Using associated top quark production to probe for new physics within the framework of effective field theory

#### CMS-TOP-19-001 DOI: <u>10.1007/JHEP03(2021)095</u>

Brent R. Yates The Ohio State University On behalf of the CMS Collaboration

brent.yates@cern.ch





# Motivation for new physics

The Standard Model (SM) is wonderfully precise, but only accounts for 5% of the universe

Shortcomings include:

- **Dark matter/energy** Invisible particles? Non-zero vacuum energy?
- Hierarchy problem Why is the Higgs mass  $\mathcal{O}(10^2)$  GeV but Plank mass  $\mathcal{O}(10^{19})$  GeV?
- Baryon asymmetry Why do we live in a universe devoid of anti-matter?

After discovering the Higgs boson in 2012, the LHC provided no definitive evidence of anything unexpected

Assume  $\Lambda_{NP} > \Lambda_{LHC}$ How might it appear at the LHC?





### Introduction to effective field theory

New physics at scales beyond what the LHC can directly probe can be approximated by expanding terms of higher dimensional (d) operators O consisting of SM fields

Operators are suppressed by powers of the energy scale  $\Lambda$ , and the strength is controlled by the Wilson coefficients (WCs)  $c_i$ 

$$\mathcal{L}_{\rm eff} = \mathcal{L}_{\rm SM} + \sum_{d,i} \frac{c_i^{(d)}}{\Lambda^{d-4}} \mathcal{O}_i^{(d)}$$

Dimension five violates lepton number

Dimension six is the focus of this analysis

Higher dimensions are suppressed by additional powers of  $\Lambda$ 

H





#### Analysis overview

A novel technique to examine data collected in 2017 Performs a global fit across all processes (signal and background)

Probe EFT effects using multiple lepton final states

Procedure helps constrain systematic uncertainties

Correlations rely on data (no assumptions made)

Using channels with  $t\bar{t}l\nu$ ,  $t\bar{t}l\bar{l}$ ,  $t\bar{l}q$ ,  $t\bar{t}H$ , and tHq production (H  $\rightarrow$  W<sup>+</sup>W<sup>-</sup>, ZZ,  $\tau^+\tau^-$ , exclude H  $\rightarrow$  bb)







# EFT parameterization

Matrix elements  $\mathcal{M}$  split into SM and EFT terms

Parameterize cross section by WCs:

- SM terms (s<sub>0i</sub>)
- Interference terms between the SM and EFT  $(s_{1ij})$
- Pure EFT terms  $(s_{2ij})$
- Interference terms between EFT  $(s_{3ijk})$

Individual events  $(d\sigma)$  have weight  $w_i \rightarrow$  can be summed to produce the predicted event yields









#### 







HEFT 2021





# Dim6TopEFT Model

EFT simulations are generated by MADGRAPH\_AMC@NLO using the dim6TopEFT[1] model

- Warsaw basis of dimension six operators
- $\Lambda = 1 \text{ TeV}$
- CKM matrix is assumed to be a unit matrix
- u, d, s, c, e,  $\mu$  masses all set to zero
- The unitary gauge is used and Goldstone bosons are removed
- Baryon and lepton number violating operators are not included
- Only tree-level simulation is possible
- Lepton universality is assumed (all flavors set to same WCs)

The 16 operators which have the largest impact on the signal processes, and relatively small impact on  $t\bar{t}$  background, are considered





#### Model operators

Only the real components are considered since the imaginary coefficients lead to CP violation, and are well constrained by EDM experiments and  $B \rightarrow X_s \gamma$  decays



