This Dissertation

entitled

SEARCH FOR NEW PHYSICS IMPACTING ASSOCIATED TOP PRODUCTION IN MULTILEPTON FINAL STATES USING THE FRAMEWORK OF EFFECTIVE FIELD THEORY

typeset with NDdiss2$_\varepsilon$ v3.2017.2 (2017/05/09) on January 25, 2023 for

Kelci Mohrman

This LaTeX 2$_\varepsilon$ classfile conforms to the University of Notre Dame style guidelines as of Fall 2012. However it is still possible to generate a non-conformant document if the instructions in the class file documentation are not followed!

> Be sure to refer to the published Graduate School guidelines at `http://graduateschool.nd.edu` as well. Those guidelines override everything mentioned about formatting in the documentation for this NDdiss2$_\varepsilon$ class file.

This page can be disabled by specifying the "`noinfo`*" option to the class invocation.* (i.e.,`\documentclass[...,noinfo]{nddiss2e}`)

This page is *NOT* part of the dissertation/thesis. It should be disabled before making final, formal submission, but should be included in the version submitted for format check.

NDdiss2$_\varepsilon$ documentation can be found at these locations:

`http://graduateschool.nd.edu`
`https://ctan.org/pkg/nddiss`

SEARCH FOR NEW PHYSICS IMPACTING ASSOCIATED TOP PRODUCTION IN MULTILEPTON FINAL STATES USING THE FRAMEWORK OF EFFECTIVE FIELD THEORY

A Dissertation

Submitted to the Graduate School
of the University of Notre Dame
in Partial Fulfillment of the Requirements
for the Degree of

Doctor of Philosophy

by

Kelci Mohrman

Kevin Lannon, Director

Graduate Program in Physics
Notre Dame, Indiana
January 2023

SEARCH FOR NEW PHYSICS IMPACTING ASSOCIATED TOP PRODUCTION IN MULTILEPTON FINAL STATES USING THE FRAMEWORK OF EFFECTIVE FIELD THEORY

Abstract

by

Kelci Mohrman

This thesis presents a search for new physics impacting top quarks within the context of an effective field theory (EFT). Parameterizing potential new physics effects in terms of 26 dimension-six EFT operators, six associated top production processes are studied. The analysis targets the leptonic signatures of these processes, requiring the selected events to contain two leptons of the same charge or three or more leptons. In order to gain sensitivity to EFT effects by differentiating the admixture of processes and effects in each category, the events are subcategorized and binned in terms of kinematical distributions. The predicted distribution in each category is compared to the experimentally observed distribution, using 138 fb^{-1} of proton-proton collision data collected by the CMS experiment from 2016 to 2018. A likelihood fit of the 26 EFT parameters to the observed data is performed to extract the one and two standard deviation confidence intervals for the EFT parameters. While the results are consistent with the standard model prediction, the extracted limits can help to constrain theoretical models of new physics; interpreted in terms of the energy scale probed by the experiment, the results can also provide information about the energy frontier beyond which new physics discoveries may yet lie.

CONTENTS

Figures vi

Tables xv

Acknowledgments xvii

Chapter 1: Introduction 1

Chapter 2: Theory 6
2.1 The standard model 6
2.2 Effective field theory 9

Chapter 3: Accelerator and detector 13
3.1 The large hadron collider 13
3.2 The CMS detector 14
3.2.1 Coordinate system 15
3.2.2 The tracker 17
3.2.3 The electromagnetic calorimeter 19
3.2.4 The hadronic calorimeter 20
3.2.5 The muon detectors 22
3.2.6 The trigger systems 23

Chapter 4: Data and Monte Carlo samples 27
4.1 Data samples and triggers 27
4.2 Monte Carlo samples 28
4.2.1 Monte Carlo generation of signal samples 29
4.2.2 Parameterization of the predicted yields in terms of the WCs . 32

Chapter 5: Object reconstruction 37
5.1 Particle flow reconstruction 37
5.2 Jets and b-tagging 39
5.3 Lepton object selection 41
5.3.1 Conceptual overview of lepton object selection 41
5.3.2 Lepton object selection stages 43
5.3.3 Definitions of variables used in lepton object selection 47

Chapter 6: Event selection 49
6.1 Event selection categorization motivation 49
6.2 Event selection category requirements 51
6.2.1 The 2ℓss category 51
6.2.2 The 3ℓ category 53
6.2.3 The 4ℓ category 53
6.3 Optimization studies 54
6.4 Event selection summary 57

Chapter 7: Data-to-MC corrections 60
7.1 Pileup reweighting 61
7.2 Trigger efficiency 61
7.3 Lepton efficiency 62
7.4 b-tag efficiency and mistag rate 63
7.5 ECAL and muon prefiring correction 64
7.6 Jet energy scale and resolution 65
7.7 Muon p_T correction 66

Chapter 8: Backgrounds 67
8.1 Nonprompt background 69
8.2 Charge misidentification background 72

Chapter 9: Systematics 76
9.1 EFT dependence of the systematic uncertainties 76
9.2 Experimental systematic uncertainties 78
9.3 Theoretical systematic uncertainties 80

Chapter 10: Statistical methods 82
10.1 Likelihood fitting 82
10.2 Statistical framework 85
10.3 Navigating false minima 86

Chapter 11: Results 91
11.1 SM expected limits 92
11.2 Observed results 92
11.3 Visualization of results: Prefit and postfit summary histograms 92

Chapter 12: Discussion of results 107
12.1 Comparison of observed limits to predicted limits 107
12.2 Exploration of correlations among WCs 109
12.2.1 Methodology of identification of correlations 109
12.2.2 Correlated pairs of WCs 112
12.2.3 Discussion of factors affecting correlations 117

12.3 Interpretation of sensitivity to WCs . . . 119
12.3.1 WCs from the two-heavy-two-lepton category of operators . . 120
12.3.2 WCs from the four-heavy category of operators . . . 121
12.3.3 WCs from the two-heavy-two-light category of operators . . . 121
12.3.4 WCs from the two-heavy-with-bosons category of operators . . 122
12.4 Impacts of uncertainties . . . 124
12.4.1 Relative importance of statistical and systematic contributions 124
12.4.2 Impacts of the systematic uncertainties . . . 125
12.5 Comparison of results to other analyses . . . 127
12.5.1 Comparisons against other CMS analyses . . . 127
12.5.2 Comparisons against global theory combination . . . 131
12.6 Interpretation of results in terms of energy scale . . . 134

Chapter 13: Summary . . . 136

Appendix A: Reference information for data and MC samples . . . 139
A.1 Data samples . . . 139
A.2 MC samples . . . 144

Appendix B: Validation of EFT samples . . . 151
B.1 Validation of leading order matching procedure for EFT samples . . . 151
B.2 Validation of starting points for EFT samples . . . 153

Appendix C: Comparison of privately generated MC samples to centrally generated MC samples . . . 158
C.1 Summary of comparisons for the $t\bar{t}H$ sample . . . 160
C.2 Summary of comparisons for the $t\bar{t}l\nu$ sample . . . 160
C.3 Summary of comparisons for the $t\bar{t}t\bar{t}$ sample . . . 160
C.4 Summary of comparisons for the $t\bar{t}l\bar{l}$ sample . . . 160
C.5 Summary of comparisons for the $tl\bar{l}q$ sample . . . 168

Appendix D: One-dimensional quadratic parameterization plots . . . 174
D.1 $t\bar{t}H$ . . . 175
D.2 $t\bar{t}l\nu$ . . . 178
D.3 $t\bar{t}l\bar{l}$ . . . 181
D.4 $tl\bar{l}q$ . . . 184
D.5 tHq . . . 187
D.6 $t\bar{t}t\bar{t}$ . . . 190

Appendix E: Control region plots . . . 194

Appendix F: Discussion of the EFT dependence of the systematics . . . 215
F.1 The "correct" approach . . . 215
F.2 The TOP-19-001 approach . . . 217

F.3 The TOP-22-006 approach 218

Appendix G: Example usage of random starting point approach for navigating false minima 221

Appendix H: Sensitivity interpretation figure of merit plots 224

Appendix I: Impact plots for selected WCs 236

Appendix J: Limit comparisons to other analyses 240
J.1 TOP-19-001 comparison 240
J.2 TOP-21-003 comparison 241

Bibliography 244

FIGURES

1.1 Example SM diagrams for the associated top processes studied in this analysis. 3

2.1 Example Feynman diagrams illustrating WCs from each of the categories listed in Table 2.2. From left to right, the diagrams show vertices associated with the c_{Qt}^{1}, c_{Qq}^{11}, $c_{t\ell}^{(\ell)}$, and c_{tG} WCs. 11

3.1 Schematic of the CMS detector with the main subdetectors labeled. [12] 15

3.2 Coordinates used by the CMS experiment. An example momentum vector $\vec{p}$ is shown. [16] 16

3.3 Diagram of a quadrant of the CMS tracker showing the pixel and strip trackers (with the interaction point at the lower lefthand corner of the schematic). [19] 17

3.4 Diagram of the CMS ECAL, with components labeled. [12] 19

3.5 A quadrant of the CMS HCAL (with the interaction point at the lower lefthand corner). The barrel detector (HB), outer barrel detector (HO), endcap detector (HE), and forward detector (HF) are shown. [12] 21

3.6 Diagram of a quadrant of the CMS detector (with the interaction point at the lower lefthand corner) illustrating the muon chambers. The components labeled "MB" represent the DTs, the components labeled "ME" represent the CSCs, and the components labeled "RE" and "RB" represent the RPCs. [22] 22

3.7 Schematic of the L1 trigger system, as described in the text. [23] . . . 24

4.1 Example $c_{\varphi t}$ diagrams for $t\bar{t}H$ without and with an extra parton. . . . 30

5.1 Schematic representation of the reconstruction of particles within CMS. An example signature for a muon, electron, charged hadron, neutral hadron, and a photon are shown. The PF algorithm is described in detail in Ref. [17] (from which this figure is taken) and summarized in Section 5.1. 38

5.2 Schematic depiction of b jet, illustrating the impact parameter "IP" with respect to the PV. [35] 40

5.3 Venn diagram illustrating the three lepton object selection stages. . . 43

6.1 Summary of the event selection subdivisions. The details for the selection requirements are listed in Sections 6.2.1, 6.2.2, and 6.2.3. . . . 50

6.2 Summary of the sensitivity provided by fits to various differential distributions. The y axis represents the percent improvement with respect to the inclusive N_{jet} fit (based on the widths of the 2σ confidence intervals from fits to Asimov data). 56

8.1 Charge flip probabilities calculated with DY and $t\bar{t}$ samples using MC truth information for UL16APV (a), UL16 (b), UL17 (c), and UL18 (d) samples. The rates are binned according to the p_T and $|\eta|$. 73

10.1 Left: NLL for the one-dimensional profiled fit. The $c_{\varphi Q}^{-}$ parameter is scanned while c_{tG} is profiled. Right: NLL for the two-dimensional scan. Here both $c_{\varphi Q}^{-}$ and c_{tG} are scanned. The color scale shows the NLL at each of the two-dimensional scan points. The black overlaid points show the path of the one-dimensional profiled fit. 88

10.2 Left: NLL for the one-dimensional profiled fit after the implementation of the random starting point approach, where $c_{\varphi Q}^{-}$ is scanned and c_{tG} is profiled. Right: NLL for the two-dimensional scan over $c_{\varphi Q}^{-}$ and c_{tG}. The color scale shows the NLL at each of the two-dimensional scan points. The black overlaid points show the path of the one-dimensional profiled fit after the implementation of the random starting point approach. 90

11.1 SM-expected 2ΔNLL values for 1-dimensional scans for each WCs. The 2ΔNLL values represented in black correspond to the case where the other WCs are profiled, while the 2ΔNLL values represented in red correspond to the case where the other WCs are fixed at their SM values of zero. 97

11.2 Summary of limits from Asimov fits. 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 at their SM values of zero (in red). 99

11.3 Observed 2ΔNLL values for 1d scans for each WCs. The 2ΔNLL values represented in black correspond to the case where the other WCs are profiled, while the 2ΔNLL values represented in red correspond to the case where the other WCs are fixed at their SM values of zero. 103

11.4 Summary of limits from fits to data. 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 at their SM values of zero (in red). 105

11.5 The observed yields and the predicted prefit (top) and postfit (bottom) yields. As explained in the text (Section 11.3), in the prefit case the predicted yields have been reweighted to the SM, while in the postfit case the predicted yields have been reweighted to the best fit point from the unblind fits. Here we have integrated over the kinematic variables, so the bins in these histogram correspond to the jet multiplicity categories of the analysis categories. [62] 106

12.1 The 2σ observed limits (solid black line) compared against the 2σ Asimov limits (dashed blue lines). Figure (a) shows the results of the likelihood fits in which a single WC is fit with all other WCs fixed to their SM values of zero. Figure (b) shows the results of the likelihood fits in which the other WCs are profiled. In these plots, $\Lambda = 1$ TeV. . 108

12.2 Plots showing profiled WC values (on the y axis) vs the value of the scanned parameter (on the x axis). In (a), the scanned parameter is $c_t^{S(\ell)}$ and the profiled WC shown is c_{tG}. In (b), the scanned WC is $c_{\varphi Q}^-$ and the profiled WC shown is c_{tt}^1. In (c), the scanned parameter is again $c_{\varphi Q}^-$, but this time the profiled parameter shown is c_{tG}. Figure (a) shows essentially no signs of interplay between the WCs, (b) shows very minimal interplay, and figure (c) shows potentially moderate interplay. 110

12.3 Plots showing profiled WC values (on the y axis) vs the value of the scanned parameter (on the x axis). In (a), the scanned parameter is c_{Qt}^8 and the profiled WC shown is c_{Qt}^1. In (b), the scanned WC is c_{tW} and the profiled WC shown is c_{tZ}. Figure (a) shows that there is some non-trivial correlation between the two WCs, and (b) also shows a correlation between the WCs (in this case the correlation is linear). 111

12.4 All 676 profiled WC vs scanned WC plots. The plots without a background color show no signs of correlation between the WCs (i.e. the plot is essentially flat), grey indicates minimal correlation, blue indicates potentially moderate correlation, orange indicates significant correlation between the WCs, and red indicates a linear correlation. The order of the WCs in the rows and columns is as follows: two-light-two-heavy WCs (c_{Qq}^{31}, c_{Qq}^{38}, c_{Qq}^{11}, c_{tq}^1, c_{Qq}^{18}, c_{tq}^8), four-heavy WCs (c_{tt}^1, c_{QQ}^1, c_{Qt}^8, c_{Qt}^1), two-heavy-plus-boson WCs (c_{tW}, c_{tZ}, $c_{t\varphi}$, $c_{\varphi Q}^-$, c_{tG}, c_{bW}, $c_{\varphi Q}^3$, $c_{\varphi tb}$, $c_{\varphi t}$), and two-heavy-two-lepton WCs ($c_{Q\ell}^{3(\ell)}$, $c_{Q\ell}^{-(\ell)}$, $c_{Qe}^{(\ell)}$, $c_{te}^{(\ell)}$, $c_t^{S(\ell)}$, $c_t^{T(\ell)}$). 113

12.5 Two-dimensional scans over the pairs of WCs from Table 12.1 involving four-heavy WCs. The other 24 WCs are profiled in the likelihood fit. The color indicates the NLL, and the overlaid black markers show the paths of the one-dimensional scan (within the 2σ range of the scan) for each of the WCs shown in the plot. The path of the one-dimensional scan along the x-axis WC is shown with square markers, and the path of the one-dimensional scan along the y-axis WC is shown with triangular markers. Figure (a) shows c^1_{QQ}-c^1_{Qt}, (b) shows c^1_{QQ}-c^8_{Qt}, (c) shows c^1_{QQ}-c^1_{tt}, (d) shows c^1_{Qt}-c^8_{Qt}, (e) shows c^1_{Qt}-c^1_{tt}, (f) shows c^8_{Qt}-c^1_{tt}, (g) shows $c_{t\varphi}$-c^8_{Qt}, and (h) shows $c_{t\varphi}$-c^1_{QQ}. 115

12.6 Two-dimensional scans over the pairs of WCs from Table 12.1 involving only two-heavy-with-boson WCs. Figure (a) shows $c^-_{\varphi Q}$-$c_{\varphi t}$, (b) shows $c^-_{\varphi Q}$-$c^3_{\varphi Q}$, (c) shows c_{tZ}-c_{tW}, and (d) shows c_{tG}-$c_{t\varphi}$. All other relevant details of the plots are described in the caption to Figure 12.5. 116

12.7 Two-dimensional scans over the pairs of WCs that do not show signs of strong correlations. Figure (a) shows $c_{t\varphi}$-$c_{\varphi t}$, (b) shows c_{tG}-$c^-_{\varphi Q}$, (c) shows $c^{-(\ell)}_{Q\ell}$-$c^{(\ell)}_{Qe}$, and (d) shows $c_{\varphi tb}$-$c^{3(\ell)}_{Q\ell}$. All other relevant details of the plots are described in the caption to Figure 12.5. 116

12.8 The 2σ individual limits of this analysis (indicated with the solid black line labeled TOP-22-006) compared against the individual limits obtained by other CMS analyses. The other CMS results for each WC are as follows: $c_{\varphi t}$ and $c^-_{\varphi Q}$ [77], c_{tW} and $c^3_{\varphi Q}$ [78], c_{tZ} [79], c_{tG} [80], c_{bW} $c_{\varphi tb}$ and $c_{t\varphi}$ [76], and all four-heavy WCs [81]. The details of each referenced analysis are discussed in the text. In this plot, $\Lambda = 1$ TeV. 128

12.9 The 2σ profiled limits of this analysis (indicated with the solid black line labeled TOP-22-006) compared against the limits obtained in the 2021 SMEFiT global theory combination [82] (indicated with the dashed blue lines). Figure (a) shows the results of the likelihood fits in which a single WC is fit with all other WCs fixed to their SM values of zero. Figure (b) shows the results of the likelihood fits in which the other WCs are profiled. In these plots, $\Lambda = 1$ TeV. 133

12.10 The 2σ profiled limits obtained by this analysis interpreted in terms of the energy scale Λ for fits to Asimov data (in grey) and the real data (in blue) for three different assumptions for the value of the WC as indicated in the legend and described in the text. The Asimov and observed limits are taken from the 2σ limits presented in Tables 11.1 and 11.2; for asymmetric $+2\sigma$ and -2σ limits, we have taken the absolute value of the looser limit. 135

B.1 DJR histograms for LO matched samples $t\bar{t}H$ (a), $t\bar{t}l\nu$ (b), and $t\bar{t}l\bar{l}$ (c). The distributions have been reweighted to a non-SM point. . . . 152

B.2 Distribution of weights for $t\bar{t}H$ at the SM (a) and a non-SM point (b), $t\bar{t}l\nu$ at the SM (c) and a non-SM point (d), and $t\bar{t}l\bar{l}$ at the SM (e) and a non-SM point (f). 156

B.3 Distribution of weights for $tl\bar{l}q$ at the SM (a) and a non-SM point (b) and tHq at the SM (c) and a non-SM point (d). 157

C.1 RECO level comparison for UL16 $t\bar{t}H$. This plot shows the privately produced LO samples (reweighted to the SM) and the centrally produced NLO samples (datasets used for the central samples are listed in Table C.1). For this comparison, we have summed over all selection categories in the SR. The shaded band represents the systematic uncertainties for the private sample. 161

C.2 RECO level comparison for UL16APV $t\bar{t}H$. All other relevant details are the same as described in Figure C.1. 161

C.3 RECO level comparison for UL17 $t\bar{t}H$. All other relevant details are the same as described in Figure C.1. 162

C.4 RECO level comparison for UL18 $t\bar{t}H$. All other relevant details are the same as described in Figure C.1. 162

C.5 RECO level comparison for UL16 $t\bar{t}l\nu$. All other relevant details are the same as described in Figure C.1. 163

C.6 RECO level comparison for UL16APV $t\bar{t}l\nu$. All other relevant details are the same as described in Figure C.1. 163

C.7 RECO level comparison for UL17 $t\bar{t}l\nu$. All other relevant details are the same as described in Figure C.1. 164

C.8 RECO level comparison for UL18 $t\bar{t}l\nu$. All other relevant details are the same as described in Figure C.1. 164

C.9 RECO level comparison for UL16 $t\bar{t}t\bar{t}$. All other relevant details are the same as described in Figure C.1. 165

C.10 RECO level comparison for UL16APV $t\bar{t}t\bar{t}$. All other relevant details are the same as described in Figure C.1. 165

C.11 RECO level comparison for UL17 $t\bar{t}t\bar{t}$. All other relevant details are the same as described in Figure C.1. 166

C.12 RECO level comparison for UL18 $t\bar{t}t\bar{t}$. All other relevant details are the same as described in Figure C.1. 166

C.13 GEN level comparison for 2017 and UL17 $t\bar{t}l\bar{l}$ for the N_{jets} (a) and H_T (b). Some basic jet cleaning has been applied. As can be seen in the plots, the central UL sample has changed in comparison to the central pre-UL samples (and this change happens to make the tension with the private $t\bar{t}l\bar{l}$ sample somewhat worse). As discussed in the text, the change in the central sample seems to be due to a change in the default shower starting scale, and does not seem to represent an improvement in the modeling of the $t\bar{t}l\bar{l}$ process. 167

C.14 RECO level comparison for UL16 $t\bar{t}l\bar{l}$. All other relevant details are the same as described in figure C.1. 168

C.15 RECO level comparison for UL16APV $t\bar{t}l\bar{l}$. All other relevant details are the same as described in figure C.1. 169

C.16 RECO level comparison for UL17 $t\bar{t}l\bar{l}$. All other relevant details are the same as described in figure C.1. 169

C.17 RECO level comparison for UL18 $t\bar{t}l\bar{l}$. All other relevant details are the same as described in figure C.1. 170

C.18 RECO level comparison for UL16 $tl\bar{l}q$. All other relevant details are the same as described in figure C.1. 171

C.19 RECO level comparison for UL16APV $tl\bar{l}q$. All other relevant details are the same as described in figure C.1. 172

C.20 RECO level comparison for UL17 $tl\bar{l}q$. All other relevant details are the same as described in figure C.1. 172

C.21 RECO level comparison for UL18 $tl\bar{l}q$. All other relevant details are the same as described in figure C.1. 173

C.22 RECO level comparison for UL17 $tl\bar{l}q$. All other relevant details are the same as described in figure C.18. As discussed in the text, there is tension between the central and private $tl\bar{l}q$ samples for this distribution (invariant mass of the leading two leptons). However, this distribution is not directly used in the analysis. 173

D.1 The one-dimensional quadratic dependence of $t\bar{t}H$ on the two-heavy-two-light WCs. 175

D.2 The one-dimensional quadratic dependence of $t\bar{t}H$ on the two-heavy-with-boson WCs. 176

D.3 The one-dimensional quadratic dependence of $t\bar{t}H$ on the two-heavy-two-lepton WCs. 177

D.4 The one-dimensional quadratic dependence of $t\bar{t}l\nu$ on the two-heavy-two-light WCs. 178

D.5 The one-dimensional quadratic dependence of $t\bar{t}l\nu$ on the two-heavy-with-boson WCs. 179

D.6 The one-dimensional quadratic dependence of $t\bar{t}l\nu$ on the two-heavy-two-lepton WCs. 180
D.7 The one-dimensional quadratic dependence of $t\bar{t}l\bar{l}$ on the two-heavy-two-light WCs. 181
D.8 The one-dimensional quadratic dependence of $t\bar{t}l\bar{l}$ on the two-heavy-with-boson WCs. 182
D.9 The one-dimensional quadratic dependence of $t\bar{t}l\bar{l}$ on the two-heavy-two-lepton WCs. 183
D.10 The one-dimensional quadratic dependence of $tl\bar{l}q$ on the two-heavy-two-light WCs. 184
D.11 The one-dimensional quadratic dependence of $tl\bar{l}q$ on the two-heavy-with-boson WCs. 185
D.12 The one-dimensional quadratic dependence of $tl\bar{l}q$ on the two-heavy-two-lepton WCs. 186
D.13 The one-dimensional quadratic dependence of tHq on the two-heavy-two-light WCs. 187
D.14 The one-dimensional quadratic dependence of tHq on the two-heavy-with-boson WCs. 188
D.15 The one-dimensional quadratic dependence of tHq on the two-heavy-two-lepton WCs. 189
D.16 The one-dimensional quadratic dependence of $t\bar{t}t\bar{t}$ on the two-heavy-two-light WCs. 190
D.17 The one-dimensional quadratic dependence of $t\bar{t}t\bar{t}$ on the two-heavy-with-boson WCs. 191
D.18 The one-dimensional quadratic dependence of $t\bar{t}t\bar{t}$ on the two-heavy-two-lepton WCs. 192
D.19 The one-dimensional quadratic dependence of $t\bar{t}t\bar{t}$ on the four-heavy WCs. 193

E.1 Control region plots for the 2lss CR (for UL16 samples). The shaded gray band indicates the systematic uncertainty. 195
E.2 Control region plots for the 2lss CR (for UL16APV samples). The shaded gray band indicates the systematic uncertainty. 196
E.3 Control region plots for the 2lss CR (for UL17 samples). The shaded gray band indicates the systematic uncertainty. 197
E.4 Control region plots for the 2lss CR (for UL18 samples). The shaded gray band indicates the systematic uncertainty. 198
E.5 Control region plots for the 3l CR (for UL16 samples). The shaded gray band indicates the systematic uncertainty. 199

E.6 Control region plots for the 3l CR (for UL16APV samples). The shaded gray band indicates the systematic uncertainty. 200
E.7 Control region plots for the 3l CR (for UL17 samples). The shaded gray band indicates the systematic uncertainty. 201
E.8 Control region plots for the 3l CR (for UL18 samples). The shaded gray band indicates the systematic uncertainty. 202
E.9 Control region plots for the 2los $t\bar{t}$ CR (for UL16 samples). The shaded gray band indicates the systematic uncertainty. 203
E.10 Control region plots for the 2los $t\bar{t}$ CR (for UL16APV samples). The shaded gray band indicates the systematic uncertainty. 204
E.11 Control region plots for the 2los $t\bar{t}$ CR (for UL17 samples). The shaded gray band indicates the systematic uncertainty. 205
E.12 Control region plots for the 2los $t\bar{t}$ CR (for UL18 samples). The shaded gray band indicates the systematic uncertainty. 206
E.13 Control region plots for the 2los Z CR (for UL16 smaples). The shaded gray band indicates the systematic uncertainty. 207
E.14 Control region plots for the 2los Z CR (for UL16APV smaples). The shaded gray band indicates the systematic uncertainty. 208
E.15 Control region plots for the 2los Z CR (for UL17 smaples). The shaded gray band indicates the systematic uncertainty. 209
E.16 Control region plots for the 2los Z CR (for UL18 smaples). The shaded gray band indicates the systematic uncertainty. 210
E.17 Charge flip control regions for the UL16 samples after applying the scaling factors listed in 8.1. The shaded gray band indicates the systematic uncertainty. 211
E.18 Charge flip control regions for the UL16APV samples after applying the scaling factors listed in 8.1. The shaded gray band indicates the systematic uncertainty. 212
E.19 Charge flip control regions for the UL17 samples after applying the scaling factors listed in 8.1. The shaded gray band indicates the systematic uncertainty. 213
E.20 Charge flip control regions for the UL18 samples after applying the scaling factors listed in 8.1. The shaded gray band indicates the systematic uncertainty. 214

F.1 Visualization of the "correct" quadratic dependence of the systematics on the WCs. 216
F.2 Visualization of one event in the TOP-19-001 approach. 217
F.3 Visualization of two events in the TOP-19-001 approach. 218

F.4 Visualization of the evaluation of the systematics at the SM in the TOP-19-001 approach. 219

G.1 Profile fits from the [4] analysis with (red) and without (blue) random starting points for the profiled POIs. 223

H.1 An example FOM plot for the -2σ limit for c^1_{QQ}. 226

H.2 An example FOM plot for the -2σ limit for c^1_{QQ}, scaled to the FOM at the best fit point. 226

H.3 The full set of 52 FOM plots (relative to the best fit point) for all 26 WCs at the $+2\sigma$ and -2σ limits. 235

I.1 Impact plot for c^{18}_{Qq} with Asimov data. The rest of the WCs are frozen to their SM values of 0. [46] 237

I.2 Impact plot for c^8_{tq} with Asimov data. The rest of the WCs are frozen to their SM values of 0. [46] 237

I.3 Impact plot for $c_{\varphi t}$ with Asimov data. The rest of the WCs are frozen to their SM values of 0. [46] 238

I.4 Impact plot for $c^-_{\varphi Q}$ with Asimov data. The rest of the WCs are frozen to their SM values of 0. [46] 238

I.5 Impact plot for c_{tG} with Asimov data. The rest of the WCs are frozen to their SM values of 0. [46] 239

I.6 Impact plot for $c_{t\varphi}$ with Asimov data. The rest of the WCs are frozen to their SM values of 0. [46] 239

J.1 The 2σ profiled limits of this analysis (indicated with the solid black lines labeled TOP-22-001) compared against the limits obtained by the direct predecessor of this analysis (TOP-19-001 [4], indicated with the dashed blue lines). Figure (a) shows the results of the likelihood fits in which a single WC is fit with all other WCs fixed to their SM values of zero. Figure (b) shows the results of the likelihood fits in which the other WCs are profiled. In these plots, $\Lambda = 1$ TeV. 242

J.2 The 2σ profiled limits of this analysis (indicated with the solid black lines labeled TOP-22-001) compared against the limits obtained by Ref. [76] (TOP-21-003), indicated with the dashed blue lines). Figure (a) shows the results of the likelihood fits in which a single WC is fit with all other WCs fixed to their SM values of zero. Figure (b) shows the results of the likelihood fits in which the other WCs are profiled. In these plots, $\Lambda = 1$ TeV. 243

TABLES

2.1 Matter fields in the SM. 8
2.2 List of Wilson Coefficients included in the analysis. 11

5.1 Muon object selection requirements summary 45
5.2 Electron object selection requirements summary 46

6.1 Summary of event selection categories. 58
6.2 Expected SM yields and observations in the analysis categories (summed over jet categories). 59

8.1 Charge flip scaling factors. 75

11.1 SM-expected 2σ uncertainty intervals for the 1d scans. 98
11.2 Observed 2σ uncertainty intervals for the 1d scans. 104

12.1 Pairs of correlated WCs based on Figure 12.4. 114
12.2 Summary of categories that provide leading contributions to the sensitivity for subsets of the WCs. 120

A.1 JSON files with certified luminosity blocks. 139
A.2 Triggers used to record the 2016 data. 141
A.3 Triggers used to record the 2017 data. 142
A.4 Triggers used to record the 2018 data. 143
A.5 Theoretical cross sections used for normalizing the signal simulation samples. 145
A.6 Privately produced UL16 signal samples. 145
A.7 Privately produced UL16APV signal samples. 145
A.8 Privately produced UL17 signal samples. 146
A.9 Privately produced UL18 signal samples. 146
A.10 Central tZq samples used for calculating the additional systematic uncertainty that is applied to the single top samples. 147
A.11 List of UL16APV background samples. 147
A.12 List of UL16 background samples. 148

A.13 List of UL17 background samples. 149
A.14 List of UL18 background samples. 150
C.1 Central samples used for comparison against our privately produced samples. 159

ACKNOWLEDGMENTS

CHAPTER 1

INTRODUCTION

The goal of this analysis is to search for new physics impacting associated top production in multi-lepton final states, using the framework of effective field theory (EFT) to parametrize the potential new physics effects.

While there are many compelling indications that the standard model (SM) of particle physics does not provide a complete description of nature (e.g. the strong evidence for dark matter [1, 2] and dark energy [3]), there is no a priori reason to assume new particles must exist in the energy range that is directly accessible at the LHC. If new physics particles are too heavy to be produced on-shell at the LHC, it may not be possible to identify their signatures with a direct search. However, an approach that indirectly probes higher energy scales may be able to discover these particles via their off-shell effects. The center of mass energy for collisions at the LHC will not significantly increase throughout its remaining years of operation, so indirect approaches may provide an exciting opportunity to extend the discovery reach of the LHC. EFT is an example of such an indirect probe; as a flexible method of systematically describing the off-shell effects of heavy new particles, EFT represents an important part of the search for new physics at the energy frontier.

In general, an effective theory is an approximation, valid under a certain energy range, for a more fundamental underlying theory. In SM effective field theory (SM EFT), the SM is treated as the lowest order term in an expansion of higher dimensional operators; the operators are constructed from products of SM fields that obey the symmetries of the SM. The EFT operators describe new physics interactions at

a mass scale Λ. The strengths of the new physics interactions are described by dimensionless parameters known as Wilson Coefficients (WCs). The EFT Lagrangian can thus be expressed as follows:

$$\mathcal{L}_{\text{EFT}} = \mathcal{L}_{\text{SM}} + \sum_i \frac{c_i^5}{\Lambda}\mathcal{O}_i^5 + \sum_i \frac{c_i^6}{\Lambda^2}\mathcal{O}_i^6 + ..., \tag{1.1}$$

where $\mathcal{L}_{\text{SM}}$ is the SM Lagrangian, $\mathcal{O}_i^d$ are the EFT operators of dimension d, and c_i^d are the WCs for the operators of dimension d. Since each order in the expansion is scaled by an additional power of Λ, the terms in the lowest orders are expected to contribute the most significantly. This analysis therefore focuses on dimension-six operators, as these are the lowest order terms that contribute. The EFT framework will be discussed in more detail in Section 2.2.

While many analyses target a specific signature predicted by a particular new physics model, the EFT approach is more general. Assuming that the SM Lagrangian is the correct and complete description of all physics that is light enough to be probed directly with current experimental capabilities, the EFT Lagrangian provides a systematic description of the off-shell effects of heavy new physics scenarios, allowing for a consistent method of describing these effects across multiple sectors. EFT is thus a complementary approach to dedicated searches; if the off-shell effects of new physics manifest in a variety of signatures across many final states, a global EFT approach may be capable of identifying a statistically significant observation of the combination of effects, even if the effects are not significant when studied individually.

Although the ultimate goal of the EFT paradigm would be a global combination across all sectors of study at the LHC, the first step towards this goal is to begin performing EFT analyses within individual sectors. The analysis described in this thesis focuses on the top sector, targeting processes in which top quarks are produced in association with additional charged leptons. In the SM, these signatures are pri-

marily produced by $t(\bar{t})X$ processes, where the X is a H, W, or Z boson. We refer to these processes as associated top production. Involving multiple heavy particles, the processes are relatively rare, and we are just now reaching the point where we have accumulated enough statistics to study these processes in detail; for these reasons, associated top processes may be an interesting venue in which to stage a search for new physics. In accordance with the global mindset of the EFT approach, this analysis aims to study all dimension-six EFT operators (involving top quarks) that can significantly impact associated top production processes.

The full set of associated top processes studied in this analysis is $t\bar{t}H$, $t\bar{t}l\nu$, $t\bar{t}l\bar{l}$, $tl\bar{l}q$, tHq, and $t\bar{t}t\bar{t}$. Figure 1.1 shows example diagrams for these processes.

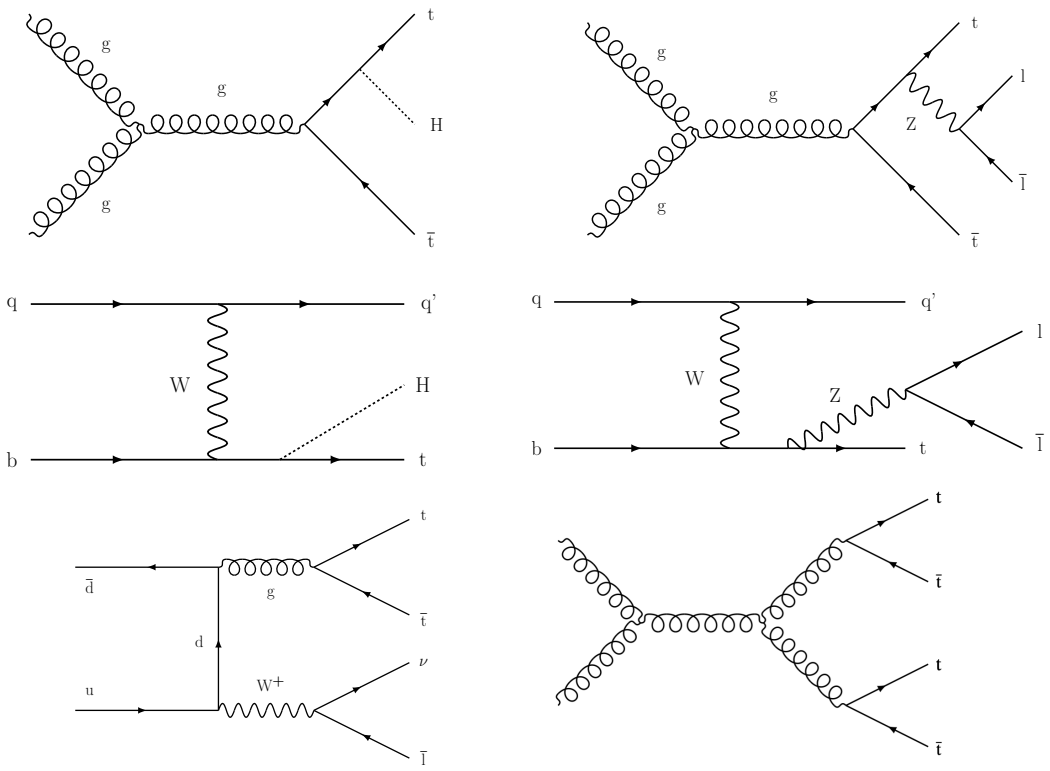


Figure 1.1. Example SM diagrams for the associated top processes studied in this analysis.

The associated top processes give rise to a variety of final state signatures; in this analysis, we choose to focus on signatures involving multiple charged leptons. Referred to as multilepton final states, these signatures contain 2 leptons of the same charge or contain 3 or more leptons. Multilepton final states have relatively few backgrounds, clean detector signatures, and efficient triggers. Despite these benefits, a multilepton EFT analysis also gives rise to several challenges. These challenges primarily stem from the fact that many different processes and effects are capable of contributing to the same final state multilepton signatures. For example, if we consider a final state with two leptons of the same charge, we would expect contributions from both SM $t\bar{t}W$ and SM $t\bar{t}H$ production (as well as a contribution from SM $t\bar{t}Z$ when one of the leptons is lost). Many different dimension-six EFT operators can impact these processes, interfering with each other and with the SM, making this final state a complicated admixture of processes and effects. Other multilepton final states will contain similarly complex admixtures of processes and EFT effects.

Because these effects cannot be isolated from each other, it is important to analyze the effects of all relevant operators across all channels simultaneously. For this reason, it would be difficult to construct this EFT analysis as a reinterpretation of inclusive or differential cross section measurements. Instead, we make use of an approach that directly targets the EFT effects at detector level. First developed in Ref. [4], the key idea of this approach is the parameterization of the predicted yields in terms of the WCs. The procedure through which we obtain this parametrization will be detailed in Section 4.2.2.

Making use of more than three times as much data as was available for [4], the analysis described in this thesis builds on the techniques and tools developed in [4], improving on [4] in several key ways. Since [4] was performed with limited statistics, only inclusive categories defined by the multiplicity of final state objects were studied. With the increased statistics, this analysis leverages differential kinematic

distributions within each inclusive bin, allowing additional sensitivity to be gained. An additional signal process ($t\bar{t}t\bar{t}$) and 10 more dimension-six EFT operators are also included, bringing the total number of WCs to 26. These improvements allow more comprehensive limits to be placed on the WCs, resulting in a better understanding of the possibility of heavy new physics effects in the top sector.

The chapters of this thesis are organized in the following order. In Chapter 2, the theoretical concepts of the SM and of the EFT framework are discussed. The CMS detector is described Chapter 3. Chapter 4 describes the data and simulated samples used in the analysis (including a discussion of the EFT parametrization of the signal samples). Chapter 5 explains how the particle reconstruction and object selection are performed. The event selection is detailed in Chapter 6. Chapter 7 describes the data to Monte Carlo corrections for simulated events. The backgrounds for this analysis are discussed in Chapter 8. Chapter 9 enumerates the systematic uncertainties of the analysis. The statistical tools used to extract the confidence intervals for the WCs are explained in Chapter 10. Chapter 11 presents the results of the analysis, which are discussed further in Chapter 12. A summary is provided in Chapter 13.

CHAPTER 2

THEORY

This chapter will discuss the theoretical background for this analysis. The SM is summarized in Section 2.1, and the relevant aspects of the SM EFT framework are discussed in Section 2.2.

2.1 The standard model

The standard model (SM) of particle physics is the mathematical framework that describes fundamental particles and their interactions. The SM is a quantum field theory with $\mathrm{SU}(3)_c \times \mathrm{SU}(2)_Y \times \mathrm{U}(1)_w$ gauge symmetry. The $\mathrm{SU}(3)_c$ component corresponds to QCD, and is non-abelian. The $\mathrm{SU}(2)_w$ component is referred to as weak isospin, and is also non-abelian. The $\mathrm{U}(1)_Y$ group is referred to as hypercharge, and it is abelian. The SM Lagrangian contains kinetic terms for each of these three gauge fields:

$$\mathcal{L}_{SM} \supset -\frac{1}{4}(G^A_{\mu\nu})^2 - \frac{1}{4}(W^a_{\mu\nu})^2 - \frac{1}{4}(B_{\mu\nu})^2, \tag{2.1}$$

where $G^A_{\mu\nu}$ is the $\mathrm{SU}(3)_c$ field strength tensor (with $A = 1...8$), $W^a_{\mu\nu}$ is the $\mathrm{SU}(2)_w$ field strength tensor (with $a = 1...3$), and $B_{\mu\nu}$ is the $\mathrm{U}(1)_Y$ field strength tensor.

In the SM, the $\mathrm{SU}(3)_c$ symmetry is exact, while the $\mathrm{SU}(2)_w \times \mathrm{U}(1)_Y$ is spontaneously broken by the Higgs mechanism. In order to preserve $\mathrm{SU}(3)_c$, the Higgs field ϕ must transform as a singlet under $\mathrm{SU}(3)_c$; in order to break $\mathrm{SU}(2)_w$ and $\mathrm{U}(1)_Y$, the Higgs field must be charged under these symmetries. The Higgs field is a doublet under $\mathrm{SU}(2)_w$, and has a hypercharge of $1/2$. This can be expressed as $\phi = (0, 2, 1/2)$,

where the first number corresponds to the $\mathrm{SU}(3)_c$ representation, the second number corresponds to the $\mathrm{SU}(2)_w$ representation, and the third number corresponds to the $\mathrm{U}(1)_Y$ representation. With the inclusion of the Higgs field, the SM Lagrangian contains:

$$\mathcal{L}_{SM} \supset |D_\mu \phi|^2 + V(\phi), \tag{2.2}$$

where $D_\mu \phi$ is the covariant derivative and V is the potential, given by

$$V(\phi) = \lambda \left(|\phi|^2 - \frac{v^2}{2} \right)^2, \tag{2.3}$$

which is minimized when $|\phi|^2 = v^2/2$. The Higgs field ϕ is a complex doublet:

$$\phi = \frac{1}{\sqrt{2}} \begin{pmatrix} h_1 + i h_2 \\ h_0 + i h_3 \end{pmatrix}, \tag{2.4}$$

so we have $\phi^\dagger \phi = 1/2 \ (h_0^2 + h_1^2 + h_2^2 + h_3^2)$. We know that $V(\phi)$ is minimized when $\phi^\dagger \phi = v^2/2$, but there are infinitely many ways to satisfy this. We chose $\langle h_0 \rangle = v$, with $\langle h_1 \rangle = \langle h_2 \rangle = \langle h_3 \rangle = 0$, breaking the symmetry. With this choice, the vacuum expectation value for ϕ is:

$$\langle \phi \rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v \end{pmatrix}. \tag{2.5}$$

Plugging equation 2.5 into the Higgs kinetic term $|D_\mu \phi|^2$ gives rise to the mass terms for the massive gauge bosons. Expanding Eq. 2.5 around the minimum, we have

$$\langle \phi \rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + h \end{pmatrix}, \tag{2.6}$$

where h is the physical Higgs boson particle.

