

YSF talk: Search for new physics in top quark production with additional leptons using the framework of effective field theory

KELCI MOHRMAN⁽¹⁾(²) on behalf of the CMS COLLABORATION

⁽¹⁾ *University of Notre Dame - Notre Dame, IN, USA*

⁽²⁾ *University of Florida - Gainesville, FL, USA*

received 10 October 2023

Summary. — This analysis presents a search for new physics impacting associated top quark production within the context of effective field theory (EFT). Making use of 138 fb^{-1} of proton-proton collisions at $\sqrt{s} = 13 \text{ TeV}$ collected by the CMS experiment at the CERN LHC during 2016-2018, the analysis selects events with multiple jets, b-tagged jets, and two same-sign leptons or three or more leptons. The data are further binned according to kinematical distributions associated with the transverse momentum of the objects in the event. The effects of 26 dimension-six EFT operators are incorporated into the simulation, allowing for the predicted yields in each observable bin to be parameterized in terms of the strengths of the 26 EFT operators. A simultaneous fit of the 26 EFT parameters to the observed data is performed, and confidence intervals are extracted. The results are consistent with the standard model prediction.

1. – Introduction

While the standard model (SM) of particle physics has been very successful in describing fundamental particles and their interactions, there are compelling indications (*e.g.*, the strong evidence for dark matter) that the SM is not complete. However, there is no guarantee that new physics particles will be light enough to be produced on-shell at the center-of-mass energies produced by the CERN LHC, which will not increase significantly throughout its remaining years of operation. For these reasons, indirect methods of probing higher energy scales are becoming increasingly important and timely in the search for new physics at the energy frontier. Standard model effective field theory (EFT) is an example of such an indirect probe. Treating the SM Lagrangian (\mathcal{L}_{SM}) as the lowest order term in an expansion of higher dimensional operators, the EFT Lagrangian can be written as follows:

$$(1) \quad \mathcal{L}_{\text{EFT}} = \mathcal{L}_{\text{SM}} + \sum_{d,i} \frac{c_i^d}{\Lambda^{d-4}} \mathcal{O}_i^d.$$

Here, the \mathcal{O}_i^d are the EFT operators (composed of products of SM fields and their derivatives) at dimension d , the c_i^d are dimensionless parameters called Wilson coefficients (WCs) that describe the strength of the interactions, and the Λ is the energy scale of the new physics.

Since EFT is a relatively model-independent method of describing the potential off-shell effects of heavy new physics, it can be useful for new physics searches across all sectors at the LHC. However, here we focus specifically on the top quark sector, studying processes in which one or more top quarks are produced in association with other heavy particles (such as Higgs, W, or Z boson, or even other top quarks). Referred to as associated top quark processes, these interactions involve multiple heavy particles and are relatively rare; as we are just now reaching the point where we have accumulated enough statistics to start probing them in detail, these processes may provide an interesting avenue through which to search for new physics. The analysis summarized in these proceedings [1] studies associated top quark processes in the context of EFT in order to probe 26 dimension-six EFT operators. The preliminary results presented in these proceedings has been superseded by [2].

2. – Analysis strategy

The analysis aims to study all dimension-six EFT operators involving top quarks that can significantly impact associated top quark processes. As shown in table I, these 26 operators can be categorized into four groups: operators involving two heavy quarks and bosons, operators involving two heavy quarks and leptons, operators involving two heavy quarks and two light quarks, and operators involving four heavy quarks. The definitions of the WCs and the corresponding operators can be found in [3].

These 26 operators can impact various associated top quark processes, which can lead to a variety of different final-state signatures. However, this analysis focuses specifically on the leptonic decays of the associated top quark processes. Referred to as multilepton final states, these include signatures with two same-sign leptons, or three or more leptons. While leptonic final states have many experimental advantages, a multilepton EFT analysis also gives rise to several challenges, primarily stemming from the fact that many processes and EFT effects can all contribute to the same final-state signatures. For example, if we consider a two lepton same-sign channel, there will be contributions from SM $t\bar{t}H$ and $t\bar{t}W$ (as well as $t\bar{t}Z$ if one of the leptons is missed); these processes can be impacted by various EFT operators, producing interference among themselves and with