Operators involving two quarks and one or more bosons						
Operator	Definition	Wilson coefficient				
$O_{u\phi}^{(ij)}$	$\overline{\mathbf{q}}_{i}\mathbf{u}_{j}\tilde{\varphi}\left(\varphi^{\dagger}\varphi\right)$	$c_{\mathrm{t}\varphi} + ic_{\mathrm{t}\varphi}^{I}$				
$O_{arphi \mathrm{q}}^{1(ij)}$	$(\varphi^{\dagger} i \overrightarrow{D}_{\mu} \varphi) (\overline{\mathbf{q}}_{i} \gamma^{\mu} \mathbf{q}_{j})$	$c_{\varphi Q}^{-} + c_{\varphi Q}^{3}$				
$O_{\varphi q}^{3(ij)}$	$(\varphi^{\dagger} \overleftrightarrow{iD}_{\mu}^{I} \varphi) (\overline{\mathbf{q}}_{i} \gamma^{\mu} \tau^{I} \mathbf{q}_{j})$	$c_{\varphi Q}^3$				
$O_{\varphi u}^{(ij)}$	$(\varphi^{\dagger} i \overrightarrow{D}_{\mu} \varphi) (\overline{\mathbf{u}}_{i} \gamma^{\mu} \mathbf{u}_{j})$	C <sub>\$\varphi\$t\$</sub> t				
$O_{\varphi ud}^{(ij)}$	$(\tilde{\varphi}^{\dagger}iD_{\mu}\varphi)(\overline{\mathbf{u}}_{i}\gamma^{\mu}\mathbf{d}_{j})$	$c_{\varphi tb} + i c^I_{\varphi tb}$				
$O_{uW}^{(ij)}$	$(\overline{\mathbf{q}}_i \sigma^{\mu\nu} \tau^I \mathbf{u}_j)  \tilde{\varphi} \mathbf{W}^I_{\mu\nu}$	$c_{\rm tW} + i c_{\rm tW}^I$				
$O_{dW}^{(ij)}$	$(\overline{\mathbf{q}}_i \sigma^{\mu\nu} \tau^I \mathbf{d}_j) \varphi \mathbf{W}^I_{\mu\nu}$	$c_{\rm bW} + i c_{\rm bW}^I$				
$O_{uB}^{(ij)}$	$(\overline{\mathbf{q}}_i \sigma^{\mu\nu} \mathbf{u}_j)  \tilde{\varphi} \mathbf{B}_{\mu\nu}$	$(c_{\rm W}c_{\rm tW} - c_{\rm tZ})/s_{\rm W} + i(c_{\rm W}c_{\rm tW}^I - c_{\rm tZ}^I)/s_{\rm W}$				
$O_{uG}^{(ij)}$	$(\overline{\mathbf{q}}_i \sigma^{\mu\nu} T^A \mathbf{u}_j)  \tilde{\varphi} G^A_{\mu\nu}$	$c_{\mathrm{tG}} + i c_{\mathrm{tG}}^{I}$				
Operators involving two quarks and two leptons						
Operator	Definition	Wilson coefficient				
$O_{\ell q}^{1(ijkl)}$	$(\overline{\ell}_i \gamma^{\mu} \ell_j) (\overline{\mathbf{q}}_k \gamma^{\mu} \mathbf{q}_\ell)$	$c_{Q\ell}^{-(\ell)} + c_{Q\ell}^{3(\ell)}$				
$O_{\ell q}^{3(ijkl)}$	$(\overline{\ell}_i \gamma^\mu \tau^I \ell_j) (\overline{\mathbf{q}}_k \gamma^\mu \tau^I \mathbf{q}_\ell)$	$c_{Q\ell}^{3(\ell)}$				
$O_{\ell \mathrm{u}}^{(ijkl)}$	$(\overline{\ell}_i\gamma^\mu\ell_j)(\overline{\mathbf{u}}_k\gamma^\mu\mathbf{u}_\ell)$	$c_{t\ell}^{(\ell)}$				
$O_{\mathrm{e}\overline{\mathrm{q}}}^{(ijkl)}$	$(\bar{\mathbf{e}}_i \gamma^{\mu} \mathbf{e}_j) (\overline{\mathbf{q}}_k \gamma^{\mu} \mathbf{q}_\ell)$	$c_{Qe}^{(\ell)}$				
$O_{\rm eu}^{(ijkl)}$	$(\bar{\mathbf{e}}_i \gamma^{\mu} \mathbf{e}_j) (\overline{\mathbf{u}}_k \gamma^{\mu} \mathbf{u}_\ell)$	$C_{\rm te}^{(\ell)}$				
$O_{\ell equ}^{1(ijkl)}$	$(\overline{\ell}_i \mathbf{e}_j) \ \varepsilon \ (\overline{\mathbf{q}}_k \mathbf{u}_\ell)$	$c_{\mathrm{t}}^{S(\ell)} + i c_{\mathrm{t}}^{SI(\ell)}$				
$O_{\ell equ}^{3(ijkl)}$	$(\overline{\ell}_i \sigma^{\mu\nu} \mathbf{e}_j) \varepsilon (\overline{\mathbf{q}}_k \sigma_{\mu\nu} \mathbf{u}_\ell)$	$c_{t}^{T(\ell)} + ic_{t}^{TI(\ell)}$				
		9				