Next, let us consider the matter fields in the SM. The matter fields and their $\mathrm{SU(3)}_c \times \mathrm{SU(2)} \times \mathrm{U(1)}$ representations are listed in Table 2.1.

TABLE 2.1

MATTER FIELDS IN THE SM.

	$\mathrm{SU(3)}_c$	$\mathrm{SU(2)}_w$	$\mathrm{U(1)}_Y$
Q	3	2	1/6
u_R	3	0	2/3
d_R	3	0	-1/3
L	0	2	-1/2
e_R	0	0	-1

In Table 2.1, the left-handed quark doublet is written as Q; the components of this $\mathrm{SU(2)}_w$ doublet can be written as u_L and d_L. The right-handed quark singlets are written as u_R and d_R. The u quarks are referred to as up-type quarks, while the d quarks are referred to as down-type quarks. The left-handed lepton doublet is written as L, with $\mathrm{SU(2)}_w$ components e_L (which is electrically charged) and ν_L (which is electrically neutral). The right-handed charged lepton is written as e_R. There is no known right-handed neutral lepton.

There are three generations of each of the matter fields listed in Table 2.1. The generations of up-type quarks are referred to as up (u), charm (c), and top (t), while the three generations of down-type quarks are referred to as down (d), strange (s), and bottom (b). The three generations of charged leptons are referred to as electrons (e),

muons (μ), and taus (τ), while the three generations of neutral leptons are referred to as electron neutrinos (ν_e), muon neutrinos (ν_μ), and tau neutrinos (ν_τ).

We can write kinetic terms (of the form $i\bar{\psi}\not{D}\psi$) for each of the fermion fields. Since these terms involve either two left-handed fields or two right-handed fields, the kinetic terms are gauge invariant. However, because the left and right hand fields transform differently under $\mathrm{SU}(2)_Y \times \mathrm{U}(1)_w$ (i.e. the left and right hand fields have different $\mathrm{SU}(2) \times \mathrm{U}(1)$ representations), we cannot write Dirac mass terms for the matter fields, since these terms would not be invariant under the SM symmetries. We can, however, use the Higgs field to write Yukawa terms for the matter fields, which gives rise to mass terms and to terms that describe the interactions between the fermions and the Higgs boson.

Putting together Eq. 2.1 (the kinetic terms for the gauge fields), Eq. 2.2 (the kinetic and potential terms for the Higgs field), the kinetic terms for the fermion fields, and the Yukawa terms for the fermion fields, we obtain the full SM Lagrangian. The SM Lagrangian is written as $\mathcal{L}_{\mathrm{SM}}$ in Eq. 1.1.

2.2 Effective field theory

As introduced in Chapter 1, EFT provides a general framework for describing the off-shell effects of heavy new physics as an expansion of higher-dimensional[1] operators. The operators are constructed of products of SM fields and their derivatives. At each order in the expansion, the operators are scaled by powers of Λ, the mass scale of the new physics. All operators of odd dimension violate baryon and/or lepton number [5], so are not considered in this analysis. Operators of dimension six thus represent the leading new physics effects and are the focus of this analysis.

[1]Here dimension refers to the mass dimension of the operator in natural units (where $\hbar = c = 1$ and all units are expressed as energy to some power). The SM operators are of dimension four, so "higher dimensional" refers to operators of dimension greater than four.

The dimension-six operators can be expressed in different bases, the most common of which is known as the Warsaw basis [6]. The number of operators at each dimension depends on the flavor symmetries that are assumed. Under the assumption that all three generations may vary independently, there are 2499 operators at dimension six; assuming flavor universality, this number reduces to 59 (assuming lepton number and baryon number conservation) [7].

Adopting a flavor symmetry assumption in between the two extremes, the model presented in Ref. [8] is the EFT model used in this thesis. This model is referred to as the dim6top model; it makes use of the Warsaw basis and provides tree-level modeling for dimension-six operators. Developed to facilitate studies focusing on third-generation effects, the dim6top model has compiled the set of 33 dimension-six operators involving two or more third-generation quarks. In the dim6top model, the operators are assumed to be invariant under $\mathrm{U}(2)_{\mathrm{Q}} \times \mathrm{U}(2)_{\mathrm{u}} \times \mathrm{U}(2)_{\mathrm{d}}$, so the couplings for operators involving third generation quarks may vary independently from the first two generations. While dim6top allows for EFT effects to vary independently for each generation of leptons, this analysis imposes the assumption that the EFT effects impact each lepton generation in the same way.

In this analysis, we aim to include all operators from the dim6top model that significantly impact processes in which one or more top quarks are produced in association with charged leptons; as listed in Table 2.2, this comes to 26 operators in total. The definitions of the WCs in Table 2.2 and the definitions of the corresponding operators can be found in Table 1 of Ref. [8]. However, for the vertices involving the c_{tG} WC, there is one important difference with respect to the definition in Ref. [8]. In order to allow the MadGraph event generator [9] 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 (as explained in Ref. [10]).

TABLE 2.2

LIST OF WILSON COEFFICIENTS INCLUDED IN THE ANALYSIS.

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

The 26 operators fall into four main categories: operators involving 4 heavy quarks, operators involving two heavy quarks and two light quarks, operators involving two heavy quarks and two leptons, and operators involving two heavy quarks and bosons. The vertices arising from these operators can impact the signal processes of this analysis, as illustrated in the example diagrams in Figure 2.1. The four four-heavy WCs only have significant impacts on the $t\bar{t}t\bar{t}$ signal process, so these WCs are not included in the modeling of the other five signal processes. The details of the Monte Carlo generation of the signal samples will be discussed in Section 4.2.1.

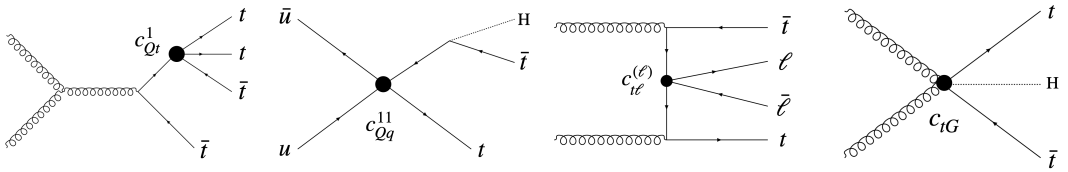


Figure 2.1. Example Feynman diagrams illustrating WCs from each of the categories listed in Table 2.2. From left to right, the diagrams show vertices associated with the c_{Qt}^{1}, c_{Qq}^{11}, $c_{t\ell}^{(\ell)}$, and c_{tG} WCs.

For each of the six signal processes, we account for diagrams with zero EFT vertices (i.e. the SM contribution) and diagrams with one EFT vertex (i.e. the new physics contribution). The amplitude for each process will thus depend linearly on the WCs, and the cross sections will depend quadratically on the WCs. Since the weight of each generated event corresponds to the event's contribution to the inclusive cross section, each event weight will also depend quadratically on the 26 WCs; the parametrization of the event weights in terms of the WCs is the key concept that allows us to obtain detector-level predictions in terms of the WCs and will be discussed in detail in Section 4.2.2.

CHAPTER 3

ACCELERATOR AND DETECTOR

Stretching 27 km in circumference as it straddles the Switzerland-France border near Geneva, Switzerland, CERN's Large Hadron Collider (LHC) is the largest and most powerful particle accelerator in the world. The LHC is described in detail in Ref. [11], which will be summarized in Section 3.1. Within the LHC, two counterrotating beams of protons are accelerated to a center of mass energy of 13 TeV before being made to collide within the four main particle detectors at the LHC. The Compact Muon Solenoid (CMS) detector is one of the two general-purpose detectors at the LHC. The CMS detector is composed of multiple subdetectors designed to measure and to record the energies, charges, and momenta of the particles produced by the high-energy collisions. A detailed description of the CMS detector and its components is provided in [12], which will be summarized in Section 3.2.

3.1 The large hadron collider

Before being injected into the LHC, the proton beams are accelerated by a series of smaller machines in the LHC accelerator complex. The beams enter the LHC with an energy of 450 GeV. Superconducting dipole magnets steer the beams around the ring. With a "twin-bore" design, the clockwise and counterclockwise beams are guided by separate coils of magnets in separate beam channels that reside in the same beampipe. Each beam is accelerated to a final energy of 6.5 TeV (during the 2016-2018 data-taking periods) by radio frequency (RF) cavities [13]. The RF cavities also help to keep the protons separated into discrete groups referred to as bunches.

There are about 10^{11} protons in each bunch, and the machine can be filled with a maximum of 2808 bunches. The nominal bunch spacing is 25 ns, meaning that the collision rate (i.e. bunch crossing rate) is about 40 MHz.

The LHC was designed to provide a instantaneous luminosity of $10^{34}\,\mathrm{cm}^{-2}\mathrm{s}^{-}1$. Throughout the 2016-2018 data-taking years, the LHC achieved a peak instantaneous luminosity of 1.5 to 2.1 times the nominal, delivering a total integrated luminosity of approximately $160\,\mathrm{fb}^{-1}$ [14]. The total number of proton-proton interactions in a given bunch crossing is referred to as the pileup (PU), and it is directly proportional to the luminosity. The average PU was 23, 33, and 32 for the 2016, 2017, and 2018 data taking years, respectively [15].

3.2 The CMS detector

The CMS detector is composed of multiple concentric subdetectors, each designed to measure the properties of different types of particles produced in the collisions. Cylindrical in shape, the detector has a length of 21.6 m, a diameter of 14.6 m, and a weight of 12,500 t. The coordinate system used by CMS is explained in Section 3.2.1. A schematic of the CMS detector (illustrating the various subdetectors) is shown in Figure 3.1.

The innermost subdetector is the tracker; it is composed of silicon and is responsible for recording the trajectories of the charged particles that emerge from the collisions. The tracker is discussed in Section 3.2.2. Surrounding the tracker, the electromagnetic calorimeter (ECAL) is a homogenous calorimeter that measures the energy of electrons and photons by detecting the scintillated light produced as the particles shower in the lead tungstate crystals that compose the ECAL. The ECAL is described in Section 3.2.3. The hadronic calorimeter (HCAL) surrounds the ECAL. Responsible for measuring the energies of charged and neutral hadrons, the HCAL is a sampling calorimeter composed of layers of brass absorber and plastic scintillator.

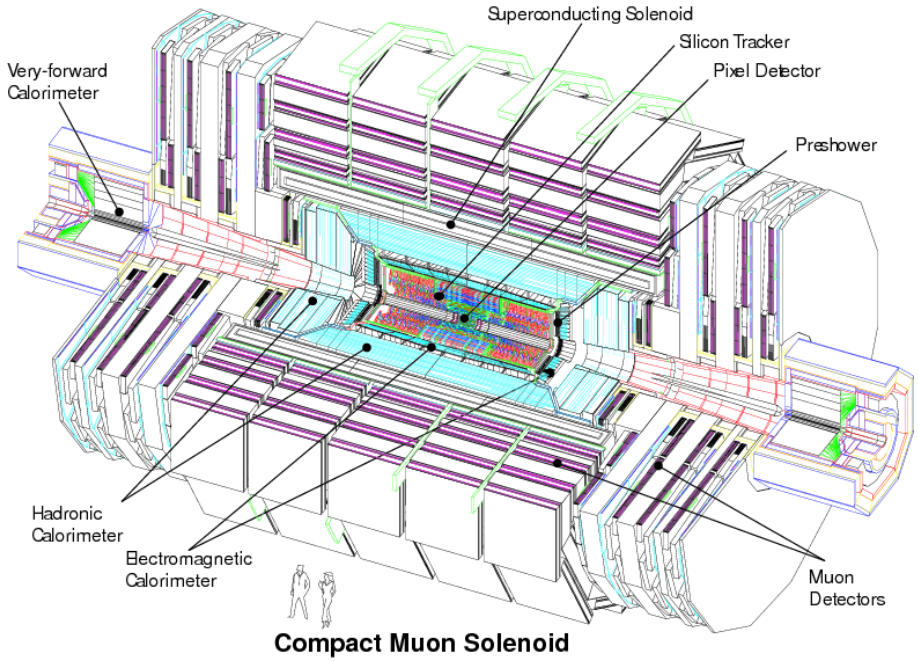


Figure 3.1. Schematic of the CMS detector with the main subdetectors labeled. [12]

The HCAL is discussed in Section 3.2.4. Beyond the HCAL lies CMS's superconducting solenoid magnet. With a magnetic field of 3.8 T, the solenoid bends the paths of charged particles as they traverse the detector, allowing their charge and momentum to be determined. Outside of the solenoid magnet are the CMS muon systems; composed of gas ionization detectors, the CMS muon system records the paths of muons as they pass through the detector. The muon detectors are described in Section 3.2.5. Responsible for parsing information from each subdetector in order to determine which collision events to record and which to reject, the CMS trigger system is described in Section 3.2.6.

3.2.1 Coordinate system

This section will summarize the coordinates and conventions used by the CMS experiment, which are illustrated in Figure 3.2. The origin of the coordinate system

is the nominal interaction point inside the CMS detector. The x axis points radially inward toward the center of the LHC accelerator, the y axis points upward, and the z axis points along the beam line in the counterclockwise direction. The azimuthal angle ϕ is measured from the x axis in the $x-y$ plane, and the polar angle θ is measured from the z axis in the $z-y$ plane.

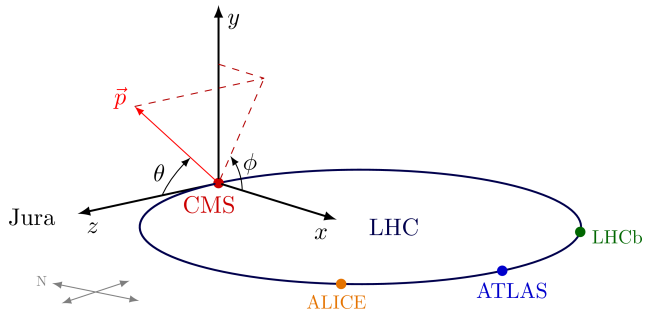


Figure 3.2. Coordinates used by the CMS experiment. An example momentum vector $\vec{p}$ is shown. [16]

The rapidity is defined as $\frac{1}{2}\ln\left(\frac{E+p_z}{E-p_z}\right)$. Differences in rapidity are preserved under boosts along the z direction, so the relative rapidity between particles is meaningful even if the particles have different boosts. Similarly, the pseudorapidity is defined as $\eta = \frac{1}{2}\ln\left(\frac{|\vec{p}|+p_z}{|\vec{p}|-p_z}\right) = -\ln\tan\left(\frac{\theta}{2}\right)$. Thus η is zero along the y axis and η approaches infinity along the beam line. In the limit that $|\vec{p}| \gg m$ (a good approximation for stable particles produced at the LHC), the rapidity and pseudorapidity are equal, so the pseudorapidity is usually used to describe angles in the $z-y$ plane.

Since the total momentum in the z direction varies for each collision, it is useful to consider the momentum in the plane transverse to the beam (the $x-y$ plane), where the total momentum of the colliding partons is zero. We thus define the transverse momentum $p_T = |\vec{p}|\sin(\theta) = \sqrt{p_x^2 + p_y^2}$.

Conservation of momentum requires that the sum of the transverse momentum vectors of the particles produced in the collision must be zero, so momentum that is carried away by neutrinos (or by any other hypothetical particles that are invisible to CMS) can be identified as the opposite of the sum of the p_T vectors of all of the visible particles produced in the collision; the magnitude of this vector is referred to as the missing transverse momentum, or $p_{\mathrm{T}}^{\mathrm{miss}}$.

3.2.2 The tracker

The CMS tracker is described in detail in Refs. [12, 17, 18]. The purpose of the tracker is to record the trajectories of charged particles without significantly disturbing their momenta. Since the tracker is the innermost component of the detector and is thus closest interaction point, it must have fine granularity yet be able to withstand high levels of radiation. Silicon provides these characteristics, so this is the material that was chosen for the CMS tracker. The tracker comprises two concentric subdetectors. Referred to as the silicon pixel detector and the silicon strip detector, these subdetectors are illustrated in in Figure 3.3.

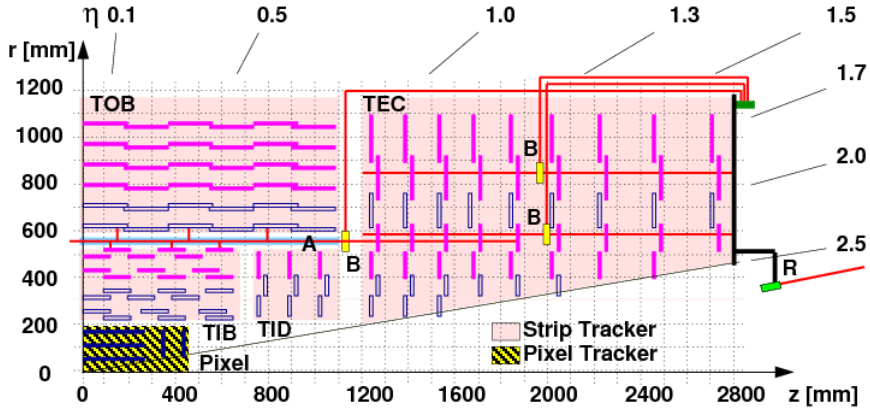


Figure 3.3. Diagram of a quadrant of the CMS tracker showing the pixel and strip trackers (with the interaction point at the lower lefthand corner of the schematic). [19]

Closest to the interaction point, the pixel detector contains 66 million silicon sensors referred to as pixels. Each pixel is 100 μm by 150 μm, so the total surface area of the pixel detector is about 1 m^2. In the barrel region, the pixel detector has several layers of sensors; the innermost layer has a radius of about 4 cm, and the outermost layer has a radius of about 11 cm. The endcaps have two layers of sensors, extending the coverage of the pixel detector over a range of $|\eta| < 2.5$. Surrounding the pixel detector is the silicon strip detector. Composed of 9.6 million silicon sensors, this detector comprises 10 layers in the barrel and 12 layers in each end cap. The silicon strip detector has a length of 5.8 m and a diameter of 2.5 m, and the total surface area of its sensors is about 200 m^2 (i.e. about the area of a singles tennis court).

The pixel tracker and the strip tracker record the paths of charged particles through the same mechanism. Silicon is a semiconductor, and its properties can be manipulated to provide sensitive detection of ionizing particles. Through the process of doping, n-type semiconductors (with excess electrons) and p-type semiconductors (with missing electrons i.e. holes) may be constructed from silicon. Interfacing p-type silicon with n-type silicon, a p-n junction is created. When an electric potential is applied such that the positive terminal is connected to the n-type region and the negative terminal is connected to the p-type region, the mobile charge carriers (electrons and holes) will move towards the positive and negative terminals, causing the region at the interface to become depleted of mobile charge. This arrangement is referred to as a reverse biased junction. An incoming particle may then ionize silicon atoms as it traverses the p-n junction in a pixel or strip sensor, resulting in free electrons that will move towards the positive terminal, producing a small pulse of current that can be amplified and recorded. This signal is referred to as a "hit" in the tracker. The hits are linked together to determine the path of the charged particle through the detector.

3.2.3 The electromagnetic calorimeter

The CMS ECAL is described in detail in Refs. [12, 17, 20]. The purpose of the ECAL is to measure precisely the energy of electrons and photons, absorbing the particles in the process. Constructed of lead tungstate ($PbWO_4$) crystals, the ECAL is a homogenous detector, which provides good energy resolution, as the entire detector is composed of active medium.

Lead tungstate has a high density and a short radiation length (a characteristic of the material corresponding to the mean distance an electron travels before losing all but $1/e$ of its energy via bremsstrahlung), allowing for a relatively compact design. The Molière radius (the radius in which approximately 90% of a shower energy is contained, a characteristic that is related to the radiation length) is correspondingly small, providing fine granularity and less overlap between the energy deposits of different particles. Figure 3.4 illustrates the components of the ECAL.

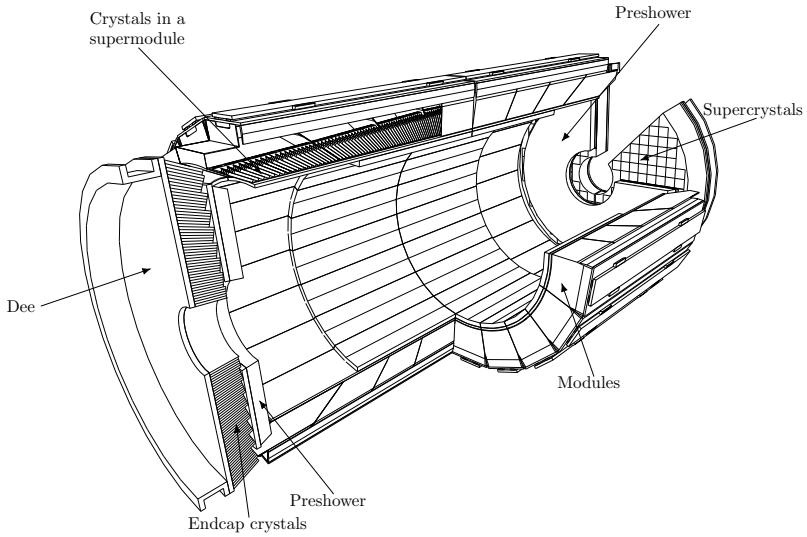


Figure 3.4. Diagram of the CMS ECAL, with components labeled. [12]

The barrel region of the ECAL contains over 60 k crystals and covers a range of $|\eta| < 1.479$. The ECAL endcaps contain over 12 k crystals and cover the $1.479 < |\eta| < 3.0$ regions. Spanning about 25 radiation lengths, the ECAL is able to contain more than 98% of the energy of the incoming electrons and photons. Incoming particles interact electromagnetically with the ECAL crystals, producing light in proportion to their energy. A photodetector is attached to each crystal to collect the scintillated light; avalanche photodiodes are used for the barrel, while vacuum phototriodes are used in the endcaps.

In front of each endcap, a finer-grained detector (referred to as the preshower detector) helps to identify pairs of photons arising from π^0 decays. With a total width of about 20 cm, the preshower detector is a sampling calorimeter composed of lead and silicon; the lead layer initiates the electromagnetic shower, while the silicon sensors measure the energy of the shower. The preshower detector covers a range of $1.653 < |\eta| < 2.6$.

3.2.4 The hadronic calorimeter

The CMS HCAL is described in detail in Refs. [12, 17, 21]. Composed of layers of brass absorber and plastic scintillator material, the HCAL is a sampling calorimeter designed to measure the energy of hadrons, absorbing them in the process. The HCAL has several subcomponents covering the barrel and endcap regions, as illustrated in Figure 3.5.

Constrained to fit between the outer radius of the ECAL (at 1.77 m) and the inner radius of the solenoid magnet (2.95 m), the main barrel region of the HCAL (HB) covers the range of $|\eta| < 3.0$. The total depth of the HB absorber is about six interaction lengths (where the interaction is the average length traveled by a particle before undergoing an interaction with the material). In the barrel region outside of the solenoid magnet, the HB is aided by an additional HCAL layer known as the

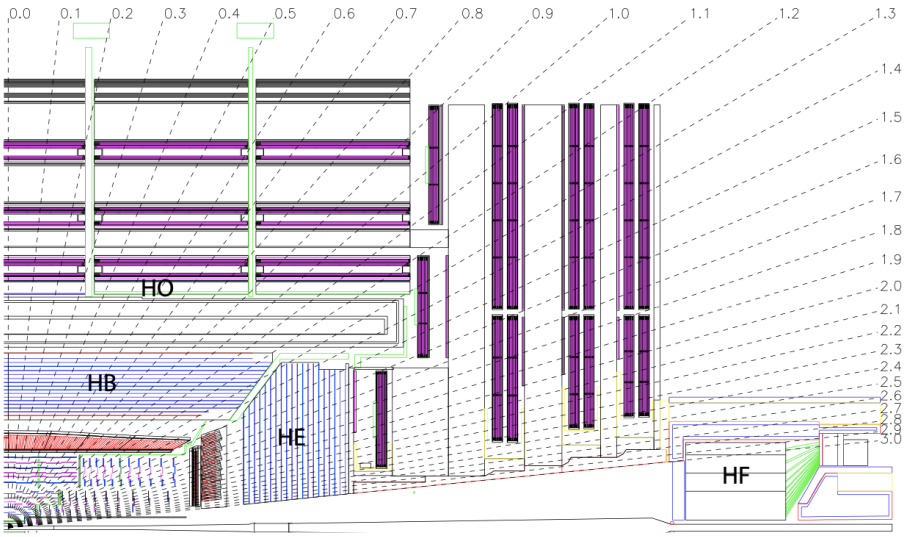


Figure 3.5. A quadrant of the CMS HCAL (with the interaction point at the lower lefthand corner). The barrel detector (HB), outer barrel detector (HO), endcap detector (HE), and forward detector (HF) are shown. [12]

outer detector (HO), which extends the total depth of the calorimeters to about 12 interaction lengths (this includes the ∼1 interaction length provided by the ECAL material). On the endcaps, the HCAL endcaps detectors (HE) cover the range of $1.3 < |\eta| < 3.0$ with a depth of about 10 interaction lengths (including the ECAL material). Like the barrel detector, the HE is composed of brass absorber and plastic scintillator.

To extend the η coverage of the HCAL to an $|\eta|$ of 5, a forward component (HF) is positioned beyond the muon systems as shown in Figure 3.5. Due to the high particle flux in the forward region, the design of the HF prioritizes radiation hardness. For this reason, quartz fibers were chosen as the active material. The HF detector is composed of steel absorber with grooves in which the quartz fibers are inserted. Signal in the HF is produced when shower particles generate Cherenkov radiation in the quartz fibers.

3.2.5 The muon detectors

The CMS muon detectors are described in detail in Refs. [12, 17, 22]. The purpose of the system is to identify muons, accurately reconstructing their momenta and charges. Composed of three subdetectors, the muon systems identify muons through the same basic mechanism; as muons pass through gaseous chambers, they ionize the gas and the free electrons drift towards positively charged wires or plates, creating signals known as hits. The muon detectors are illustrated in Figure 3.6.

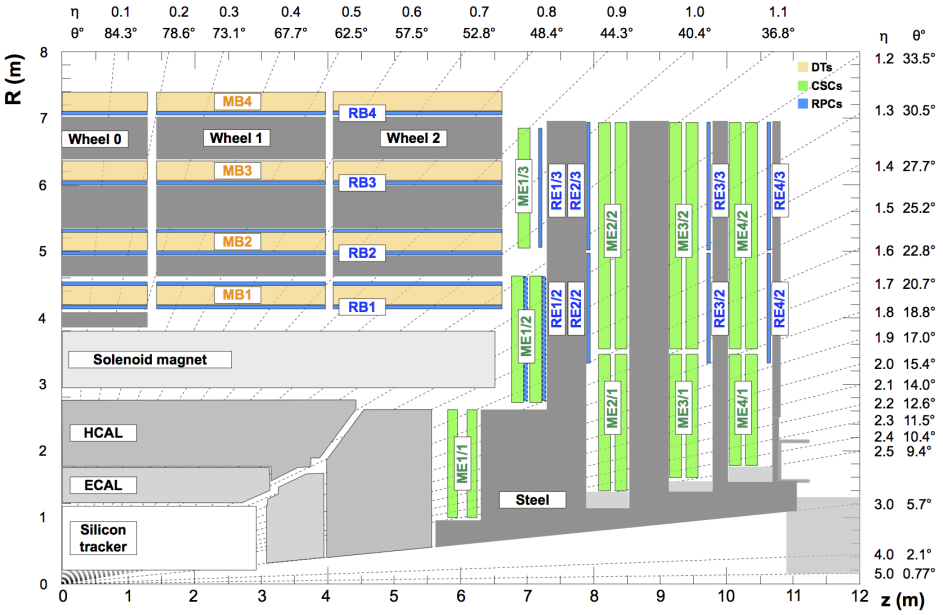


Figure 3.6. Diagram of a quadrant of the CMS detector (with the interaction point at the lower lefthand corner) illustrating the muon chambers. The components labeled "MB" represent the DTs, the components labeled "ME" represent the CSCs, and the components labeled "RE" and "RB" represent the RPCs. [22]

In the barrel region, 250 drift tube chambers (DTs) cover the $|\eta| < 1.2$ range. Each DT is filled with a mixture of Ar and CO_2 and contains a positively charged wire running through the chamber. When muons ionize the gas, the electrons drift towards the positively charged wire, producing a signal that is recorded as a hit. The location of the muon in the DT the chamber can be determined based the observed drift time and the known drift velocity.

In the endcaps, over 500 cathode strip chambers (CSCs) cover a range of $0.9 < |\eta| < 2.4$. The CSCs are finely segmented, with negatively charged cathode strips running radially outward (towards which the positive ions drift) and positively charged wires running approximately perpendicularly to the strips (towards which the electrons drift). This high resolution is important for the CSCs, as the endcaps are subject to high muon fluxes and a non-uniform magnetic field.

In both the barrel and endcap regions, about 1000 additional detectors known as resistive plate chambers (RPCs) cover a range of $\eta < 1.9$. Unlike the DTs and CSCs, the RPCs do not contain wire anodes. Rather, positively and negatively charged parallel plates lining the chamber form the anode and cathode for the detector. The RPCs provide fine timing resolution, but coarse position resolution; the precision timing information is useful for the muon triggers.

3.2.6 The trigger systems

Bunch crossings take place within CMS at 40 MHz, producing data at about 40 TB/s. This is much more data than could be recorded or stored for future analysis; however, only a small fraction of these collisions contain events likely to be associated with new or interesting phenomena. In order to select as many of the potentially interesting events as possible while reducing the rate to a more manageable 1 kHz, CMS employs a two-tiered trigger system consisting of a hardware based Level-1 (L1) trigger [23] and a software based high level trigger (HLT).

The L1 trigger is responsible for reducing the rate to approximately 100 kHz. Since collisions take place at 40 MHz, the L1 trigger must on average make a decision every 25 ns. The detector is capable of buffering data from 160 collisions [24], so the L1 trigger is able to have a latency of 4 μs; however, this is not enough time for the data from the tracker to be read out, so the L1 trigger uses information from only the calorimeters and the muon detectors. Figure 3.7 shows a schematic of the L1 trigger system architecture.

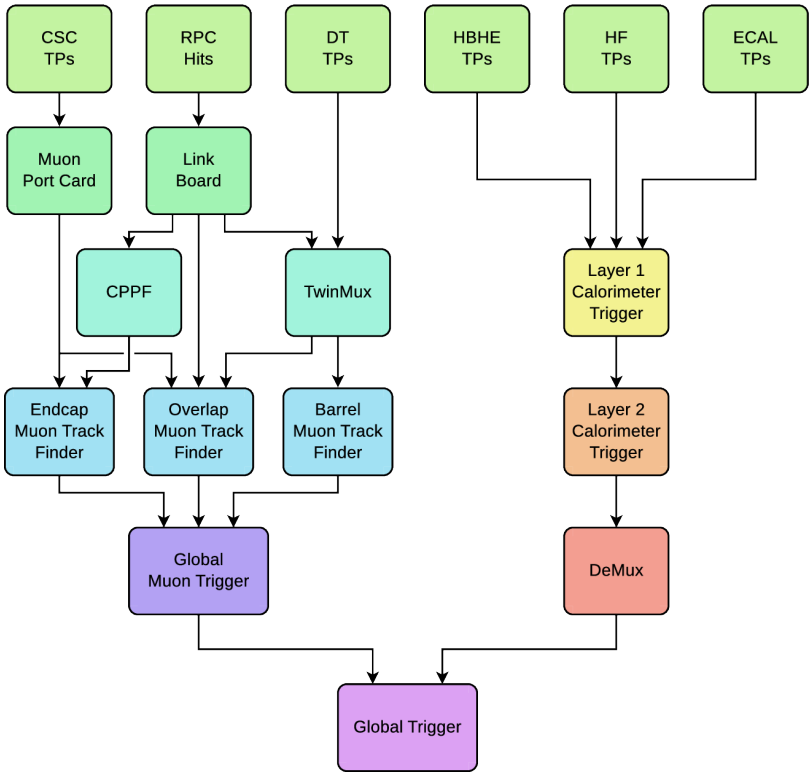


Figure 3.7. Schematic of the L1 trigger system, as described in the text. [23]

As shown in the Figure 3.7, the data from the calorimeters are received by the Layer 1 Calorimeter Trigger; these data correspond to the energy deposits recorded

by the calorimeters and are referred to as trigger primitives (TPs). The Layer 2 Calorimeter Trigger reconstructs physics objects from the TPs, and the demultiplexer board ("DeMux") serializes and formats the events to be sent to the L1 Global Trigger. From the muon systems, the muon track finder (MTF) algorithms combine information from the DTs, CSCs, and RPCs, using the concentrator and preprocessor fanout (CPPF) to cluster hits from the endcap RPCs into TPs and the TwinMux [25] to combine information from the DTs and RPCs. The MTFs pass the candidate objects to the Global Muon Trigger, which selects up to 8 objects to pass to the L1 Global Trigger.

The L1 Global Trigger receives the calorimeter and muon objects and executes a set of algorithms on these objects; the algorithms are designed to identify events involving potentially interesting phenomena, e.g. events with multiple muons or events with high p_T jets. The set of algorithms is referred to as the L1 trigger menu, and the total number of algorithms in the menu is typically about 400. If an event passes the selection criteria for any of the algorithms in the menu, the L1 Global Trigger accepts the event. The event is otherwise discarded.

If the L1 Global Trigger accepts an event, the data acquisition system reads out all of the data associated with the event and sends the data to the HLT. The HLT consists of a farm of processors comprising approximately 30,000 CPU cores. Responsible for reducing the rate of accepted events to about 1 kHz (on average over the course of an LHC fill), the HLT has a latency of about 300 ms. This latency allows the HLT to execute sophisticated reconstruction algorithms that are similar to offline reconstruction, using information from all of the subdetectors including the tracker. The HLT uses a simplified version of the particle flow (PF) reconstruction algorithm used for offline reconstruction (described in Chapter 5). Optimized for CPU speed, the online reconstruction does not incorporate electron identification into PF, and does not perform reconstruction for low p_T tracks in the tracker [17].

Similar to the L1 menu, the HLT menu consists of a set of selection criteria targeting potentially interesting signatures. The reconstruction steps are performed in a set order, with faster steps being executed first and slower steps executed last. This ordering allows events that do not pass the selection criteria for any of the paths in the menu to be rejected as quickly as possible. If an event passes the selection requirements for any of the algorithms in the HLT menu, it is transferred to the CMS Tier-0 computing center to be processed offline and stored permanently.

CHAPTER 4

DATA AND MONTE CARLO SAMPLES

The data samples used in this analysis are described in Section 4.1. The Monte Carlo (MC) simulated samples are described in Section 4.2, with the generation of the EFT samples (with the necessary EFT weights) covered in 4.2.1, and the quadratic parametrization of the weights detailed in 4.2.2. All samples used in this analysis are in the v9 NanoAOD format [26] with Ultra Legacy (UL) reconstruction [27].

4.1 Data samples and triggers

This analysis uses data from proton-proton collisions at $\sqrt{s} = 13$ TeV collected by the CMS experiment from 2016 to 2018, using the subset of lumisections that have been certified by CMS as good for physics analysis (Table A.1). The total integrated luminosity is 138 fb^{-1} with an uncertainty of 1.6% fb^{-1} [28].

The data used in this analysis are collected with a combination of single, double, and triple-lepton triggers. The p_T thresholds for the various single-lepton triggers range from 22 to 35 GeV. The p_T thresholds for the double and triple-lepton triggers are generally not as high as the single-lepton thresholds, since events with multiple leptons are more rare, so the p_T thresholds may be lowered without resulting in too high of a trigger rate; for example, the triple-muon trigger has p_T thresholds of 12, 10, and 5 GeV. Sets of related triggers are grouped into categories referred to as datasets. The datasets and associated triggers used in this analysis are listed in Table A.2, A.3, and A.4 (for 2016, 2017, and 2018, respectively).

While the triggers in a given dataset are exclusive (i.e. an event may never pass more than one trigger in a given dataset), triggers from different datasets may overlap. This overlap must be accounted for in order to avoid double counting. The following procedure is used:

- An arbitrary order of the datasets from a given year is chosen.
- An event that is from the first dataset (dataset A) is never discarded.
- An event that is from the second dataset (dataset B) is discarded if it passes any of the triggers from dataset A (since it was already accounted for in A).
- An event that is from the third dataset (dataset C) is discarded if it passes any of the triggers from dataset A or dataset B (since it was already accounted for).
- The procedure continues for all of the datasets that are included in the given data-taking period.

The orders of the datasets listed in Tables A.2, A.3, and A.4 correspond to the order used in the overlap removal procedure implemented in this analysis, and Appendix A.1 steps through the procedure for an example event.

4.2 Monte Carlo samples

This analysis aims to study dimension-six EFT effects on processes in which one or more top quarks are produced in association with additional charged leptons; processes which lead to the same multilepton final-state signatures but are not impacted by these EFT operators are backgrounds for this analysis. The expected background contributions are estimated using a combination of simulated samples and data-driven techniques, discussed in Chapter 8 (with the simulated samples used in the background estimation listed in Appendix A).

The expected yield for a given selection is calculated as

$$\text{Expected yield} \;=\; \sigma\,\mathcal{L}\;\frac{\sum_{\text{Pass}} w}{\sum_{\text{Gen}} w}, \tag{4.1}$$

where σ is the inclusive SM cross section for the given process, $\mathcal{L}$ is the integrated luminosity, and w are the event weights. Conceptually, the event weights represent how much a given event contributes to the overall cross section. The sum in the numerator ($\sum_{\text{Pass}} w$) runs over the events that pass the given selection criteria, and the sum in the denominator ($\sum_{\text{Gen}} w$) runs over all generated events. The ratio of these sums corresponds to the acceptance times efficiency. As will be explained in the following sections, the weights of the signal samples are functions of the WCs; the details of the signal sample generation are described in Section 4.2.1, and 4.2.2 covers the details that are specific to the EFT weights. After the EFT weights have been explained, we will revisit Eq. 4.1 (in Section 4.2.2 Eq. 4.6), discussing the subtleties of the normalization that are specific to this analysis.

4.2.1 Monte Carlo generation of signal samples

The signal processes for this analysis are $t\bar{t}H$, $t\bar{t}l\nu$, $t\bar{t}l\bar{l}$, $tl\bar{l}q$, tHq, and $t\bar{t}t\bar{t}$. The signal samples are produced at leading order (LO) with the MadGraph [9] event generator (version 2.6.5). As discussed in Chapter 2, the dim6top UFO model [8] is used to incorporate the EFT effects. Parton showering and hadronization for the samples are performed with the Pythia generator [29], which also handles the decays of the top quark and the Higgs boson. In order to avoid overlap between the $t\bar{t}l\bar{l}$ and $t\bar{t}H$ samples, we specify in the MadGraph process card that the $t\bar{t}l\bar{l}$ process should not include an intermediate H; the same requirement is made for the $tl\bar{l}q$ process in order to avoid overlap with tHq. All simulated signal processes are normalized to theoretical SM cross sections at next-to-leading order (NLO) in QCD, as listed in Table A.5. The EFT samples produced for this analysis are stored at the Notre Dame T3. For reference, their file paths are listed in Tables A.6, A.7, A.8, and A.9.

For the $t\bar{t}X$ processes ($t\bar{t}H$, $t\bar{t}l\nu$, and $t\bar{t}l\bar{l}$), we include an additional final state parton in the matrix element (ME) generation. The inclusion of the additional par-

ton can improve the modeling at high jet multiplicities, and can also significantly impact the dependence of the $t\bar{t}X$ processes on the WCs [10]. The primary factors contributing to the modification of the cross section's EFT dependence are related to the new quark-gluon initiated diagrams that become available when an additional final state parton is included in the ME calculation. For example, without an extra parton, $c_{\varphi t}$ can only contribute to $t\bar{t}H$ via quark-anti-quark initiated diagrams (e.g. the diagram in the lefthand side of Figure 4.1); however, when an extra parton is included, $c_{\varphi t}$ can contribute via quark-gluon initiated diagrams like the one shown in the righthand side of Figure 4.1.

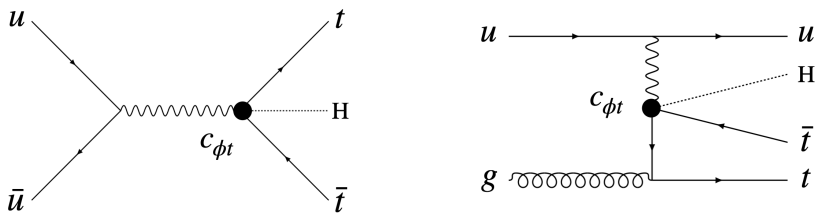


Figure 4.1. Example $c_{\varphi t}$ diagrams for $t\bar{t}H$ without and with an extra parton.

The single top processes ($t\bar{l}\bar{l}q$ and tHq) and the $t\bar{t}t\bar{t}$ sample are not produced with an additional parton. The single top processes have technical complications associated with correctly performing the jet matching between the ME and the parton shower (PS) for t-channel single top processes that currently do not allow a valid matched sample to be produced. In the case of $t\bar{t}t\bar{t}$, an additional parton is not included because the generation of the MadGraph gridpack is very computationally expensive. It would not be feasible to produce enough $t\bar{t}t\bar{t}$ samples to perform thorough validations of the starting point and matching parameters. However, the effect of an additional parton for $t\bar{t}t\bar{t}$ is not expected to be very significant, since the

$t\bar{t}t\bar{t}$ process already naturally populates the high-jet multiplicity bins of the analysis. Furthermore, $t\bar{t}t\bar{t}$ is already dominated by gluon-gluon initiated diagrams, so quark-gluon initiated diagrams would be expected to have a smaller impact.

Since we are unable to include an additional parton for the single top samples, and in these cases the extra parton may potentially have a significant effect on the high jet multiplicity categories (since these single top processes would not generally produce as many jets as our other signal processes), we apply an additional uncertainty to these processes, described in Chapter 9. This uncertainty is determined by comparing the jet multiplicity distribution of our EFT samples (reweighted to the SM) against SM NLO samples, listed in Table A.10.

For the samples produced with an additional parton, a matching procedure must be applied to account for the overlap in phase space between the contributions of the ME and PS. For this analysis, the matching is implemented using the MLM scheme [30], an event-rejection based approach that matches ME partons to jets clustered by Pythia, discarding events in which the jets are not successfully matched to partons in order to avoid double counting.

It should be noted that the matching procedure can lead to complications when applied to EFT samples; since EFT effects are included in the ME contribution, but not in the PS contribution, it is possible that an inconsistency could arise. Specifically, if an EFT vertex produces a significant soft and collinear contribution, the events removed by the matching procedure will never be replaced by corresponding events generated by the PS, causing this contribution to be missed. However, of the WCs considered in this study, the operator associated with the c_{tG} WC is the most prone to these effects, and its contributions to the soft and collinear regime are suppressed; thus, the phase space overlap with the SM contribution from the PS is small, and the effects of this potential issue are negligible [10].

In addition to the theoretical justification outlined above, we can validate the matching procedure empirically by examining differential jet rate (DJR) distributions for the simulated samples. The resulting DJR distributions provide further evidence that the matching is working properly. The details of the validation of the DJR distributions may be found in Appendix B.1, and a more detailed discussion of the validation of matched $\mathrm{t\bar{t}X}$ samples is presented in [10].