TABLE I. – List of WCs included in this analysis. The definitions of the WCs and the definitions of the corresponding operators can be found in table 1 of ref. [3]. Note that in order to allow MadGraph to properly handle the emission of gluons from the vertices involving the c_{tG} WC, an extra factor of the strong coupling is applied to the c_{tG} coefficients. [1]

Operator category	WCs
Two heavy quarks	$c_{t\varphi}, c_{\varphi Q}^-, c_{\varphi Q}^3, c_{\varphi t}, c_{\varphi tb}, c_{tW}, c_{tZ}, c_{bW}, c_{tG}$
Two heavy quarks two leptons	$c_{Q\ell}^{3(\ell)}, c_{Q\ell}^{-3(\ell)}, c_{Qe}^{(\ell)}, c_{t\ell}^{(\ell)}, c_{te}^{(\ell)}, c_t^{S(\ell)}, c_t^{T(\ell)}$
Two light quarks two heavy quarks	$c_{Qq}^{31}, c_{Qq}^{38}, c_{Qq}^{11}, c_{Qq}^{18}, c_{tq}^1, c_{tq}^8$
Four heavy quarks	$c_{QQ}^1, c_{Qt}^1, c_{Qt}^8, c_{tt}^1$

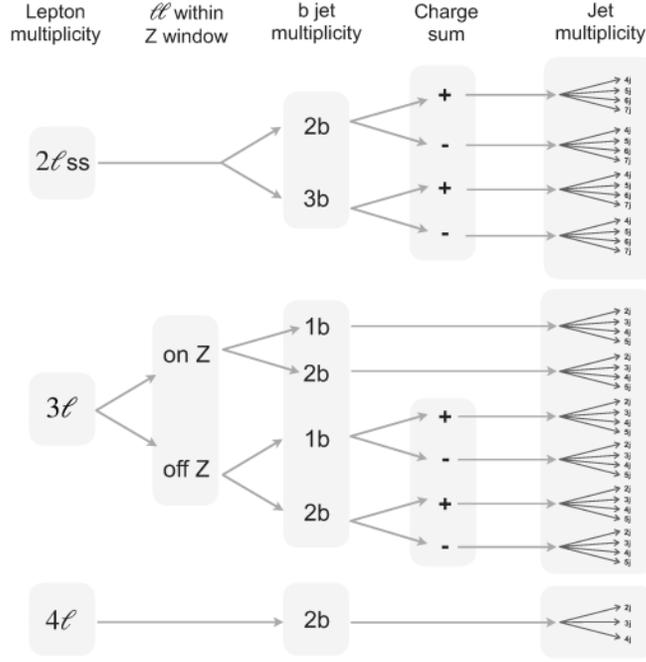


Fig. 1. – Summary of the event selection categorization for the analysis presented in [1].

the SM. The final states are thus complicated admixtures of processes and effects. Since it is not possible to isolate the contributions, it is important to ensure that the EFT effects are studied consistently across all channels simultaneously.

To this end, the analysis studies the EFT effects directly at detector level, using an approach developed in [4]. With this approach, the EFT effects are incorporated into the event weights of each of the simulated events (using the MadGraph event reweighting technique [5]), allowing the each of the event weights to be parametrized as a 26-dimensional quadratic function in terms of the 26 WCs. Once the parametrization for each event has been obtained, the parametrization of the yield in an arbitrary observable bin can be obtained by summing the weight functions of each simulated event that passes the selection criteria for the given bin. In this way, the predicted yield across all of the bins in the analysis can be obtained as a function of the WCs, allowing the prediction to be compared to the observed data the confidence intervals for the WCs to be extracted.

3. – Samples and event selection

The analysis makes use of 138 fb^{-1} of proton-proton collision data collected by the CMS detector [6] during the years 2016-2018. The analysis selects events with two leptons with the same charge or three or more leptons; the selection also requires multiple jets (at least one of which must be b tagged). In the SM, these final states are mainly populated by associated top quark processes ($t\bar{t}H$, $t\bar{t}l\nu$, $t\bar{t}l\bar{l}$, $t\bar{t}lq$, tHq , and $t\bar{t}t$), which

constitute the signal processes for this analysis. Any process that can populate the same final state categories but is not significantly impacted by the 26 WCs studied by this analysis is considered to be a background contribution. The primary backgrounds arise from diboson processes and from processes with misidentified leptons. The backgrounds are modeled with a combination of Monte Carlo simulation and data-driven techniques.

