important that, even if the experimental groups want to account for other constraints, their results should also always contain the bounds derived from their analysis, in isolation, to avoid double-counting in future combinations.

2 Fitting experience in the EFT by the LHC collaborations

In what follows we briefly comment on the current efforts from the ATLAS and CMS collaborations that were presented in the meeting of February 22, 2021, in terms of re-interpreting experimental measurements in the frame of the SMEFT formalism. One general remark that applies to most of these studies, is that they are typically restricted to incomplete sets of dimension-6 interactions modifying the corresponding process. Therefore, this may be problematic from the point of view of combining these results with other measurements in a global fit. In principle, although including all contributions may lead to flat directions, these can be tractable, e.g., presenting the results in terms of the combinations of operators constrained by the analysis at the linear level, as already illustrated in the ATLAS contribution.

2.1 ATLAS experience

ATLAS is currently performing EFT interpretations in different types of processes, covering the EW, Higgs and Top sectors. (See, e.g., Refs. [9], [10, 11] and [12], respectively.) All the most recent analyses are performed consistently in the CP-even sector of the dimension-6 SMEFT, using the Warsaw basis [13] and the (G_F, m_W, m_Z) input scheme.[3]. The analyses aim to include the complete set of the operators in the set described above, making use of principal component analysis to identify the operator combinations to which the measurements are sensitive, and to isolate "flat" directions.

Current plans for future work include defining the setup for a first ATLAS EFT global combination this year. The goal is to combine the constraints from the Higgs Simplified Template Cross Section (STXS) measurements with the results from EW diboson and VBF Z processes and, at least one process from the Top sector. The combination could also include non-LHC constraints (e.g. EW precision tests from LEP/SLD) to help reduce the number of flat directions.

2.2 CMS experience

CMS discussed the experience in global EFT fits in the Higgs and Top sectors, while no recent global fits of the existing EFT limits put by single analyses in the EW sector – as in the ATLAS case – were reported.

Higgs studies include both the reinterpretation of the STXS measurements in the EFT formalism, as well as dedicated optimal analyses (e.g. [15]). While current parameterization of the STXS bins is based in the HEL model [16] – which relies in the conventions of the SILH effective Lagrangian [17] – the plan is to switch to the Warsaw basis as well and investigate similar approaches to ATLAS. The EFT studies in the Top sector are based

[3]Previous EW studies were usually performed using the HISZ parameterization [14] but newer studies are also moving towards the use of the Warsaw basis.

on the analysis of the associated production of top quark pairs and charged leptons [18] in terms of a set of 16 Wilson coefficients, for which CMS is able to provide results both individually, as well as in a global analysis floating all parameters.

2.3 ATLAS and CMS Combinations

The experience in combining results from ATLAS and CMS was discussed in the context of Higgs studies, where a major effort was pursued in the preparation of the Run 1 combination of the ATLAS and CMS measurements of the Higgs properties [19]. This combination study, however, did not include any re-interpretation in the context of the EFT framework.

Currently, ATLAS and CMS are in the initial preparatory phase of the Run 2 combination, where it is expected to have EFT-interpreted results based on the combination of the STXS measurements. A further combination covering other sectors as well (top, EW...) is eventually foreseen on a longer timescale. These efforts will require first the agreement on common assumptions of the EFT interpretation, such as:

- The operator basis and the subset of operators to be considered, if not including the whole basis (e.g. a restriction to CP-even terms only).
- The Standard Model EW input scheme.
- Truncation of the EFT results, e.g. inclusion of C^2/Λ^4 terms, ...
- The scheme describing the uncertainties on the EFT theory predictions

All these aspects are part of the items treated in Area 1 of this *LHC EFT WG*. While an early agreement on these would be beneficial, it is not clear whether this will arrive in time to influence the setup of this Run 2 combination.

3 On the presentation of experimental results and their use as input in global EFT studies

The way public LHC experimental results are presented is sometimes not ideal for the utilization/re-interpretation of such measurements by external groups. In particular, in order to be able to use these results in global combinations, it should be possible to reconstruct from public data some approximation to the likelihood leading to the experimental results. [4] Some recommendations were presented for this purpose:

- At the very least, the information needed to construct the full covariance matrix should be provided.

[4]Note that, in some cases, e.g. dedicated EFT analyses, even this may not be enough, if one attempts to use such likelihoods under different assumptions that those used in their construction.

- To help in combining different results, such covariance matrices should be provided with as much granularity as possible, separating contributions from different sources of experimental errors that could be correlated across experiments, and separating also the theory uncertainties.

- Ideally, the full likelihood could be made public in some numerical format. These are very complicated functions involving many parameters, but they could be provided either fully or, e.g., using machine learning techniques in the form of deep-neutral networks trained to reproduce the true likelihood.