As an additional form of validation, our EFT signal samples are reweighted to the SM and compared against SM samples that are centrally produced by the CMS collaboration. The details of this comparison are presented in Appendix C. The comparisons show that the level of agreement is generally good, providing further evidence that the reweighting is working properly and that the LO modeling (and associated uncertainties) are sufficient for this analysis.

4.2.2 Parameterization of the predicted yields in terms of the WCs

This section will describe the method through which the predicted yields are parameterized in terms of the WCs. In order to write the predicted yields as a function of the WCs, it is first necessary to understand how the cross section depends on the WCs. Starting with the ME, we can write the amplitude for a given process as the sum of the SM and new physics components:

$$\mathcal{M} = \mathcal{M}_{\mathrm{SM}} + \sum_i \frac{c_i}{\Lambda^2}\mathcal{M}_i, \tag{4.2}$$

where $\mathcal{M}_{\mathrm{SM}}$ is the SM ME, $\mathcal{M}_i$ are the MEs corresponding to the new physics components, and c_i are the WCs. Since the cross section (inclusive or differential) is proportional to the square of the ME, it will depend quadratically on the WCs:

$$\sigma \propto |\mathcal{M}|^2 \propto s_0 + \sum_j^N s_{1j}\frac{c_j}{\Lambda^2} + \sum_j^N s_{2j}\frac{c_j^2}{\Lambda^4} + \sum_{j\neq k}^N s_{3jk}\frac{c_j}{\Lambda^2}\frac{c_k}{\Lambda^2}, \tag{4.3}$$

where s_{jk} are structure constants of the N-dimensonal quadratic function for N WCs. The number of structure constants (K) required to describe an N-dimensional quadratic can be written as the following:

$$K = \frac{(N+1)\cdot(N+1) - (N+1)}{2} + (N+1). \tag{4.4}$$

This analysis considers 26 WCs, so by Eq. 4.4, there are 378 structure constants required to fully describe the 26-dimensional quadratic. In principle, we could solve for these structure constants if the cross section at 378 points in the 26-dimensional WC space were known. However, this would require generating 378 unique simulated samples at 378 unique points in the 26-dimensional WC space. In practice, it would not be feasible to generate this many simulated samples.

Instead of attempting to determine the parametrization for the inclusive cross section, we parametrize each event's weight in terms of the WCs. Since each weight corresponds to the event's contribution to the inclusive cross section, the event weight essentially represents a differential cross section, which can be described by a 26-dimensional quadric in terms of the WCs, as written in equation 4.3. In order to determine the 378 structure constants of the event weight's quadratic parameterization, we need to know the event weight at 378 distinct points in the 26 dimensional space. This is feasible to do using the MadGraph event reweighting procedure [31].

Given an event generated under a specific theoretical scenario, the MadGraph event reweighting procedure computes additional weights associated with the same event under alternative theoretical scenarios. In the case of EFT reweighting, the original theoretical scenario corresponds to a particular point in the 26-dimensional WC space, provided to MadGraph by the user. We refer to this as the "starting point" for the sample. The alternative theoretical scenarios correspond to other distinct points in the 26-dimensional WC space (i.e. other sets of values for the 26 WCs),

also provided to MadGraph by the user. From the matrix-element computations, MadGraph calculates the weight at the starting point and at each of the additional reweight points. With at least 378 weights corresponding to 378 independent points in the 26-dimensional WC space, we can solve for the 378 structure constants, and fully determine the 26-dimensional quadric function that describes the event's weight in terms of the WCs.

Once we have obtained each event's 26-dimensional quadratic parametrization $w_i(\vec{c}/\Lambda^2)$, we can find the dependence of any observable bin on the WCs by summing the quadratic parameterizations for each of the events that passes the selection criteria for the given bin. The sum of the weights for the passing events ($\sum_{\text{Pass}} w$ from Eq. 4.1) can thus be written follows:

$$\begin{aligned}\sum_{\text{Pass}} w &= \sum_i w_i \left(\frac{\vec{c}}{\Lambda^2}\right) \\ &= \sum_i \left(s_{0i} + \sum_j s_{1ij}\frac{c_j}{\Lambda^2} + \sum_j s_{2ij}\frac{c_j^2}{\Lambda^4} + \sum_{j\neq k} s_{3ijk}\frac{c_j}{\Lambda^2}\frac{c_k}{\Lambda^2}\right),\end{aligned} \tag{4.5}$$

where the sum over i corresponds to the sum over all of the events that pass the selection criteria for the given bin. Performing a similar sum over all generated events in the sample, we can obtain the $\sum_{\text{Gen}} w$ term from Eq. 4.1 as well. Since the sum of multiple quadratic functions is also quadratic, both $\sum_{\text{Pass}} w$ and $\sum_{\text{Gen}} w$ are quadratic in terms of the WCs.

In this analysis, we evaluate $\sum_{\text{Gen}} w$ at the SM (effectively canceling the LO cross section). We then normalize the predicted yield to a more precise (i.e. NLO or better) SM cross section calculation. Rewriting Eq. 4.1 with these nuances included, the expected yield in any bin can be expressed as a function of the WCs as follows:

$$\text{Expected yield } (\vec{c}) \;=\; \sigma_{SM}\,\mathcal{L}\;\frac{\sum_{\text{Pass}} w(\vec{c})}{\sum_{\text{Gen}} w(SM)}, \tag{4.6}$$

Where $\vec{c}$ are the WCs, σ_{SM} is the inclusive SM cross section from an NLO (or better) calculation, $\sum_{\text{Pass}} w(\vec{c})$ is the sum of the event weight parameterizations of the passing events, and $\sum_{\text{Gen}} w(SM)$ is the sum of the event weight parameterizations for all generated events (evaluated at the SM).

Since we are thus able to write the predicted yield of any observable bin as a function of the 26 WCs, we can obtain detector-level predictions at any arbitrary point in the 26-dimensional EFT space. This is the key enabling concept of this analysis, as it allows for all EFT effects across all analysis bins to be simultaneously accounted for when performing the likelihood fitting with the statistical framework (which will be described in Chapter 10).

We generate all of our signal processes using the reweighting procedure described in this section. However, we do not include all WCs for all processes (since some of the WCs do not impact some of the processes), so the number of reweight points included in the MC generation varies by sample. The $t\bar{t}t\bar{t}$ process incorporates the full set of 26 WCs. By Eq. 4.4, a total of 378 weights are required to fully determine the 26-dimensional quadratic parameterization. However, in order to ensure that a good fit can be found, we over-constrain the fit by including approximately 20% more points than the minimum number required, for a total of 454 reweight points. As discussed in Section 2.2, the other five signal samples have a negligible dependence on the four four-heavy operators to which $t\bar{t}t\bar{t}$ is sensitive, so these samples incorporate only 22 WCs. This means a minimum of 276 reweight points are required to determine the 22-dimensional quadratic fit, but we again ensure the fit is over-constrained by generating additional reweight points, for a total of 332 reweight points for each event.

The MadGraph reweighting procedure is powerful because it allows different regions of EFT phase space to be probed with a single MC sample; however, there is an important caveat to the procedure that should be highlighted. Since MadGraph

produces unweighted samples of events, the events generated by MadGraph mainly correspond to phase space occupied by original event. Thus, the reweighting procedure does not work unless the original point in phase space (i.e. the starting point) and the alternative points in phase space (i.e. the reweight point) have some overlap. EFT operators lead to new diagrams that may populate areas in phase space that are not present in the SM, therefore the SM cannot be used as a valid starting point for the reweighting procedure. Instead, a point that is relatively far from the SM should be chosen.

Nevertheless, even for non-SM starting points, there is still no guarantee that the chosen point will allow MadGraph to properly reweight to all areas of relevant phase space. Therefore, it is important to validate reweighted samples to ensure that they are able to be consistently reweighted to as much of the relevant phase space as possible. For example, we check that the samples are able to be consistently reweighted to other points in EFT phase space (by comparing against dedicated samples produced at the given point in phase space), as well as checking the distribution of event weights for samples generated at different starting points. Details regarding the validation checks performed for this analysis are provided in Appendix B.2.

CHAPTER 5

OBJECT RECONSTRUCTION

This chapter describes the reconstruction and selection of the final state objects (leptons, jets, and b-tagged jets) used in this analysis. The reconstruction algorithm is explained in Section 5.1. The identification of jets and b-tagged jets is discussed in Section 5.2. The lepton selection is described in Section 5.3.

5.1 Particle flow reconstruction

In order to identify and reconstruct the particles produced in each collision, CMS uses a holistic reconstruction technique to correlate the elements from each subdetector (tracks and clusters) and construct a global picture of each event. This approach is referred to as particle flow (PF) reconstruction [17]. Since a particle generally interacts with multiple subdetectors, there are expected to be several PF elements associated with a given particle. The first step in the PF reconstruction is to link together the PF elements from the different subdetectors into sets of elements referred to as PF blocks.

In each block, the identification and reconstruction of the particles is performed in a specific order, and as each particle is identified, the associated elements are removed from the block. First, muons are identified based on tracks in the muon chambers, tracks in the inner tracker, and lack of energy deposits in the calorimeters. Next, electrons and isolated photons are reconstructed from energy clusters in the ECAL, lack of significant energy clusters in the HCAL, and tracks (or lack thereof) in the inner tracker. Finally, charged hadrons, neutral hadrons, and non-isolated photons

are identified through several iterative steps using information from the inner tracker and both calorimeters. A schematic representation of PF reconstruction is shown in Figure 5.1, with example signatures displayed for each particle type (muon, electron, charged hadron, neutral hadron, and photon).

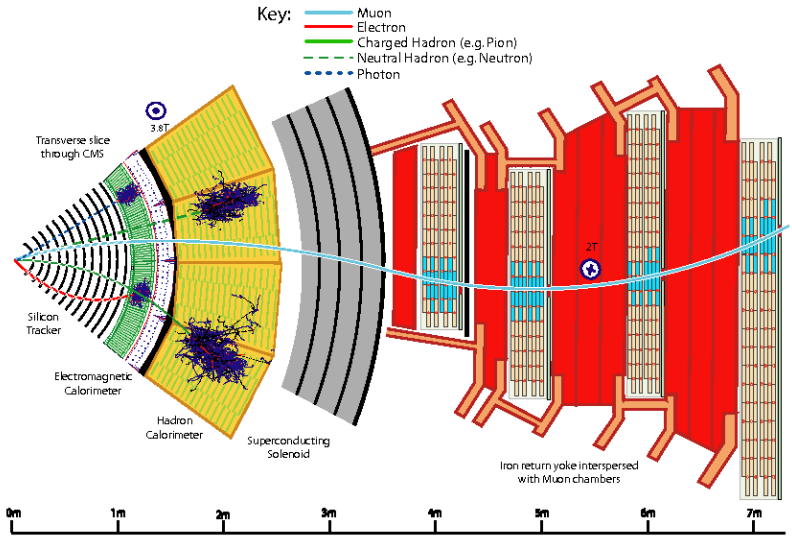


Figure 5.1. Schematic representation of the reconstruction of particles within CMS. An example signature for a muon, electron, charged hadron, neutral hadron, and a photon are shown. The PF algorithm is described in detail in Ref. [17] (from which this figure is taken) and summarized in Section 5.1.

Despite this thorough reconstruction algorithm, there may still be rare cases in which particles are misreconstructed, so a post-processing step is performed to mitigate these potential issues. The post-processing algorithm primarily aims to identify events with artificially high $p_{\mathrm{T}}^{\mathrm{miss}}$, usually due to the misidentification or misreconstruction of a high-p_T muon. The reconstruction of such muons is modified a posteriori and the $p_{\mathrm{T}}^{\mathrm{miss}}$ is recomputed; if the $p_{\mathrm{T}}^{\mathrm{miss}}$ is consequently reduced by at least a factor of two, the modified reconstruction is used.

5.2 Jets and b-tagging

After the PF reconstruction has identified all particles in the event, jets (collimated sprays of particles representing the experimental signature of quarks and gluons) are reconstructed using the anti-k_{T} jet clustering algorithm [32]. As explained in [32], the algorithm first defines a distance measure between the objects (d_{ij}), and between objects and the beam (d_{iB}); if the smallest distance measure is between two objects, they are combined into a single object, while if the smallest distance measure is between an object and the beam, the object is called a jet and removed from the list of objects. The distances are then recalculated and the process continues until all objects have been clustered into jets. For the anti-k_{T} algorithm, the distance measures are defined as follows:

$$d_{ij} = \min\left(\frac{1}{p_{Ti}^2}, \frac{1}{p_{Tj}^2}\right)\frac{\Delta R_{ij}^2}{R^2} \tag{5.1}$$

$$d_{iB} = \frac{1}{p_{Ti}^2} \tag{5.2}$$

where p_{Ti} is the p_T of particle i, p_{Tj} is the p_T of particle j, and R is a radius parameter; $\Delta R_{ij}^2 = (y_i - y_j)^2 + (\phi_i - \phi_j)^2$, where y is the rapidity, and ϕ is the azimuthal angle. The anti-k_{T} clustering algorithm is implemented with the FASTJET package [33]. The distance parameter R in the anti-k_{T} algorithm is set to 0.4.

Charged particles arising from pileup interactions are excluded from the jet clustering algorithm using the charged hadron subtraction (CHS) technique [34]. With this approach, charged hadrons with tracks associated with pileup vertices (i.e. vertices other than the primary vertex with the highest Σp_T^2) are removed from the list of particles to be used in the object reconstruction.

In this analysis, we require all jets to have $p_T > 30$GeV and $|\eta| < 2.4$. The jets are cleaned using the loose (or better) leptons with a ΔR requirement of greater than

0.4. Jets that overlap with objects in the fakeable lepton collection (electrons and muons) are also removed. The definitions of loose and fakeable leptons are provided in Section 5.3.

In order to identify jets originating from b quarks, the DeepJet b-tagging algorithm is used [35]. The relatively long lifetimes of b hadrons can lead to displaced secondary vertices that produce tracks with large impact parameters with respect to the primary vertex (PV); as shown in Figure 5.2, the impact parameter is defined as the distance at the closest point of approach between the track and the PV. The relatively large mass of b hadrons can lead to decay particles with relatively large p_T with respect to the jet axis. These distinguishing characteristics are leveraged in the identification of b jets with the DeepJet algorithm.

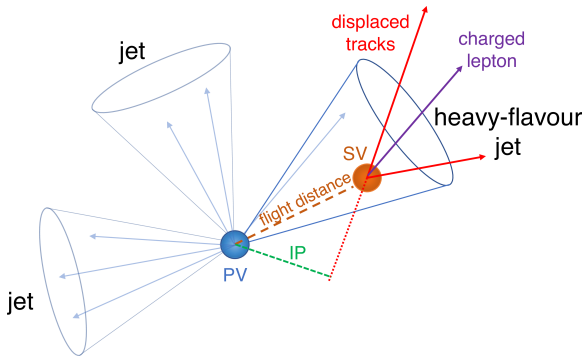


Figure 5.2. Schematic depiction of b jet, illustrating the impact parameter "IP" with respect to the PV. [35]

The b-tagging efficiency and mistag rates depend on the working point utilized (i.e. the cut on the discriminant output, which ranges from zero to one). For this analysis, we make use of the loose and medium working points of the DeepJet algorithm, as defined by the CMS b-tagging and vertexing (BTV) physics object group (POG) for the UL datasets [36–39].

5.3 Lepton object selection

This section will describe the lepton object selection used in this analysis, which is based on the selection used in the $t\bar{t}H$ multilepton analysis [40]. Section 5.3.1 provides an overview of the selection strategy and a conceptual introduction to the three stages of the lepton selection. A more detailed description of each selection stage is provided in Section 5.3.2, the requirements for which are summarized in Tables 5.1 and 5.2. Finally, the technical definitions of the observables that are used in the selection requirements are described in Section 5.3.3.

5.3.1 Conceptual overview of lepton object selection

The event selection for the signal regions of this analysis (Chapter 6) is defined primarily by the multiplicity of charged leptons (usually referred to as simply "leptons"). For this reason, it is important to ensure that the objects that are classified as leptons are likely to correspond to true leptons of interest. To this end, the lepton object selection aims to identify leptons that arise promptly from the PV.

To understand which leptons are to be classified as prompt, let us consider a $t\bar{t}W$ event. Each top quark will decay to a W boson and a b quark[1]. The event thus leads to three W bosons and two b quarks. The W bosons will decay either leptonically or hadronically; for the purposes of this discussion, let us consider leptonic decays, and let us assume the leptons are electrons or muons [2]. The bottom quarks will hadronize, producing collimated sprays of hadrons (i.e. b jets). At this point, the collision has produced three charged leptons, three neutrinos, and two b jets. From an experimental perspective, all of these particles are produced essentially instantaneously, directly at the location of the PV.

[1]While it is possible for a top quark to decay to other flavors of quarks, the branching fraction for a top quark to decay to a W and a bottom quark is nearly 100% [41]

[2]Electrons and muons that arise from the decays of prompt taus are also considered to be prompt.

The charged leptons will travel outward from the PV, eventually encountering the CMS detector (which will measure their properties as described in Section 5.1); these leptons are considered to be prompt. The b hadrons in the b jets, however, may travel some observable distance before decaying; the decays may lead to leptons, but these leptons will not have originated directly from the PV, so these leptons are considered to be nonprompt. The goal of the lepton object selection is to distinguish prompt leptons from nonprompt leptons. The leptons that are identified as prompt are the leptons we use for the event selection.

Although a $t\bar{t}W$ event was used in this example, the same logic is applied to all processes. Any lepton that arises promptly from the PV is considered to be prompt. For example, a lepton that arises from the decay of a Z boson from the decay of a Higgs boson in a $t\bar{t}H$ event is also considered to be a prompt lepton. As one more example, a lepton that arises directly from an EFT vertex (e.g. the $c_{t\ell}^{(\ell)}$ vertex in Figure 2.1) would also be considered to be a prompt lepton.

The identification of prompt leptons is performed in three stages, referred to as the loose, fakeable, and tight selections. As illustrated in Figure 5.3, the stages apply increasingly stringent requirements, so the collection identified in each stage is a subset of the previous collection. The loose stage aims for high efficiency; this baseline set of objects is used to veto events with leptons from light resonances (i.e. from the decays of J/ψ and Υ particles, as explained in Section 6.2) and in the training of the MVA that is used in the tight selection. Building on the requirements of the loose selection, the fakeable selection is used to identify the set of leptons that are used in the estimation of the nonprompt background (discussed in Sectoin 8.1). Finally, the tight selection is applied in order to select the final set of prompt leptons that will be used in the event selection.

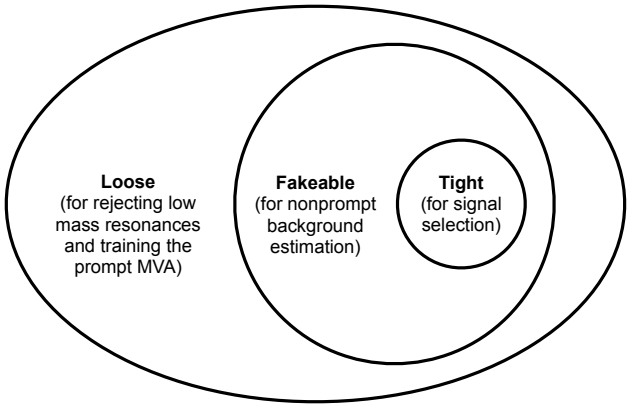


Figure 5.3. Venn diagram illustrating the three lepton object selection stages.

5.3.2 Lepton object selection stages

With the goal of separating electrons from jets, the loose stage of the electron object selection makes use of a boosted decision tree (BDT) multivariate algorithm, trained by the CMS EGamma POG [42–44]. For this loose selection, we use the loose working point (which has an effeciency of 98% [43]). Loose muons are required to pass the loose identification PF requirements defined by the CMS Muon POG [45]. The loose selection requirements for electrons and muons also include cuts on the isolation of the lepton and the impact parameter of the lepton's track with respect to the PV. Since signal leptons are expected to be relatively isolated from hadronic activity and to originate from the hard scatter process, these requirements can help to distinguish the leptons of interest from background objects. All of the requirements for the loose selection are are summarized in the "Loose" columns of Tables 5.1 and 5.2. The loose leptons are used to identify and veto events with a low-mass resonance (as described in Section 6.2), and are also used for training the prompt-e and prompt-μ MVAs described below.

The next stage in the object selection is referred to as the fakeable selection. Building on the loose collection, the objects that pass the fakeable requirements are a subset of the objects that pass the loose selection. Listed in the "Fakeable" columns of Tables 5.1 and 5.2, these selection requirements result in a collection of leptons that are very close to the final selection of signal leptons; this collection is used for the estimation of the nonprompt background contribution (as described in Section 8.1).

The final stage in the lepton object selection is the tight selection, which is used to identify the set of leptons that are considered to be signal leptons. Building on the fakeable and loose selections described above, the objects that pass the tight selection are a subset of the objects that pass the fakeable selection. As introduced in Section 5.3.1, these are referred to as "prompt" leptons. The tight selection criteria likewise aims to reject leptons that are produced in other ways (e.g. leptons arising from the decays of hadrons produced in the hadronization of a b jet), which are referred to as "nonprompt" leptons. This separation of prompt from nonprompt leptons is accomplished with a BDT that is trained by the $t\bar{t}H$ multilepton group. The BDT is described in detail in [40, 46, 47], a brief overview of which is provided below.

The training of the BDT is performed with simulated $t\bar{t}H$ and $t\bar{t}$ samples; the electrons and muons used in the training are required to pass the loose selection criteria defined in Tables 5.1 and 5.2. A separate BDT is trained for electrons and for muons. For electrons, MC samples with detector conditions corresponding to each each ultra-legacy (UL) period (UL16APV, UL16, UL17, and UL18) were used. For muons, the training from Ref. [47] (which was performed with end-of-year samples) was used, as there was no improvement in performance observed with the ultra-legacy datasets. The variables used in the training include the p_T of the lepton, the η of the lepton, the impact parameters, the isolation of the lepton with respect to other charged particles, and the output of the BDT trained by the CMS EGamma POG.

The BDTs for electrons and for muons are referred to as the "prompt-e MVA" and "prompt-μ MVA", respectively.

TABLE 5.1

MUON OBJECT SELECTION REQUIREMENTS SUMMARY

Observable	Loose	Fakeable	Tight
p_T	> 5 GeV	> 10 GeV[4]	> 10 GeV
$\|\eta\|$	< 2.4	< 2.4	< 2.4
$\|d_{xy}\|$	< 0.05 cm	< 0.05 cm	< 0.05 cm
$\|d_z\|$	< 0.1 cm	< 0.1 cm	< 0.1 cm
d/σ_d	< 8	< 8	< 8
I_μ	$< 0.4 \times p_T$	$< 0.4 \times p_T$	$< 0.4 \times p_T$
PF muon	>WP-loose[1]	>WP-loose[1]	>WP-medium[1]
Deep Jet of nearby jet	–	<WP-interp. (<WP-medium)[2]	<WP-medium[2]
Jet relative isolation[3]	–	<0.5 (–) †	–
Prompt-μ MVA	–	< 0.85 (> 0.85)	> 0.85

[1] WPs as defined by Muon POG (see Section 5.3.2).
[2] Upper cut on the Deep Jet score defined with a linear interpolation from Deep Jet WP-medium at cone-p_T 20 GeV to Deep Jet WP-loose at cone-p_T 45 GeV, taking the Deep Jet WPs as defined by JetMET POG (see Section 5.2).
[3] Defined as 1/`jetPtRatio`-1 if the muon is matched to a jet within $\Delta R < 0.4$ or as the PF relative isolation with ΔR=0.4 otherwise.
[4] Here cone-p_T is used.
† Fails (passes) the requirement prompt-μ MVA > 0.85.

TABLE 5.2

ELECTRON OBJECT SELECTION REQUIREMENTS SUMMARY

Observable	Loose	Fakeable	Tight
p_T	> 7 GeV	> 10 GeV[5]	> 10 GeV
$\|\eta\|$	< 2.5	< 2.5	< 2.5
$\|d_{xy}\|$	< 0.05 cm	< 0.05 cm	< 0.05 cm
$\|d_z\|$	< 0.1 cm	< 0.1 cm	< 0.1 cm
d/σ_d	< 8	< 8	< 8
I_e	$< 0.4 \times p_T$	$< 0.4 \times p_T$	$< 0.4 \times p_T$
$\sigma_{i\eta i\eta}$	–	$<$ { 0.011 / 0.030 }[1]	$<$ { 0.011 / 0.030 }[1]
H/E	–	< 0.10	< 0.10
1/E - 1/p	–	> -0.04	> -0.04
Conversion rejection	–	✓	✓
Missing hits	≤ 1	$= 0$	$= 0$
EGamma POG MVA	>WP-loose[2]	>WP-90 (>WP-loose)[2]†	>WP-loose[2]
Deep Jet of nearby jet	–	<WP-interp. (<WP-medium)[2]	<WP-medium[2]
Jet relative isolation[4]	–	< 1.0 (–) †	–
Prompt-e MVA	–	< 0.90 (> 0.90)	> 0.90

[1] Barrel / endcaps.
[2] WPs as defined by EGamma POG (see Section 5.3.2).
[3] WPs as defined by BTV POG (see Section 5.2).
[4] Defined as $1/p_T^{ratio}$-1 if the electron is matched to a jet within $\Delta R < 0.4$ or as the PF relative isolation with ΔR=0.4 otherwise.
[5] Here cone-p_T is used.
† Fails (passes) the requirement prompt-e MVA > 0.80.

5.3.3 Definitions of variables used in lepton object selection

The loose, fakeable, and tight lepton selection is summarized in Tables 5.2 and 5.1. This section will define and describe the observables that are used in each stage of the selection. Where relevant, we specify parenthetically the name of the property (in the NanoAOD) corresponding to the given variable.

- p_T (`pt`): The transverse momentum of the lepton, as defined in Section 3.2.1. For the fakeable selection, the cone-p_T is used. The cone-p_T is designed to provide a characterization of the p_T of the parton that led to the nonprompt lepton. The lepton isolation and the p_T of the nearest jet are incorporated into the cone-p_T definition, which is provided in [46]. For fakeable leptons, the cone-p_T generally exceeds the reconstructed p_T of the lepton; for leptons that pass the tight selection criteria, the p_T and cone-p_T are equal.
- η (`eta`): The lepton's pseudorapidity, as defined in Section 3.2.1.
- $|d_{xy}|$ (`dxy`): This corresponds to the lepton track's transverse impact parameter with respect to the PV.
- $|d_z|$ (`dx`): This corresponds to the lepton track's longitudinal impact parameter with respect to the PV.
- d/σ_d (`sip3d`): This refers to the signed 3-dimensional impact parameter (with respect to the PV) divided by its uncertainty.
- I_e, I_μ (`miniPFRelIso_all`): This is a measure of the isolation of the lepton (corresponding to the sum of the p_T of the objects reconstructed within a cone centered on the lepton direction, where cone size is scaled inversely with the p_T, which also helps to mitigate the effects of PU).
- $\sigma_{i\eta i\eta}$ (`sieie`): The $\sigma_{i\eta i\eta}$ of the supercluster in the ECAL, a measure of the energy distribution within the crystal cluster.
- H/E (`hoe`): A measure of the energy deposited in the HCAL to the energy deposited in the ECAL.
- 1/E - 1/p (`eInvMinusPInv`): This corresponds to the difference between the reciprocal of the electron cluster energy and the reciprocal of its track momentum.
- PF muon (`looseId`): This requires the muon to pass the loose requirements specified by the Muon POG [45].
- Conversion rejection (`convVeto`): Requires the electron's `convVeto` property to be `True`.

- Missing hits (`lostHits`): The number of missing hits in the tracker.
- Jet relative isolation (`jetRelIso`): In the case where there is a matched jet, this corresponds to the relative isolation, defined as the difference between the matched jet p_T and the lepton p_T, with respect to the lepton-p_T. In the case where there is not a matched jet, the `jetRelIso` is equal to `pfRelIso04_all`.
- DeepJet of nearby jet: The output of the DeepJet discriminant (`btagDeepFlavB`) for the nearest jet (`matched_jet`).
- EGamma POG MVA: The output of the BDT trained by the EGamma POG. As described in Section 5.3.2, this helps to distinguish real electrons from jets.
- Prompt-e MVA: The output of the prompt lepton MVAs trained by the $t\bar{t}H$ multilepton team, as discussed in Section 5.3.2.

CHAPTER 6

EVENT SELECTION

This chapter will describe the event selection for the signal regions of this analysis. The conceptual motivation for the categorization is discussed in Section 6.1. The details of the selection requirements are enumerated in Section 6.2. The optimization of the sensitivity is described in Section 6.3. A summary is provided in Section 6.4.

6.1 Event selection categorization motivation

As introduced in Section 4.2, any process that is significantly impacted by the 26 WCs listed in Table 2.2 is considered to be a signal process for this analysis. The signal processes comprise $t\bar{t}H$, $t\bar{t}l\nu$, $t\bar{t}l\bar{l}$, $tl\bar{l}q$, tHq, and $t\bar{t}t\bar{t}$.

As outlined in Chapter 1, this analysis focuses on multilepton signatures of the $t(\bar{t})X$ processes; events with two same-sign leptons are categorized as 2ℓss, events with three leptons are categorized as 3ℓ, and events with four or more leptons are categorized as 4ℓ. All events are also required to contain jets, with one or more of them b-tagged. The events are further subdivided based on b-tag multiplicity, jet multiplicity, the lepton charge sum, and whether or not there is a same-flavor-opposite-sign pair of leptons with an invariant mass close to the Z mass. Aiming to isolate subsamples of events with distinct admixtures of each contribution, these subdivisions improve the sensitivity of the analysis by allowing the effects of the WCs (which impact each signal process differently) to be distinguished more distinctly.

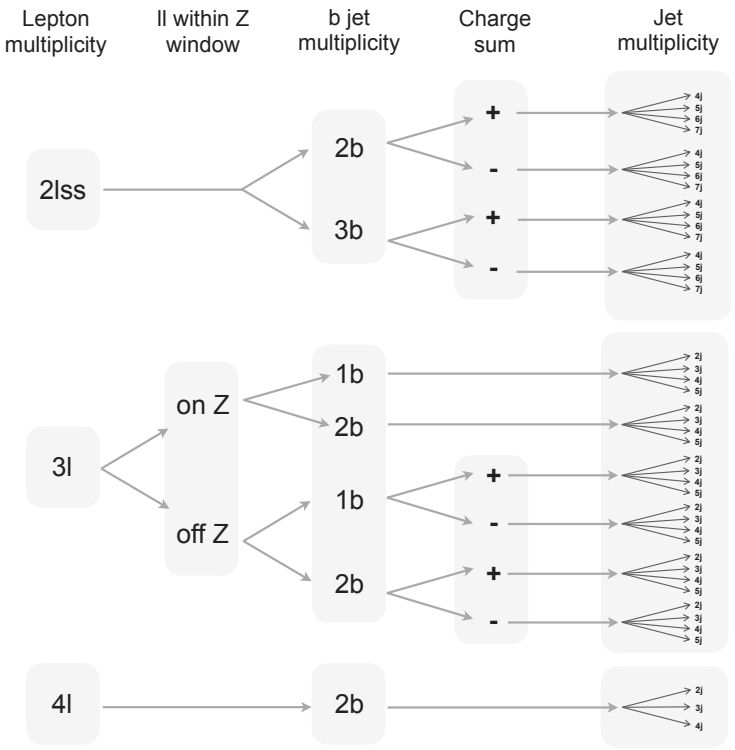


Figure 6.1. Summary of the event selection subdivisions. The details for the selection requirements are listed in Sections 6.2.1, 6.2.2, and 6.2.3.

The categorization scheme for this analysis is summarized in Figure 6.1, and the motivation for each subdivision is described below:

- Same-flavor-opposite-sign pair with an invariant mass close to the Z: This categorization helps to distinguish $t\bar{t}l\bar{l}$ and $tl\bar{l}q$ from the other processes that do not involve a Z. This distinction is also important for the identification of effects from the 2-quark-2-lepton WCs. These WCs are associated with vertices that directly produce a same-flavor-opposite-sign pair of leptons (without an intermediate Z), so the 3ℓ off-Z categories provide important sensitivity to these effects. The on-Z vs off-Z distinction is not applied for 2ℓss because $t\bar{t}l\bar{l}$ and $tl\bar{l}q$ do not naturally lead to 2ℓss final states.
- Multiplicity of b-tags: In the 3ℓ category, this distinction helps to separate single top processes from $t\bar{t}X$ processes. In the 2ℓss category, we use a high b-tag multiplicity selection to isolate a subsample that is enriched in $t\bar{t}t\bar{t}$.

- Charge sum of leptons: In categories that are naturally populated by $t\bar{t}l\nu$ (2ℓss and 3ℓ off-Z), we distinguish events with a positive charge sum from events with a negative charge sum. This helps to distinguish $t\bar{t}l\nu$ from $t\bar{t}H$ (since the LHC is a pp collider, $t\bar{t}l^{+}\nu$ will have a larger cross section than $t\bar{t}l^{-}\nu$, so $t\bar{t}l\nu$ will contribute more significantly to the + categories while $t\bar{t}H$ will populate the + and - categories symmetrically).
- Multiplicity of jets: This helps to distinguish processes which tend to produce more jets (e.g. $t\bar{t}H$ and $t\bar{t}t\bar{t}$) from processes which tend to produce fewer jets (e.g. $t\bar{t}l\nu$, $t\bar{t}l\bar{l}$, and the single top processes).

6.2 Event selection category requirements

The details of the event selection requirements for the 2ℓss, 3ℓ, and 4ℓ categories will be enumerated in Sections 6.2.1, 6.2.2, and 6.2.3. In addition to the category-specific requirements, some requirements are applied to all categories:

- To remove background contributions from light resonances, events that contain a pair of leptons (passing the loose lepton requirements outlined in Section 5.3) with an invariant mass less than or equal to 12 GeV are vetoed.
- Events with anomalously large $p_{\mathrm{T}}^{\mathrm{miss}}$ (caused by e.g. detector noise) are removed with the CMS MET group $p_{\mathrm{T}}^{\mathrm{miss}}$ Filters [48].
- To ensure that electrons are well measured, some requirements are implemented on top of the requirements outlined in Chapter 5; the number of missing hits (the `lostHits` property) is required to be 0, and and the electrons are required to pass the conversion veto (the `convVeto` property must be `True`).
- For simulated events, the leptons in the event are further required to pass MC truth requirements to ensure that they are prompt (the `genPartFlav` property is required to be 1 or 15). This ensures that we do not include any MC contributions in the nonprompt estimation (which is estimated with a data-driven approach, as described in Chapter 8).

6.2.1 The 2ℓss category

The 2ℓss category requires at least two leptons to pass the fakeable lepton object selection requirements defined in Section 5.3. Ordered by cone-p_T, the leading two leptons must also pass the tight selection requirements defined in Section 5.3, and

these must be the only leptons in the event that pass the tight requirements (i.e. there must not be more than two tight leptons in the event). The leptons are required to have the same charge. The p_T of the leading lepton is required to be greater than 25 GeV, and the p_T of the second lepton is required to be greater than 15 GeV.

For both electrons and muons, additional requirements are implemented to ensure that the charges are well measured. This helps to reduce the charge flip contribution. These requirements are applied on top of the object selection requirements outlined in Chapter 5. For muons, the `tightCharge` property is required to be greater than or equal to 1 (this requires that the ratio of the uncertainty on the p_T to the p_T is less than 0.2). For electrons, the `tightCharge` property is required to be greater than or equal to 2 (this requires that the multiple methods of calculating the sign of the electron charge [49] yield consistent results). Furthermore, events where the two tight leptons are electrons with an invariant mass within 10 GeV of the Z mass are vetoed; this also helps to reduce the contribution from charge flip events. For simulated samples, electrons are required to pass MC truth requirements to ensure that the charge has not been mismeasured (the electron's `matched_gen gen_pdgId` is required to have the same sign as the electron's `pdgId`); this ensures that MC does not contribute to the charge flip background (which is estimated with a data-driven approach, as described in Chapter 8).

At least four jets (with $p_T > 30$ GeV and $|\eta| < 2.4$) are required. Of these jets, at least two must pass the loose working point for the DeepJet algorithm, and at least one of these must also pass the medium working point for the DeepJet algorithm; i.e. there must be at least two loose b-tagged jets, at least one of which must also be a medium b-tagged jet. A subcategory is defined for events with at least three medium b-tagged jets; this allows us to isolate a collection of events that is relatively enriched in $t\bar{t}t\bar{t}$.

6.2.2 The 3ℓ category

The 3ℓ category requires at least three leptons to pass the fakeable lepton object selection requirements outlined in Section 5.3. Ordered by cone-p_T, the leading three leptons must also pass the tight requirements defined in Section 5.3, and these must be the only leptons in the event that pass the tight requirements (i.e. there must not be more than three tight leptons in the event). The p_T of the leading lepton is required to be greater than 25 GeV, and the p_T of the second lepton is required to be greater than 15 GeV. If the third lepton is an electron, $p_T > 15$GeV is required; if the third lepton is a muon, $p_T > 10$ GeV is required.

At least two jets (with $p_T > 30$ GeV and $|\eta| < 2.4$) are required. Of these jets, at least one must pass the medium working point for the DeepJet algorithm; i.e. there must be at least one medium b-tagged jet. The events are subdivided based on whether there is exactly medium b-tagged jet, or more than one medium b-tagged jet. This helps to distinguish between the single top processes and the $t\bar{t}X$ processes.

6.2.3 The 4ℓ category

The 4ℓ category requires at least four leptons to pass the fakeable lepton object selection requirements defined in Section 5.3. Ordered by cone-p_T, the leading four leptons must also pass the tight requirements defined in Section 5.3. The p_T of the leading lepton is required to be greater than 25 GeV, and the p_T of the second lepton is required to be greater than 15 GeV. For the trailing leptons, the requirements are $p_T > 15$ GeV for electrons and $p_T > 10$ GeV for muons.

At least four jets (with $p_T > 30$ GeV and $|\eta| < 2.4$) are required. Of these jets, at least two must pass the loose working point for the DeepJet algorithm, and at least one of these must also pass the medium working point for the DeepJet algorithm; i.e. there must be at least two loose b-tagged jets, at least one of which must also be a medium b-tagged jet.

6.3 Optimization studies

Based on the multiplicity of leptons, multiplicity of b-jets, sum of lepton charges, and the invariant mass of dilepton pairs, the binning described in Sections 6.2.1, 6.2.2, and 6.2.3 results in 11 independent categories. Further subdividing the categories by the jet multiplicity leads to 43 independent categories. We refer to this binning as inclusive N_{jet} binning, where "inclusive" refers to the fact that N_{jet} is the finest subcategorization. The inclusive N_{jet} binning provides good sensitivity to many of the WCs studied in this analysis, and is similar to the categories used in Ref. [4] (the predecessor to this analysis, which made use of only 2017 data). However, this analysis makes use of more than three times the data that was available for Ref. [4], and the additional statistics allow a more differential approach to be applied. In order to gain additional sensitivity to EFT effects, we bin the events in each of the 43 categories according to a kinematical variable.

In principle, a kinematic distribution could be fit for each category, resulting in $43 \cdot n$ total bins, where n is the number of bins in each kinematic distribution, assuming n is the same for all categories. However, it is not necessary to use the same kinematic binning in every category, as the binning may be adjusted to account for varying statistics. In the limit where a single bin is used for the differential variable, the inclusive N_{jet} binning would be recovered for the given category. For the categories defined in this analysis, we have found that using 4 or 5 differential bins provides a good increase in sensitivity while maintaining reasonable statistics in each category.

While the same kinematical variable may be used across all categories, it would also be possible to use a different variable in every category; any case in between these two extremes may also be implemented. However, it is important to keep in mind that the WCs cannot be fully isolated or associated with a single category, so it is not possible to choose a particular variable for each WC. Nevertheless, since some

WCs impact certain categories more strongly than others, it is possible to target WCs by choosing specific variables for categories that may be particularly sensitive to the given WCs. The goal is thus to find a variable that provides good sensitivity to all WCs, or to find a combination of different variables (to use in different categories) that may improve the sensitivity to the target WCs without degrading the sensitivity to other WCs. To assess the sensitivity provided by a given variable, we compare the limits obtained from the differential fit against the limits obtained from the more inclusive N_{jet} fits. For these optimization studies, Asimov data (i.e. simulated data that is equal to the SM prediction) was used; signals, backgrounds, and systematic uncertainties were included in the fit.

Since the contributions of many EFT vertices scale with energy, a variable that is related to the center of mass energy of the collision may provide generally good sensitivity for many WCs. For this reason, we investigated several variables related to the overall energy of the event. For example, we studied the S_T variable, which is defined as the scalar sum of the p_T of all of the leptons and jets in the event. It is also interesting to consider variables associated with the highest p_T object in an event, as it is possible that these objects could be associated with the EFT vertex in the process. For example, we can consider the leading lepton p_T, or the p_T of the leading object (lepton or jet). It is also interesting to consider variables that combine multiple high-p_T objects. For example, we may consider the p_T of the leading pair of objects from the collection of leptons and jets in the event, a variable we refer to as $p_T(lj)_0$. Testing our sensitivity to these and other similar variables, we found an improvement of approximately 50% for most of our WCs (compared to the more inclusive approach of fitting to the N_{jet} distributions).

For the on-Z categories, it is also interesting to consider the p_T of the same flavor opposite sign lepton pair. Using this variable (referred to as $p_T(Z)$) for all on-Z categories while using one of the other variables (e.g. $p_T(lj)_0$) for all other categories,

we observe improvements for some WCs (most notably WCs from among the "two-heavy" category); however, we observed a significant decrease in sensitivity to c_{Qq}^{31} and c_{Qq}^{38}. Members of the "two-light-two-heavy" category of WCs, these two WCs are unique in that they give rise to q-q'-t-b vertices. These WCs can thus contribute to 3ℓ on-Z 2b 2j (and 3j) final states. In cases where c_{Qq}^{31} and c_{Qq}^{38} contribute to these signatures, the Z boson is not part of the EFT vertex, so it is not optimal to use $p_T(Z)$ as the differential variable, thus explaining the loss in sensitivity to these WCs when using $p_T(Z)$ for all onZ categories. To mitigate this effect, we tested the scenario where $p_T(Z)$ was used for all on-Z categories except the on-Z 2b 2j/3j final states. This modification indeed mitigated the degradation of the sensitivity to c_{Qq}^{31} and c_{Qq}^{38}, while still providing improvements in sensitivity for the "two-heavy" WCs. Figure 6.2 summarizes the sensitivity observed for several of the differential variables considered during the optimization studies.

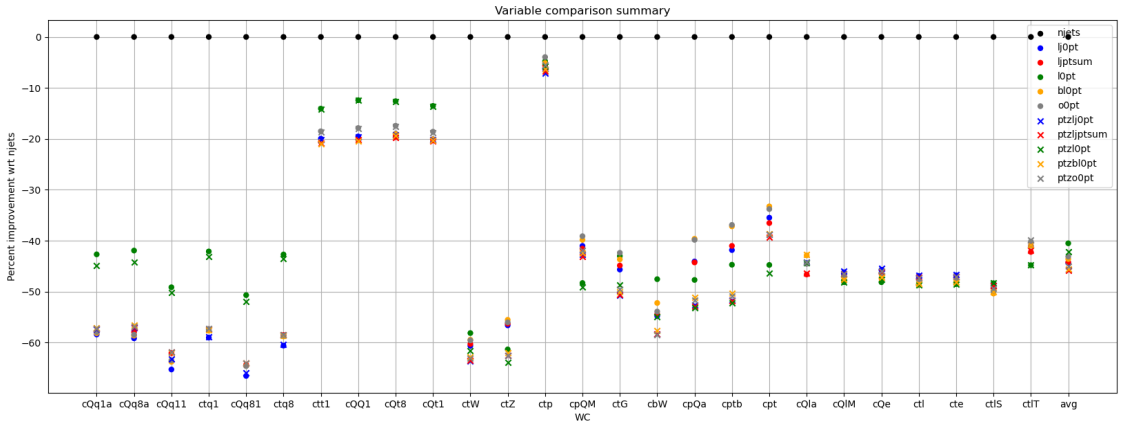


Figure 6.2. Summary of the sensitivity provided by fits to various differential distributions. The y axis represents the percent improvement with respect to the inclusive N_{jet} fit (based on the widths of the 2σ confidence intervals from fits to Asimov data).