The events in each lepton multiplicity category are further subdivided according to the jet multiplicity, the b-tag multiplicity, the sum of the charges of the leptons, and by whether or not there is a pair of leptons consistent with the decay of a Z boson. The goal of the event selection categorization is to isolate the different processes as much as possible in order to obtain a set of bins with unique admixtures of the contributions. Since the EFT operators affect the processes in different ways, this separation helps to provide sensitivity to the EFT effects. This event selection categorization results in 43 independent categories, which are visualized in the schematic shown in fig. 1.

To gain additional sensitivity to the EFT effects, the events in each of the 43 sub-categories are binned according to kinematical distribution. In general, EFT affects tend to grow with energy, so studying variables related to the objects with the highest transverse momentum (p_T) in the event can help to increase the sensitivity to the EFT. The specific kinematical distribution that is used depends on the category. In the majority of the categories, a variable referred to as $p_T(\ell j)_0$ is used. To construct this variable, pairs of objects are formed from the collection of leptons and jets in the event; the pairs may consist of either two leptons, two jets, or a lepton and a jet. The p_T of the objects in each pair are summed vectorially, and the $p_T(\ell j)_0$ corresponds to the p_T of the pair with the leading p_T . In most of the on-Z categories, a variable corresponding to the p_T of the two leptons from the Z (referred to as $p_T(Z)$) is used. This variable helps to provide sensitivity to EFT operators involving the Z boson. However, it should be noted that the $p_T(Z)$ variable is not used in the 3ℓ -on-Z 2b categories with low jet multiplicity (2j and 3j bins) because two operators from the two-heavy-two-light group (c_{Qq}^{31} , c_{Qq}^{38}) derive significant sensitivity from these event selection categories, but do not involve a Z boson in the EFT vertex. Thus, in order to retain sensitivity to these two WCs, the $p_T(\ell j)_0$ is used in these two on-Z categories [1]. After binning the events in the 43 selection categories according to $p_T(\ell j)_0$ or $p_T(Z)$, there are a total of 178 bins in the analysis.

4. – Results

In each of the 178 analysis bins, the prediction for each signal sample is parameterized as a quadratic in terms of the WCs, allowing the prediction to be compared to the data and the confidence intervals for the WCs to be extracted. The 26 WCs are the parameters of interest in the likelihood fit, and systematic uncertainties are treated as nuisance parameters. One-dimensional scans are performed in which we step through a series of values for the given WC, performing a maximum likelihood fit at each step. In the likelihood fit, the scanned WC is held at the given value of the step in the scan, while the other 25 WCs are either profiled or fixed to their SM values of zero. In this way, the 1σ and 2σ confidence intervals can be extracted for all 26 WCs (for both the profiled and fixed scenarios). The results of the likelihood fits are listed in table II and are shown graphically in fig. 2. Since there is no reason to assume that only a single WC may be non-zero (as is enforced in the case where the other WCs are fixed to their SM values of zero), the profiled confidence intervals represent the more general results. As shown in fig. 2, all of the 2σ profiled confidence intervals overlap with the SM prediction of zero, indicating that the analysis has not identified any signs of new physics.

TABLE II. – The 2σ uncertainty intervals extracted from the likelihood fits described in sec. 4. The intervals are shown for the case where the other WCs are profiled, and the case where the other WCs are fixed to their SM values of zero. [1]