Other issues concerning the re-interpretability of the experimental results were also discussed, e.g. treatment of flat directions in the presentation of results, the treatment of EFT contribution in the backgrounds, etc. Regarding the re-interpretability of these experimental analyses, it should be noted the significant overlap between the interest of this WG and the activities of the *Forum on the Interpretation of the LHC Results for BSM studies*, where this and other issues are covered, see e.g. [20].

4 LHC EFT WG exercise: Towards a first combination of ATLAS and CMS EFT constraints

A global fit of experimental results yielding from an unprecedented set of independent analyses attained by at least two independent Collaborations, from several different working groups inside each of them, with a theory model with a wide number of parameters of interest with respect to the past experiences will demand for a very solid technical infrastructure. In particular, such a tool will have to behave well in presence of:

- several different analysis implementations, which may be based for example on observed event counts, histograms, or fitted analytical shapes;

- a very large set of systematic uncertainties, differently correlated among initial states, final states, data sets;

- different ways of simulating the EFT behaviour.

Is is therefore one prime objective of the Area 4 to bring the ATLAS and CMS Collaborations to the level where a robust procedure for the combination is built. Hence, it is proposed to perform an exercise to combine ATLAS and CMS constraints in the EFT framework. The scope of this exercise is not to provide a comprehensive combination of the available information from all LHC process, rather to identify key aspects that may need to be addressed in such a study in the future. To be more precise, the purpose of the exercise is:

1. To implement the conventions and recommendations of the WG.

2. To identify possible technical difficulties that may arise in the process of combining ATLAS and CMS inputs, and determine the optimal configuration of the fitting techniques to be used in future combinations.

The details of this exercise will be discussed during future meetings.

References

[1] N. Castro, J. Erdmann, C. Grunwald, K. Kröninger, and N.-A. Rosien, *EFTfitter—A tool for interpreting measurements in the context of effective field theories*, *Eur. Phys. J. C* **76** (2016), no. 8 432, [arXiv:1605.05585].

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

[3] 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.

[4] J. de Blas et al., ***HEPfit**: a code for the combination of indirect and direct constraints on high energy physics models*, *Eur. Phys. J. C* **80** (2020), no. 5 456, [arXiv:1910.14012].

[5] A. Coccaro, M. Pierini, L. Silvestrini, and R. Torre, *The DNNLikelihood: enhancing likelihood distribution with Deep Learning*, *Eur. Phys. J. C* **80** (2020), no. 7 664, [arXiv:1911.03305].

[6] R. Lafaye, T. Plehn, and D. Zerwas, *SFITTER: SUSY parameter analysis at LHC and LC*, hep-ph/0404282.

[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] A. Falkowski, B. Fuks, K. Mawatari, K. Mimasu, F. Riva, and V. Sanz, *Rosetta: an operator basis translator for Standard Model effective field theory*, *Eur. Phys. J. C* **75** (2015), no. 12 583, [arXiv:1508.05895].

[9] **ATLAS** Collaboration, G. Aad et al., *Differential cross-section measurements for the electroweak production of dijets in association with a Z boson in proton–proton collisions at ATLAS*, *Eur. Phys. J. C* **81** (2021), no. 2 163, [arXiv:2006.15458].

[10] **ATLAS Collaboration** Collaboration, *Interpretations of the combined measurement of Higgs boson production and decay*, https://cds.cern.ch/record/2743067.

[11] **ATLAS Collaboration** Collaboration, *Combined effective field theory interpretation of $H \to WW^*$ and WW measurements using ATLAS data*, https://cds.cern.ch/record/2758785.

[12] **ATLAS Collaboration** Collaboration, *Inclusive and differential measurement of the charge asymmetry in $t\bar{t}$ events at 13 TeV with the ATLAS detector*, https://cds.cern.ch/record/2682109.

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

[14] K. Hagiwara, S. Ishihara, R. Szalapski, and D. Zeppenfeld, *Low-energy effects of new interactions in the electroweak boson sector*, *Phys. Rev. D* **48** (1993) 2182–2203.

[15] **CMS** Collaboration, *Constraints on anomalous Higgs boson couplings to vector bosons and fermions in production and decay in the $H \to 4\ell$ channel*, .

[16] C. Hays, V. Sanz Gonzalez, and G. Zemaityte, *Constraining EFT parameters using simplified template cross sections*, .

[17] G. F. Giudice, C. Grojean, A. Pomarol, and R. Rattazzi, *The Strongly-Interacting Light Higgs*, *JHEP* **06** (2007) 045, [hep-ph/0703164].

[18] **CMS** Collaboration, A. M. Sirunyan et al., *Search for new physics in top quark production with additional leptons in proton-proton collisions at $\sqrt{s} = 13$ TeV using effective field theory*, *JHEP* **03** (2021) 095, [arXiv:2012.04120].

[19] **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].

[20] **LHC Reinterpretation Forum** Collaboration, W. Abdallah et al., *Reinterpretation of LHC Results for New Physics: Status and Recommendations after Run 2*, *SciPost Phys.* **9** (2020), no. 2 022, [arXiv:2003.07868].