Most of the differential variables in Figure 6.2 performed similarly to each other, providing a significant improvement in sensitivity compared to the inclusive N_{jet} fit. In general, the fits in which we use $p_T(Z)$ for the selected on-Z categories (indicated with "x" shaped markers in Figure 6.2) provide better sensitivity than the cases where the same variable is used in all categories (shown with circular markers in Figure 6.2). The case where $p_T(Z)$ is used for the selected on-Z categories and $p_T(lj)_0$ is used for all other categories (i.e. "ptzlj0pt" in the plot, denoted with blue "x" shaped markers) shows consistently good sensitivity across all categories, as well as providing the best sensitivity to the "two-light-two-heavy" WCs. For these reasons, we choose to use the $p_T(Z)$-$p_T(lj)_0$ distributions as the kinematic distributions for this analysis. For the fit, we use 4 bins in $p_T(lj)_0$ and 5 bins in $p_T(Z)$, resulting in 178 total bins.

6.4 Event selection summary

Targeting the multilepton signatures of t($\bar{\text{t}}$)X processes, the event selection categories in this analysis constitute 2ℓss, 3ℓ, and 4ℓ. The events are further subdivided into 43 unique categories designed to differentiate as much as possible between the different t($\bar{\text{t}}$)X contributions. To gain additional sensitivity, the events in each of the 43 categories are binned according to a differential kinematical distribution, resulting in 178 total bins. The $p_T(Z)$ variable is used for all of the on-shell Z categories, except for the 2 and 3 jet categories with 2 b-tagged jets; the $p_T(Z)$ variable is thus used in 6 total categories. In the remaining 37 categories, the $p_T(lj)_0$ variable is used. Binning the 43 analysis categories in terms of the $p_T(lj)_0$ and $p_T(Z)$ variables provides an improvement in sensitivity of a factor of about 2 (compared to the case where the 43 analysis bins are not further subdivided). Table 6.1 summarizes the selection requirements for each of the 43 categories in this analysis. Requirements separated by commas indicate a division into subcategories. The differential kinematical variable that is used in the category is also listed.

TABLE 6.1

SUMMARY OF EVENT SELECTION CATEGORIES.

Category	Leptons	$m_{\ell\ell}$	b-tags	Lepton charge sum	Jets	Differential variable
2ℓss 2b	2	No requirement	2	>0, <0	4,5,6,$\geq$7	$p_T(lj)_0$
2ℓss 3b	2	No requirement	≥ 3	>0, <0	4,5,6,$\geq$7	$p_T(lj)_0$
3ℓ off-Z 1b	3	$\lvert m_Z - m_{\ell\ell}\rvert > 10$GeV	1	>0, <0	2,3,4,$\geq$5	$p_T(lj)_0$
3ℓ off-Z 2b	3	$\lvert m_Z - m_{\ell\ell}\rvert > 10$GeV	≥ 2	>0, <0	2,3,4,$\geq$5	$p_T(lj)_0$
3ℓ on-Z 1b	3	$\lvert m_Z - m_{\ell\ell}\rvert \leq 10$GeV	1	No requirement	2,3,4,$\geq$5	$p_T(Z)$
3ℓ on-Z 2b	3	$\lvert m_Z - m_{\ell\ell}\rvert \leq 10$GeV	≥ 2	No requirement	2,3	$p_T(lj)_0$
3ℓ on-Z 2b	3	$\lvert m_Z - m_{\ell\ell}\rvert \leq 10$GeV	≥ 2	No requirement	4,$\geq$5	$p_T(Z)$
4ℓ	$\geq$4	No requirement	≥ 2	No requirement	2,3,$\geq$4	$p_T(lj)_0$

Applying this selection to the data and simulated samples described in Chapter 4, Table 6.2 shows the resulting event yield in each category (summed over jet bins) for the data and for the SM prediction. The observed event yields are generally larger than the predicted event yields across all of the categories; overall, the observed yield (3927 events) is about 14% higher than the prediction (3440.0 events). However, it should be noted that there are significant systematic effects (described in Chapter 9) that can influence many bins in a correlated way. For this reason, the agreement between the prediction and the observation should not be judged until after a likelihood fit incorporating the systematic uncertainties (as described in Chapter 10) has been performed.

TABLE 6.2

EXPECTED SM YIELDS AND OBSERVATIONS IN THE ANALYSIS CATEGORIES (SUMMED OVER JET CATEGORIES).

	2ℓss 3b -	2ℓss 3b +	2ℓss 2b -	2ℓss 2b +	3ℓ 1b -	3ℓ 1b +	3ℓ 2b -	3ℓ 2b +	3ℓ on-Z 1b	3ℓ on-Z 2b	4ℓ 2b
tWZ	0.46	0.47	6.66	6.7	4.78	4.77	1.69	1.69	63.34	20.3	2.72
Diboson	0.1	0.25	12.08	15.79	30.45	29.64	2.02	3.19	338.24	34.35	4.81
Triboson	0.04	0.07	2.16	3.15	0.95	1.33	0.1	0.17	16.02	2.61	0.45
Charge flips	1.62	1.57	17.5	17.4	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Nonprompt	6.72	8.94	112.61	120.56	56.95	56.5	11.86	10.47	55.31	10.06	0.0
Conversions	0.96	0.84	11.45	9.57	3.37	2.93	3.25	3.07	0.78	0.68	0.01
Sum bkg	9.9	12.14	162.45	173.17	96.52	95.16	18.92	18.59	473.69	68.0	7.99
$t\bar{t}l\nu$	12.54	23.7	144.18	272.76	25.69	47.38	27.26	50.67	10.18	11.42	0.03
$t\bar{t}l\bar{l}$	12.29	12.31	119.02	119.64	51.58	51.04	48.44	49.78	320.72	295.81	40.22
$t\bar{t}H$	9.5	9.48	83.3	83.48	23.51	23.24	22.92	22.71	9.62	9.69	3.5
$tl\bar{l}q$	0.47	0.87	6.51	11.75	5.46	9.54	2.4	4.23	111.51	48.11	0.01
tHq	0.12	0.23	1.42	2.61	0.47	0.83	0.35	0.61	0.35	0.23	0.03
$t\bar{t}t\bar{t}$	9.61	9.53	7.58	7.46	0.87	0.85	4.88	4.92	0.21	1.3	0.55
Sum sig	44.53	56.12	362.01	497.71	107.59	132.88	106.23	132.93	452.59	366.56	44.35
Sum expected	54 ± 6	68 ± 7	524 ± 50	671 ± 63	204 ± 23	228 ± 24	125 ± 11	152 ± 13	926 ± 132	435 ± 46	52 ± 6
Observation	71.0	68.0	608.0	781.0	233.0	270.0	148.0	158.0	1074.0	466.0	50.0

CHAPTER 7

DATA-TO-MC CORRECTIONS

The simulated MC samples that are used to estimate the signal and the background contributions are known to differ from data samples in several key ways. The differences are usually due to detector or reconstruction effects that are difficult to properly model with simulations. In order to account for these differences, corrections are applied to the simulated samples.

The corrections usually take the form of weights that are applied event-by-event to the simulated samples. These types of corrections do not change the total number of MC events that pass the selection, and they do not change which bin (kinematic or otherwise) the events fall into; rather, the weights simply affect how much each individual event counts toward the total contribution. Usually referred to as scale factors (SFs), the weights may be larger or smaller than one. A weight of one would imply no change with respect to the nominal simulation, while a weight further from one would suggest a significant adjustment with respect to the nominal weight of the event. The types of corrections that fall under this category include pileup reweighting (Section 7.1), trigger efficiency (7.2), lepton identification efficiency (7.3), b-tagging efficiency (7.4), and prefiring corrections (7.5).

Other forms of corrections involve adjusting the kinematic properties (such as the p_T) of particular objects in the event. Since the event selection (described in Chapter 6) depends on the multiplicity of the final state objects, and the object selection (described in Chapter 5) depends on the kinematic properties of the objects, these types of corrections impact not only the total number of events that pass the

event selection, but also the distribution of events across the categories and bins within the categories. In this analysis, we consider two such corrections: jet energy scale and resolution corrections (7.6) and muon scale corrections (7.7).

7.1 Pileup reweighting

The total number of proton-proton interactions per bunch crossing is referred to as the pileup (PU). In the data, the PU can be estimated based on the luminosity and the proton-proton cross section. The generation of the simulated samples incorporates PU, but there are differences between the simulated PU and the PU present in the data. To account for these differences, each simulated event is weighted by a SF. The SFs are derived from centrally produced histograms of the PU distributions for data [50] and MC [51]. The total proton-proton inelastic cross section is taken to be 69.2 mb with an uncertainty of 4.6% [50].

7.2 Trigger efficiency

The trigger efficiencies for simulated events differ from the trigger efficiencies for data. To calculate the efficiency in data, we need to study a set of events that contain leptons; however, since we are trying to calculate the efficiency of the lepton triggers in this analysis, we cannot use these triggers of interest to identify the collection of events for the study. Instead, $p_{\mathrm{T}}^{\mathrm{miss}}$ triggers are used (since $p_{\mathrm{T}}^{\mathrm{miss}}$ triggers target neutrinos, and charged leptons are often produced alongside neutrinos). The efficiency is then calculated as the number of events that pass the analysis triggers and the $p_{\mathrm{T}}^{\mathrm{miss}}$ triggers divided by the number of events that pass the $p_{\mathrm{T}}^{\mathrm{miss}}$ triggers. This efficiency is calculated for both data and MC, and the ratio of the efficiencies is taken as the SF. The measurements are performed separately for two and three lepton final states, and are measured as a function of lepton flavor and p_T. The MC samples used for the measurement are centrally produced $\mathrm{t\bar{t}}$ and $\mathrm{t\bar{t}W}$ samples (for the two lepton and

three lepton final states, respectively). However, for the three lepton categories, the statistics were very low and the SFs were very close to one, so it was decided that the SFs and corresponding SF uncertainties would not be applied for these categories.

7.3 Lepton efficiency

SFs are applied to account for differences between data and MC in lepton object reconstruction and selection. For the reconstruction of electrons and muons, the centrally derived SFs (from the CMS Egamma and Muon POGs) are used [52, 53]. The efficiencies for leptons to pass the loose and tight selection criteria used in this analysis (defined in Section 5.3) are computed privately using the tag and probe approach [54]; the full details of the calculation are provided in Appendix A of [46], a brief summary of which is provided below.

The electron and muon efficiencies are measured separately using $\mathrm{Z}/\gamma^* \to ee$ and $Z/\gamma^* \to \mu\mu$ events. The selection requires the pair of leptons to have an invariant mass near the Z mass peak. Both leptons are required to pass the loose selection. One lepton is required to pass the tight selection; this lepton corresponds to the "tag". The other lepton (which may or may not pass the tight selection) is referred to as the "probe". The events are separated based on whether the probe passes or fails the tight selection, and for each selection a fit (incorporating signal and background shapes) is performed. For data, this allows the yield for passing and failing events to be obtained by integrating over the signal shape. For MC, the yields can be read directly without a fit. With the passing and failing signal yields, the efficiency can be calculated as $N_{\mathrm{pass}}/(N_{\mathrm{pass}} + N_{\mathrm{fail}})$. The efficiency is calculated in bins of p_T and η and the measurement is performed separately for each data-taking period.

The SF for each lepton is computed as the ratio of the measured efficiency of data events to the measured efficiency of simulated events. The product of each of the per-lepton SFs is applied to each simulated events as a weight.

7.4 b-tag efficiency and mistag rate

The probability for a b jet to pass the medium or loose working point of the DeepJet b-tagging algorithm differs between data and MC. To account for these effects, we follow method 1a defined by the CMS BTV POG [55]. With this approach, events do not migrate between b-tag multiplicity bins. Rather, a per-event weight is constructed from a product of per-jet efficiencies; this weight is applied to each simulated event in order to account for the differences in b-tagging efficiency between data and MC.

The efficiency for simulated jets to pass the b-tagging working points are measured as a function of p_T and η using the selection criteria for this analysis. The centrally produced $\mathrm{t\bar{t}H}$ samples (for each data-taking period) are used for this measurement, and the resulting efficiency histograms are provided in Appendix C of Ref. [46]. From this per-jet efficiency, we can write down an expression for the per-event b-tagging efficiency based on the number of jets tagged at each working point:

$$P(\mathrm{MC}) = \prod_{i=\text{tagged } M} \epsilon_i^M \prod_{j=\text{tagged } M, \text{ not } L} (\epsilon_j^L - \epsilon_j^M) \prod_{k=\text{not tagged}} (1 - \epsilon_k^L), \tag{7.1}$$

where the ϵ^L and ϵ^M are the per jet efficiencies calculated as described above (for the loose and medium working points, respectively). We can write down a similar expression for data (using scale factors provided by the BTV POG to scale the per-jet efficiencies measured in MC to approximate the per-jet efficiencies in data [55]):

$$P(\mathrm{DATA}) = \prod_{i=\text{tagged } M} \mathrm{SF}_i^M \epsilon_i^M \prod_{j=\text{tagged } M, \text{ not } L} (\mathrm{SF}_j^L \epsilon_j^L - \mathrm{SF}_j^M \epsilon_j^M) \prod_{k=\text{not tagged}} (1 - \mathrm{SF}_k^L \epsilon_k^L), \tag{7.2}$$

where SF are the scale factors provided by the BTV POG. The ratio of the per-event data efficiency $P(\mathrm{DATA})$ to the per-event MC efficiency $P(\mathrm{MC})$ becomes the weight by which we scale each simulated event.

For the UL16 samples, an issue with the centrally provided SFs for light flavor jets has been observed. The disagreement was visible in b-tag multiplicity distributions in the DY CRs and has similarly been observed by Ref. [56]. In order to work around this issue, we follow the approach used in [56]: the UL16APV SFs are used for the light-flavor jets (u,d,s) in UL16.

7.5 ECAL and muon prefiring correction

In 2016 and 2017 there was a gradual shift in the timing calibration of the ECAL that was not corrected for in the online data taking. The underlying cause of the issue was related to radiation-induced damage to the ECAL crystals, so the crystals at high $|\eta|$ were most strongly affected. The result of the issue was that some trigger primitives (TPs) for a given bunch crossing (BX 0) were incorrectly associated with the previous bunch crossing (BX-1). Due to this incorrect assignment of energy deposits, the L1 trigger would occasionally fire on the BX-1 event (when it would not have otherwise fired). This BX-1 event would usually get rejected by the HLT (as it is unlikely to contain any physics objects of interest), but there is a more problematic consequence for the BX 0 event. The L1 trigger rules specify that no more than one event may be accepted from three consecutive bunch crossings, so if the BX-1 event is accepted, the BX 0 event is not allowed to be accepted. This issue therefore may cause some events containing potentially interesting physics to be lost.

The ECAL prefireing issue is described in more detail in Appendix A of Ref. [23]. A similar issue arises in the muon detectors (due to the timing resolution of the detectors) [57]. The effects of these issues are highly dependent on p_T and η of the objects in the event, and are not modeled in the MC simulation. Instead, the effects are accounted for via event weights that are computed centrally by CMS [57] and included in the `NanoAOD` v9 datasets.

7.6 Jet energy scale and resolution

Reconstructed jets must be corrected to account for various detector effects. The jet energy corrections [58] are performed in sequential stages. First, the contributions of PU to the jet energy and momentum are estimated and removed; this is referred to as the PU offset correction, and it is applied to both data and MC. Next, corrections are applied to the detector response to the PU-corrected simulated jets. Determined by comparing particle-level jets (i.e. jets clustered from stable, visible final-state particles) to reconstructed jets, these simulated jet response corrections are not applied to data.

After the PU and response corrections have been applied, the simulated jets are compared against jets in data, and SFs are applied to the simulated jets to account for the residual differences. The SFs depend on p_T and η and are determined with various techniques. The p_T corrections are determined with processes that do not involve $p_{\mathrm{T}}^{\mathrm{miss}}$ (e.g. $\mathrm{Z}(\rightarrow \ell\ell)$+jet events) using a momentum-balance approach. By comparing the momentum of the jet to the momentum of a reference object (e.g the lepton pair from the Z), residual issues in the jet p_T may be identified. Dijet events are used to determine residual η dependent corrections via a tag and probe approach (in which a jet in the barrel region serves as the tag).

The corrections described above are included in the `NanoAOD` data tier used in this analysis. After the jets have been corrected, there remain residual differences between the p_T resolution of jets in MC and data. To account for these differences, the p_T of the simulated jets is smeared such that the resolution matches the data. The hybrid method defined by the CMS JERC subgroup [59] is used to perform this smearing. When a particle-level jet is present, the jet's p_T is scaled using a factor constructed from reconstructed p_T, the particle-level p_T, and a data-to-simulation resolution SF; otherwise, a stochastic smearing of the corrected jet's p_T is applied.

7.7 Muon p_T correction

Differences between the simulated p_T of muons and the p_T of muons in observed data lead to a visible shift in the dimuon invariant mass peak [60]. To correct for this mismodeling, the p_T of muons in simulated events is adjusted according to the so called "Rochester correction" procedure [61]. For this analysis, the up and down variations of the Rochester correction are found to have a negligible effect (the up/down variations differed from the nominal by $< 0.05\%$), so a systematic uncertainty is not included for this correction.

CHAPTER 8

BACKGROUNDS

This chapter describes the backgrounds for this analysis. Backgrounds are defined as contributions that populate the signal regions (SRs) described in Chapter 6 but are not significantly impacted by the WCs enumerated in Table 2.2. The backgrounds are categorized as either irreducible or reducible.

A background is categorized as irreducible if all of the final-state leptons are prompt. The dominant irreducible backgrounds are diboson processes, but smaller contributions also arise from triboson processes and the tWZ process. It is interesting to note that the tWZ process in principle could be affected by some of the WCs included in this analysis. However, preliminary studies indicated that the effect is not large, so this process is not included as a signal for this analysis. It may be interesting to revisit this process in the future and explore if it would be possible to improve the sensitivity to this and other rare processes. The irreducible backgrounds are modeled using MC simulation. The datasets (produced centrally by CMS) are listed in Appendix A, in Tables A.11, A.12, A.13, and A.14.

Reducible backgrounds result from the misreconstruction or misidentification of objects. The primary source of reducible backgrounds arises when objects that are not genuine prompt leptons pass the tight selection criteria defined in Section 5.3. This contribution is referred to as the "nonprompt" background, and it is estimated with a data-driven technique, which will be described in Section 8.1. In the 2ℓss category, there is also a contribution from two-lepton-opposite-sign (2ℓos) events where the charge of one of the leptons is mismeasured. This background is referred

to as the "charge flip" background. The charge flip background is also estimated with a data-driven technique, which will be described in Section 8.2.

Though the nonprompt and charge flip backgrounds represent the largest contributions to the reducible backgrounds, we also account for an additional, smaller source of reducible background. Referred to as the "conversion" background, this backgrounds arises from $\gamma \to e^+e^-$ conversions where one of the leptons carries most of the energy of the photon (meaning the other may fail to be reconstructed). The contribution of this background is modeled with MC simulation, using the datasets listed in Appendix A. In the event selection, MC truth requirements are applied for this sample to ensure that at least one lepton is associated with a conversion (the `genPartFlav` property is required to be 22).

In order to validate the handling of the various sources of backgrounds, several control regions (CRs) are studied. To examine the nonprompt and flip backgrounds, we define a 2ℓss CR. This CR is similar to the 2ℓss SR, except that it requires exactly one medium b-tagged jet and fewer jets than the 2ℓss SR (which guarantees that this region will not overlap with the SR). This CR is dominated by the nonprompt and flip backgrounds, and plots for various kinematic distributions in this CR are shown in Appendix E. A dedicated CR for the charge flips is also defined (described in Section 8.2). In order to study the diboson background, a 3ℓ CR is defined. This CR is similar to the 3ℓ SR, except that we require exactly zero medium b tags (to guarantee that there is no overlap with the SR). This CR is dominated by diboson events, primarily WZ, and plots for various kinematic distributions in this CR are also shown in Appendix E. We additionally define a 2ℓos CR that is dominated by Drell-Yan (DY) events, as well as a 2ℓos CR that is dominated by $\mathrm{t\bar{t}}$ events. While these processes do not represent significant backgrounds for our SRs, it is useful to study these relatively pure CRs as a cross check of the data-to-MC corrections. Plots for various kinematic distributions in these CRs are also shown in Appendix E.

8.1 Nonprompt background

The nonprompt background arises when leptons that are not prompt pass the tight selection criteria defined in Section 5.3. As discussed in Section 5.3, a prompt lepton is a lepton that is produced promptly at the PV (e.g. from the decay of W boson in a $t\bar{t}W$ event); a nonprompt lepton is a lepton that is produced at a later point that is removed from the PV (e.g. in the decay of a hadron associated with a jet arising from the hadronization of a b quark). The procedure of estimating the nonprompt contribution is developed by the $t\bar{t}H$ multilepton team, and the measured probabilities are shared between the $t\bar{t}H$ analysis [40] and the analysis described in this thesis [62], which also share a synchronized object selection. The nonprompt estimation involves two main steps: the measurement of the probability for nonprompt leptons to pass the tight selection, and the application of these probabilities to a set of events in a sideband of the signal region in order to estimate the contribution in the signal region.

To measure the probability for a nonprompt lepton to pass the tight selection, we first identify a set of events that is dominated by nonprompt leptons. Referred to as the measurement region (MR), the data for this sample is collected with a set of single lepton nonisolated triggers (listed in [46]). The selected events are required to contain exactly one lepton that passes the fakeable selection criteria defined in Section 5.3 and at least one jet. The leptons in this collection are then subdivided based on whether the fakeable lepton passes or fails the tight selection criteria. If this sample were composed entirely of nonprompt leptons, the probability f for a nonprompt lepton to pass the tight selection would be simply

$$f = N_{\text{pass}}/(N_{\text{pass}} + N_{\text{fail}}), \tag{8.1}$$

where N_{pass} and N_{fail} are the number of events in which the fakeable lepton passes

or fails the tight selection, respectively. However, the MR also contains a small contribution from processes that produce genuine prompt leptons (e.g. W+jets). In order to account for this contamination, the sample is binned according to a variable (referred to as m_T^{fix} and defined in [46]) that is designed to discriminate between the processes of interest (the multijet contribution) and the background processes (e.g. W+jets). A fit is performed in different p_T and η regions in order to extract the N_{pass} and N_{fail} for each region. We can then use Eq. 8.1 to obtain the probability f for a nonprompt lepton to pass the tight selection.

Once the probabilities f have been measured, the next step is to use these probabilities to estimate the nonprompt contribution to the signal region. In order to obtain this estimation, the probabilities are applied to events in a sideband of the signal region, which is referred to as the application region (AR). Orthogonal to the signal region, the AR requirements are identical to signal region categories (defined in Chapter 6) except that at least one of the leptons is required to fail the tight requirements. From the number of events observed in the AR and the measured probability f for a nonprompt lepton to pass the tight selection, we can work backwards to obtain the estimation for the contribution in the signal region.

For example, let us consider the two-lepton case. The total number of events that make it into the signal region (i.e. have two tight leptons) where at least one of the leptons is nonprompt can be written as

$$N_{\text{SR}} = fN_{\text{1np}} + f^2 N_{\text{2np}}, \tag{8.2}$$

where N_{1np} is the number of events with exactly one nonprompt lepton and one prompt lepton, N_{2np} is the number of events with exactly two nonprompt leptons, and f is the probability for a nonprompt lepton to pass the tight selection. Here it is assumed that the probability for a true prompt lepton to pass the tight selection is

1. We do not know $N_{1\text{np}}$ or $N_{2\text{np}}$, but we do know the number of events with exactly two fakeable leptons where one of the leptons passes the tight selection (which we can call N_{tf}) and the number of events with exactly two fakeable leptons where neither pass the tight selection (which we can call N_{ff}), as these are the events we observe in the AR. We can then write N_{tf} and N_{ff} in terms of $N_{1\text{np}}$, $N_{2\text{np}}$, and f (which will allow us to solve for $N_{1\text{np}}$ and $N_{2\text{np}}$):

$$N_{\text{tf}} = (1-f)N_{1\text{np}} + 2f(1-f)N_{2\text{np}} \tag{8.3}$$

$$N_{\text{ff}} = (1-f)(1-f)N_{2\text{np}} \tag{8.4}$$

Solving Eq. 8.4 for $N_{2\text{np}}$, we can plug the result into Eq. 8.3 and solve for $N_{1\text{np}}$. Since we now have expressions for $N_{2\text{np}}$ and $N_{1\text{np}}$ in terms of known quantities (N_{tf}, N_{ff}, and f), we can plug these expressions into Eq. 8.2 in order to express N_{SR} in terms of known quantities:

$$\begin{aligned} N_{\text{SR}} &= f\left(\frac{N_{\text{tf}}}{1-f} - \frac{2fN_{\text{ff}}}{(1-f)^2}\right) + f^2\left(\frac{N_{\text{ff}}}{(1-f)^2}\right) \\ &= \frac{f}{1-f}\,N_{\text{tf}} - \left(\frac{f}{1-f}\right)^2 N_{\text{ff}}, \end{aligned} \tag{8.5}$$

where the quantity $\frac{f}{1-f}$ is usually referred to as F. We have thus written the estimation for the nonprompt contribution to the signal region in terms of the probability f for a nonprompt lepton to pass the tight selection (which is measured in the MR as described above) and the number of data events observed in the AR. In this example calculation, we have made the simplifying assumption that all leptons have the same f, and have only considered the two-lepton case. The expressions resulting from the full calculation (for the two-lepton and three-lepton cases) are shown in Eq. 13 and 14 of [46]; the weights expressed in these equations are then applied to the events in the AR in order to estimate the nonprompt contribution to each of the SR categories.

In order to account for the fact that the probability of a prompt lepton to pass the tight selection is not actually one, the size of the contribution to the AR is estimated with MC samples and subtracted from the data contribution. As described in Section 6.2, MC truth information is utilized to require that the leptons in simulated samples are prompt; if the probability for prompt leptons to fail the tight selection were actually one, no simulated events would populate the AR. Events from these simulated samples that fall into the AR therefore represent an estimation of the contribution of events containing prompt leptons that fail the tight selection criteria. The contribution of these samples in the AR is subtracted from the total nonprompt yield, allowing us to make the assumption that this contribution is zero when working out the above combinatorics. This procedure is known as prompt subtraction.

8.2 Charge misidentification background

When the charge of one of the leptons in 2ℓos event is mismeasured, the event can enter the 2ℓss signal region, contributing to a reducible background referred to as the "charge flip" background. The charges of particles traversing through the detector are determined by the curvature of the tracks, so charge flips occur when the curvature of the track is incorrectly identified. One reason why this may occur is if a radiated photon converts into an electron-positron pair, complicating the reconstruction and possibly resulting in an incorrectly reconstructed charge. Since the charge misidentification rates for muons are negligibly small, the charge flip background is only relevant for electrons.

The charge flip probabilities are expected to be larger for electrons of higher p_T, since the tracks are straighter, making it more difficult to determine the direction of curvature. The flip probabilities are also expected to be larger in the endcap than in the barrel region since the electrons must traverse more detector material, so the probability to emit Bremsstrahlung is larger. For these reasons, we measure

the charge flip probabilities in bins of p_T and $|\eta|$, following the approach outlined in Ref. [63].

The measurement of the charge flip probability is performed with MC DY and $t\bar{t}$ samples. We count the number of electrons that pass the tight requirements (defined in section 5.3) that have had their charges mismeasured (according to the MC truth information). In addition to the tight lepton requirement, we also apply the tight charge requirement (as discussed in 6.2.1), since this requirement is applied to all electrons in our 2ℓss categories. The measurement is performed for each UL period separately, using the DY and $t\bar{t}$ samples listed in tables A.11, A.12, A.13, and A.14. The measured charge flip probabilities for each year are shown in figure 8.1.

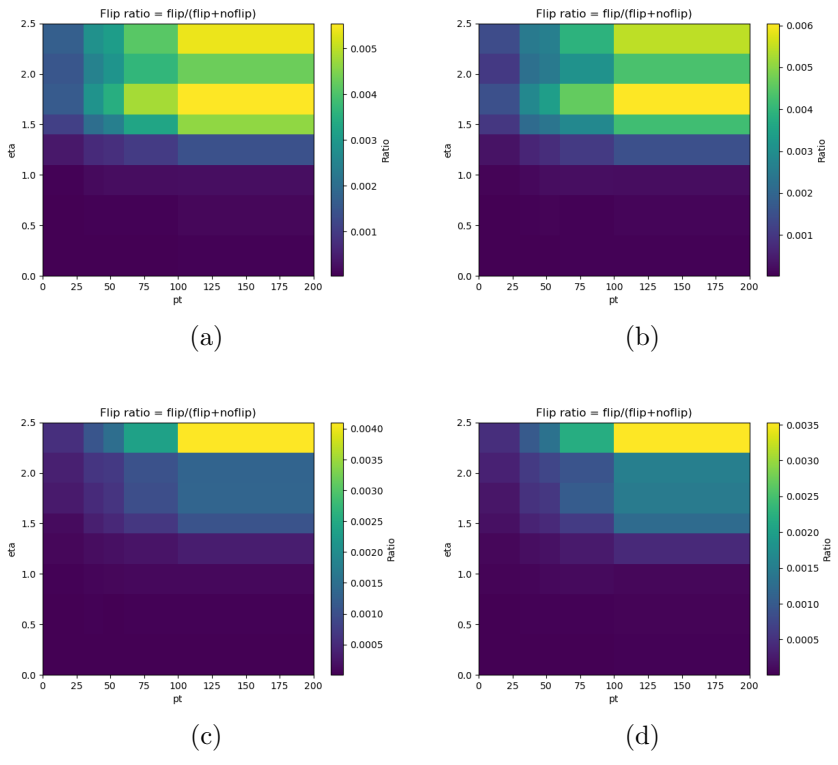


Figure 8.1. Charge flip probabilities calculated with DY and $t\bar{t}$ samples using MC truth information for UL16APV (a), UL16 (b), UL17 (c), and UL18 (d) samples. The rates are binned according to the p_T and $|\eta|$.

Once the charge flip probabilities have been measured, we assess their validity in a dedicated charge flip CR. The charge flip CR is designed to be dominated by charge flip events. We require two tight electrons within 30 GeV of the Z peak. We make no requirement on the number of b tags, but we require fewer than four jets (to maintain orthogonality to the 2ℓss SR). The charge flip contribution is determined by scaling opposite-signed events in this region by the measured charge flip probabilities. The charge flip factor for the event corresponds to the probability that the charge of the first electron is mismeasured (and that the charge of the second electron is not mismeasured), or that the charge of the second electron is mismeasured (and the charge of the first electron is not mismeasured). We therefore have the following event weight w:

$$w = p_1(1 - p_2) + p_2(1 - p_1), \tag{8.6}$$

where p_1 is the probability that the charge of the first electron is mismeasured, and p_2 is the probability that the charge of the second electron is mismeasured. Assuming that the probabilities are small enough that terms of order $p_i \cdot p_j$ may be neglected, the charge flip probability for the event becomes the following:

$$w = p_1 + p_2. \tag{8.7}$$

The charge flip probabilities p_i for electrons are taken from the measurements described above. Since the charge flip probabilities for muons are assumed to be negligible, the flip probabilities are taken to be zero for all muons. In principle, some same-sign events should also migrate into the opposite-sign categories, but since the number of opposite-sign events is much large than same-sign events, this contribution may be neglected.

The prediction we obtain from applying the charge flip rates to opposite-sign events can then be compared to the actual same-sign data in the charge flip CR. For

completeness, we include all MC backgrounds as well as the nonprompt background in the comparison. Similar to what was observed in [63], we see that our charge flip estimation over-predicts in the UL16 and UL16APV periods, and under-predicts in the UL17 and UL18 periods. To account for these differences, we apply per-year scaling factors to the charge flip probabilities (as was similarly implemented in [63]). The scaling factors are shown in Table 8.1:

TABLE 8.1

CHARGE FLIP SCALING FACTORS.

Year	Scaling factor
UL16APV	0.79
UL16	0.81
UL17	1.22
UL18	1.12

Plots from the charge flip CRs (after applying the scaling factors in Table 8.1) are shown in Appendix E in Figures E.18, E.17, E.19, and E.20. To account for the uncertainty in the measurement of the charge flip contribution, we apply a 30% flat rate uncertainty on the charge flip contribution.

CHAPTER 9

SYSTEMATICS

This chapter describes the systematic uncertainties included in the analysis. As will be discussed in Chapter 10, the systematics are handled as nuisance parameters in the likelihood fit. Systematics can affect either the normalization of the contributions (i.e. a flat scaling of all of the bins in a histogram) or both the normalization and shape of the contributions (i.e. each bin in the histogram may be affected differently). As introduced in Section 10.2, the `Combine` tool is used to perform the statistical analysis; systematics that only impact the normalization are handled as "rate" systematics (i.e. numbers in the `Combine` datacard), while template histograms (for the up and down variations) are used to handle the "shape" systematics.

A discussion of the EFT dependence of the shape systematics is provided in Section 9.1. The full set of systematic uncertainties included in this analysis are listed in the subsequent sections. In Section 9.2 the systematics arising from experimental sources are discussed; these include uncertainties on the data-to-MC corrections described in Chapter 7, uncertainties on the background estimation, and various other sources. The systematics associated with the theoretical aspects and modeling are listed in Section 9.3.

9.1 EFT dependence of the systematic uncertainties

Like the histograms that correspond to the nominal contribution, the histograms that correspond to the up and down variations of the shape systematics will carry 26-dimensional quadratic dependence on the WCs. However, it should be noted that

the quadratic dependence of the up and down histograms may be different from the nominal histogram. To understand why this is true, let us consider an "up" variation of a SF. We recall from Chapter 4 that each generated event is characterized by a unique 26-dimensional quadratic function. In the case of a nominal SF, the function for a given event will be scaled by the nominal SF value for that event; in the case of an "up" variation, the function will be scaled by the "up" value of that SF for that event. This will be similarly true for each of the subsequent events, and in general the difference between nominal and "up" will vary from event to event. For a given bin, the quadratic parametrization corresponds to the sum of the quadratics of each of the events that pass the selection criteria for the bin. Since each of the quadratics in the nominal sum will differ from each of the quadratics in the "up" sum, the nominal and "up" quadratic parameterization will in general have different shapes.

For the systematic uncertainties associated with the generation of the MC samples (described in Section 9.3 below), there is an additional complication to consider. The up and down weights for these systematics are calculated by MadGraph at the starting point of the sample. In principle, there is no reason to assume that the up and down weights at any other point in the WC space must be the same as at the starting point. For this reason, it would be most correct to generate a dedicated sample at every point in space that we are interested in (or to modify MadGraph to calculate the up and down weights at each reweight point). These approaches are currently infeasible at scale, but small-scale studies suggest that the dependence of the systematics on the WC space is not large [10]. Therefore, this effect is neglected in this analysis.

A more detailed discussion of the EFT dependence of the systematic uncertainties is provided in Appendix F; this appendix includes a discussion of how the approach implemented in this analysis differs from the approach utilized in the predecessor to this analysis (Ref. [4]), and how both of these approaches differ from the fundamentally correct approach.

Luminosity: The total uncertainty on the total luminosity for the 2016, 2017, and 2018 data-taking periods is 1.6% [64–66]. This uncertainty affects only the rate (not the shape). The systematic is correlated across years; this means that if the luminosity is high or low by a certain amount in one data-taking year, it is assumed to be high or low by the same amount for all data-taking years.

PU correction: The uncertainty for the PU correction (Section 7.1) is obtained by varying the proton-proton cross section used to estimate the data PU histograms by 4.6% (which corresponds to a 1σ variation). This systematic is correlated across years.

Trigger efficiency correction: The uncertainty on the trigger efficiency corrections (Section 7.2) is taken to be 2%, which is a conservative estimate that includes the effects of the dependence on the phase space used to perform the measurement and the effects of the correlation between the $p_{\mathrm{T}}^{\mathrm{miss}}$ triggers and the analysis triggers. This systematic is uncorrelated across years.

Lepton identification efficiency: As described in Section 7.3 lepton identification efficiencies are computed with the tag and probe approach. There are several sources of uncertainty that contribute to this measurement, including the statistical uncertainty, the uncertainty on the signal modeling (which is estimated by comparing the results obtained with LO and NLO DY samples) and the uncertainty on the functions used in the fitting (which is estimated by performing the fits with alternative models). The total uncertainty for this correction is the quadrature sum of these sources. There is a separate uncertainty for electrons and for muons; the systematics are correlated across years. The full details regarding the systematic uncertainties on the lepton SFs are available in Appendix A of [46].

b-tagging correction: The per-jet b-tagging SFs (Section 7.4) are different for heavy (b,c) and light jets, so a separate systematic uncertainty is included for each. For both heavy and light uncertainties, a component that is correlated and uncorrelated (across years) is considered, bringing the total number of b-tagging uncertainties to 10.

Prefire correction: The uncertainty on the L1 prefire correction (Section 7.5) takes into account the uncertainties on the prefire probabilities (20%) and the statistical uncertainty of the given bin. The uncertainty is correlated across years.

Jet energy corrections: There are numerous uncertainties associated with the corrections to the jet energy scale described in Section 7.6. These uncertainties can be grouped into categories defined by the CMS JERC group [67]. Of these groupings of uncertainty sources, we include the `Absolute` uncertainty (a combination of sources from the PU offset correction and simulated response

correction), the `BBEC1` uncertainty (a combination of sources from the PU offset correction, relative p_T differences, relative resolution differences, and statistical uncertainty), the `RelativeBal` uncertainty (which accounts for differences in different methods of the p_T balance calculations), the `RelativeSample` uncertainty (which accounts for differences in the residual data to simulation differences obtained with dijet and Z+jet approaches), and the `FlavorQCD` uncertainty (which accounts for differences in responses to different jet flavors). We treat these sources of uncertainty as correlated across all data-taking periods. In total, we thus include five nuisance parameters for the jet energy scale corrections.

The jet energy resolution (JER) uncertainty is obtained by shifting the resolution applied and is uncorrelated across data-taking periods, for a total of four nuisance parameters [59].

Charge misidentification: As described in Section 8.2, a flat 30% rate uncertainty is applied to the charge flip contribution, and the systematic is correlated across years.

Nonprompt estimation: Several sources of uncertainties are associated with the nonprompt estimation. The first source corresponds to the uncertainty in the measurement of the probability for nonprompt leptons to pass the tight selection. As described in Section 8.1, this measurement involves performing a fit to a discriminating variable; the fit is performed in three ways, and the envelope of the results (including the statistical uncertainties) is taken as the uncertainty. This envelope uncertainty is fully correlated across all p_T and η bins in the fit. To account for effects that are not fully correlated across these variables, we also include two additional sources of uncertainty that cover the most extreme variations across the p_T bins and across the η bins. After the measurement has been performed, the results are compared against results obtained with MC simulations, and the residual differences are covered with an additional closure uncertainty (a separate closure uncertainty is considered for each UL period).

In summary, the systematic uncertainty on the nonprompt estimation includes seven total nuisance parameters corresponding to the envelope uncertainty, the envelope p_T uncertainty, the envelope η uncertainty, and the four closure uncertainties. In addition to these seven systematic uncertainties, the statistical uncertainty of the AR is also included (using the `Combine` tool's `autoMCStats` functionality [68]).

Additional radiation: As described in Section 4.2.1, an additional uncertainty is applied to the single-top processes, which are not generated with an additional parton. This uncertainty is computed by comparing the N_{jet} distribution of the LO EFT samples (reweighted to the SM) against the centrally produced NLO samples listed in Table A.10. This systematic is taken to be the up/down shift required to cover the difference between the samples (such that when this uncertainty is included in quadrature with the other systematics, the difference between the samples is fully covered by the total uncertainty). The systematic is correlated across years.

Diboson N_{jet} uncertainty: This systematic is derived from the 3ℓ control region to account for the under-prediction of the MC for the high-jet-bin yields. To derive the uncertainty for each jet bin, we calculate the factor by which the diboson contribution would need to be scaled in order for the prediction to match the data; a linear function is then fit to this set of points, and we evaluate the linear function to find the uncertainty factor in each jet bin. For a given bin, the size of the shift is determined by taking the difference between the diboson contribution and the diboson contribution scaled by the uncertainty factor. The uncertainty is then applied symmetrically (i.e. the up and down variations are of equal magnitude). This systematic is applied only to the diboson process. This systematic is correlated across years.

Renormalization and factorization: The renormalization and factorization scales (μ_{R} and μ_{F}, respectively) are fluctuated up and down by a factor of 2. The weights for each variation are computed by the event generator during the production of the sample. As described in Section 9.1, these weights are calculated at the starting point of the sample.

The effects of the $\mu_{\mathrm{R}}/\mu_{\mathrm{F}}$ systematics are handled somewhat differently than the experimental systematics described above. In order to understand this difference, let us first recall Eq. 4.1. The up and down variations of the experimental SFs affect how likely an event is to pass the given selection, so these weights are only applied to the sum in the numerator of Eq. 4.1. However, the μ_{R} and μ_{F} variations are applied to both the numerator and denominator of Eq. 4.1 since these weights correspond to how often certain parts of the phase space are populated by MC events (so these variations affect all generated events, not just the ones passing the selection). Another way to think about this is to recall that we already include a systematic uncertainty on the cross section of the simulated samples; including the μ_{R} and μ_{F} variations in both the numerator and denominator effectively cancels the overall normalization effect (which is already covered by the cross section uncertainty), leaving us with the relevant shape effect of the systematics.

In the event generation, the μ_R and μ_F are independently varied (i.e. μ_R is varied up/down while μ_F is held at nominal and vice versa). These systematics are handled independently (i.e. there are separate nuisance parameters for each) and are uncorrelated across processes (i.e. there are separate nuisance parameters for each process). Each systematic is correlated across all data-taking periods.

Parton shower: The initial and final state radiation (ISR and FSR, respectively) scales are fluctuated up and down by a factor of $\sqrt{2}$. Similar to the μ_R and μ_F systematics, these variations are applied to both the numerator and denominator of Eq. 4.1. The FSR systematic is treated as correlated across all data-taking periods. The ISR systematic is also treated as correlated across all data-taking periods; however, the ISR systematic is decorrelated among processes with different initial states.

Cross section: As discussed in Section 4.2, the simulated samples are normalized to NLO or NNLO cross sections where available. The uncertainties associated with the cross section calculations are applied as rate uncertainties to the relevant processes and are correlated across data-taking periods.

An exception to the above procedure is applied for the $t\bar{t}\gamma$ samples (used to estimate the conversion contribution); these samples are normalized to LO cross sections, as there is not an appropriate NLO calculation available. In order to account for the LO cross section uncertainty, the μ_R/μ_F, ISR, and FSR systematic variations are applied only to the numerator of Eq. 4.1 for this process. This allows the uncertainty on both the shape and normalization to be incorporated.

CHAPTER 10

STATISTICAL METHODS

Once the event selection has been performed and the selected events are binned according to the differential kinematical distributions (as described in Chapter 6), the next step is to perform the statistical analysis in order to extract the confidence intervals (CIs) for the WCs. In Section 10.1, we will step through the relevant statistical concepts and tools. In Section 10.2, the details of the statistical framework will be described. Finally, in Section 10.3, one of the challenges of multi-dimensional EFT fits will be discussed, and the workaround developed to mitigate this challenge will be explained.