WC/ Λ^2 [TeV^{-2}]	2σ Interval (others profiled)	2σ Interval (others fixed to SM)
$c_t^{T(\ell)}$	[-0.37, 0.37]	[-0.40, 0.40]
$c_t^{S(\ell)}$	[-2.60, 2.59]	[-2.80, 2.80]
$c_{te}^{(\ell)}$	[-1.76, 2.20]	[-1.90, 2.39]
$c_{t\ell}^{(\ell)}$	[-1.78, 2.10]	[-2.01, 2.20]
$c_{Qe}^{(\ell)}$	[-1.89, 1.94]	[-2.04, 2.12]
$c_{Q\ell}^{-\ell}$	[-1.56, 2.27]	[-1.80, 2.33]
$c_{Q\ell}^{3(\ell)}$	[-2.81, 2.54]	[-2.68, 2.58]
$c_{\varphi t}$	[-10.76, 7.91]	[-4.95, 3.19]
$c_{\varphi tb}$	[-3.23, 3.23]	[-3.15, 3.19]
$c_{\varphi Q}$	[-0.81, 2.01]	[-0.84, 1.91]
c_{bW}	[-0.75, 0.76]	[-0.75, 0.75]
c_{tG}	[-0.27, 0.24]	[-0.22, 0.25]
$c_{\varphi Q}^-$	[-6.09, 8.20]	[-2.66, 2.95]
$c_{t\varphi}$	[-8.98, 2.85]	[-7.68, 2.15]
c_{tZ}	[-0.70, 0.63]	[-0.58, 0.59]
c_{tW}	[-0.54, 0.45]	[-0.47, 0.41]
c_{Qt}^1	[-2.71, 2.66]	[-2.75, 2.62]
c_{Qt}^8	[-5.15, 5.74]	[-5.24, 5.66]
c_{QQ}^1	[-3.03, 3.28]	[-3.04, 3.28]
c_{tt}^1	[-1.56, 1.60]	[-1.54, 1.63]
c_{tq}^8	[-0.67, 0.25]	[-0.68, 0.24]
c_{Qq}^{18}	[-0.68, 0.21]	[-0.67, 0.21]
c_{tq}^1	[-0.21, 0.21]	[-0.22, 0.20]
c_{Qq}^1	[-0.19, 0.19]	[-0.19, 0.19]
c_{Qq}^{38}	[-0.17, 0.16]	[-0.17, 0.16]
c_{Qq}^{31}	[-0.08, 0.07]	[-0.08, 0.07]

To visualize the results, predicted yields before and after the likelihood fit may be plotted. The predicted yields before the likelihood fit (*i.e.*, where the predictions are set to the SM values) is referred to as the “prefit” scenario, while the predicted yields after the likelihood fit (*i.e.*, where the free parameters are set to their best fit values) is referred to as the “postfit” scenario. Figure 3 shows the prefit and postfit yields for the 178 analysis bins.

Pairs of WCs are also scanned simultaneously (with the remaining 24 WCs either profiled or fixed to their SM values of zero) in order to explore correlations among the WCs. Figure 4 shows two example scans in which correlations among the WCs are visible.

For most of the WCs studied in this analysis, the results are limited by the statistical uncertainty rather than systematic uncertainties. However, for the cases where systematic uncertainties are also important, the leading contributions arise from the uncertainties associated with the higher-order (next-to-leading order or better) cross sections to which the signal samples are normalized.