# Event selection

The analysis is split into lepton  $(\ell)$  categories as well as jet multiplicity (both light and b-tagged jets)

A BDT is applied to separate prompt from non-prompt leptons

Final-state observables are an **admixture** of processes (the method does **not** require we **separate** processes)

 Each analysis bin stores the sum of the quadratic coefficients → event yields are fully parametrized by the WCs







### Event categorization



loose = 85% efficiency, 10% light/quark jets medium = 70% efficiency, 1% light/quark jets





### Misidentified lepton background



Lepton production

Probability of a non-prompt lepton passing prompt cuts is measured in a multijet enriched region Data-driven





# Fitting procedure

Each bin is treated as a Poisson experiment with a probability of obtaining the observed data

Profiled likelihood simultaneously fits all bin and extract the  $2\sigma$  confidence intervals of the various WCs

Two fitting procedures are used:

Scan single WC, other 15 are unconstrained nuisance parameters

More physical of the two, no reason for new physics to only favor one WC

Scan a single WC, other 15 are fixed to their SM value of zero

 Extreme scenario where nature has a single WC; the ability to fit one is limited by the lack of knowledge of 15 others







HEFT 2021







#### predicted yields (WCs = 1/6 final value)









#### EFT enhances predicted yields (WCs = 2/6 final value)









#### EFT enhances

predicted yields (WCs = 3/6 final value)









#### predicted yields (WCs = 4/6 final value)









predicted yields (WCs = 5/6 final value)











HEFT 2021





# Single WC scan



HEFT 2021





# Two-dimensional WC scans

Pairs of WCs are also scanned to help investigate the correlations between WCs, as visualizing the full 16-dimensional hypersurface in not feasible

Other 14 WCs treated as unconstrained nuisance parameters

Other 14 WCs are fixed to zero (SM)







#### Important systematic uncertainties

#### Analysis specific

 Misidentified lepton rate estimate – Contamination from non-prompt leptons Overcome by examining the analysis side-bands
 Data-driven → statistically limited

#### Monte Carlo simulation modeling

- Matrix element parton shower matching Matching extra partons to final-state jets
- Missing parton uncertainty Extra partons cannot be added for tHq and tllq Compare LO EFT without extra partons to NLO SM simulations, assign uncertainty to cover any discrepancies These issues will not be present in SMEFT@NLO, and we are very interested in the development
- Scale uncertainties FSR and ISR





### Wilson coefficient CIs

The  $1\sigma$  and  $2\sigma$  CIs are given

When the other WCs are fixed to zero, the fit can produce degenerate minima in  $c_{tW}$ ,  $c_{t\varphi}$ ,  $c_{tG}$ , and  $c_{\varphi t}$ 

Degenerate minima are due to the quadratic nature of the parameterization

None of the WCs exclude the SM point of zero by a statistically significant amount



HEFT 2021



#### Conclusion

The production of or more t quarks in association with additional leptons were used to measure the confidence intervals of 16 dimension-six EFT operators using data collected in 2017  $_{\tilde{c}}$ 