10.1 Likelihood fitting

Let us start by defining the likelihood L as the probability to have measured the observed number of data events, given some theory, i.e. $L = P(\text{data}|\text{theory})$. The number of observed events should follow a Poisson distribution, with a mean corresponding to the number of predicted events. Since each bin is statistically independent, the likelihood will be given by the product of the Poisson probabilities for all of the bins in the analysis.

For many analyses, the predicted number of events in a given bin i can be written as $\mu s_i + b_i$, where b_i is the expected number of background events, s_i is the expected number of signal events (according to the SM prediction), and μ is a free parameter. The μ parameter is usually referred to as the signal strength; it must be greater than or equal to 0, and it is constant across all bins. However, in this analysis, we cannot

write the predicted number of events as $\mu s_i + b_i$, because the prediction does not scale linearly with a universal signal strength μ. Rather, the predicted number of events in each bin depends quadratically on the 26 WCs, and the quadratic parameterization is different in each bin. We will write this prediction as $m(\theta)_i$, where θ are the values of the WCs. This function $m(\theta)_i$ represents the prediction in each bin (which in our case is a 26-dimensional quadratic in terms of the WCs), and should not be confused with the signal strength μ. With this notation, we can write the likelihood as follows:

$$L = \prod_{i=1}^{N} \frac{m(\theta)_i^{n_i}}{n_i!} e^{-m(\theta)_i}, \tag{10.1}$$

where N is the number of bins, n_i is the number of observed events in bin i, and $m(\theta)_i$ is the number of predicted events in bin i (as a function of the WCs θ).

In Eq. 10.1, the θ represents the set of 26 dimension-six WCs studied in this analysis. To understand the likelihood's dependence on the WCs, we could in principle perform a 26-dimensional grid scan. To perform such a scan, we would choose a reasonable range for each WC (based on the estimated sensitivity to the WC), choose a granularity with which to scan, and then proceed to record the likelihood at each point on the 26-dimensional grid. However, this "brute-force" approach scales exponentially with the number of dimensions, and becomes prohibitively expensive when more than a few dimensions are considered. Let us step through an example for 26 dimensions. Even if we chose a very sparse grid with only 5 scan points in each dimension, we would still need to scan 5^{26} points. Assuming one hour per scan point (a typical length of time for the fits in this analysis) and 10 k CPU cores (a reasonable amount of resources we could utilize with an opportunistic pool such as Notre Dame's CRC), it would take about 17 billion years to perform the scan. This brute force approach is thus not feasible for our analysis.

Instead of a 26-dimensional scan, we perform a one-dimensional scan for each WC, profiling the other 25 WCs. Continuing to refer to the scanned parameter as

θ, let us refer to the profiled parameters as ν. In the one-dimensional scan, we step along one axis θ in the 26-dimensional space (i.e. we step through a set of values for one WC). At each of the steps along the θ axis, the profiled parameters ν are set to the values that cause the likelihood to be maximized at this given value of the scanned parameter θ. The profile likelihood L_p is thus written as follows:

$$L_p(\theta) = L(\theta, \hat{\hat{\nu}}(\theta)), \tag{10.2}$$

where the double-hat notation denotes the values of the profiled parameters ν that maximize the likelihood for the given θ. The profiled likelihood is thus a function of θ only; it is not a function of ν, since the profiled parameters ν do not freely vary (they are a function of θ). As these scans are only one-dimensional, they are much less computationally expensive than the 26-dimensional scan described above. We can thus perform a one-dimensional scan for each of the WCs in order to determine the one-dimensional profile likelihood for each dimension. In principle, we can extend this concept to scan over any number of the WCs (profiling the remaining WCs); however, in practice, the largest number of parameters we can scan is two, since even for a three-dimensional scan, the space is too large to fully explore with our current computational abilities.

Next, we would like to understand how the profile ratio compares to the maximum likelihood as a function of θ. We will refer to the values of θ and ν that globally maximize the likelihood as $\hat{\theta}$ and $\hat{\nu}$, respectively. We can then write the profile likelihood ratio $\lambda_p(\theta)$ as follows:

$$\lambda_p(\theta) = \frac{L(\theta, \hat{\hat{\nu}}(\theta)}{L(\hat{\theta}, \hat{\nu})}, \tag{10.3}$$

where the numerator is the profile likelihood from Eq. 10.2, and the denominator is the value of the likelihood at its global maximum. From the profile likelihood ratio,

we can form the test statistic $-2\ln\lambda_p(\theta)$. Wilks' theorem [69] states that $-2\ln\lambda_p(\theta)$ should approach a χ^2 distribution in the limit where the data sample is large, where the degrees of freedom correspond to the number of free parameters in the $\lambda_p(\theta)$. For example, for a one-dimensional scan, there is one degree of freedom. To find the one-dimensional confidence intervals for a given WC, we would thus need to perform a scan for the WC, finding the $-2\ln\lambda_p(\theta)$ at each scan point; since the $-2\ln\lambda_p(\theta)$ is assumed to follow a χ^2 distribution with one degree of freedom, we can read off the one and two σ confidence intervals by observing where the $-2\ln\lambda_p(\theta)$ crosses one and four, respectively [70].

This section has discussed how the predicted yield in each bin depends on the parameters of interest (the WCs), but the prediction also depends on the systematic uncertainties (enumerated in Chapter 9). The systematic uncertainties are taken into account via additional free parameters in the fit; these degrees of freedom are referred to as nuisance parameters. When finding the profile likelihood $L(\theta, \hat{\hat{\nu}}(\theta))$, the nuisance parameters are profiled (i.e. they may be included in the ν in Eq. 10.3).

10.2 Statistical framework

The CMS Higgs `Combine` software tool [71] is used to perform the likelihood fits. The `Combine` tool uses the `ROOT` framework's `RooFit` tools [72] and the `MINUIT2` software library [73] to numerically minimize the negative log of the profile likelihood function described in Section 10.1. As discussed in Section 4.2.2, the expected yield in each bin is parameterized as a 26-dimensional quadratic in terms of the WCs, as given in Eq. 4.5. In order for `Combine` to calculate the profile likelihood described in Section 10.1, this quadratic dependence must be made known to `Combine`. In principle, a template histogram could be defined for each of the 378 structure constants of the 26-dimensional quadratic, with the appropriate normalizations of the templates specified by the `Combine PhysicsModel`. However, in practice this is not possible

with `Combine`, since the interference terms of the quadratic may be negative, and `Combine` does not handle histograms with negative yields.

To work around this challenge, we use the approach developed in [74]. With this approach, the 378 terms of the quadratic parametrization are rearranged into 378 linear combinations of the original terms, defined such that each term is positive by construction. We can then create a template histogram for each of the rearranged terms, encoding the appropriate normalization of each histogram in the `Combine PhysicsModel`. The template histograms and normalizations encoded in the `PhysicsModel` contain the full description of the 26-dimensional quadratic function, so `Combine` is able to appropriately handle this dependence while performing the likelihood fits.

The expected yields also depend on the experimental and theoretical systematic uncertainties (enumerated in Chapter 9), the effects of which are taken into account by a set of nuisance parameters. As mentioned in Section 10.1, the nuisance parameters are profiled in the likelihood fit. The systematic uncertainties may affect either the normalization of the template histograms, or both the normalization and the shape of the template histograms. The former are accounted for via rate systematics in `Combine`, and the latter are accounted for via shape systematics in `Combine`. The systematic uncertainties that affect the shape of the template histograms also carry a 26-dimensional quadratic dependence on the WCs (discussed in Section 9.1), which is accounted for in the same way as the nominal templates.

10.3 Navigating false minima

Multi-dimensional EFT fits may have features in the likelihood "surface" that are difficult for profile fits to navigate. Specifically, if these features include local minima, the profile fits may be susceptible to incorrectly identifying a local minimum as the true minimum (i.e. the fit may become "stuck" in the local minimum). Cases

where the fit becomes "stuck" in local minima may lead to false best fit points, discontinuities in the negative log likelihood (NLL) scans, and inaccurate confidence intervals. It is thus important to identify and to mitigate this issue.

Symptoms of this issue had been observed in the predecessor to this analysis (Ref. [4]) in the form of discontinuities in the NLL values obtained in the one-dimensional scans. To work around this issue, Ref. [4] performed two-dimensional scans for pairs of WCs that had been identified as problematic, avoiding the discontinuities by making use of the NLL values obtained in the two-dimensional scans. However, this approach is not only computationally expensive, but it also does not guarantee that the correct minima will be found (as the fit may encounter similar challenges with false minima while profiling the remaining $n-2$ WCs); a more general approach would be beneficial. Since local minima can arise as a result of the interference terms in the n-dimensional quadratic parameterizations, the challenge of navigating local minima seems to be an inherent feature of multi-dimensional EFT likelihood fitting. As the EFT community continues to explore simultaneous fits to larger sets of WCs, this pernicious issue may become increasingly problematic. For these reasons, one of the intermediate goals of this analysis was to develop a more general approach to the navigation of local minima within the `Combine` framework.

Before developing an approach to address these issues, we wanted to first gain a better understanding of the underlying cause. To this end, we worked to reproduce the issue in a much simpler case. In this simplified model, we only considered two WCs ($c_{\varphi Q}^{-}$ and c_{tG}). We then performed a profiled likelihood fit, scanning over $c_{\varphi Q}^{-}$ and profiling c_{tG}; in other words, we asked the fit to step through a set of given points for $c_{\varphi Q}^{-}$, and at each of those points to find the value of c_{tG} that would minimize the NLL. The result of this likelihood fit is shown in the lefthand side of Figure 10.1.

At a value of approximately $c_{\varphi Q}^{-} = 17$, a discontinuity is observed. To understand this discontinuity, we performed a two-dimensional scan over both WCs. This allows

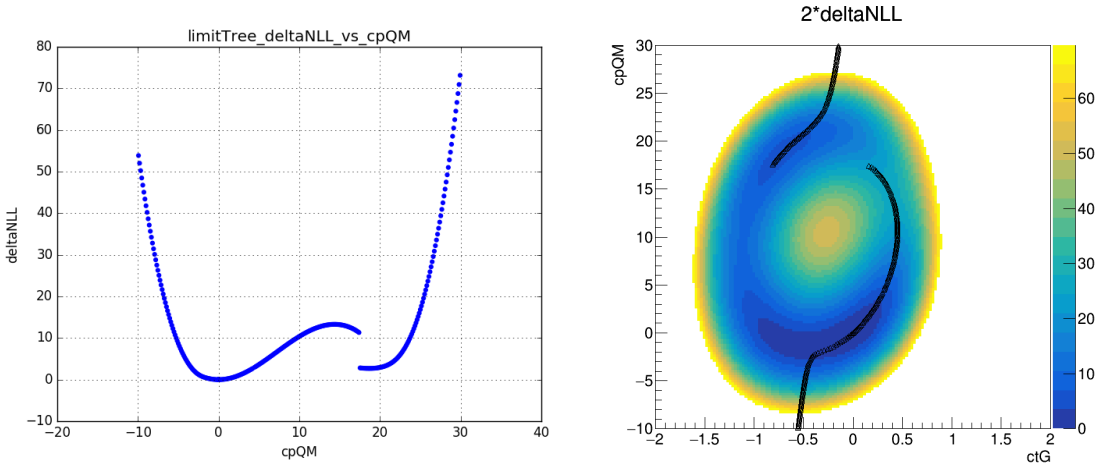


Figure 10.1. Left: NLL for the one-dimensional profiled fit. The $c_{\varphi Q}^{-}$ parameter is scanned while c_{tG} is profiled. Right: NLL for the two-dimensional scan. Here both $c_{\varphi Q}^{-}$ and c_{tG} are scanned. The color scale shows the NLL at each of the two-dimensional scan points. The black overlaid points show the path of the one-dimensional profiled fit.

us to see the complete picture of the space, helping us to visualize why the profiled fit fails. The two-dimensional scan is shown on the righthand side of Figure 10.1, with the path of the one-dimensional profiled fit overlaid in black points. In the one-dimensional profiled fit, $c_{\varphi Q}^{-}$ was scanned while c_{tG} was profiled, meaning that for every point along the y direction (i.e. the $c_{\varphi Q}^{-}$ direction) of the two-dimensional scan, the fit profiles along the x direction (i.e. the c_{tG} direction) in order to find the c_{tG} value at which the NLL (represented by the color scale) is minimized. Following the path of the profiled fit (the overlaid black points), we see that the fit was correctly identifying the c_{tG} point that minimized the NLL from $c_{\varphi Q}^{-} = 0$ until approximately $c_{\varphi Q}^{-} = 5$. At this point, the true minimum lies on the left of the"hill" in the NLL, but the fit continues around the right of this "hill", subsequently incorrectly identifying these local minima as the best fit points. Once the scan reaches approximately $c_{\varphi Q}^{-} = 17$, the fit suddenly jumps to the deeper minimum on the left side of the "hill", resulting in the discontinuity observed in the one-dimensional NLL plot.

To avoid discontinuities, the ideal solution would be to perform a simultaneous scan over all parameters. However, as discussed in Section 10.1, this approach scales exponentially with the number of parameters, and is infeasible for the case of 26 parameters of interst. As an alternative approach, we introduced an element of randomness into the fit in order to sample from the 26-dimensional space. To motivate this approach, let us revisit the profiled fit discussed above. During a fit, `Combine` always uses the same starting value for the profiled parameters; this is useful for reproducibility, but it means that if the starting point happens to be near a local minimum (where the global minimum lies on the other side of a "hill" in NLL), the fit will always find the local minimum, and never find the correct global minimum. If we instead allow the starting point for the profiled fit to be chosen randomly, the starting point will sometimes lie on the other side of the "hill" in NLL, allowing the fit to find the correct minimum.

To test the random starting point method, we modified the `Combine` tool to incorporate random starting points for the profiled parameters. After first generating a list of random starting points, the modified version of the `Combine` script loops through the random starting points, finding the NLL at each, and keeping track of the point that gives rise to the lowest NLL. After trying each of the random starting points, the point that produced the lowest NLL is taken to be the profiled value of the parameter at that point. Applying this approach to the two-dimensional case described above, the NLL is found to be continuous, as shown in Figure 10.2, where the profiled fit finds the correct global minimum at each scan point. For example, at approximately $c^{-}_{\varphi Q} = 5$, the fit is able to identify the deeper minimum on the left side of the "hill" in NLL, jumping to that minimum immediately instead of erroneously continuing around the right side of the "hill" (as had been observed in Figure 10.1).

After the random starting point approach was shown to be successful in the simple two-dimensional case, we generalized the method and tested it on the 16-dimensional

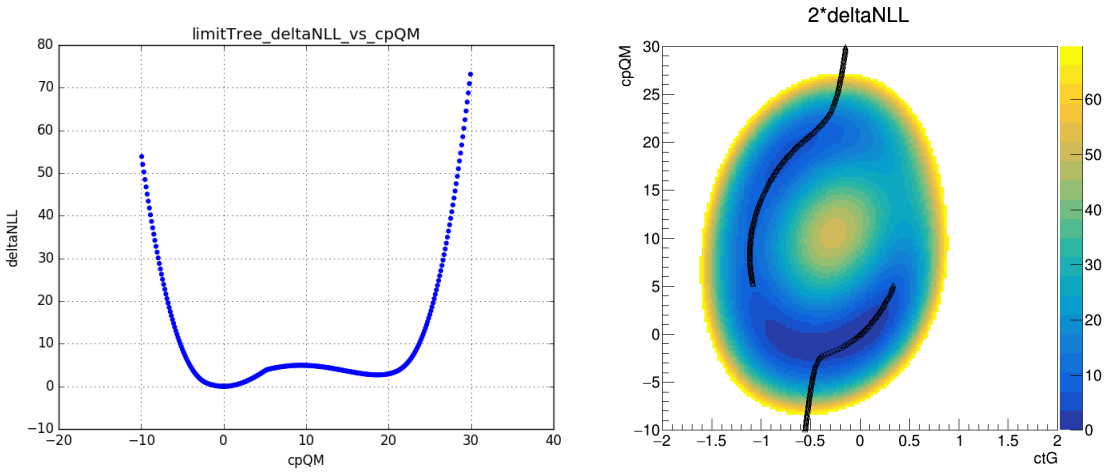


Figure 10.2. Left: NLL for the one-dimensional profiled fit after the implementation of the random starting point approach, where $c_{\varphi Q}^{-}$ is scanned and c_{tG} is profiled. Right: NLL for the two-dimensional scan over $c_{\varphi Q}^{-}$ and c_{tG}. The color scale shows the NLL at each of the two-dimensional scan points. The black overlaid points show the path of the one-dimensional profiled fit after the implementation of the random starting point approach.

fits from Ref. [4]. Although the number of random starting points required in order to obtain a smooth NLL curve was larger in the 16-dimensional case than in the two-dimensional case (greater than 50 as opposed to less than 10), the method successfully avoided discontinuities in NLL. The results of this test are shown in Appendix G.

Although this approach is computationally feasible, it is still relatively computationally expensive (as the likelihood fit must be run x times for each scan point, where x is the number of random starting points). In the future, it would be interesting to optimize the approach by considering methods of sampling the space more efficiently. For example, one idea would be to first identify a set of distinct local minima in the space, and use a set of points from these local minima as the starting values instead of choosing the starting points randomly from the full space.

CHAPTER 11

RESULTS

As described in Chapter 10, the `Combine` tool is used to perform the likelihood fits from which we extract the results of this analysis. Before performing the likelihood fits to the data, we first perform the fits with Asimov data (i.e. simulated data that corresponds to the SM expectation). These blinded Asimov fits allow us to determine the expected sensitivity and perform optimization studies without introducing biases; for example, the sensitivity optimization described in Section 6.3 was performed using only Asimov data. The results of the Asimov fits are shown in Section 11.1. The unblinded results of this analysis (from the fits to the observed data) are reported in Section 11.2. The results are found to be consistent with the SM expectation. A visualization of the results (in the form of post-fit histograms) is included in Section 11.3.

It should be noted that after unblinding this analysis, there were no signs of problematic discontinuities or false minima, so the random starting point approach (described in Section 10.3) was not used to obtain the final results. The 1-dimensional profile scans in this analysis exhibit fewer double minima than were observed in [4] (the predecessor to this analysis), which could be due to degeneracies that have been broken by the improvements implemented in this iteration; the fact that there are fewer local minima may help to explain why the profiled fits in this analysis seem to be able to successfully navigate the space and find the true minima without the need for random starting points. However, it is difficult to know a priori whether or not the fits will be susceptible to local minima, so future analyses should keep this

challenge in mind when performing profiled likelihood scans.

11.1 SM expected limits

The 1-dimensional likelihood scans are shown for each WC in Figure 11.1; the other 25 WCs are either profiled (shown by the black markers) or fixed at their SM values of zero (shown by the red markers). As explained in Section 10.1, the 1σ and 2σ confidence intervals can be read off of these scans by observing where the 2ΔNLL curve crosses 1 and 4, respectively. The resulting 1σ and 2σ confidence intervals are listed in Table 11.1 and displayed in Figure 11.2. Since these scans are performed with Asimov data, they represent the expected sensitivity of the analysis under the SM hypothesis.

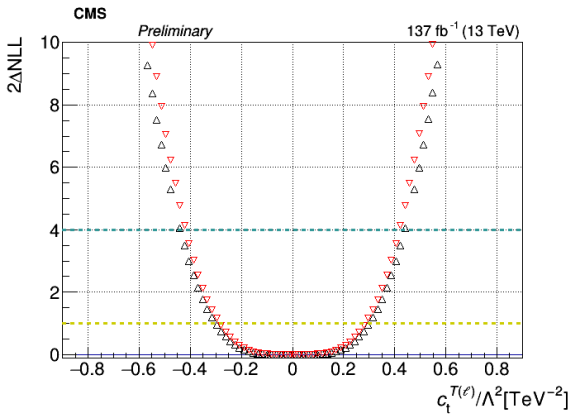
11.2 Observed results

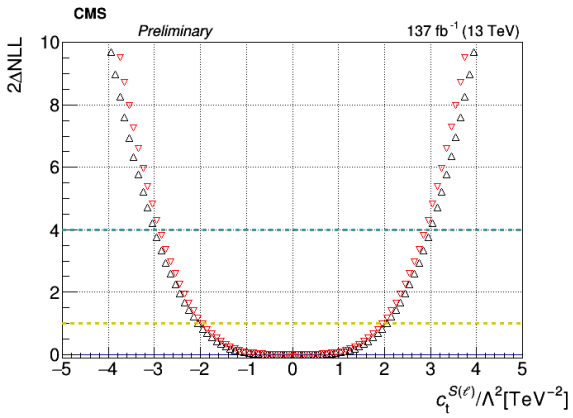
The 1-dimensional likelihood scans are shown for each WC in Figure 11.3; the other 25 WCs are either profiled (shown by the black markers) or fixed at their SM values of zero (shown by the red markers). The resulting 1σ and 2σ confidence intervals are listed in Table 11.1 and displayed in Figure 11.2. The results are consistent with the SM prediction.

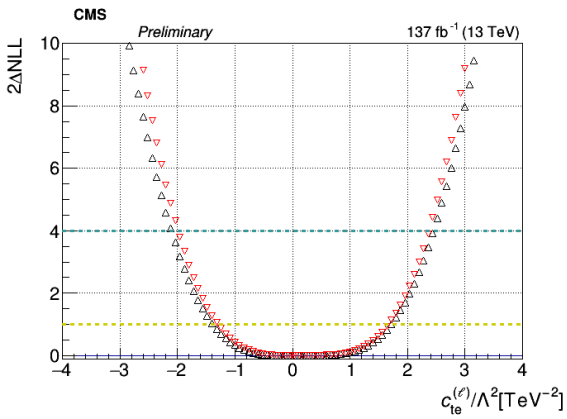
11.3 Visualization of results: Prefit and postfit summary histograms

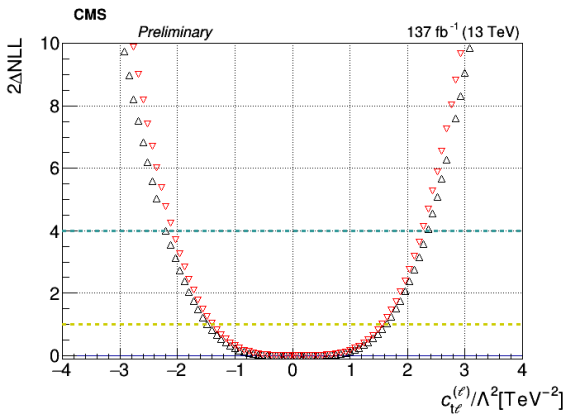
To visualize the results, we can compare histograms (for the categories and bins in the analysis) of the observed yields and the predicted yields. The top figure in Figure 11.5 shows the observed data compared against the SM prediction. In this "prefit" case, we have evaluated the 26-dimensional quadratic (that parametrizes the predicted yields in terms of the WCs) at the SM value of zero for each WC. In other words, we have reweighted the prediction to the SM. In the bottom figure of

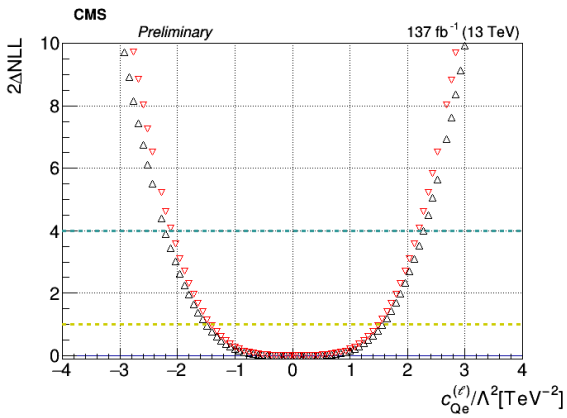
Figure 11.5, the prediction has been reweighted to the best fit point (as obtained from the likelihood fits). This is referred to as the "postfit" scenario. The shaded bands in the plots represent the total uncertainty, and it can be seen that in both the prefit and postfit scenarios, the observation is within the uncertainty on the prediction for most of the bins in the analysis; this visualization is consistent with the results of the statistical analysis, which found the SM point to be consistent with the data.

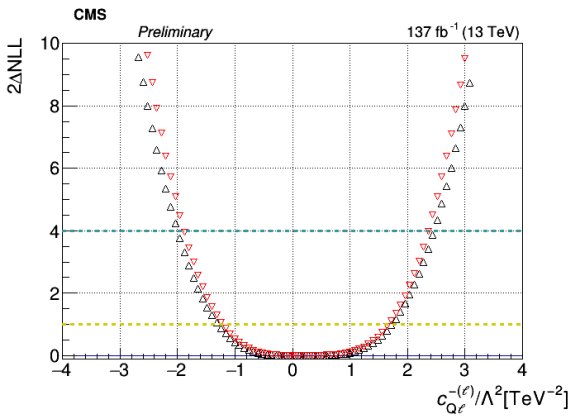

CMS
Preliminary
137 fb^{-1} (13 TeV)
2ΔNLL
$c_t^{T(\ell)}/\Lambda^2[\text{TeV}^{-2}]$

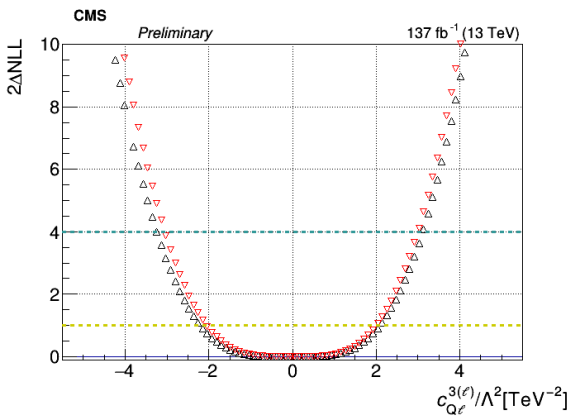

CMS
Preliminary
137 fb^{-1} (13 TeV)
2ΔNLL
$c_t^{S(\ell)}/\Lambda^2[\text{TeV}^{-2}]$

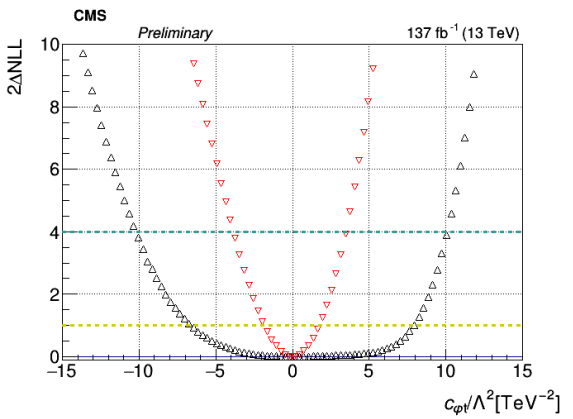

CMS
Preliminary
137 fb^{-1} (13 TeV)
2ΔNLL
$c_{te}^{(\ell)}/\Lambda^2[\text{TeV}^{-2}]$

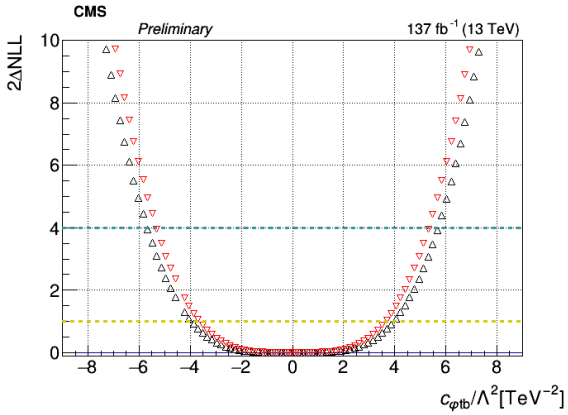

CMS
Preliminary
137 fb^{-1} (13 TeV)
2ΔNLL
$c_{t\ell}^{(\ell)}/\Lambda^2[\text{TeV}^{-2}]$

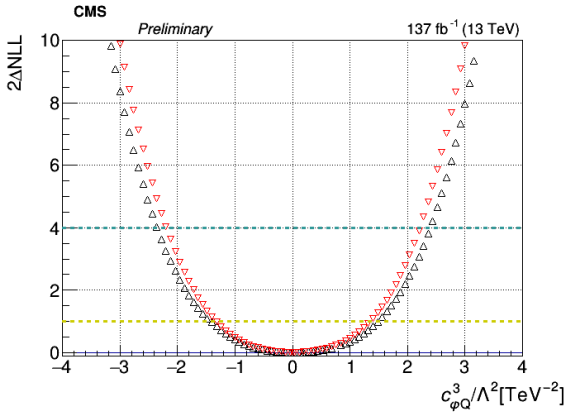

CMS
Preliminary
137 fb^{-1} (13 TeV)
2ΔNLL
$c_{Qe}^{(\ell)}/\Lambda^2[\text{TeV}^{-2}]$

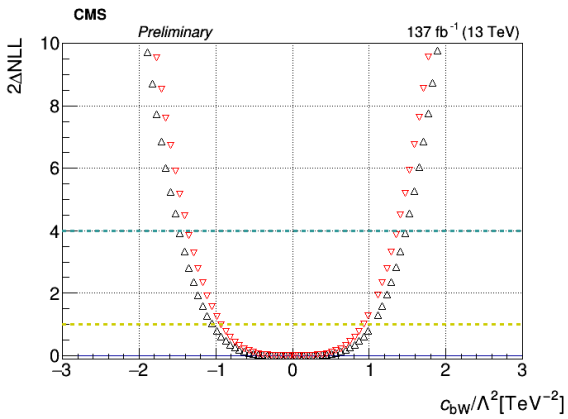

CMS
Preliminary
137 fb^{-1} (13 TeV)
2ΔNLL
$c_{Q\ell}^{-(\ell)}/\Lambda^2[\text{TeV}^{-2}]$

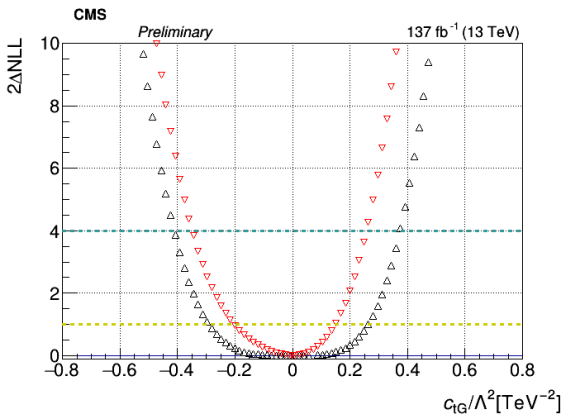

CMS
Preliminary
137 fb^{-1} (13 TeV)
2ΔNLL
$c_{Q\ell}^{3(\ell)}/\Lambda^2[\text{TeV}^{-2}]$

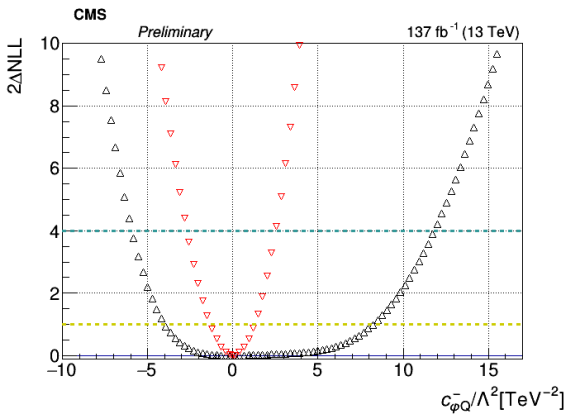

CMS
Preliminary
137 fb^{-1} (13 TeV)
2ΔNLL
$c_{\varphi t}/\Lambda^2[\text{TeV}^{-2}]$

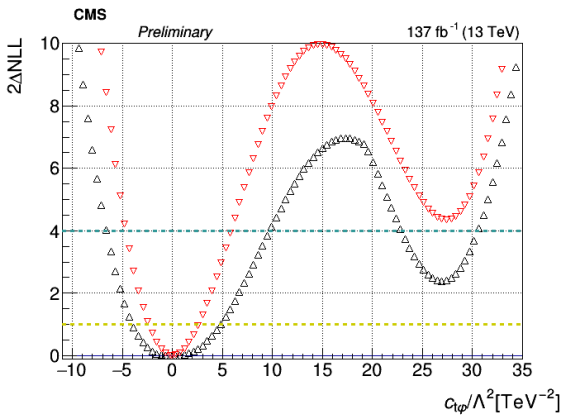

CMS
Preliminary
137 fb^{-1} (13 TeV)
2ΔNLL
$c_{\varphi tb}/\Lambda^2$[TeV^{-2}]

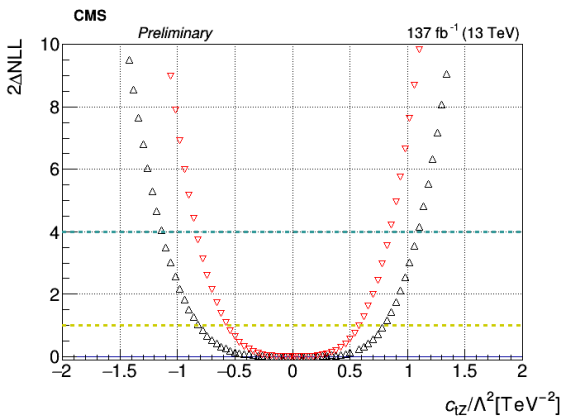

CMS
Preliminary
137 fb^{-1} (13 TeV)
2ΔNLL
$c^3_{\varphi Q}/\Lambda^2$[TeV^{-2}]

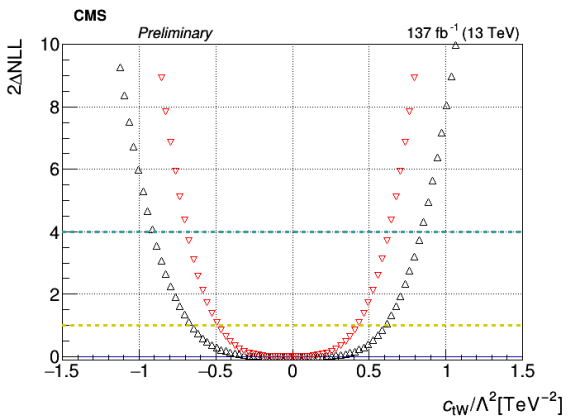

CMS
Preliminary
137 fb^{-1} (13 TeV)
2ΔNLL
c_{bW}/Λ^2[TeV^{-2}]

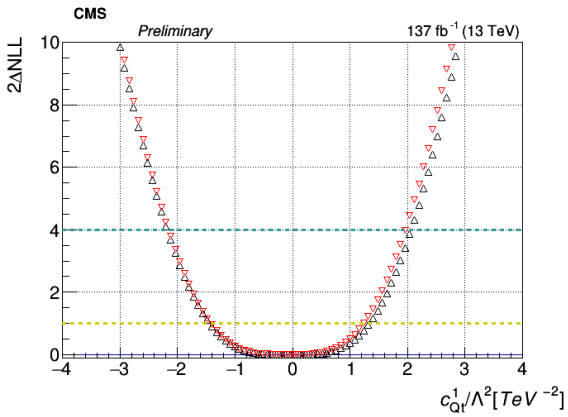

CMS
Preliminary
137 fb^{-1} (13 TeV)
2ΔNLL
c_{tG}/Λ^2[TeV^{-2}]

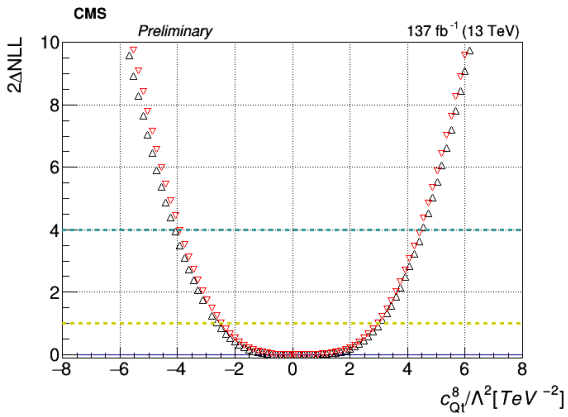

CMS
Preliminary
137 fb^{-1} (13 TeV)
2ΔNLL
$c^-_{\varphi Q}/\Lambda^2$[TeV^{-2}]

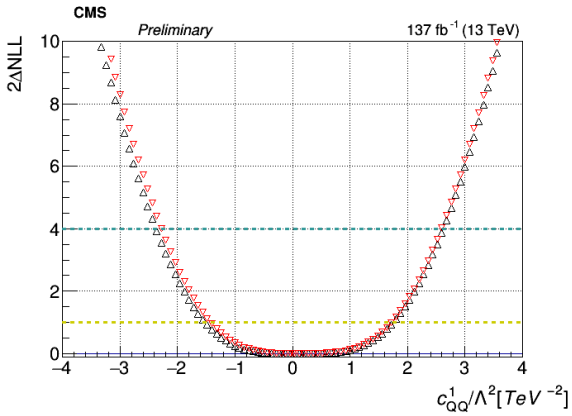

CMS
Preliminary
137 fb^{-1} (13 TeV)
2ΔNLL
$c_{t\varphi}/\Lambda^2$[TeV^{-2}]

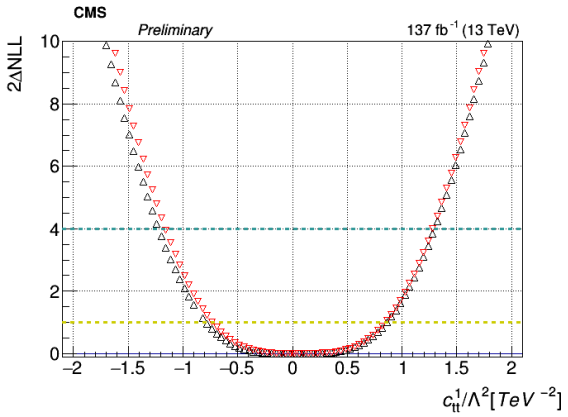

CMS
Preliminary
137 fb^{-1} (13 TeV)
2ΔNLL
c_{tZ}/Λ^2[TeV^{-2}]

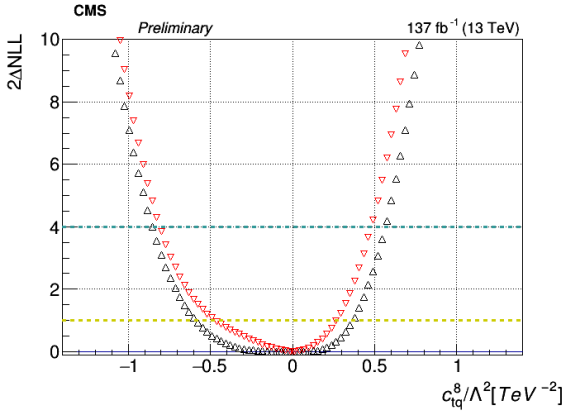

CMS
Preliminary
137 fb^{-1} (13 TeV)
2ΔNLL
c_{tW}/Λ^2[TeV^{-2}]

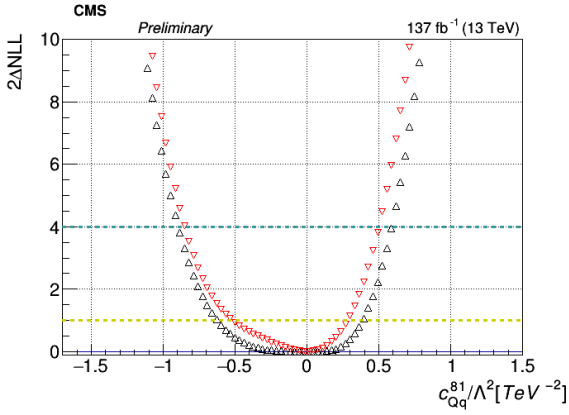

CMS
Preliminary
137 fb^{-1} (13 TeV)
2ΔNLL
$c^{1}_{Qt}/\Lambda^2[TeV^{-2}]$

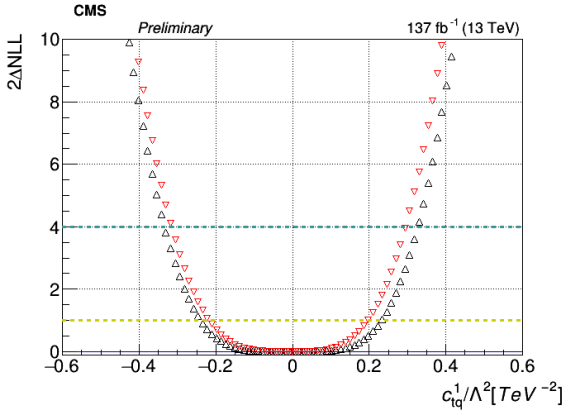

CMS
Preliminary
137 fb^{-1} (13 TeV)
2ΔNLL
$c^{8}_{Qt}/\Lambda^2[TeV^{-2}]$

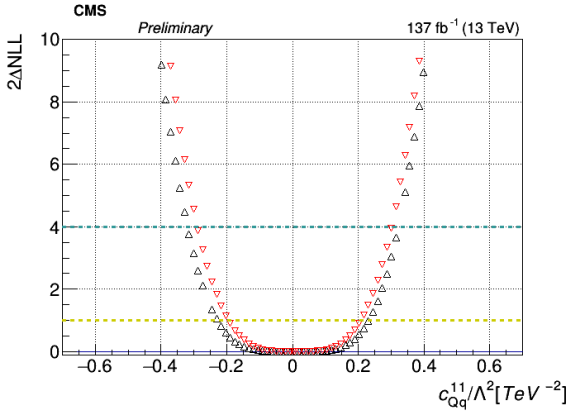

CMS
Preliminary
137 fb^{-1} (13 TeV)
2ΔNLL
$c^{1}_{QQ}/\Lambda^2[TeV^{-2}]$


CMS
Preliminary
137 fb^{-1} (13 TeV)
2ΔNLL
$c^{1}_{tt}/\Lambda^2[TeV^{-2}]$


CMS
Preliminary
137 fb^{-1} (13 TeV)
2ΔNLL
$c^{8}_{tq}/\Lambda^2[TeV^{-2}]$


CMS
Preliminary
137 fb^{-1} (13 TeV)
2ΔNLL
$c^{81}_{Qq}/\Lambda^2[TeV^{-2}]$


CMS
Preliminary
137 fb^{-1} (13 TeV)
2ΔNLL
$c^{1}_{tq}/\Lambda^2[TeV^{-2}]$


CMS
Preliminary
137 fb^{-1} (13 TeV)
2ΔNLL
$c^{11}_{Qq}/\Lambda^2[TeV^{-2}]$

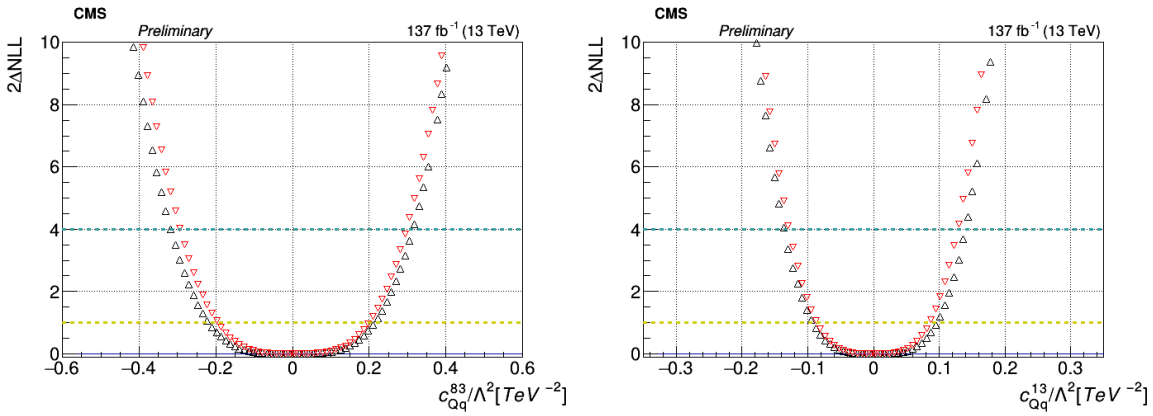


Figure 11.1. SM-expected 2ΔNLL values for 1-dimensional scans for each WCs. The 2ΔNLL values represented in black correspond to the case where the other WCs are profiled, while the 2ΔNLL values represented in red correspond to the case where the other WCs are fixed at their SM values of zero.