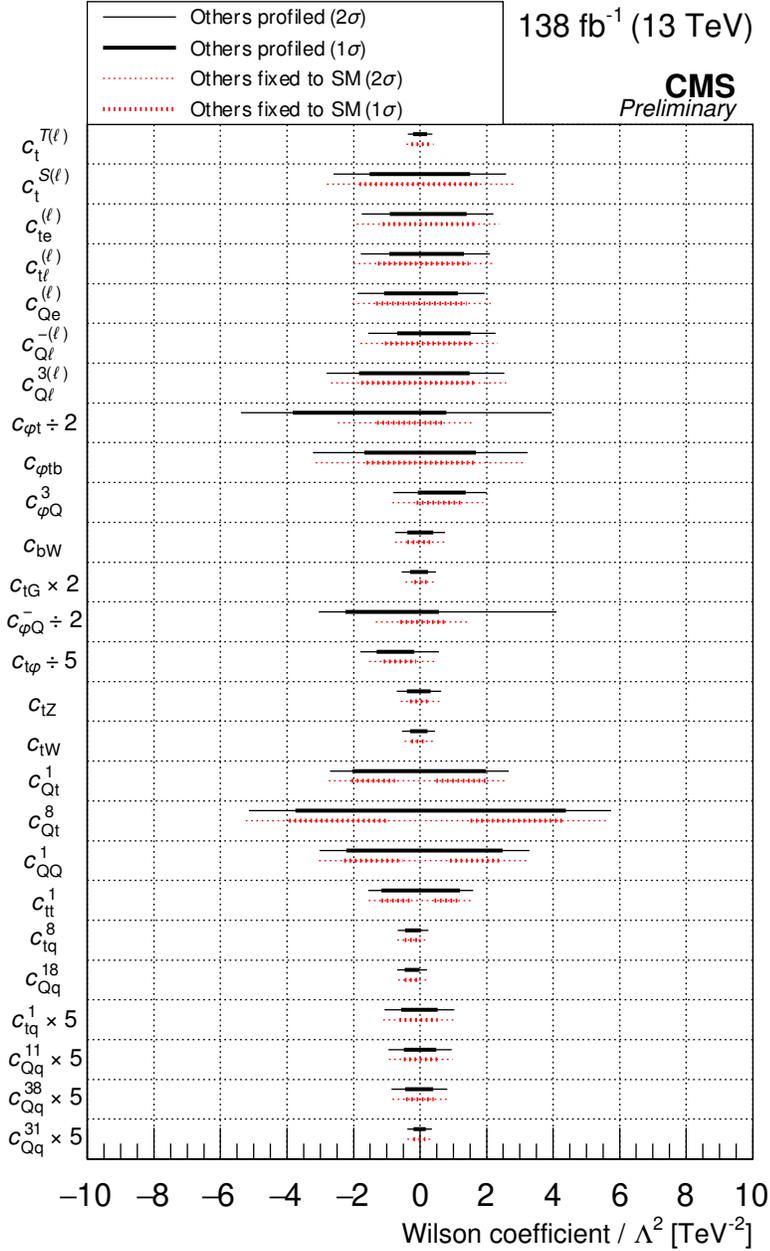


Fig. 2. – Summary of CIs extracted from the likelihood fits described in sec. 4. WC 1σ (thick line) and 2σ (thin line) uncertainty intervals are shown for the case where the other WCs are profiled (in black), and the case where the other WCs are fixed to their SM values of zero (in red). To make the figure more readable, the intervals for $c_{t\varphi}$ were scaled by $1/5$, the intervals for $c_{\varphi t}$ and $c_{\varphi Q}^-$ were scaled by $1/2$, the intervals for c_{tG} were scaled by 2 , and the intervals for c_{tq}^1 , c_{Qq}^{11} , c_{Qq}^{38} , and c_{Qq}^{31} were all scaled by 5 [1].