The EFT yields are parameterized using a quadratic function of event weights

This novel technique allows us to extract EFT from difficult data

The  $2\sigma$  CIs were extracted for these operators Intervals are compatible with the SM and other analyses [1]

With the full Run II data set (almost triple the integrated luminosity) more sophisticated analyses may be performed, including using differential distributions



[1] <u>https://twiki.cern.ch/twiki/bin/view/CMSPublic/PhysicsResultsTOPSummaryFigures</u>

HEFT 2021





# Backup





# Selecting operators of interest

There are 59 total dimension-6 operators that conserve baryon and lepton number

We only consider **16** operators:

- Operators must appear in signal processes (top + boson) at tree level
- Ignore operators that strongly affect background processes
- Imaginary parts of non-Hermitian operators are set to zero

'Two heavy + boson':  $c_{t\phi}$ ,  $c_{\phi Q}^{-}$ ,  $c_{\phi Q}^{3}$ ,  $c_{\phi t}$ ,  $c_{\phi tb}$ ,  $c_{tW}$ ,  $c_{tZ}$ ,  $c_{bW}$ , and  $c_{tG}$ 

'Two heavy + two lepton':  $c_{0l}^{3(l)}$ ,  $c_{0l}^{-(l)}$ ,  $c_{Qe}$ ,  $c_{tl}$ ,  $c_{te}$ ,  $c_{t}^{S(l)}$ , and  $c_{t}^{T(l)}$ 

- These operators have three copies that couple to the different lepton flavors
- We assume equal coupling to all flavors to reduce the number of operators to these seven





 $\mathcal{A}$ 

0.0566

0.0562

0.0363

0.0119

0.0064

### Lepton identification

$I_{\ell} = \sum_{\text{charged}} p_{\text{T}} + \max\left(0, \sum_{\text{neutral}} p_{\text{T}} - \rho \mathcal{A}\left(\frac{R}{0.3}\right)\right)$					
Floating	$R = \begin{cases} 0.0\\ 10\\ 0.2 \end{cases}$	$\begin{array}{ll} 05 & \text{if } p_{\mathrm{T}} > \\ 0  \mathrm{GeV} / p_{\mathrm{T}} & \text{if } 50 < \\ 20 & \text{if } p_{\mathrm{T}} < \end{array}$	200 GeV <i>p</i> <sub>T</sub> < 200 GeV 50 GeV		
Electrons	4		Muons		
Pseudorapidity range	$\mathcal{A}$		Pseudorapidity range		
$ 0.0 <  \eta  < 1.0$	0.1500		$0.0 <  \eta  < 0.8$		
$  1.0 <  \eta  < 1.479$	0.1020		$0.8 <  \eta  < 1.3$		
$ 1.479 <  \eta  < 2.0$	0.1075		$1.3 <  \eta  < 2.0$		
$ 2.0 <  \eta  < 2.2$	0.0854		2.0 <  n  < 2.2		
$ 2.2 <  \eta  < 2.3$	0.1051		2.2 <  n  < 2.5		
$ 2.3 <  \eta  < 2.4$	0.1204				
24 <  n  < 25	0 1524				





Fakeable

 $> 10 \, GeV$ 

< 2.4

< 0.05 cm

< 0.1 cm

< 8

 $< 0.4 imes p_{
m T}$ 

> 0.3

> 0.6

< 0.07

< 0.90

Tight

> 10 GeV

< 2.4

< 0.05 cm

< 0.1 cm

< 8

 $< 0.4 \times p_{T}$ 

 $\checkmark$ 

\_

< 0.4941

> 0.90

#### Inputs to BDT

- Lepton  $p_{\mathrm{T}}$  and  $\eta$
- $I_{\ell}^{\text{charged}}$
- $I_{\ell}^{\text{neutral}}$
- $p_{\mathrm{T}}^{\ell}/p_{\mathrm{T}}^{\mathrm{jet}}$
- CSVv2 b-tagging algorithm
- $N_{\text{charged}}$  of charged particles within the jet
- $p_{\mathrm{T}}^{\mathrm{rel}} = p_{\ell} \sin(\theta)$
- $d_{xy}$  and  $d_z$  w.r.t. the PV
- $d/\sigma_d$  signed 3D impact parameter significance w.r.t the PV
- Lepton MVA from EGamma POG
- · Compatibility of track segments with the muon system