TABLE 11.1

SM-EXPECTED 2σ UNCERTAINTY INTERVALS FOR THE 1D SCANS.

WC/Λ^2 [TeV^{-2}]	2σ Interval (others profiled)	2σ Interval (others fixed to SM)
$c_t^{T(\ell)}$	[-0.44, 0.44]	[-0.42, 0.42]
$c_t^{S(\ell)}$	[-3.00, 3.00]	[-2.89, 2.89]
$c_{te}^{(\ell)}$	[-2.11, 2.46]	[-1.99, 2.38]
$c_{t\ell}^{(\ell)}$	[-2.20, 2.35]	[-2.08, 2.26]
$c_{Qe}^{(\ell)}$	[-2.22, 2.28]	[-2.11, 2.19]
$c_{Q\ell}^{-(\ell)}$	[-2.00, 2.46]	[-1.89, 2.36]
$c_{Q\ell}^{3(\ell)}$	[-3.24, 3.11]	[-3.05, 3.00]
$c_{\varphi t}$	[-10.21, 10.10]	[-3.86, 3.49]
$c_{\varphi tb}$	[-5.69, 5.70]	[-5.33, 5.33]
$c_{\varphi Q}^3$	[-2.35, 2.40]	[-2.20, 2.22]
c_{bW}	[-1.48, 1.48]	[-1.36, 1.36]
c_{tG}	[-0.41, 0.37]	[-0.35, 0.26]
$c_{\varphi Q}^-$	[-5.92, 11.83]	[-2.71, 2.51]
$c_{t\varphi}$	[-6.61, 9.72] and [22.88, 30.56]	[-4.67, 5.87]
c_{tZ}	[-1.14, 1.09]	[-0.84, 0.85]
c_{tW}	[-0.91, 0.84]	[-0.69, 0.63]
c_{Qt}^1	[-2.18, 2.07]	[-2.16, 1.96]
c_{Qt}^8	[-4.10, 4.54]	[-3.93, 4.44]
c_{QQ}^1	[-2.38, 2.63]	[-2.28, 2.59]
c_{tt}^1	[-1.22, 1.30]	[-1.16, 1.28]
c_{tq}^8	[-0.85, 0.56]	[-0.81, 0.48]
c_{Qq}^{18}	[-0.89, 0.59]	[-0.85, 0.50]
c_{tq}^1	[-0.33, 0.33]	[-0.32, 0.30]
c_{Qq}^{11}	[-0.32, 0.32]	[-0.29, 0.30]
c_{Qq}^{38}	[-0.32, 0.31]	[-0.29, 0.30]
c_{Qq}^{31}	[-0.14, 0.14]	[-0.13, 0.13]

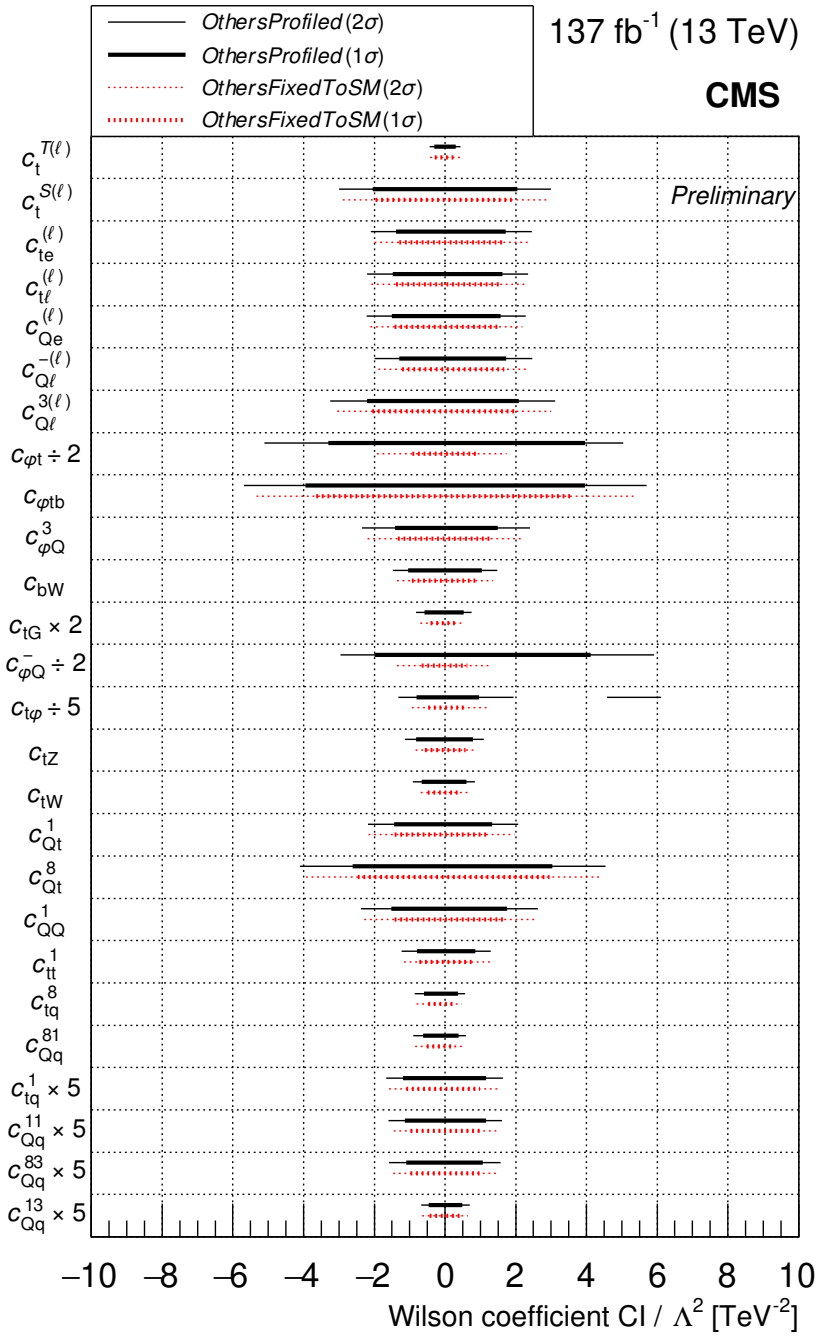
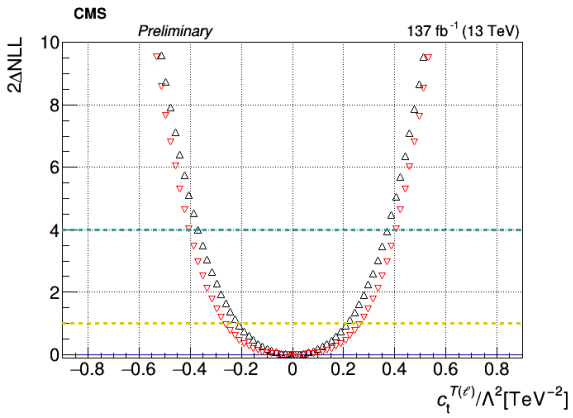
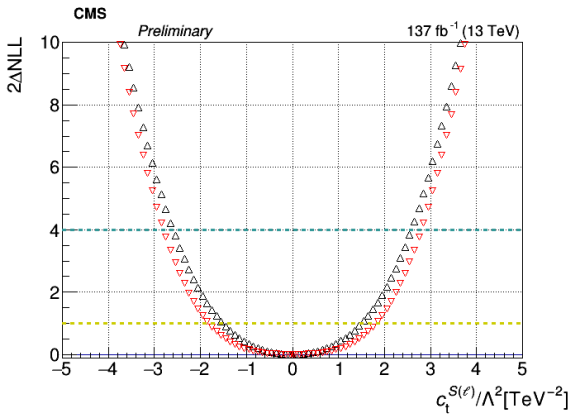
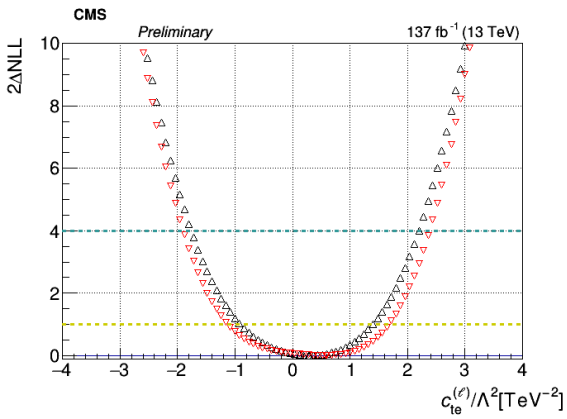
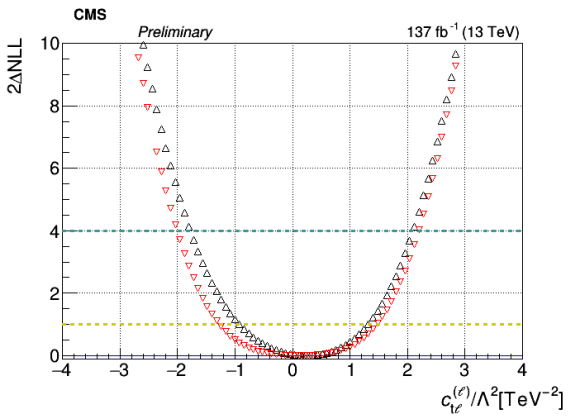


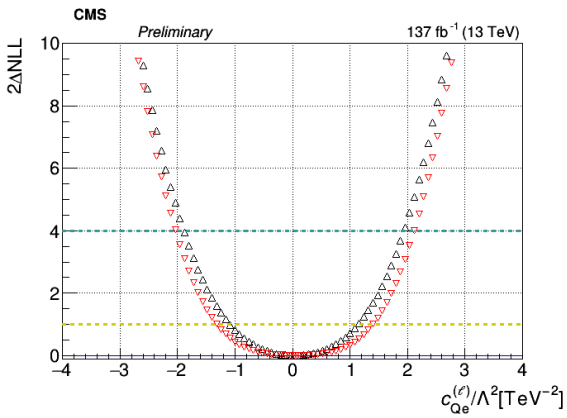
Figure 11.2. Summary of limits from Asimov fits. WC 1σ (thick line) and 2σ (thin line) uncertainty intervals are shown for the case where the other WCs are profiled (in black), and the case where the other WCs are fixed at their SM values of zero (in red).

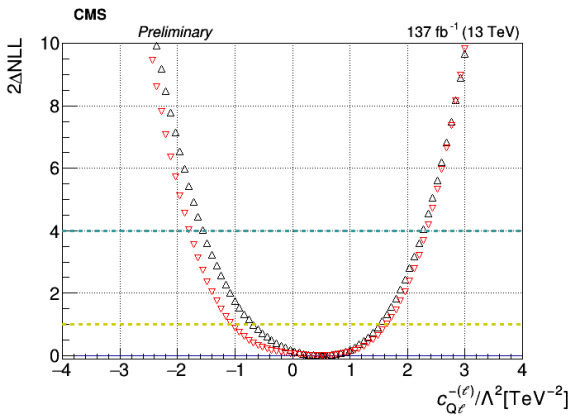

CMS
Preliminary
137 fb^{-1} (13 TeV)
2ΔNLL
$c_t^{T(\ell)}/\Lambda^2[\text{TeV}^{-2}]$

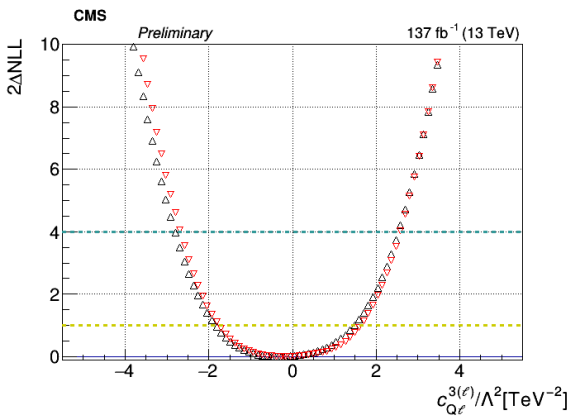

CMS
Preliminary
137 fb^{-1} (13 TeV)
2ΔNLL
$c_t^{S(\ell)}/\Lambda^2[\text{TeV}^{-2}]$

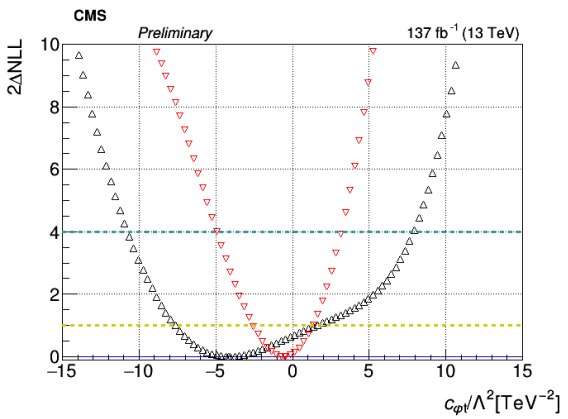

CMS
Preliminary
137 fb^{-1} (13 TeV)
2ΔNLL
$c_{te}^{(\ell)}/\Lambda^2[\text{TeV}^{-2}]$

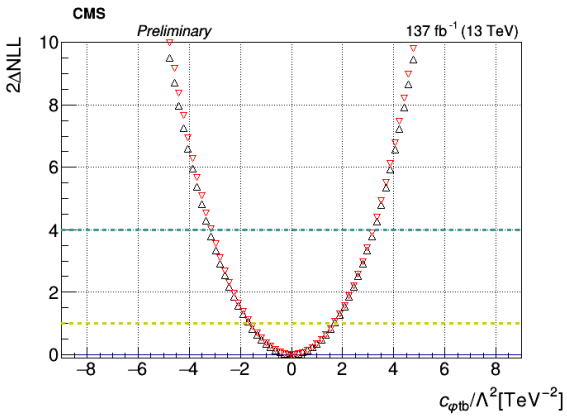

CMS
Preliminary
137 fb^{-1} (13 TeV)
2ΔNLL
$c_{t\ell}^{(\ell)}/\Lambda^2[\text{TeV}^{-2}]$

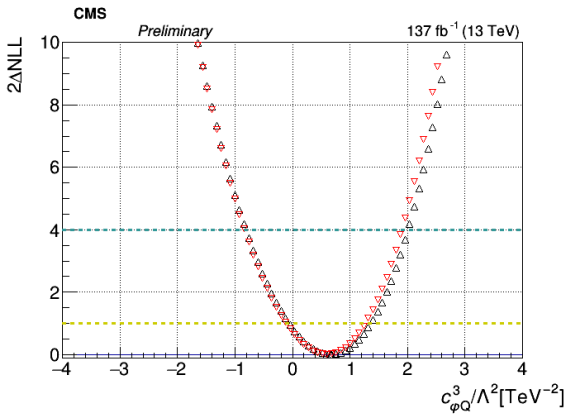

CMS
Preliminary
137 fb^{-1} (13 TeV)
2ΔNLL
$c_{Qe}^{(\ell)}/\Lambda^2[\text{TeV}^{-2}]$

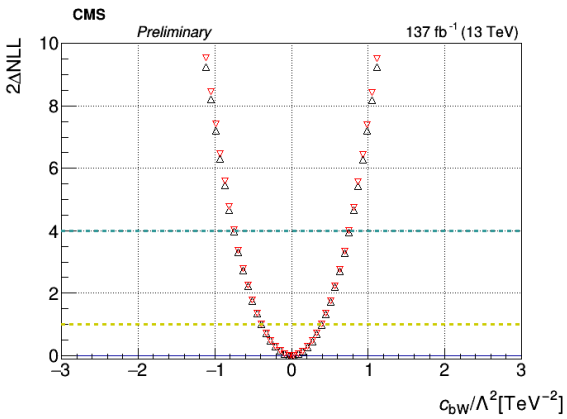

CMS
Preliminary
137 fb^{-1} (13 TeV)
2ΔNLL
$c_{Q\ell}^{-(\ell)}/\Lambda^2[\text{TeV}^{-2}]$

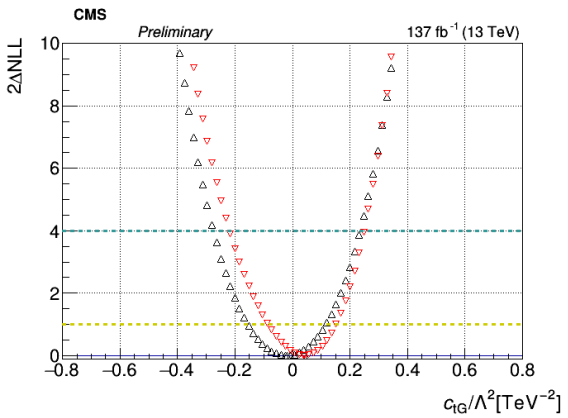

CMS
Preliminary
137 fb^{-1} (13 TeV)
2ΔNLL
$c_{Q\ell}^{3(\ell)}/\Lambda^2[\text{TeV}^{-2}]$

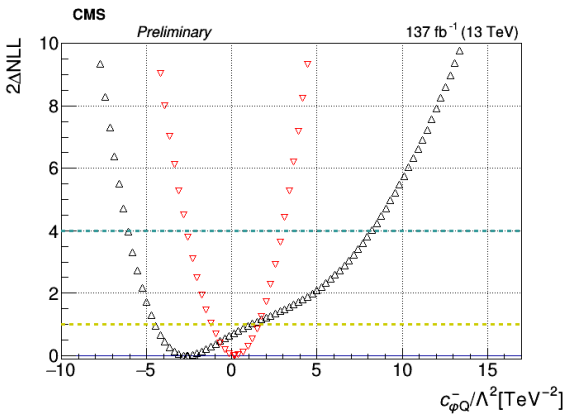

CMS
Preliminary
137 fb^{-1} (13 TeV)
2ΔNLL
$c_{\varphi t}/\Lambda^2[\text{TeV}^{-2}]$

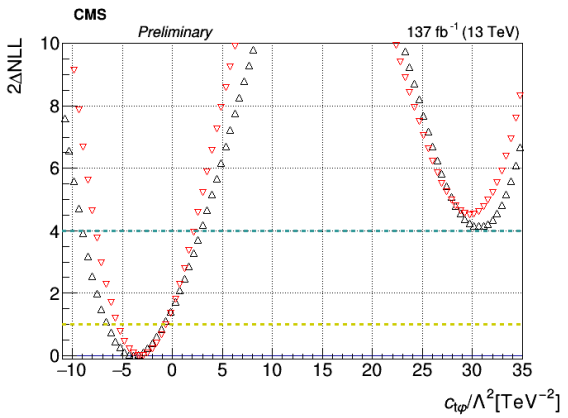

CMS
Preliminary
137 fb^{-1} (13 TeV)
2ΔNLL
$c_{\varphi tb}/\Lambda^2$[TeV^{-2}]

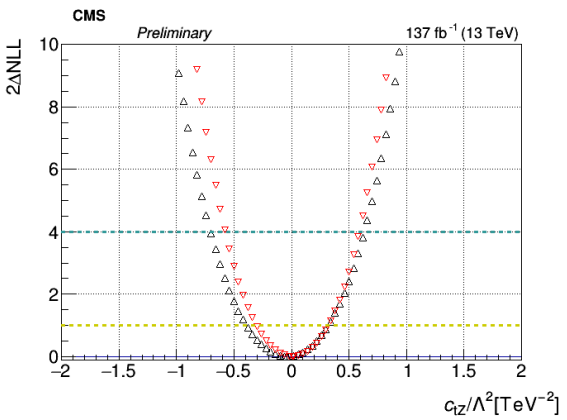

CMS
Preliminary
137 fb^{-1} (13 TeV)
2ΔNLL
$c^3_{\varphi Q}/\Lambda^2$[TeV^{-2}]

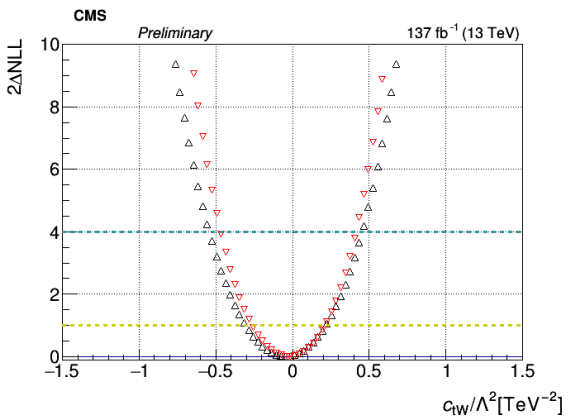

CMS
Preliminary
137 fb^{-1} (13 TeV)
2ΔNLL
c_{bW}/Λ^2[TeV^{-2}]

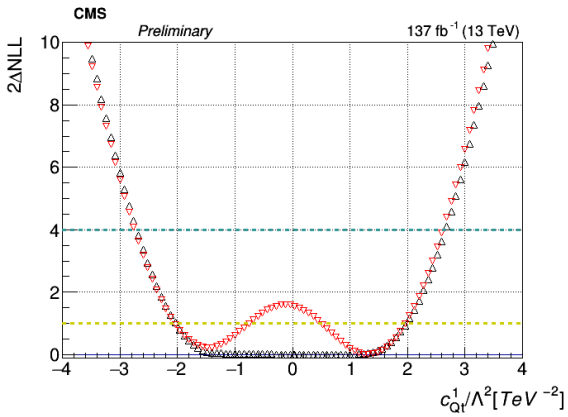

CMS
Preliminary
137 fb^{-1} (13 TeV)
2ΔNLL
c_{tG}/Λ^2[TeV^{-2}]

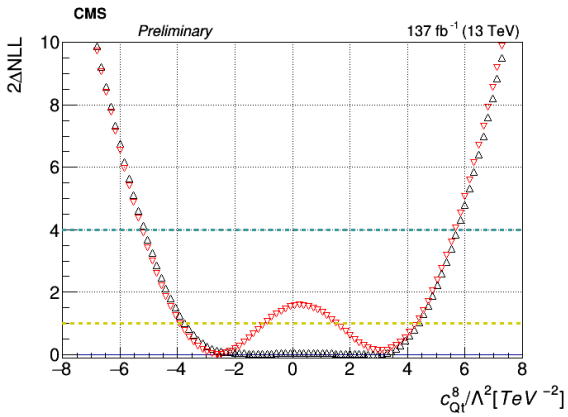

CMS
Preliminary
137 fb^{-1} (13 TeV)
2ΔNLL
$c^-_{\varphi Q}/\Lambda^2$[TeV^{-2}]

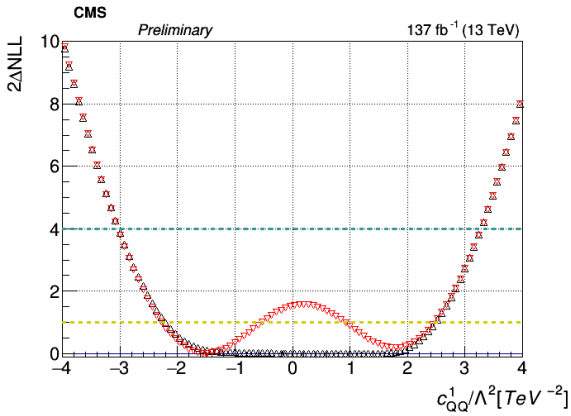

CMS
Preliminary
137 fb^{-1} (13 TeV)
2ΔNLL
$c_{t\varphi}/\Lambda^2$[TeV^{-2}]

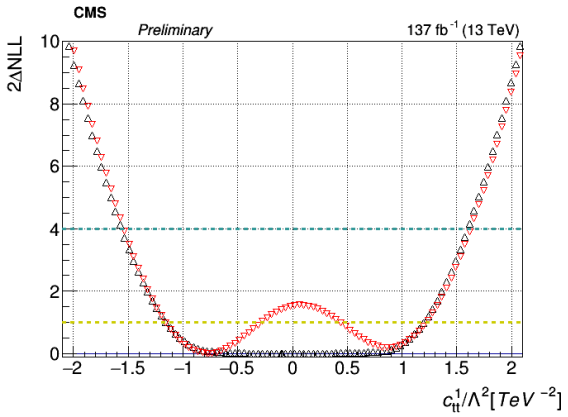

CMS
Preliminary
137 fb^{-1} (13 TeV)
2ΔNLL
c_{tZ}/Λ^2[TeV^{-2}]

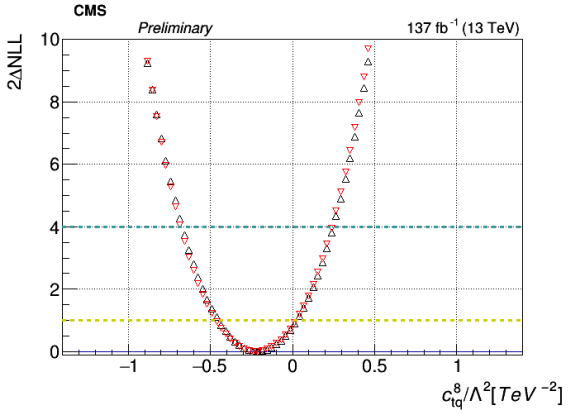

CMS
Preliminary
137 fb^{-1} (13 TeV)
2ΔNLL
c_{tW}/Λ^2[TeV^{-2}]

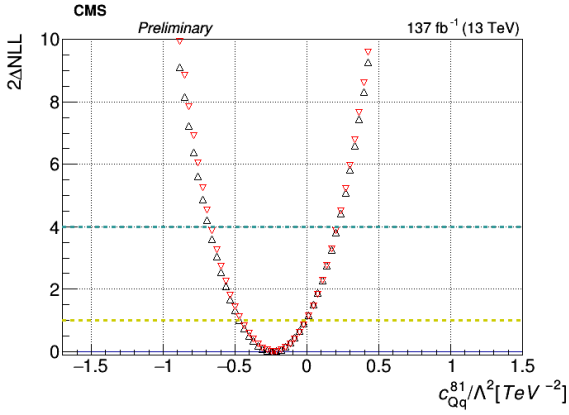

CMS
Preliminary
137 fb^{-1} (13 TeV)
2ΔNLL
$c^{1}_{Qt}/\Lambda^2[TeV^{-2}]$

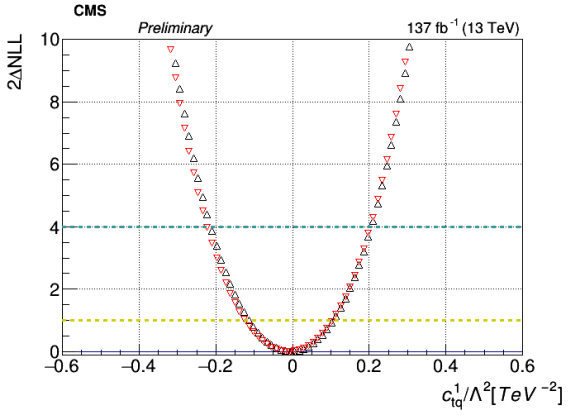

CMS
Preliminary
137 fb^{-1} (13 TeV)
2ΔNLL
$c^{8}_{Qt}/\Lambda^2[TeV^{-2}]$

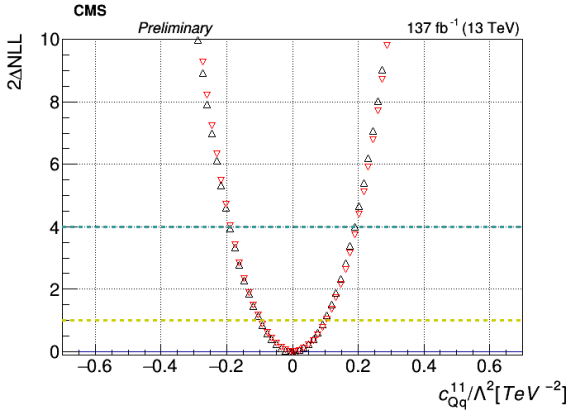

CMS
Preliminary
137 fb^{-1} (13 TeV)
2ΔNLL
$c^{1}_{QQ}/\Lambda^2[TeV^{-2}]$


CMS
Preliminary
137 fb^{-1} (13 TeV)
2ΔNLL
$c^{1}_{tt}/\Lambda^2[TeV^{-2}]$


CMS
Preliminary
137 fb^{-1} (13 TeV)
2ΔNLL
$c^{8}_{tq}/\Lambda^2[TeV^{-2}]$


CMS
Preliminary
137 fb^{-1} (13 TeV)
2ΔNLL
$c^{81}_{Qq}/\Lambda^2[TeV^{-2}]$


CMS
Preliminary
137 fb^{-1} (13 TeV)
2ΔNLL
$c^{1}_{tq}/\Lambda^2[TeV^{-2}]$


CMS
Preliminary
137 fb^{-1} (13 TeV)
2ΔNLL
$c^{11}_{Qq}/\Lambda^2[TeV^{-2}]$

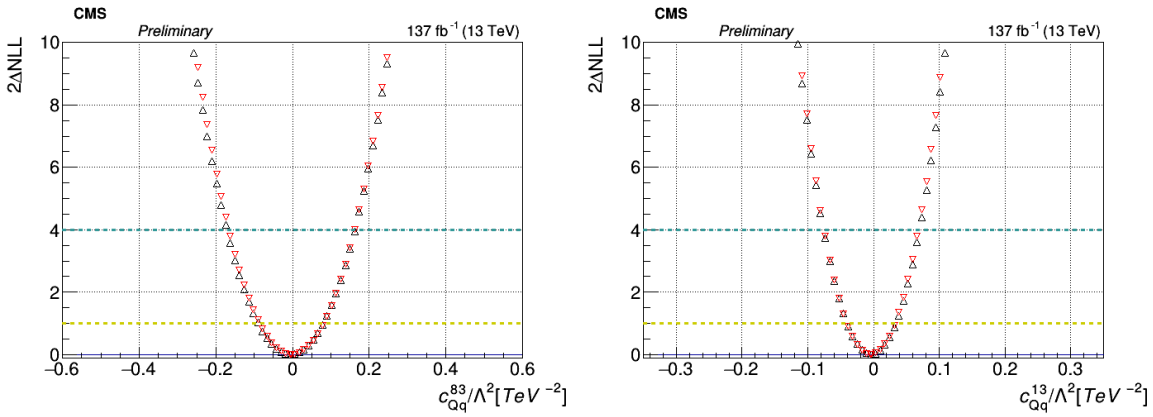


Figure 11.3. Observed 2ΔNLL values for 1d scans for each WCs. The 2ΔNLL values represented in black correspond to the case where the other WCs are profiled, while the 2ΔNLL values represented in red correspond to the case where the other WCs are fixed at their SM values of zero.

TABLE 11.2

OBSERVED 2σ UNCERTAINTY INTERVALS FOR THE 1D SCANS.

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]
$c_{\varphi Q}^{3}$	[-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}^{11}	[-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]

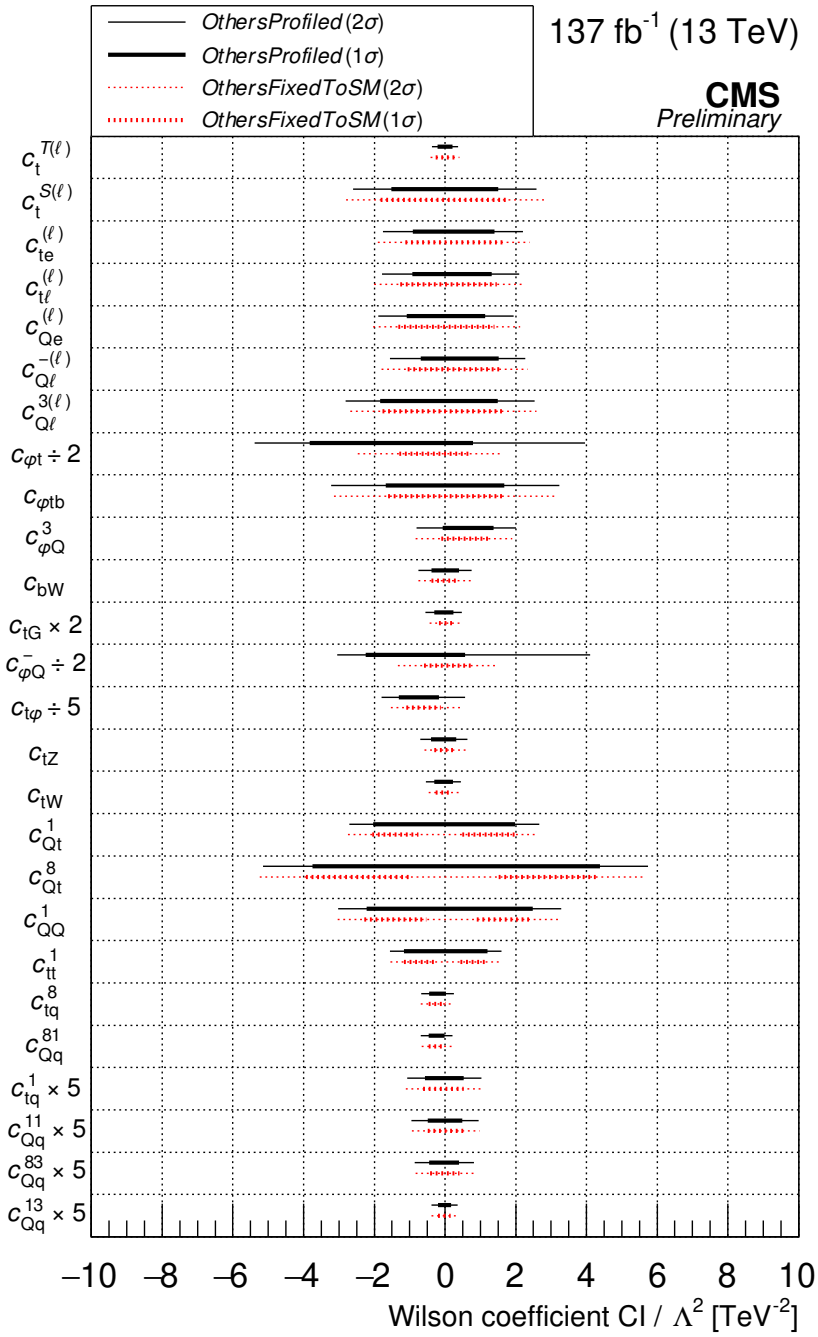


Figure 11.4. Summary of limits from fits to data. WC 1σ (thick line) and 2σ (thin line) uncertainty intervals are shown for the case where the other WCs are profiled (in black), and the case where the other WCs are fixed at their SM values of zero (in red).

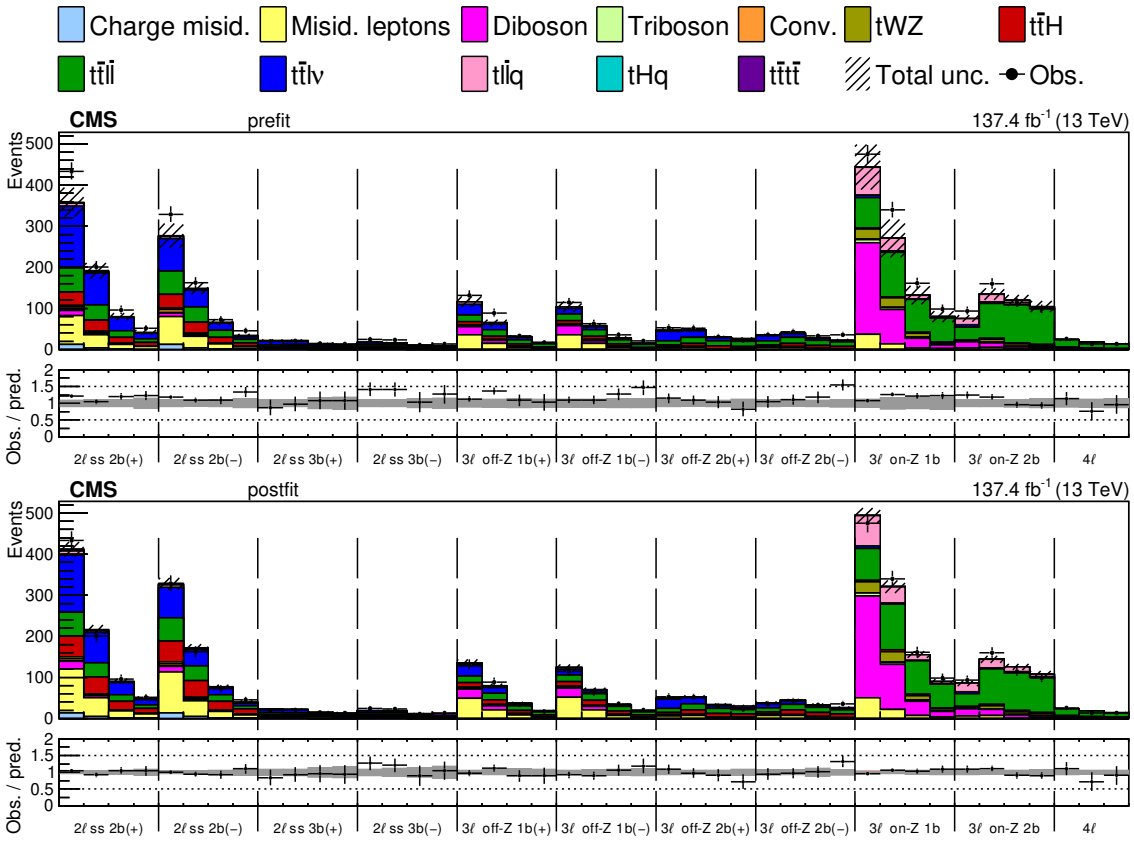


Figure 11.5. The observed yields and the predicted prefit (top) and postfit (bottom) yields. As explained in the text (Section 11.3), in the prefit case the predicted yields have been reweighted to the SM, while in the postfit case the predicted yields have been reweighted to the best fit point from the unblind fits. Here we have integrated over the kinematic variables, so the bins in these histogram correspond to the jet multiplicity categories of the analysis categories. [62]

CHAPTER 12

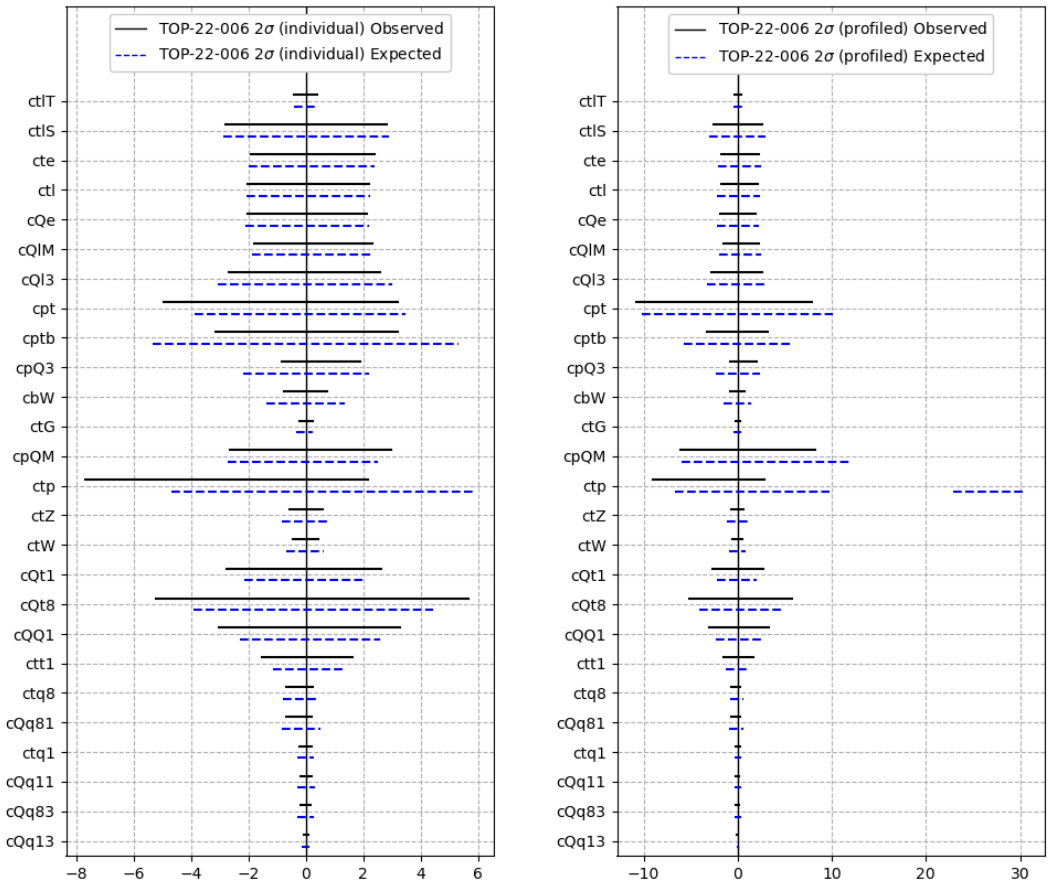
DISCUSSION OF RESULTS

This chapter will explore various aspects of the results presented in Chapter 11. First, Section 12.1 will compare the observed 2σ limits to the expected 2σ limits and discuss relevant differences. In Section 12.2, correlations among the WCs will be studied, and two-dimensional scans for selected pairs of WCs will be presented. Section 12.3 will discuss the relative importance of the various bins and categories to the sensitivity to each WC. Next, Section 12.4 will explore the impacts of the uncertainties on the limits. Section 12.5 will compare the observed limits to the results of other analyses. Lastly, the limits will be interpreted in terms of the energy scale Λ in Section 12.6.

12.1 Comparison of observed limits to predicted limits

Figure 12.1 shows a comparison between the observed 2σ limits and the expected SM 2σ limits (obtained by fitting to Asimov data). We expect statistical fluctuations to cause differences between these limits. The differences between the observed and the expected 2σ ranges vary by WC from about 2% to about 70%. The anomalously large 70% difference corresponds to the $c_{t\varphi}$ fit in which the other WCs are profiled. As can be seen in the one-dimensional scans presented in Chapter 11, the second minimum for $c_{t\varphi}$ crosses the 2σ threshold in the Asimov fit, but not in the fit to the observed data; the widths of the 2σ intervals for this comparison are calculated from the extrema of the 2σ limits, thus explaining the large difference for this WC. Another noticeable difference is the $c_{t\varphi}$ limit in the case where all other WCs are

fixed to 0. While the widths of the intervals are similar, the ranges are significantly shifted with respect to each other (since the minimum in the Asimov fit is at $c_{t\varphi} = 0$, while the minimum in the data fit is near $c_{t\varphi} = -4$). This feature is similarly visible in the one-dimensional scans presented in Chapter 11.



(a) (b)

Figure 12.1. The 2σ observed limits (solid black line) compared against the 2σ Asimov limits (dashed blue lines). Figure (a) shows the results of the likelihood fits in which a single WC is fit with all other WCs fixed to their SM values of zero. Figure (b) shows the results of the likelihood fits in which the other WCs are profiled. In these plots, $\Lambda = 1$ TeV.