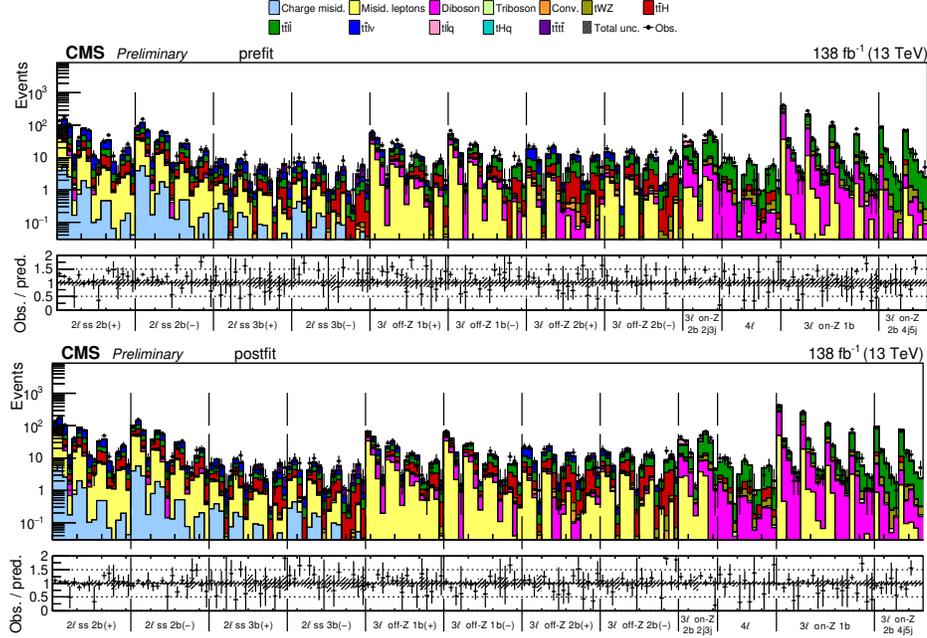


Fig. 3. – Expected yields prefit (above) and postfit (below). All kinematic variables are shown. Each event category (*e.g.*, $2lss$) is sub-divided into its jet multiplicity components. For example, the first four sub-bins of the $2lss$ bin are the $p_T(\ell_j)_0$ variable for four jets, the next four sub-bins are the $p_T(\ell_j)_0$ variable for 5 jets, etc. The postfit values are obtained by simultaneously fitting all 26 WCs and the nuisance parameters. The process labeled “Conv.” corresponds to the photon conversion background, “Misid. leptons” corresponds to misidentified leptons, and “Charge misid.” corresponds to leptons with a mismeasured charge. The lower panel contains the ratios of the observed yields over the expected. The error bands correspond to nuisance parameter and WC uncertainties added in quadrature. All yields are plotted on a log scale [1].

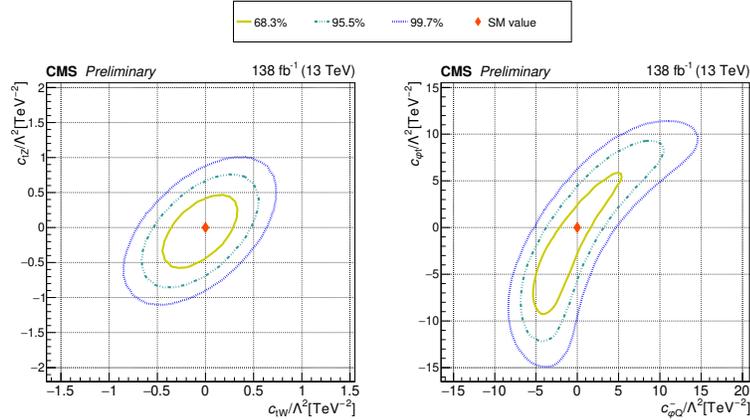


Fig. 4. – The observed 68.3%, 95.5%, and 99.7% confidence contours of a 2D scan for c_{tW} and c_{tZ} with the other WCs profiled (left) and for $c_{\varphi Q}^-$ and $c_{\varphi t}$ with the other WCs profiled (right). Diamond markers show the SM prediction [1].

5. – Summary

The analysis presented in [1] performs a search for new physics impacting associated top quark production in multilepton final states using the framework of EFT. In order to gain sensitivity to the EFT effects, the events in each category are binned according to kinematical distributions related to the p_T of the objects in the event. Studying 26 dimension-six EFT operators, the analysis parametrizes the predicted yields in the analysis bins as 26-dimensional quadratic functions in terms of the WCs. A simultaneous fit of the 26 WCs to the data is performed in order to extract the 1σ and 2σ confidence intervals for the WCs. Although no signs of new physics are identified, the results provide information about correlations among the 26 WCs and represent the most global detector-level EFT analysis in the top quark sector to date.

REFERENCES

- [1] CMS COLLABORATION, CMS-PAS-TOP-22-006, <https://cds.cern.ch/record/2851651/>.
- [2] CMS COLLABORATION, CMS-TOP-22-006, submitted to *JHEP*, arXiv:2307.15761.
- [3] BARDUCCI D. *et al.*, *Interpreting top-quark LHC measurements in the standard-model effective field theory*, arXiv:1802.07237.
- [4] CMS COLLABORATION, *JHEP*, **03** (2021) 095, arXiv:2012.04120.
- [5] MATTELAER O., *Eur. Phys. J. C*, **76** (2016) 674, arXiv:1607.00763.
- [6] CMS COLLABORATION, *JINST*, **3** (2008) S08004.