EGamma POG MVA cuts					
	$ \eta  < 0.8$	$0.8 <  \eta  < 1.44$	$ \eta  > 1.44$		
$p_{\rm T} < 10  GeV$	-0.13	-0.32	-0.08		
$p_{\rm T} > 10  GeV$	-0.86	-0.81	-0.72		

Observable

Loose PF muon

Prompt-*µ* MVA

Medium PF muon

Segment compatibility

DeepCSV(b) of nearby jet

 $p_{\rm T}$ 

 $|\eta|$ 

 $d_{xy}$ 

 $|d_z|$ 

 $d/\sigma_d$ 

 $p_T^{\mu}/p_T^{j}$ 

Tight

> 10 GeV

< 2.5

< 0.05 cm

 $< 0.1 \, \text{cm}$ 

< 8

 $< 0.4 \times p_{T}$ 

= 0

See Table 10

{0.011 / 0.011 / 0.030

< 0.10

 $> -0.04 \frac{1}{\text{GeV}}$ 

< 0.4941

> 0.90

Loose

 $>7 \, \text{GeV}$ 

< 2.5

 $< 0.05 \, \text{cm}$ 

< 0.1 cm

< 8

 $< 0.4 \times p_{\rm T}$ 

See Table 10

<1>

 $^{1} > \{-0.13 / -0.32 / -0.08\}$  if  $p_{\rm T} < 10$  GeV  $^{2} > 0.50$  if prompt-e MVA < 0.90

Fakeable

> 10 GeV

< 2.5

 $< 0.05 \, \text{cm}$ 

 $< 0.1 \, \text{cm}$ 

< 8

 $< 0.4 \times p_{\rm T}$ 

= 0

> 0.50

 $\checkmark$ 

< {0.011 / 0.011 / 0.030}

 $> -0.04 \frac{1}{\text{GeV}}$ 

> 0.6

< 0.07

< 0.90

< 0.10

Observable

Missing hits

EGamma POG MVA

Conversion rejection

DeepCSV(b) of nearby jet

 $p_{\rm T}$ 

 $|\eta|$ 

 $d_{xy}$ 

 $|d_z|$ 

 $d/\sigma_d$ 

 $\sigma_{inin}$ 

H/E

 $p_{\rm T}^{\rm e}/p_{\rm T}^{\rm J}$ 

1/E - 1/p

Prompt-e MVA

 $\mu$  cuts

Loose

 $> 5 \, GeV$ 

< 2.4

< 0.05 cm

< 0.1 cm

< 8

 $< 0.4 \times p_{\rm T}$ 





#### **Event selection**

	2ℓss	3ℓ	$\geq 4\ell$
Tight Leptons	2	3	$\geq 4$
Jets	$\geq 4$	$\geq 2$	≥2
DeepCSV(b) Medium Jets	$\geq 1$	1, ≥2	$\geq 1$
DeepCSV(b) Loose Jets	≥2	-	≥2
nJet Subcategories	4, 5, 6, ≥7	2, 3, 4, ≥5	2, 3, ≥4
Other Subcategories	Lepton Charge	Sign of Net Lepton Charge	-
		"SFOS Z"	-