12.2 Exploration of correlations among WCs

In Section 11.2, we explored the likelihood along one-dimensional directions of the full 26-dimensional surface. In principle, it would be interesting to study the shape of the full 26-dimensional surface; however, in practice it is very difficult to explore high-dimensional spaces. In lieu of the full 26-dimensional visualization, we can at least explore two-dimensional slices of the space. Similar to the one-dimensional scans described in Chapter 11, we can scan over two WCs and profile the remaining 24 WCs. However, there would be a total of 325 unique pairs of WCs to consider, and two-dimensional scans are significantly more computationally expensive than one-dimensional scans. For example, each the one-dimensional scans shown in Figure 11.3 was performed with 100 scan points, so a two-dimensional scan with the same granularity would require $100 \times 100 = 10,000$ scan points.

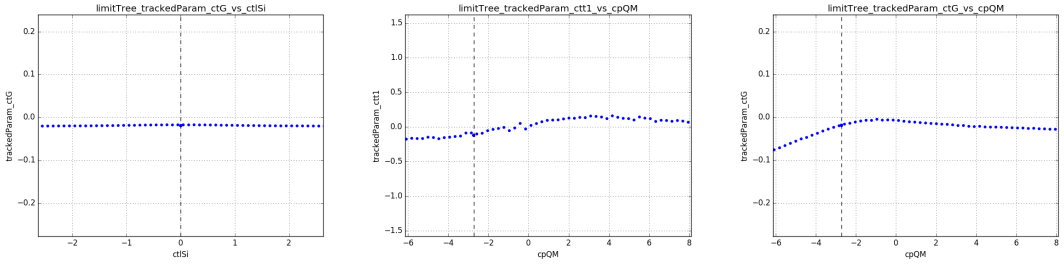
For these reasons, it would be beneficial to determine ahead of time a subset of the pairs of WCs that would be interesting to explore. Pairs of WCs are considered to be interesting if they have some correlation (i.e. if the value of one WC has some non-trivial relationship to the value of another WC). In Section 12.2.1, we will step through the method by which we identify potentially interesting pairs of correlated WCs. Then, in Section 12.2.2, the pairs of correlated WCs will be discussed and the two-dimensional scans for the pairs will be presented. Finally, Section 12.2.3 will discuss various factors that can affect the correlations.

12.2.1 Methodology of identification of correlations

In order to identify pairs of potentially interesting WCs, we might first look to the correlation matrix from the likelihood fit. While this provides information about correlations near the best fit point, the correlations may in principle vary across the 26-dimensional WC space. Thus, in order to explore the relationships among WCs throughout the full range of values explored within the 2σ intervals, we will examine

the values of the profiled WCs for each scan.

For example, let us consider the one-dimensional scan over the $c_t^{S(\ell)}$ WC, with the other 25 WCs profiled. While there are 25 profiled WCs to explore, let us first consider c_{tG}. For each of the scan points along the $c_t^{S(\ell)}$ direction, we can plot the value taken on by the profiled parameter c_{tG}, as shown in Figure 12.2 (a). In this plot, the profiled values of c_{tG} are plotted on the y axis for each of the points along the $c_t^{S(\ell)}$ likelihood scan, which are shown on the x axis. For a given value of the scanned parameter $c_t^{S(\ell)}$, the profiled value of c_{tG} corresponds to the value of c_{tG} that maximizes the likelihood at the given value of $c_t^{S(\ell)}$. Since the profiled values of c_{tG} are zero across the full range of the $c_t^{S(\ell)}$ scan, we may conclude that c_{tG} and $c_t^{S(\ell)}$ are not correlated.

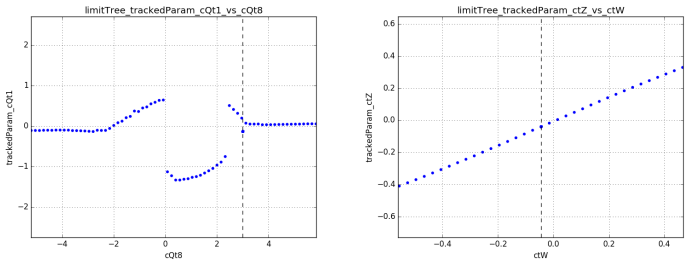


(a) (b) (c)

Figure 12.2. Plots showing profiled WC values (on the y axis) vs the value of the scanned parameter (on the x axis). In (a), the scanned parameter is $c_t^{S(\ell)}$ and the profiled WC shown is c_{tG}. In (b), the scanned WC is $c_{\varphi Q}^-$ and the profiled WC shown is c_{tt}^1. In (c), the scanned parameter is again $c_{\varphi Q}^-$, but this time the profiled parameter shown is c_{tG}. Figure (a) shows essentially no signs of interplay between the WCs, (b) shows very minimal interplay, and figure (c) shows potentially moderate interplay.

We may then repeat this procedure for each of the other 24 profiled WCs in the $c_t^{S(\ell)}$ scan, and similarly repeat the procedure for the scans over the other 25 WCs as well. In total, this results in $26 \times 26 - 26 = 650$ total plots to consider (not including the 26 plots that would show a WC vs itself). The majority of the 650 plots are either entirely flat (e.g. Figure 12.2 (a)), nearly flat (e.g. Figure 12.2 (b)), or show only moderate signs of correlation (e.g. Figure 12.2 (c)).

However, some pairs of WCs show clearly significant indications of interplay between the WCs. For example, in Figure 12.3 (a), we see that the value of the profiled parameter has a feature in its shape; a feature like this may arise if the profile fit is avoiding a "hill" in the likelihood surface, as discussed in Chapter 10. Figure 12.3 (b) shows an example of two WCs with a linear correlation. The pairs of WCs with features like the examples shown in Figure 12.3 are cases that may be interesting to explore in more detail (i.e. with two-dimensional scans).



(a) (b)

Figure 12.3. Plots showing profiled WC values (on the y axis) vs the value of the scanned parameter (on the x axis). In (a), the scanned parameter is c_{Qt}^8 and the profiled WC shown is c_{Qt}^1. In (b), the scanned WC is c_{tW} and the profiled WC shown is c_{tZ}. Figure (a) shows that there is some non-trivial correlation between the two WCs, and (b) also shows a correlation between the WCs (in this case the correlation is linear).

All 650 of the profiled WC vs scanned WC plots are shown together in Figure 12.4. In this "matrix" of plots, the first six rows/columns correspond to the two-light-two-heavy WCs, the next four correspond to the four-heavy WCs, the next nine correspond to the two-heavy-plus-boson WCs, and the final seven correspond to the two-heavy-two-lepton WCs. The diagonal of the matrix contains the plots that show a WC vs itself; by definition a WC must be correlated with itself, so the plots along the diagonal do not provide us with any new information. However, examining the off-diagonal plots, we can identify pairs of WCs that have potentially interesting correlations by looking for plots that show non-trivial relationships between the WCs.

In Figure 12.4, the background color of the individual plots indicates the approximate degree to which the WCs are correlated. Plots with a white background color show essentially no signs of correlation, as exemplified in Figure Figures 12.2 (a). The light grey and light blue plots show minimal or moderate signs of interplay, as exemplified by Figures 12.2 (b) and (c), respectively. The orange plots show pairs of WCs where there is clearly interplay between the WCs, e.g. as shown in Figure 12.3 (a), while the red plots show pairs of WCs with a linear correlation, e.g. as shown in Figure 12.3 (b). In the following section, we will explore the pairs of WCs that show significant signs of correlation.

12.2.2 Correlated pairs of WCs

The pairs of WCs that have clear signs of interplay are indicated with red or orange backgrounds in Figure 12.4. In this section, we will discuss these WCs and examine two-dimensional scans for these pairs. The full set of these correlated pairs of WCs is listed in Table 12.1.

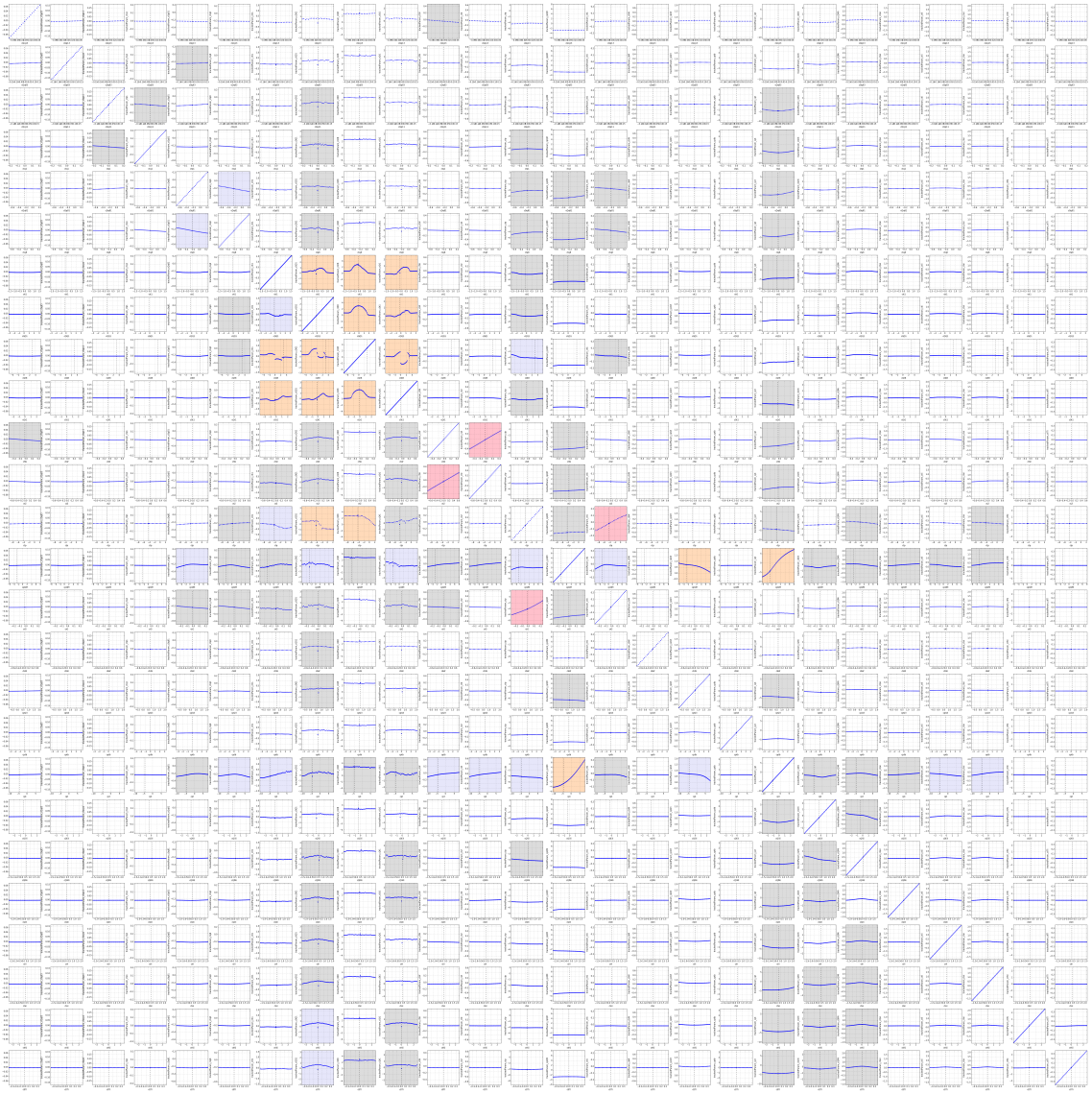


Figure 12.4. All 676 profiled WC vs scanned WC plots. The plots without a background color show no signs of correlation between the WCs (i.e. the plot is essentially flat), grey indicates minimal correlation, blue indicates potentially moderate correlation, orange indicates significant correlation between the WCs, and red indicates a linear correlation. The order of the WCs in the rows and columns is as follows: two-light-two-heavy WCs (c_{Qq}^{31}, c_{Qq}^{38}, c_{Qq}^{11}, c_{tq}^{1}, c_{Qq}^{18}, c_{tq}^{8}), four-heavy WCs (c_{tt}^{1}, c_{QQ}^{1}, c_{Qt}^{8}, c_{Qt}^{1}), two-heavy-plus-boson WCs (c_{tW}, c_{tZ}, $c_{t\varphi}$, $c_{\varphi Q}^{-}$, c_{tG}, c_{bW}, $c_{\varphi Q}^{3}$, $c_{\varphi tb}$, $c_{\varphi t}$), and two-heavy-two-lepton WCs ($c_{Q\ell}^{3(\ell)}$, $c_{Q\ell}^{-(\ell)}$, $c_{Qe}^{(\ell)}$, $c_{te}^{(\ell)}$, $c_{t}^{S(\ell)}$, $c_{t}^{T(\ell)}$).

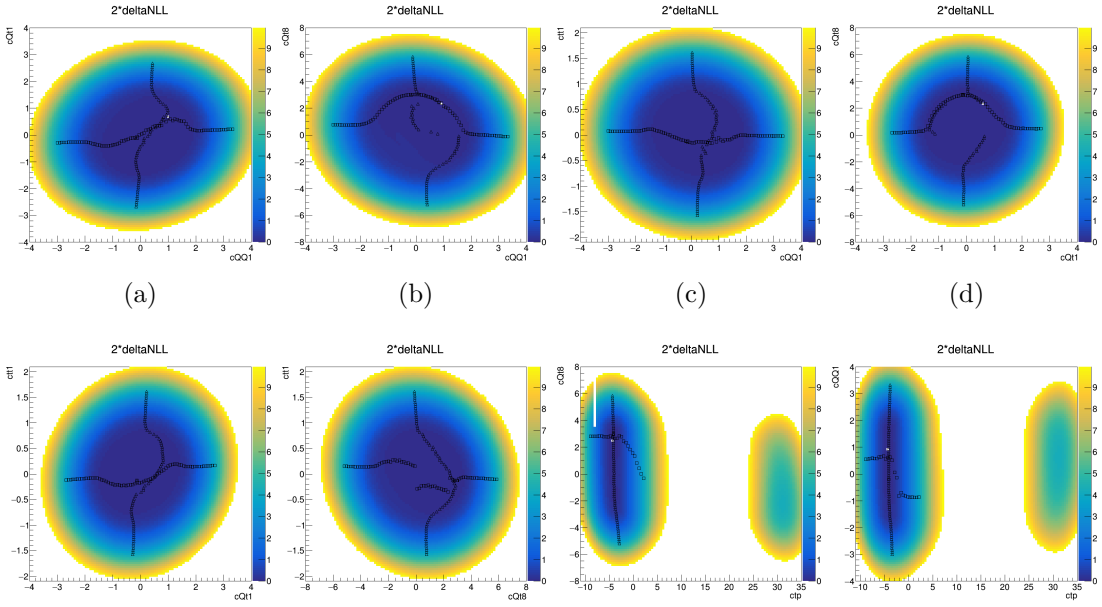
TABLE 12.1

PAIRS OF CORRELATED WCS BASED ON FIGURE 12.4.

Four heavy - four heavy:
c_{QQ}^{1}- c_{Qt}^{1} , c_{QQ}^{1}- c_{Qt}^{8} , c_{QQ}^{1}- c_{tt}^{1} , c_{Qt}^{1}- c_{Qt}^{8} , c_{Qt}^{1}- c_{tt}^{1} , c_{Qt}^{8}- c_{tt}^{1}
Four heavy - two heavy with bosons:
c_{QQ}^{1}- $c_{t\varphi}$, c_{Qt}^{8}- $c_{t\varphi}$
Two heavy with bosons - two heavy with bosons:
$c_{\varphi Q}^{-}$- $c_{\varphi t}$, $c_{\varphi Q}^{-}$- $c_{\varphi Q}^{3}$, c_{tZ}- c_{tW} , c_{tG}- $c_{t\varphi}$

All six pairs of four-heavy WCs show signs of correlations. Since these WCs all impact $t\bar{t}t\bar{t}$, it is not unexpected to observed correlations. The two-dimensional scans over each of these pairs of WCs is shown in Figure 12.5 (a) through (f). The color scale on the plots indicates the NLL (as explained in Chapter 10), and the overlaid black markers show the paths of the one-dimensional scans (within the 2σ range of the scan) for each of the WCs shown in the plot. Two of the four-heavy WCs also have some interplay with $c_{t\varphi}$. The $c_{t\varphi}$ WC affects $t\bar{t}t\bar{t}$ (as can be seen in Figure D.17), so correlation with the four-heavy WCs is not unexpected. The two-dimensional scans for each of these pairs of WCs is shown in (g) and (h) of Figure 12.5.

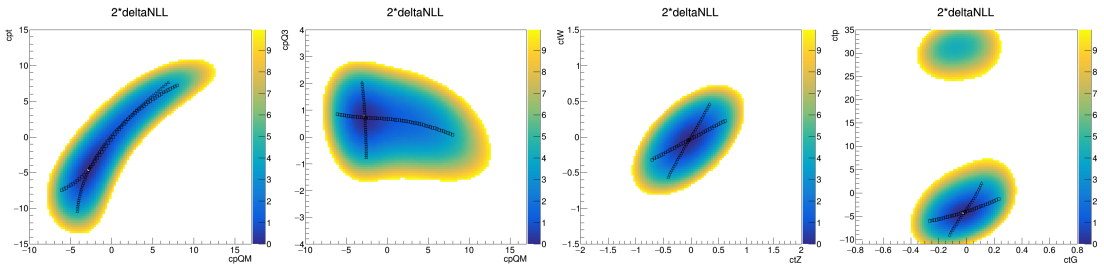
There are also many correlations among the two-heavy-with-boson WCs. The correlations seem to arise primarily among WCs that have a strong effect on $t\bar{t}Z$ (e.g. $c_{\varphi Q}^{-}$, $c_{\varphi t}$, and c_{tZ}), but c_{tG} and $c_{t\varphi}$ (which primarily affect $t\bar{t}H$) also show correlations. The two-dimensional scans over each of these pairs of WCs are shown in Figure 12.6.



(a) (b) (c) (d)

(e) (f) (g) (h)

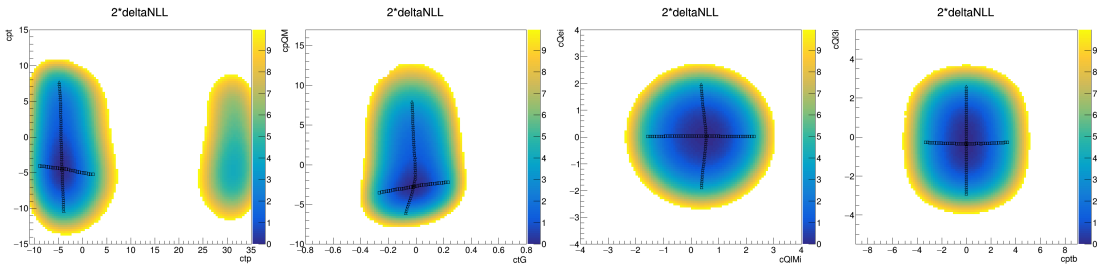
Figure 12.5. Two-dimensional scans over the pairs of WCs from Table 12.1 involving four-heavy WCs. The other 24 WCs are profiled in the likelihood fit. The color indicates the NLL, and the overlaid black markers show the paths of the one-dimensional scan (within the 2σ range of the scan) for each of the WCs shown in the plot. The path of the one-dimensional scan along the x-axis WC is shown with square markers, and the path of the one-dimensional scan along the y-axis WC is shown with triangular markers. Figure (a) shows c^1_{QQ}-c^1_{Qt}, (b) shows c^1_{QQ}-c^8_{Qt}, (c) shows c^1_{QQ}-c^1_{tt}, (d) shows c^1_{Qt}-c^8_{Qt}, (e) shows c^1_{Qt}-c^1_{tt}, (f) shows c^8_{Qt}-c^1_{tt}, (g) shows $c_{t\varphi}$-c^8_{Qt}, and (h) shows $c_{t\varphi}$-c^1_{QQ}.



(a) (b) (c) (d)

Figure 12.6. Two-dimensional scans over the pairs of WCs from Table 12.1 involving only two-heavy-with-boson WCs. Figure (a) shows $c_{\varphi Q}^{-}$-$c_{\varphi t}$, (b) shows $c_{\varphi Q}^{-}$-$c_{\varphi Q}^{3}$, (c) shows c_{tZ}-c_{tW}, and (d) shows c_{tG}-$c_{t\varphi}$. All other relevant details of the plots are described in the caption to Figure 12.5.

Although this section has focused primarily on correlated WCs, we should keep in mind that the vast majority of pairs show little to no signs of correlation. Figures 12.7 (a) and (b) show examples of WCs with only a small amount of interplay (i.e. examples of pairs that are indicated with a grey or blue background color in Figure 12.4). The pairs shown in this example are $c_{t\varphi}$-$c_{\varphi t}$ and c_{tG}-$c_{\varphi Q}^{-}$. Examples of WCs with essentially no correlation (i.e. a white background color in Figure 12.4) are shown in Figure 12.7 (c) and (d). The pairs in this example are $c_{Q\ell}^{-(\ell)}$-$c_{Qe}^{(\ell)}$ and $c_{\varphi tb}$-$c_{Q\ell}^{3(\ell)}$.



(a) (b) (c) (d)

Figure 12.7. Two-dimensional scans over the pairs of WCs that do not show signs of strong correlations. Figure (a) shows $c_{t\varphi}$-$c_{\varphi t}$, (b) shows c_{tG}-$c_{\varphi Q}^{-}$, (c) shows $c_{Q\ell}^{-(\ell)}$-$c_{Qe}^{(\ell)}$, and (d) shows $c_{\varphi tb}$-$c_{Q\ell}^{3(\ell)}$. All other relevant details of the plots are described in the caption to Figure 12.5.

12.2.3 Discussion of factors affecting correlations

In the previous section, we identified and plotted pairs of correlated WCs. In this section, we will explore what these correlations may imply, and discuss various factors that can lead to such correlations.

First, to specify the terminology, a correlation between two WCs simply means that there is some relationship between their values in the likelihood fit. This relationship could be caused by a number of factors. For example, if two WCs have primarily linear effects on the same set of bins, they may exhibit a linear correlation or anti-correlation (depending on the signs of the terms) since as the fit increases the value of one WC, the fit is free to simultaneously increase or decrease the value of the other WC to compensate for the effects of the first WC. Significant interference between two WCs may also lead to a correlation between the WCs when the interference term (which may be positive or negative) is able to compensate for the quadratic pieces (which may only be positive). The c_{tW} and c_{tZ} WCs represent an example of this type of correlation, which will be discussed further in Section 12.3.

It is also interesting to consider cases in which the WCs in question are dominated by purely quadratic components (i.e. cases in which all interference terms are small compared to the purely quadratic contributions). In this scenario, the WCs will not be able to compensate for one another, as turning any WC to a non-zero value can only increase the yield. Thus, in the SM scenario (where the prediction is exactly equal to the observation), the likelihood can never be improved by turning a second WC to a non-zero value, so the WCs will appear to be completely uncorrelated in the likelihood fit. However, it is interesting to note that this absence of a correlation does not necessarily imply that the fit is able to distinguish between the effects of the WCs in question; when the purely quadratic terms dominant, the WCs will be uncorrelated even if their effects are completely degenerate. Thus, we should be careful to avoid the assumption that uncorrelated WCs are necessarily non degenerate.

As a likelihood fit scans across a WC, the relative contributions of the linear and quadratic pieces will change. For some scans, we may begin in a region where linear and interference terms are relevant, but eventually transition into a region where the purely quadratic contributions dominate. The scans over the four-heavy WCs provide examples of this effect. Near zero, interference terms are relevant and the WCs may play off of one another (as illustrated in the two-dimensional scans shown in Figure 12.5). However, as the value of the scanned parameter grows larger, the purely quadratic terms become dominant, causing the yield to increase monotonically, which ultimately drives the likelihood past the 2σ threshold. This explains why the profiled vs scanned parameter plots for the four-heavy WCs (e.g. Figure 12.3 (a)) show signs of correlations near zero, but do now show signs of correlations at values far from zero.

It is also important to consider cases where the observation deviates from the prediction, as these effects can further complicate the interpretation of the correlations. For example, if the observed yield is larger than the predicted yield, effects from multiple WCs may combine to match the observation. Thus, even if two WCs have no interference (i.e. their effects may only increase the predicted yield), they may become correlated with one another in cases where the observation is larger than the prediction. If the observation is less than the prediction, interference effects (both with the SM and among WCs) will be important as well, since these terms can be negative and are thus able to compensate for positive contributions arising from purely quadratic effects.

In conclusion, there are multifarious factors that may influence correlations among WCs. When exploring and interpreting the sensitivity to the WCs (as we will do in the following section, 12.3), it is important to keep in mind interferences, degeneracies, and how the observation compares to the prediction for the relevant bins in the likelihood fit.

12.3 Interpretation of sensitivity to WCs

In this section, we will discuss the sensitivity to the WCs (defined by the expected 2σ limits from the profiled likelihood fits) with the goal of understanding which bins or categories of bins provide the most important contributions to the sensitivity for each WC. While all 178 analysis bins contribute to the sensitivity to the 26 WCs, the relative contribution of each bin varies by WC. Organizing the WCs based on the relative contributions of the categories of analysis bins, the WCs may be classified into seven main groups, summarized in Table 12.2.

In order to arrive at the categorization presented in Table 12.2, we have employed two complimentary approaches. Starting from the groups of operators, we can first consider their vertices and the processes they affect in order to gain insight into the analysis bins that may be particularly impacted by the given WCs. In parallel to this "top-down" approach, we also study individual bins, comparing the prediction to the observation at the 2σ limits and taking into account the uncertainty on the prediction in order to identify the bins that contribute most significantly. While this section will not focus on the methodology through which the categorization was determined, the technical details of the "bottom-up" approach can be found in Appendix H. This section will instead aim to discuss the conclusions of the study.

It should be emphasized that the conclusions summarized in Table 12.2 represent a simplified picture of the interpretation of the sensitivity; while there are indeed some cases where the majority of the sensitivity to a WC is derived from a relatively clear subset of the analysis bins, the sensitivity to many of the WCs is provided by a diverse combination of bins across all selection categories. Furthermore, when characterizing relevant bins, it is also important to keep in mind interference and correlations among WCs. The following sections will step through each of the groups of WCs outlined in Table 12.2, discussing the subsets of bins that provide the leading contributions to the sensitivity.

TABLE 12.2

SUMMARY OF CATEGORIES THAT PROVIDE LEADING CONTRIBUTIONS TO THE SENSITIVITY FOR SUBSETS OF THE WCS.

Grouping of WCs	WCs	Lead categories
Two heavy two leptons	$c_{Q\ell}^{3(\ell)}$, $c_{Q\ell}^{-(\ell)}$, $c_{Qe}^{(\ell)}$, $c_{t\ell}^{(\ell)}$, $c_{te}^{(\ell)}$, $c_{t}^{S(\ell)}$, $c_{t}^{T(\ell)}$	3ℓ off-Z
Four heavy	c_{QQ}^{1}, c_{Qt}^{1}, c_{Qt}^{8}, c_{tt}^{1}	2ℓss
Two heavy two light "$t\bar{t}l\nu$-like"	c_{Qq}^{11}, c_{Qq}^{18}, c_{tq}^{1}, c_{tq}^{8}	2ℓss
Two heavy two light "$t\bar{l}\bar{l}q$-like"	c_{Qq}^{31}, c_{Qq}^{38}	3ℓ on-Z
Two heavy with bosons "$t\bar{t}l\bar{l}$-like"	c_{tZ}, $c_{\varphi t}$, $c_{\varphi Q}^{-}$	3ℓ on-Z and 2ℓss
Two heavy with bosons "tXq-like"	$c_{\varphi Q}^{3}$, $c_{\varphi tb}$, c_{bW}	3ℓ on-Z
Two heavy with bosons with significant impacts on many processes	c_{tG}, $c_{t\varphi}$, c_{tW}	3ℓ and 2ℓss

12.3.1 WCs from the two-heavy-two-lepton category of operators

Beginning with the WCs in the two-heavy-two-lepton group, the 3ℓ off-Z channels provide the majority of the sensitivity for these WCs. To quantify the contributions of the off-Z channels, a fit is performed with only this subset of bins included; the resulting 2σ profiled limits show that the expected sensitivity is only degraded by

about 6% compared to the results when all bins are included. It is perhaps unsurprising that the off-Z bins provide the dominant contributions to our sensitivity to this group of WCs, as these WCs are associated with four-fermion vertices that produce pairs of leptons without an intermediate Z.

12.3.2 WCs from the four-heavy category of operators

The next group of WCs are those associated with the four-heavy operators. The sensitivity to these WCs is provided primarily by the 2ℓss bins, with leading contributions from the bins requiring at least three b tags. Since the $t\bar{t}t\bar{t}$ process contributes significantly to these bins and the four-heavy WCs strongly impact $t\bar{t}t\bar{t}$, it is expected that these bins would contribute significantly to the sensitivity. To obtain a quantitative characterization of the sensitivity provided by the 2ℓss bins, we performed a fit with only these bins included; the resulting 2σ limits are only degraded by about 5% (with respect to a fit with all bins included), showing that the 2ℓss bins indeed represent the dominant source of sensitivity to the four-heavy WCs.

12.3.3 WCs from the two-heavy-two-light category of operators

The next set of WCs are those associated with the two-heavy-two-light category of operators. Four of these WCs (c_{Qq}^{11}, c_{tq}^{1}, c_{Qq}^{18}, and c_{tq}^{8}) primarily affect the $t\bar{t}l\nu$ process, so bins populated significantly by $t\bar{t}l\nu$ are expected to provide important contributions to the sensitivity to these WCs. Performing a fit with only the 2ℓss bins included, the expected 2σ limits are degraded by only about 5-15%. The 2ℓss bins thus provide the primary source of sensitivity for these WCs, though other bins (e.g. from the off-Z channels) also contribute to the sensitivity.

The remaining two WCs from the two-heavy-two light group (c_{Qq}^{31} and c_{Qq}^{38}) are distinct from the other two-heavy-two-light WCs in that they feature $t-b-q-q'$ vertices. These vertices allow c_{Qq}^{31} and c_{Qq}^{38} to significantly impact the $tl\bar{l}q$ process in

3ℓ on-Z bins with two b tags and low jet multiplicity (as discussed in Section 6.3). The on-Z bins thus contribute significant sensitivity to these WCs. While the 2ℓss and off-Z categories also contribute to the sensitivity to these WCs, the 3ℓ on-Z bins provide the leading contribution; the expected 2σ intervals for these WCs loosen by more than 30% when the 3ℓ on-Z bins are excluded from the fit.

12.3.4 WCs from the two-heavy-with-bosons category of operators

The final set of WCs are those associated with the two-heavy-with-bosons category of operators. These nine WCs impact a broad range of processes, leading to diverse effects across the full set of 178 analysis bins and making it challenging to definitively characterize subsets of bins that provide dominant contributions to the sensitivity. However, the WCs can be classified into three main groups (as listed in Table 12.2) based on the processed they impact most significantly.

The c_{tZ}, $c_{\varphi Q}^{-}$, and $c_{\varphi t}$ WCs feature t$-$t$-$Z EFT vertices and primarily affect the $t\bar{t}l\bar{l}$ process; the on-Z bins are thus important for these WCs. However, these WCs also impact other processes (e.g. $t\bar{t}t\bar{t}$), meaning other categories of bins can also provide important sensitivity. Furthermore, $t\bar{t}l\bar{l}$ also significantly populates the 2ℓss bins (making up about 20% of the total expected yield), so the $t\bar{t}l\bar{l}$ effects can also be relevant in the 2ℓss bins. Thus, the 3ℓ on-Z bins and 2ℓss bins are important for these WCs. The 3ℓ off-Z bins provide a smaller (though non-zero) contribution to the sensitivity; performing a fit with these bins excluded results in an approximately 6% degradation of the expected 2σ confidence intervals.

Next, let us consider $c_{\varphi Q}^{3}$, $c_{\varphi tb}$, and c_{bW}. These WCs primarily impact $tl\bar{l}q$ and tHq, and their sensitivity arises from multiple categories of analysis bins. The 3ℓ on-Z bins represent the leading (though not overwhelmingly dominant) contribution. Performing a fit with only the 3ℓ on-Z bins included, the expected 2σ limits for these WCs loosen by an average of about 10% compared to a fit with all bins included.

The final three WCs from the two-heavy-with-bosons group are c_{tG}, $c_{t\varphi}$, and c_{tW}. Impacting multiple processes, these WCs gain sensitivity from the full spectrum of analysis bins. For example, c_{tG} impacts $t\bar{t}H$ (so the 2ℓss and 3ℓ off-Z bins are important as $t\bar{t}H$ significantly populates these bins) but also strongly impacts $t\bar{t}l\bar{l}$ (so the on-Z and 2ℓss bins also play an important role). The $c_{t\varphi}$ WC significantly impacts $t\bar{t}H$, tHq, and $t\bar{t}t\bar{t}$; most of the analysis bins provide sensitivity to this WC, though the on-Z bins provide only minor contributions (dropping the on-Z bins only results in about a 5% effect on the expected 2σ profiled limits for $c_{t\varphi}$). Finally, the c_{tW} WC impacts all signal processes and derives important sensitivity from a variety of analysis bins. Further complicating the picture, c_{tW} has significant interference with c_{tZ}, and the two WCs have a strong linear correlation in the profiled fit (as shown in Figure 12.6). Thus, when we consider the 2σ profiled limits for c_{tW}, it is important to recall that the c_{tZ} operator is also set to a non-zero value, so bins that c_{tZ} affects can also be important when considering the sensitivity to c_{tW}.

12.4 Impacts of uncertainties

The precision of this analysis is limited by both statistical and systematic uncertainties, with the relative importance of each source of uncertainty varying by WC. In Section 12.4.1, we will discuss how the total effect of the systematic uncertainties compares to the statistical uncertainty. In Section 12.4.2, the dominant sources of systematic uncertainty will be explored.

12.4.1 Relative importance of statistical and systematic contributions

To understand whether the analysis is dominated by statistics or systematics, we will begin by exploring the likelihood scans with Asimov data, in the case where all of the other 25 WCs are fixed to their SM values of zero. This simple case will allow us to probe the effects of the uncertainties without complications due to correlations among WCs or statistical fluctuations in the data. The resulting 1σ confidence intervals extracted for each WC essentially correspond to the 1σ error bars on the best fit value of the WC (which must be zero in this case, since we are fitting to Asimov data). Let us refer to this total uncertainty as σ_{Tot}.

The total uncertainty σ_{Tot} includes both statistical and systematic components; the goal of this study is to determine the relative contribution of statistical and systematic sources to σ_{Tot}. Assuming the statistical and systematic components are independent, σ_{Tot} should correspond to the quadrature sum of the components:

$$\sigma_{\text{Tot}} = \sqrt{\sigma_{\text{Stat}}^2 + \sigma_{\text{Syst}}^2}, \tag{12.1}$$

where the statistical uncertainty is σ_{Stat} and systematic uncertainty is σ_{Syst}. Although σ_{Tot} is known, σ_{Stat} and σ_{Syst} are not yet known. Fortunately, it is relatively straightforward to obtain σ_{Stat}, and once this is known we can simply solve Eq. 12.1 to obtain σ_{Syst} as well. To find σ_{Stat}, we can perform the likelihood fit without the

systematic uncertainties (i.e. nuisance parameters) included in the fit. Since the statistical uncertainties are the only sources of uncertainty in this fit, the resulting 1σ confidence intervals for each WC from will correspond to the σ_{Stat} for each WC. We can then solve Eq. 12.1 to find σ_{Syst} for each WC as well.

Following this procedure, we obtain σ_{Stat} and σ_{Syst} for all 26 WCs. We can thus compare the size of σ_{Stat} and σ_{Syst} for each WC in order to understand if the analysis is dominated by statistical or systematic uncertainties. The results of this comparison show that for the majority of the WCs, σ_{Stat} is at least as large or larger that σ_{Syst} (by up to a factor of $\sim$4, depending on the WC). However, there are also cases where σ_{Syst} is larger than σ_{Stat} (by up to a factor of about $\sim$2 to about $\sim$3, depending on the WC). The WCs for which the systematic uncertainty is significantly larger than the statistical uncertainty are c_{Qq}^{18} and c_{tq}^{8} (from the two-heavy-two-light category of WCs) and $c_{t\varphi}$, $c_{\varphi Q}^{-}$, c_{tG}, and $c_{\varphi t}$ (from the two-heavy-with-bosons category of WCs). Section 12.4.2 will explore the dominant sources of systematic uncertainty for each of these WCs.

12.4.2 Impacts of the systematic uncertainties

In this section, we will step through the leading systematic uncertainties for each WC (as determined by the impact of the systematic on the best fit value of the WC, described in Appendix I).

For all four-heavy WCs, the leading systematic is the uncertainty on the NLO $t\bar{t}t\bar{t}$ cross section to which the LO $t\bar{t}t\bar{t}$ samples are normalized. Other systematics that have large impacts for the four-heavy WCs are the ISR systematic and the renormalization systematic. However, as discussed in Section 12.4.1, the four-heavy WCs are not dominated by systematics. Rather, for these WCs, the statistical uncertainty is larger than the total systematic uncertainty by a factor of about two.

For the two-heavy-two-lepton WCs, the leading systematic uncertainties vary by

WC, with the uncertainty on the fake estimation, the electron scale factor systematic, and the diboson N_{jet} systematic frequently making the largest impacts. However, like the four-heavy WCs, this group of WCs is not dominated by systematic uncertainties. The statistical uncertainty is generally larger than the systematic contribution by a factor of approximately three.

Next, let us consider the two-heavy-two-lepton WCs. One of the leading systematic uncertainties for these WCs is the uncertainty on the NLO $t\bar{t}l\nu$ cross section (to which we normalize the LO $t\bar{t}l\nu$ sample). Other theoretical uncertainties, such as renormalization and factorization, rank fairly high as well. Some experimental sources of uncertainty (primarily the uncertainty on the fake estimation and the diboson N_{jet} uncertainty) also have relatively large impacts. While the statistical uncertainty represents the dominant contribution for most of the two-heavy-two-light WCs, the systematic contribution is dominant for c_{Qq}^{18} and c_{tq}^{8} (by approximately a factor of two). For c_{Qq}^{18} and c_{tq}^{8}, the leading systematic is the uncertainty on the NLO $t\bar{t}l\nu$ cross section.

The final group of WCs are those associated with the two-heavy-with-bosons category of operators. The dominant systematic uncertainty varies by WC, but the uncertainty on the NLO $t\bar{t}l\nu$ or $t\bar{t}l\bar{l}$ cross section often represents the leading impact. The diboson N_{jet} uncertainty, the uncertainty on the electron SF, and the uncertainty on the fake estimation also have relatively large impacts. This group of WCs is evenly split between cases where the statistical contribution is larger, and cases where the systematic contribution is larger. As reported in Section 12.4.1, the systematic uncertainty dominates for $c_{\varphi t}$, $c_{\varphi Q}^{-}$, c_{tG}, and $c_{t\varphi}$. For the $c_{\varphi t}$, $c_{\varphi Q}^{-}$, and c_{tG} WCs, σ_{Syst} is larger than σ_{Stat} by a factor of approximately three, and the leading systematic is the NLO $t\bar{t}l\bar{l}$ cross-section uncertainty. For the $c_{t\varphi}$ WC, σ_{Syst} is larger than σ_{Stat} by a factor of almost two, and the leading systematic is the $t\bar{t}l\nu$ cross-section uncertainty.

12.5 Comparison of results to other analyses

In this section, the WC limits obtained by the analysis presented in this thesis will be compared against the results of other analyses. Section 12.5.1 will discuss CMS results, while Section 12.5.2 will explore the results presented by a global theory combination. For ease of notation throughout this section, the analysis presented in this thesis will be referred to as TOP-22-006.

12.5.1 Comparisons against other CMS analyses

A summary of EFT limits obtained by CMS analyses is collected in Ref. [75]. For each WC, this section will compare the tightest limit from Ref. [75] against the limit obtained by TOP-22-006. All comparisons will use the 2σ limits from the likelihood fits in which all other WCs have been fixed at their SM value of zero. The goal of the comparison is to contextualize the results of TOP-22-006 in order to help us to understand what value this analysis brings to the field and to identify opportunities for potentially beneficial combinations with other analyses.

Before examining the results of the comparison, there are two caveats that should be mentioned. First, this comparison will include the results of the recent Ref. [76] analysis, which are not yet included in [75]. Secondly, results from Ref. [4] (the direct predecessor of TOP-22-006) will be excluded from the comparison presented in this section; TOP-22-006 supersedes the results of [4], and there is no possibility of performing a combination with [4], so including [4] would not align with the goals of the comparison presented in this section; rather, Appendix J.1 provides a direct comparison of the TOP-22-006 results against the Ref. [4] results and discusses the factors contributing to the improvements in the limits.

Figure 12.8 summarizes the comparison of the TOP-22-006 results to the best results of other CMS analyses. TOP-22-006 is the only analysis to probe the two-heavy-two-lepton WCs (apart from Ref. [4], which is excluded from this comparison)

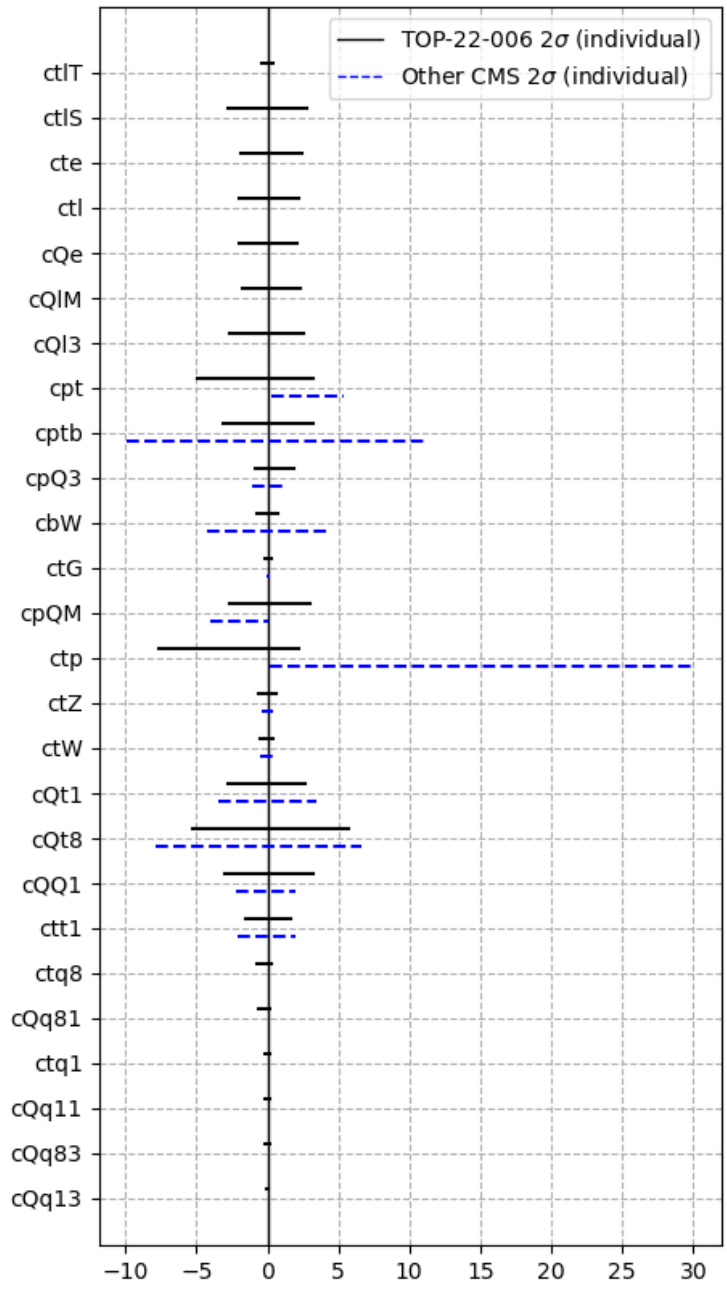


Figure 12.8. The 2σ individual limits of this analysis (indicated with the solid black line labeled TOP-22-006) compared against the individual limits obtained by other CMS analyses. The other CMS results for each WC are as follows: $c_{\varphi t}$ and $c^{-}_{\varphi Q}$ [77], c_{tW} and $c^{3}_{\varphi Q}$ [78], c_{tZ} [79], c_{tG} [80], c_{bW} $c_{\varphi tb}$ and $c_{t\varphi}$ [76], and all four-heavy WCs [81]. The details of each referenced analysis are discussed in the text. In this plot, $\Lambda = 1$ TeV.

and is the only analysis to probe the two-heavy-two-light WCs, so Figure 12.9 does not include comparisons for these WCs. For the other WCs, the TOP-22-006 2σ limits are generally comparable to or better than the best limits from other CMS analyses (by up to a factor of $\sim$6). However, for five of WCs (c^1_{QQ}, c_{tZ}, $c^-_{\varphi Q}$, $c_{\varphi t}$, and c_{tG}), the TOP-22-006 limits are approximately 40-60% looser than the best limits from other CMS analyses.