	2lss ("p")	2ℓss ("m")	3ℓ (1b "p")	3ℓ (1b "m")	3ℓ (≥ 2b "p")	3ℓ (≥ 2b "m")	3ℓ (SFOS Z,1b)	$3\ell$ (SFOS Z, $\geq$ 2b)	$\geq 4\ell$
Diboson	1.6	1.2	5.9	4.7	0.4	0.3	52.1	4.1	0.6
Triboson	0.5	0.5	0.2	0.2	0.0	0.1	3.5	0.6	0.1
Charge Flips	8.5	8.5	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Fakes	25.6	26.8	11.3	13.0	3.3	2.5	16.9	3.8	0.0
Conversions	10.9	9.2	2.3	2.6	1.7	1.9	0.8	0.4	0.0
Sum Background	47.0	46.2	19.7	20.5	5.4	4.8	73.3	8.9	0.7
tτlν	68.7	37.1	14.4	8.0	10.8	5.9	2.9	2.3	0.0
tīlī	19.3	19.0	12.7	13.3	9.1	8.5	95.5	63.2	9.4
tīH	24.7	24.1	7.9	7.6	5.1	5.2	3.2	2.2	1.0
tllq	2.7	1.5	3.5	1.8	1.2	0.6	39.8	13.3	0.0
tHq	0.8	0.4	0.3	0.2	0.2	0.1	0.1	0.1	0.0
Sum Signal	116.2	82.2	38.9	30.8	26.3	20.3	141.6	81.1	10.4
Total Expected	$163\pm20$	$128\pm15$	$59 \pm 7$	$51\pm 6$	$32 \pm 4$	25±3	$215 \pm 25$	90 ± 13	$11\pm 2$
Data	192	171	85	64	32	28	239	95	12





### **Background Estimation**

Misidentified leptons

- Data-driven by dividing data into measurement and application regions (analysis side-bands)
- Measurement region contains QCD multijet background dominated by non-prompt leptons
- A fake rate is derived by comparing looser lepton cuts to tight lepton cuts used in the main analysis
- Fake rate is applied to the application region

Lepton charge mismeasurement

- Also data-driven
- Charge mismeasurement rate is extracted from the  $2\ell ss$  region using  $Z/\gamma^* \rightarrow ee$
- Only applied to the *ee* region of the analysis





#### Lepton charge mismeasurement



2ℓss region using  $Z/\gamma^*$  → ee used to estimate rate at which the CMS detector incorrectly measures lepton charge

#### Data-driven

Only applied to the ee region of the analysis





### Systematic uncertainties

The systematic uncertainties are:

standard and analysis specific

Luminosity – vary simulation by integrated luminosity estimate uncertainty

Jet energy scale (JES) – account for pileup, nonuniform detector response, and any residual differences between the data and simulation

b jet tagging scale factors – account for tagging inefficiencies and charm jet contamination

Cross section theoretical uncertainty – vary cross sections in simulation by uncertainties

PDF shape variations – reweighting the spectra according to the 100 replica sets given by the NNPDF31\_NLO\_as\_0118 PDF parameterization

Renormalization and factorization scale – vary scales by 1/2 - 2

Parton shower – vary ISR by 2 and FSR by  $\sqrt{2}$ 

Matching uncertainty – vary the matching scale between MADGRAPH\_AMC@NLO and PYTHIA

Muon and electron ID isolation – vary corrections by their uncertainties

Trigger efficiency – vary corrections by their uncertainties

Pileup – vary the pp inelastic cross section by 5%

Missing parton uncertainty – cover differences between LO EFT w/o extra partons and NLO SM

Uncertainty on the misidentified lepton rate estimate (data-driven) – account for non-prompt contamination

Uncertainty on the charge mismeasurement estimate (data-driven) – account for  $Z/\gamma^* \rightarrow e^{\pm}e^{\mp}$  becoming  $Z/\gamma^* \rightarrow e^{\pm}e^{\pm}$ 





#### Complete set of 2D contours

CMS Preliminary





41.5 fb<sup>-1</sup> (13 TeV)











#### Complete set of 2D contours









- tHq 15

### Changes in event yields over the $2\sigma$ CI

→ ttll

-ttlv

Examining the minimum and maximum yield changes within the  $2\sigma$  CI of various WCs

--- tllq

→ ttH