There are many differences between the analyses that may contributed to the variations in the observed results (e.g. the amount of data included, final states considered, EFT approach used, etc.). In cases where the analyses study a non-overlapping set of data events, a combination between the analyses may be interesting to pursue, as the combined result would be stronger than the results obtained by either analysis independently. A brief description of the specifics of each of the analyses featured in Figure 12.8 is provided below (where the WCs for which the analysis provides limits in Figure 12.8 are indicated parenthetically):

Ref. [81] (c^1_{Qt}, c^1_{QQ}, c^1_{tt}, c^1_{tt}): This analysis presents a search for $t\bar{t}t\bar{t}$ production in multilepton final states using data collected during 2016 (35.8 fb^{-1}). The limits from this analysis are looser than the TOP-22-006 limits (by about 30%) for all but the c^1_{QQ} WC (for which TOP-22-006 is looser by about 50%). This analysis makes use of single-lepton events and events with pairs of opposite-sign leptons, so its signal selection is orthogonal to TOP-22-006 and a combination would be potentially useful. However, the EFT approach used in this analysis (a reinterpretation of the cross section measurement) is very different than the detector-level approach used in TOP-22-006, so a combination may be technically difficult to implement.

Ref. [80] (c_{tG}): This analysis presents a measurement of the top quark polarization and $t\bar{t}$ spin correlations using data collected in 2016 (35.9 fb^{-1}). The analysis finds a better constraint on c_{tG} than TOP-22-006 (by about 50%). Opposite-sign pairs of leptons are selected, so the data does not overlap with TOP-22-006 and a combination would be potentially interesting, though the EFT approach of this analysis (a reinterpretation of a differential cross section measurement) is significantly different from the EFT approach used in TOP-22-006, so a combination may be technically challenging.

However, in future iterations of the TOP-22-006 analysis, it may be interesting to explore ways of exploiting the effect of c_{tG} on $t\bar{t}$ in order to improve the sensitivity to this WC (as well as to other WCs that strongly impact $t\bar{t}$, e.g. the two-light-two-heavy WCs).

Ref. [79] (c_{tZ}): This analysis presents a measurement of $t\bar{t}\gamma$ inclusive and differential cross sections using data collected in 2016-2018 (137 fb^{-1}). The limit on c_{tZ} from this analysis is better than the limit obtained by TOP-22-006 (by about 40%). The analysis selects events with a single lepton, so the data does not overlap with the data used in TOP-22-006, and a combination could be potentially interesting. However, the EFT approach of this analysis (a reinterpretation of the cross section measurement) is significantly different from the TOP-22-006 approach, so a combination may be technically challenging. Instead, it may be interesting to target $t\bar{t}\gamma$ events in a future iteration of the TOP-22-006 analysis in order to potentially gain additional sensitivity to the c_{tZ} WC (as well as other WCs that impact $t\bar{t}\gamma$).

Ref. [77] ($c^{-}_{\varphi Q}$, $c_{\varphi t}$): This analysis presents a measurement of top pair production in association with a Z in multilepton final states. The analysis uses data collected during 2016 and 2017 (77.5 fb^{-1}). The limits on $c^{-}_{\varphi Q}$ and $c_{\varphi t}$ obtained by this analysis are better than the limits obtained in TOP-22-006 by about 40%. The EFT approach of this analysis differs from TOP-22-006; in this analysis, the EFT effects are simulated at generator level and applied to SM samples at the detector level. The event selection overlaps with TOP-22-006, so a combination may not be useful. Instead, we may consider techniques presented in this analysis that may be benificial to implement in subsequent iterations of the TOP-22-006 analysis. For example, this analysis makes use of an angular differential variable (corresponding to the cosine of the angle between the direction of the Z in the detector frame and the direction of the negatively charged lepton in the Z frame), which may help to provide additional sensitivity to these WCs.

Ref. [78] ($c^{3}_{\varphi Q}$, c_{tW}): This analysis presents a search for new physics impacting top quarks produced in association with a Z. The analysis uses data collected durring 2016-2018 (138 fb^{-1}). The limits for $c^{3}_{\varphi Q}$ and c_{tW} are similar to the limits obtained by TOP-22-006 (within $\sim$15%). Like TOP-22-006, this analysis parametrizes the predicted yields in terms of the WCs in order to study the EFT effects directly at detector level. The event selection overlaps with TOP-22-006, so a combination may not be useful. However, since the two analyses study essentially the same data, use similar techniques, and derive similar results, the analyses represent a useful cross-check of one another.

Ref. [76] (c_{bW}, $c_{\varphi tb}$, $c_{t\varphi}$): This analysis presents a search for new physics impacting events in which top quarks are produced in association with a boosted Z or Higgs. The analysis uses data collected from 2016-2018 (138 fb^{-1}).

For $c_{t\varphi}$ and $c_{\varphi tb}$, the limits obtained by TOP-22-006 are tighter than the limits obtained by this analysis by a factor of $\sim$3; for c_{bW}, the TOP-22-006 limit is tighter by a factor of $\sim$6. This analysis targets events with a single lepton, so the selection is orthogonal to the TOP-22-006 selection. Furthermore, this analysis makes use of the same EFT approach as TOP-22-006 (parametrizing the event weights in terms of the WCs in order to target EFT effects directly at detector level). A combination between this analysis and TOP-22-006 would thus be useful and potentially fairly straightforward to implement; in fact, such a combination is being pursued currently. For a direct comparison of the TOP-22-006 results to the results of Ref. [76], please see Appendix J.2.

12.5.2 Comparisons against global theory combination

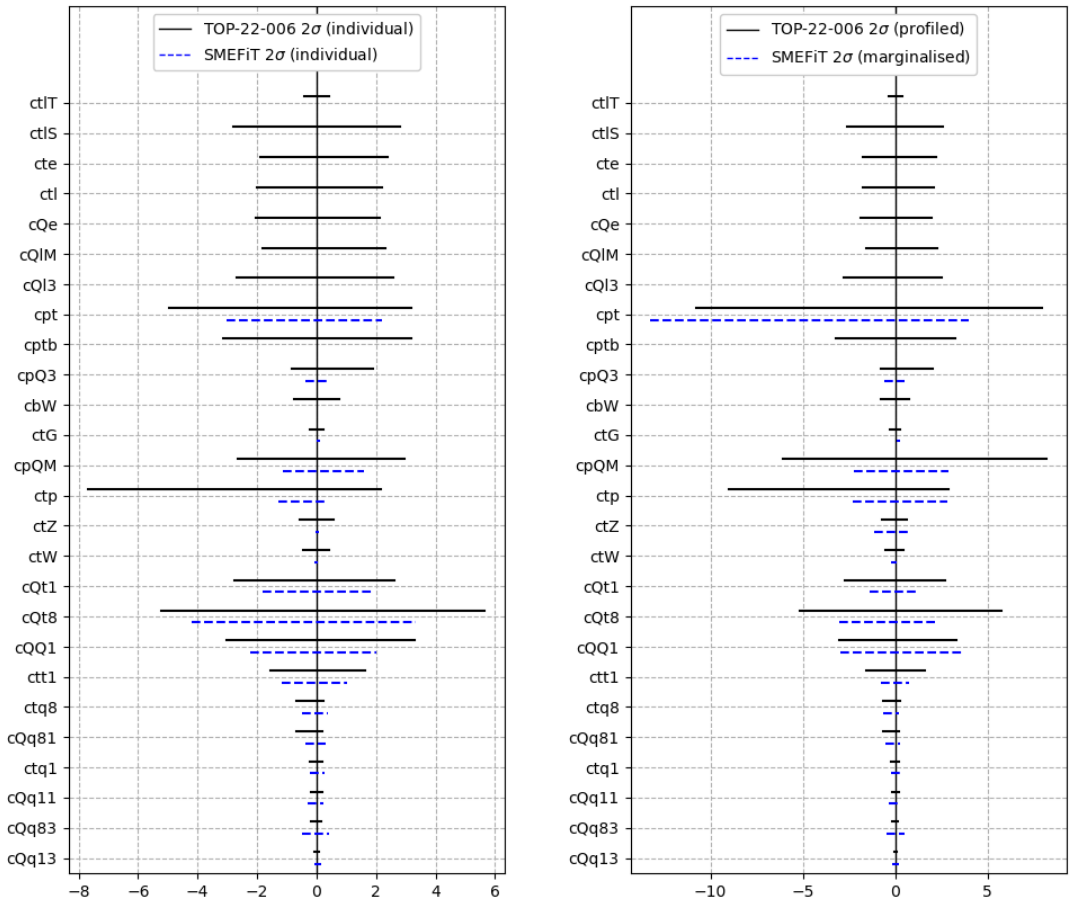
In this section, we will compare the TOP-22-006 results against the limits presented in the 2021 SMEFiT global theory combination [82]. Combining more than 300 measurements from approximately 50 papers, the SMEFiT analysis integrates a large number of processes and final states. Although global combinations can provide powerful constraints on a wide variety of EFT effects, there are many challenging aspects of such analyses. For example, it is important to ensure that a given dataset is used to extract only one measurement, or to ensure that cases of double counting are handled with a proper statistical treatment. Furthermore, it is important to keep in mind how the WCs impact processes that are considered to be backgrounds of the cross section measurements included in the combination; many WCs impact multiple processes, so challenges may be presented by cross section measurements that assume all background processes (including processes affected by the WCs of interest) are equal to the SM prediction.

Figure 12.9 shows the 2σ limits from TOP-22-006 compared against the 2σ limits obtained by the SMEFiT combination. In the case where the other WCs are fixed to zero (Figure 12.9 (a)), we see that the TOP-22-006 limits are comparable (approximately within a factor of two) to or tighter than the SMEFiT analysis for the majority of the WCs that are common to both analyses. However, for several of the WCs from the two-heavy-with-bosons category (c_{tZ}, c_{tW}, $c_{t\varphi}$, c_{tG}, and $c^3_{\varphi Q}$),

the SMEFiT limits are significantly tighter than the TOP-22-006 limits (by factors ranging from approximately three to approximately eight).

The largest difference between the TOP-22-006 and SMEFiT individual limits is associated with the c_{tZ} WC. For this WC, the SMEFiT limit is tighter than the TOP-22-006 limit by a factor of eight. The cause this large difference in sensitivity is not clear. However, it is interesting to note that the sensitivity of the SMEFiT combination to this WC improved significantly from the 2019 version of the analysis [83] to the 2021 iteration of the analysis [82]; the limit obtained by the 2021 combination is more than a factor of 50 tighter than the limit obtained by the 2019 combination. For other similar WCs (e.g. $c_{\varphi t}$ and $c^{-}_{\varphi Q}$, which also primarily impact the $t\bar{t}Z$ process), the improvement from the 2019 to 2021 SMEFiT results is not as large (a factor of $\sim$3). It would be interesting to explore the factors that contribute to the 2021 SMEFiT combination's sensitivity to c_{tZ}.

Moving from the individual likelihood fits to the profiled likelihood fits, Figure 12.9 (b) shows that the largest difference between the profiled limits is a factor of $\sim$3 (for the c_{tW} WC). It is interesting that the TOP-22-006 and SMEFiT limits are more similar for the profiled fits than for the individual fits. Since the profiled fits are in principle less comparable than the individual fits, it is possible that multiple competing effects work together to wash out the differences.



(a) (b)

Figure 12.9. The 2σ profiled limits of this analysis (indicated with the solid black line labeled TOP-22-006) compared against the limits obtained in the 2021 SMEFiT global theory combination [82] (indicated with the dashed blue lines). Figure (a) shows the results of the likelihood fits in which a single WC is fit with all other WCs fixed to their SM values of zero. Figure (b) shows the results of the likelihood fits in which the other WCs are profiled. In these plots, $\Lambda = 1$ TeV.

12.6 Interpretation of results in terms of energy scale

While EFT limits are often discussed in terms of the value of the WC c, we should recall that the variable on which we set a limit is actually c/Λ^2, where Λ is the scale of the new physics, as defined in Eq. 1.1. The scale Λ is often conventionally set to 1 TeV (e.g. in the dim6top model used in this analysis discussed in Chapter 4); this choice of Λ makes c/Λ^2=c (at least in terms of the numerical value, the units of c and c/Λ^2 of course still differ), explaining why the Λ^2 is sometimes dropped.

In fixing Λ to some assumed value, we are able to look at the implications for c; however, we can also approach the situation from the opposite perspective, assuming a value for c and exploring the implications for the scale Λ. We can express Λ in terms of c and the observed limit on c/Λ^2 as follows:

$$\text{Observed limit} = \frac{c}{\Lambda^2} \quad \Longrightarrow \quad \Lambda = \sqrt{\frac{c}{\text{Observed limit}}} \ . \tag{12.2}$$

Taking the 2σ limits from Tables 11.1 and 11.2, we can thus solve for Λ under various assumptions for c. Figure 12.10 shows the resulting values of Λ for three different assumptions for the value of c. The darkest colored bars correspond to a "small" value of c, taken to be 0.01. The medium bars correspond to $c = 1$. Finally, the lightest colored bars correspond to $c = (4\pi)^2$.

The bars in Figure 12.10 essentially show the energy scale to which this analysis has probed. Since the observed limits are inversely proportional to Λ, stronger limits correspond to longer bars on the plot (i.e. a larger excluded region). For example, for a c of 1, Figure 12.10 shows that energies up to about 1 TeV have been probed by this analysis, implying that if new physics (impacting top quarks produced in association with additional charged leptons) were to exist below a range of $\sim$1 TeV, this analysis should have been sensitive to its effects. In other words, the unexplored region beyond the $\sim$1 TeV frontier may yet hold new physics to be discovered.

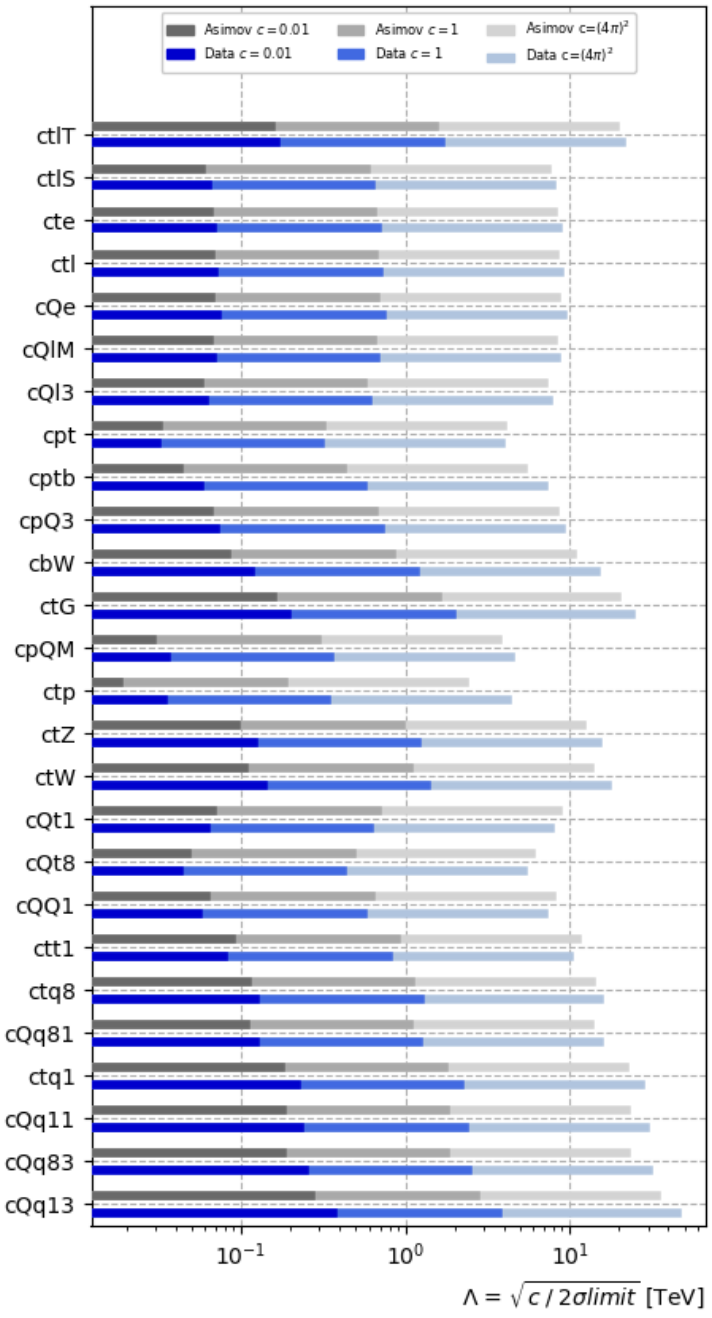


Figure 12.10. The 2σ profiled limits obtained by this analysis interpreted in terms of the energy scale Λ for fits to Asimov data (in grey) and the real data (in blue) for three different assumptions for the value of the WC as indicated in the legend and described in the text. The Asimov and observed limits are taken from the 2σ limits presented in Tables 11.1 and 11.2; for asymmetric $+2\sigma$ and -2σ limits, we have taken the absolute value of the looser limit.

CHAPTER 13

SUMMARY

Using 138 fb^{-1} of pp collisions at $\sqrt{s} = 13$ TeV collected by the CMS experiment during 2016, 2017, and 2018, this thesis has presented a search for new physics in the top sector within the context of an EFT framework. The goal of the analysis is to extract limits for the WCs of 26 dimension-six EFT operators involving top quarks. Processes that are significantly impacted by these operators correspond to the signal for this analysis; these processes constitute $t\bar{t}H$, $t\bar{t}l\nu$, $t\bar{t}l\bar{l}$, $tl\bar{l}q$, tHq, and $t\bar{t}t\bar{t}$.

This analysis focuses on multilepton signatures involving two same-sign leptons or three or more leptons. Additional jets and b-tagged jets are also required, resulting in 43 signal-region categories; to improve the sensitivity to the EFT effects, the events in each category are binned according to a differential kinematical distribution, resulting in 178 total bins. Any process that contributes to these multileptonic final-state signatures but is not significantly impacted by the 26 WCs is considered to be a background for the analysis. The largest background contributions arise from $t\bar{t}$ events with misidentified leptons and from diboson processes.

Multiple signal processes impacted by various EFT effects may all contribute to the same final-state signature, making it important to analyze all EFT effects across all channels simultaneously. To this end, we obtain detector-level EFT predictions by parameterizing the weight of each simulated event as a 26-dimensional quadratic function in terms of the WCs. For any observable bin, we are consequently able to obtain the dependence of the predicted yield on the WCs by summing the quadratic parameterizations for each of the events passing the bin's selection criteria.

Each of the 178 bins is treated as an independent Poisson measurement with a mean corresponding to the predicted yield (where the yield is a function of the 26 WCs). The likelihood corresponds to the product over the 178 bins. A statistical analysis is performed to extract the confidence intervals for the WCs by numerically minimizing the negative log of the likelihood function; scanning over one WC at a time, the other WCs are profiled in the likelihood fit. The 2σ confidence intervals for all WCs are consistent with the SM prediction. While this analysis did not identify any indications of new physics beyond the SM, the extracted limits for the WCs could be used to help to constrain theoretical models of new physics. Furthermore, the extracted limits can be interpreted in terms of the energy scale Λ probed by the experiment, providing information about the energy frontier beyond which new physics discoveries may yet lie.

There are many directions in which this analysis could be improved and expanded. Further optimization of the categorization, binning, and differential distributions would help to increase the sensitivity. It would also be beneficial to improve the EFT modeling (e.g. by generating the simulated samples at NLO), and the incorporation of an uncertainty to account for the neglected dimension-eight effects would lead to more accurate limits. To broaden the analysis, additional signal regions could be explored and more rare processes could be studied. Performing combinations with other analyses that target different final states or different EFT effects would further expand the sensitivity to potential new physics.

Looking ahead, a global combination of EFT analyses across all sectors would provide a powerful and comprehensive probe of heavy new physics at the LHC. However, before such a combination could be accomplished, many computational challenges will need to be overcome. The generation of EFT samples, the processing/histogramming of the data, and the statistical analysis of the data will all become increasingly expensive as additional data is collected and as additional processes, WCs, and categories

are studied. Each step in the analysis workfow involves unique challenges that will require innovative solutions. As our field pushes towards global EFT combinations in the pursuit of new physics at the LHC, it will be important to not only continue exploring improvements in theoretical modeling and novel analysis techniques, but also to proactively confront the navigation of computational challenges.

APPENDIX A

REFERENCE INFORMATION FOR DATA AND MC SAMPLES

This appendix includes additional details about the data samples (Section A.1) and MC samples (Section A.2) used in this analysis.

A.1 Data samples

Table A.1 lists the CMS JSON files that specify the luminosity blocked used in the analysis. The lumi blocks listed in the files (known as the "golden JSONs") exclude data that is affected by known detector issues.

TABLE A.1

JSON FILES WITH CERTIFIED LUMINOSITY BLOCKS.

Year	Lumi (fb^{-1})	Golden JSON file
2016	36.33	/afs/cern.ch/cms/CAF/CMSCOMM/COMM_DQM/certification/Collisions16/13TeV/Legacy_2016/Cert_271036-284044_13TeV_Legacy2016_Collisions16_JSON.txt
2017	41.48	/afs/cern.ch/cms/CAF/CMSCOMM/COMM_DQM/certification/Collisions17/13TeV/Legacy_2017/Cert_294927-306462_13TeV_UL2017_Collisions17_GoldenJSON.txt
2018	59.83	/afs/cern.ch/cms/CAF/CMSCOMM/COMM_DQM/certification/Collisions18/13TeV/Legacy_2018/Cert_314472-325175_13TeV_Legacy2018_Collisions18_JSON.txt

The triggers used for the 2016 data are listed in Table A.2, the triggers used for the 2017 data are listed in Table A.3, and the triggers used for the 2018 data are listed in Table A.4.

The orders of the datasets listed in tables corresponds to the order used in the dataset overlap removal procedure described in Section 4.1. As a concrete example of how this procedure works, let us step though the procedure for an example event containing two muons and an electron. Let us say the event is from 2018, so the dataset list is `DoubleMuon`, `EGamma`, `MuonEG`, `SingleMuon`; to relate this to the general description of the overlap removal procedure in Section 4.1, these datasets correspond to datasets A, B, C, and D (respectively). Depending on the specifics of the event (e.g. the p_T of the leptons), the event may pass a trigger from all of the datasets. For this example, let us assume the event is indeed present in all four datasets. This means that the analysis code will encounter this event four times (once while processing each of the four datasets).

When processing the `DoubleMuon` dataset (dataset A), the event will be accepted (as this is the first dataset in the list, so all events from this dataset are accepted). When processing `EGamma` (dataset B), we will check to see if the event passes any triggers in `DoubleMuon` (dataset A). Since the event does pass a trigger from dataset A, this event will be rejected. When processing `MuonEG` (dataset C), we will check to see if the event passes any triggers in `DoubleMuon` and `EGamma` (datasets A and B). Since the event does pass a trigger from one of these datasets (in this case, it actually passes a trigger from both datasets), the event will be rejected. The procedure will follow similarly when processing `SingleMuon` (dataset D), and the event will consequently be rejected. This procedure thus avoids double counting in order to ensure that the event will only pass the event selection once.

TABLE A.2

TRIGGERS USED TO RECORD THE 2016 DATA.

Dataset	2016 Triggers
DoubleMuon	Mu17_TrkIsoVVL_Mu8_TrkIsoVVL_DZ
	Mu17_TrkIsoVVL_Mu8_TrkIsoVVL
	Mu17_TrkIsoVVL_TkMu8_TrkIsoVVL
	Mu17_TrkIsoVVL_TkMu8_TrkIsoVVL_DZ
	TripleMu_12_10_5
DoubleEG	Ele23_Ele12_CaloIdL_TrackIdL_IsoVL
	Ele23_Ele12_CaloIdL_TrackIdL_IsoVL_DZ
	Ele16_Ele12_Ele8_CaloIdL_TrackIdL
MuonEG	Mu23_TrkIsoVVL_Ele12_CaloIdL_TrackIdL_IsoVL
	Mu23_TrkIsoVVL_Ele12_CaloIdL_TrackIdL_IsoVL_DZ
	Mu12_TrkIsoVVL_Ele23_CaloIdL_TrackIdL_IsoVL_DZ
	Mu8_DiEle12_CaloIdL_TrackIdL
	Mu8_TrkIsoVVL_Ele23_CaloIdL_TrackIdL_IsoVL
	Mu8_TrkIsoVVL_Ele23_CaloIdL_TrackIdL_IsoVL_DZ
	Mu23_TrkIsoVVL_Ele8_CaloIdL_TrackIdL_IsoVL
	Mu23_TrkIsoVVL_Ele8_CaloIdL_TrackIdL_IsoVL_DZ
	DiMu9_Ele9_CaloIdL_TrackIdL
SingleMuon	IsoMu24
	IsoTkMu24
	IsoMu22_eta2p1
	IsoTkMu22_eta2p1
	IsoMu22
	IsoTkMu22
	IsoMu27
SingleElectron	Ele27_WPTight_Gsf
	Ele25_eta2p1_WPTight_Gsf
	Ele27_eta2p1_WPLoose_Gsf

TABLE A.3

TRIGGERS USED TO RECORD THE 2017 DATA.

Dataset	2017 Triggers
DoubleMuon	Mu17_TrkIsoVVL_Mu8_TrkIsoVVL_DZ
	Mu17_TrkIsoVVL_Mu8_TrkIsoVVL_DZ_Mass3p8
	TripleMu_12_10_5
DoubleEG	Ele23_Ele12_CaloIdL_TrackIdL_IsoVL
	Ele23_Ele12_CaloIdL_TrackIdL_IsoVL_DZ
	Ele16_Ele12_Ele8_CaloIdL_TrackIdL
MuonEG	Mu23_TrkIsoVVL_Ele12_CaloIdL_TrackIdL_IsoVL
	Mu23_TrkIsoVVL_Ele12_CaloIdL_TrackIdL_IsoVL_DZ
	Mu12_TrkIsoVVL_Ele23_CaloIdL_TrackIdL_IsoVL_DZ
	Mu8_TrkIsoVVL_Ele23_CaloIdL_TrackIdL_IsoVL_DZ
	Mu8_DiEle12_CaloIdL_TrackIdL
	Mu8_DiEle12_CaloIdL_TrackIdL_DZ
	DiMu9_Ele9_CaloIdL_TrackIdL_DZ
SingleMuon	IsoMu24
	IsoMu27
SingleElectron	Ele32_WPTight_Gsf
	Ele35_WPTight_Gsf

TABLE A.4

TRIGGERS USED TO RECORD THE 2018 DATA.

Dataset	2018 Triggers
DoubleMuon	Mu17_TrkIsoVVL_Mu8_TrkIsoVVL_DZ
	Mu17_TrkIsoVVL_Mu8_TrkIsoVVL_DZ_Mass3p8
	TripleMu_12_10_5
EGamma	Ele32_WPTight_Gsf
	Ele35_WPTight_Gsf
	Ele23_Ele12_CaloIdL_TrackIdL_IsoVL
	Ele23_Ele12_CaloIdL_TrackIdL_IsoVL_DZ
	Ele16_Ele12_Ele8_CaloIdL_TrackIdL
MuonEG	Mu23_TrkIsoVVL_Ele12_CaloIdL_TrackIdL_IsoVL
	Mu23_TrkIsoVVL_Ele12_CaloIdL_TrackIdL_IsoVL_DZ
	Mu12_TrkIsoVVL_Ele23_CaloIdL_TrackIdL_IsoVL_DZ
	Mu8_TrkIsoVVL_Ele23_CaloIdL_TrackIdL_IsoVL_DZ
	Mu8_DiEle12_CaloIdL_TrackIdL
	Mu8_DiEle12_CaloIdL_TrackIdL_DZ
	DiMu9_Ele9_CaloIdL_TrackIdL_DZ
SingleMuon	IsoMu24
	IsoMu27

A.2 MC samples

The privately generated signal samples for the UL16, UL16APV, UL17, and UL18 periods are listed in Tables A.6, A.7, A.8, and A.9. Table A.5 lists the NLO cross sections used to normalize the signal samples. The files are located at the Notre Dame T3.

Table A.10 lists the central samples used for comparison against our privately produced $t\bar{l}lq$ EFT samples in order to calculate the additional systematic uncertainty that is applied to the single top samples (for which we are unable to include an additional parton in the matrix element), as described in Chapter 9.

The centrally produced background samples (CMSSW_10_6_26) used in this analysis are listed in Tables A.11, A.12, A.13, and A.14. The first section of the tables lists the samples for the processes for which we use the simulation to estimate the contribution. The second section the samples for processes that are relevant to control region, or for contributions that are estimated from data. The "TTJets*" sample was only used for the estimation of the charge-flip contributions.

TABLE A.5

THEORETICAL CROSS SECTIONS USED FOR NORMALIZING THE SIGNAL SIMULATION SAMPLES.

Process	cross section (pb)	order
$t\bar{t}H$	0.2151	NLO
$t\bar{t}l\bar{l}$	0.281	NLO
$t\bar{t}l\nu$	0.2352998	NLO
$tl\bar{l}q$	0.0758	NLO
tHq	0.07096	NLO
$t\bar{t}t\bar{t}$	0.012	NLO

TABLE A.6

PRIVATELY PRODUCED UL16 SIGNAL SAMPLES.

Process	Xsec (pb)	Events	Location
$t\bar{t}H$	0.2151	8.0M	/store/user/kmohrman/FullProduction/FullR2/UL16/Round1/Batch1/naodOnly_step/v2/nAOD_step_ttHJet_all22WCsStartPtCheckdim6TopMay20GST_run0
$t\bar{t}l\nu$	0.2353	9.1M	/store/user/kmohrman/FullProduction/FullR2/UL16/Round1/Batch1/naodOnly_step/v2/nAOD_step_ttlnuJet_all22WCsStartPtCheckdim6TopMay20GST_run0
$t\bar{t}l\bar{l}$	0.281	8.1M	/store/user/kmohrman/FullProduction/FullR2/UL16/Round1/Batch1/naodOnly_step/v2/nAOD_step_ttllNuNuJetNoHiggs_all22WCsStartPtCheckdim6TopMay20GST_run0
$tl\bar{l}q$	0.0758	7.5M	/store/user/kmohrman/FullProduction/FullR2/UL16/Round1/Batch1/naodOnly_step/v2/nAOD_step_tllq4fNoSchanWNoHiggs0p_all22WCsStartPtCheckV2dim6TopMay20GST_run0
tHq	0.07096	7.5M	/store/user/kmohrman/FullProduction/FullR2/UL16/Round1/Batch1/naodOnly_step/v2/nAOD_step_tHq4f_all22WCsStartPtCheckdim6TopMay20GST_run0
$t\bar{t}t\bar{t}$	0.0120	7.5M	/store/user/kmohrman/FullProduction/FullR2/UL16/Round1/Batch1/naodOnly_step/v2/nAOD_step_tttt_FourtopsMay3v1_run0

TABLE A.7

PRIVATELY PRODUCED UL16APV SIGNAL SAMPLES.

Process	Xsec (pb)	Events	Location
$t\bar{t}H$	0.2151	8.0M	/store/user/kmohrman/FullProduction/FullR2/UL16APV/Round1/Batch1/naodOnly_step/v2/nAOD_step_ttHJet_all22WCsStartPtCheckdim6TopMay20GST_run0
$t\bar{t}l\nu$	0.2353	9.1M	/store/user/kmohrman/FullProduction/FullR2/UL16APV/Round1/Batch1/naodOnly_step/v2/nAOD_step_ttlnuJet_all22WCsStartPtCheckdim6TopMay20GST_run0
$t\bar{t}l\bar{l}$	0.281	8.1M	/store/user/kmohrman/FullProduction/FullR2/UL16APV/Round1/Batch1/naodOnly_step/v2/nAOD_step_ttllNuNuJetNoHiggs_all22WCsStartPtCheckdim6TopMay20GST_run0
$tl\bar{l}q$	0.0758	7.5M	/store/user/kmohrman/FullProduction/FullR2/UL16APV/Round1/Batch1/naodOnly_step/v2/nAOD_step_tllq4fNoSchanWNoHiggs0p_all22WCsStartPtCheckV2dim6TopMay20GST_run0
tHq	0.07096	7.5M	/store/user/kmohrman/FullProduction/FullR2/UL16APV/Round1/Batch1/naodOnly_step/v2/nAOD_step_tHq4f_all22WCsStartPtCheckdim6TopMay20GST_run0
$t\bar{t}t\bar{t}$	0.0120	7.5M	/store/user/kmohrman/FullProduction/FullR2/UL16APV/Round1/Batch1/naodOnly_step/v2/nAOD_step_tttt_FourtopsMay3v1_run0

TABLE A.8

PRIVATELY PRODUCED UL17 SIGNAL SAMPLES.

Process	Xsec (pb)	Events	Location
$t\bar{t}H$	0.2151	15.8M	/store/user/kmohrman/FullProduction/FullR2/UL17/Round1/Batch1/naodOnly_step/v4/nAOD_step_ttHJet_all22WCsStartPtCheckdim6TopMay20GST_run0
			/store/user/kmohrman/FullProduction/FullR2/UL17/Round1/Batch2/naodOnly_step/v3/nAOD_step_ttHJet_all22WCsStartPtCheckdim6TopMay20GST_run0
			/store/user/kmohrman/FullProduction/FullR2/UL17/Round1/Batch3/naodOnly_step/v4/nAOD_step_ttHJet_all22WCsStartPtCheckdim6TopMay20GST_run0
$t\bar{t}l\nu$	0.2353	18.0M	/store/user/kmohrman/FullProduction/FullR2/UL17/Round1/Batch1/naodOnly_step/v4/nAOD_step_ttlnuJet_all22WCsStartPtCheckdim6TopMay20GST_run0
			/store/user/kmohrman/FullProduction/FullR2/UL17/Round1/Batch2/naodOnly_step/v3/nAOD_step_ttlnuJet_all22WCsStartPtCheckdim6TopMay20GST_run0
			/store/user/kmohrman/FullProduction/FullR2/UL17/Round1/Batch4/naodOnly_step/v4/nAOD_step_ttlnuJet_all22WCsStartPtCheckdim6TopMay20GST_run0
$t\bar{t}l\bar{l}$	0.281	16.0M	/store/user/kmohrman/FullProduction/FullR2/UL17/Round1/Batch1/naodOnly_step/v4/nAOD_step_ttllNuNuJetNoHiggs_all22WCsStartPtCheckdim6TopMay20GST_run0
			/store/user/kmohrman/FullProduction/FullR2/UL17/Round1/Batch2/naodOnly_step/v3/nAOD_step_ttllNuNuJetNoHiggs_all22WCsStartPtCheckdim6TopMay20GST_run0
			/store/user/kmohrman/FullProduction/FullR2/UL17/Round1/Batch3/naodOnly_step/v4/nAOD_step_ttllNuNuJetNoHiggs_all22WCsStartPtCheckdim6TopMay20GST_run0
$t\bar{l}lq$	0.0758	14.7M	/store/user/kmohrman/FullProduction/FullR2/UL17/Round1/Batch1/naodOnly_step/v4/nAOD_step_tllq4fNoSchanWNoHiggs0p_all22WCsStartPtCheckV2dim6TopMay20GST_run0
			/store/user/kmohrman/FullProduction/FullR2/UL17/Round1/Batch2/naodOnly_step/v3/nAOD_step_tllq4fNoSchanWNoHiggs0p_all22WCsStartPtCheckV2dim6TopMay20GST_run0
			/store/user/kmohrman/FullProduction/FullR2/UL17/Round1/Batch3/naodOnly_step/v4/nAOD_step_tllq4fNoSchanWNoHiggs0p_all22WCsStartPtCheckV2dim6TopMay20GST_run0
tHq	0.07096	14.8M	/store/user/kmohrman/FullProduction/FullR2/UL17/Round1/Batch1/naodOnly_step/v4/nAOD_step_tHq4f_all22WCsStartPtCheckdim6TopMay20GST_run0
			/store/user/kmohrman/FullProduction/FullR2/UL17/Round1/Batch2/naodOnly_step/v3/nAOD_step_tHq4f_all22WCsStartPtCheckdim6TopMay20GST_run0
			/store/user/kmohrman/FullProduction/FullR2/UL17/Round1/Batch3/naodOnly_step/v4/nAOD_step_tHq4f_all22WCsStartPtCheckdim6TopMay20GST_run0
$t\bar{t}t\bar{t}$	0.0120	15.0M	/store/user/kmohrman/FullProduction/FullR2/UL17/Round1/Batch4/naodOnly_step/v2/nAOD_step_tttt_FourtopsMay3v1_run0

TABLE A.9

PRIVATELY PRODUCED UL18 SIGNAL SAMPLES.

Process	Xsec (pb)	Events	Location
$t\bar{t}H$	0.2151	15.6M	/store/user/kmohrman/FullProduction/FullR2/UL18/Round1/Batch1/naodOnly_step/v5/nAOD_step_ttHJet_all22WCsStartPtCheckdim6TopMay20GST_run0
			/store/user/kmohrman/FullProduction/FullR2/UL18/Round1/Batch2/naodOnly_step/v2/nAOD_step_ttHJet_all22WCsStartPtCheckdim6TopMay20GST_run0
			/store/user/kmohrman/FullProduction/FullR2/UL18/Round1/Batch3/naodOnly_step/v2/nAOD_step_ttHJet_all22WCsStartPtCheckdim6TopMay20GST_run0
$t\bar{t}l\nu$	0.2353	18.1M	/store/user/kmohrman/FullProduction/FullR2/UL18/Round1/Batch1/naodOnly_step/v5/nAOD_step_ttlnuJet_all22WCsStartPtCheckdim6TopMay20GST_run0
			/store/user/kmohrman/FullProduction/FullR2/UL18/Round1/Batch2/naodOnly_step/v2/nAOD_step_ttlnuJet_all22WCsStartPtCheckdim6TopMay20GST_run0
			/store/user/kmohrman/FullProduction/FullR2/UL18/Round1/Batch4/naodOnly_step/v2/nAOD_step_ttlnuJet_all22WCsStartPtCheckdim6TopMay20GST_run0
$t\bar{t}l\bar{l}$	0.281	16.2M	/store/user/kmohrman/FullProduction/FullR2/UL18/Round1/Batch1/naodOnly_step/v5/nAOD_step_ttllNuNuJetNoHiggs_all22WCsStartPtCheckdim6TopMay20GST_run0
			/store/user/kmohrman/FullProduction/FullR2/UL18/Round1/Batch2/naodOnly_step/v2/nAOD_step_ttllNuNuJetNoHiggs_all22WCsStartPtCheckdim6TopMay20GST_run0
			/store/user/kmohrman/FullProduction/FullR2/UL18/Round1/Batch3/naodOnly_step/v2/nAOD_step_ttllNuNuJetNoHiggs_all22WCsStartPtCheckdim6TopMay20GST_run0
$t\bar{l}lq$	0.0758	15.0M	/store/user/kmohrman/FullProduction/FullR2/UL18/Round1/Batch1/naodOnly_step/v5/nAOD_step_tllq4fNoSchanWNoHiggs0p_all22WCsStartPtCheckV2dim6TopMay20GST_run0
			/store/user/kmohrman/FullProduction/FullR2/UL18/Round1/Batch2/naodOnly_step/v2/nAOD_step_tllq4fNoSchanWNoHiggs0p_all22WCsStartPtCheckV2dim6TopMay20GST_run0
			/store/user/kmohrman/FullProduction/FullR2/UL18/Round1/Batch3/naodOnly_step/v2/nAOD_step_tllq4fNoSchanWNoHiggs0p_all22WCsStartPtCheckV2dim6TopMay20GST_run0
tHq	0.07096	15.0M	/store/user/kmohrman/FullProduction/FullR2/UL18/Round1/Batch1/naodOnly_step/v5/nAOD_step_tHq4f_all22WCsStartPtCheckdim6TopMay20GST_run0
			/store/user/kmohrman/FullProduction/FullR2/UL18/Round1/Batch2/naodOnly_step/v2/nAOD_step_tHq4f_all22WCsStartPtCheckdim6TopMay20GST_run0
			/store/user/kmohrman/FullProduction/FullR2/UL18/Round1/Batch3/naodOnly_step/v2/nAOD_step_tHq4f_all22WCsStartPtCheckdim6TopMay20GST_run0
$t\bar{t}t\bar{t}$	0.0120	14.9M	/store/user/kmohrman/FullProduction/FullR2/UL18/Round1/Batch4/naodOnly_step/v2/nAOD_step_tttt_FourtopsMay3v1_run0

TABLE A.10

CENTRAL TZQ SAMPLES USED FOR CALCULATING THE ADDITIONAL SYSTEMATIC UNCERTAINTY THAT IS APPLIED TO THE SINGLE TOP SAMPLES.

Year	Sample
UL16APV	/tZq_ll_4f_ckm_NLO_TuneCP5_PSweights_13TeV-amcatnlo-pythia8/RunIISummer16NanoAODv6-PUMoriond17_Nano25Oct2019_102X_mcRun2_asymptotic_v7-v1/NANOAODSIM
UL16	/tZq_ll_4f_ckm_NLO_TuneCP5_PSweights_13TeV-amcatnlo-pythia8/RunIISummer16NanoAODv7-PUMoriond17_Nano02Apr2020_102X_mcRun2_asymptotic_v8-v1/NANOAODSIM
UL17	/tZq_ll_4f_ckm_NLO_TuneCP5_13TeV-amcatnlo-pythia8/RunIISummer19UL17NanoAODv2-106X_mc2017_realistic_v8-v1/NANOAODSIM
UL18	/tZq_ll_4f_ckm_NLO_TuneCP5_13TeV-amcatnlo-pythia8/RunIISummer19UL18NanoAODv2-106X_upgrade2018_realistic_v15_L1v1-v1/NANOAODSIM

TABLE A.11

LIST OF UL16APV BACKGROUND SAMPLES.

UL16APV Background Samples	Xsec (pb)
/TTGJets_TuneCP5_13TeV-amcatnloFXFX-madspin-pythia8/RunIISummer20UL16NanoAODAPVv9-106X_mcRun2_asymptotic_preVFP_v11-v2/NANOAODSIM	3.697
/TTTo2L2Nu_TuneCP5_13TeV-powheg-pythia8/RunIISummer20UL16NanoAODAPVv9-106X_mcRun2_asymptotic_preVFP_v11-v1/NANOAODSIM	87.31
/TTToSemiLeptonic_TuneCP5_13TeV-powheg-pythia8/RunIISummer20UL16NanoAODAPVv9-106X_mcRun2_asymptotic_preVFP_v11-v1/NANOAODSIM	364.351
/TTZToLL_M-1to10_TuneCP5_13TeV-amcatnlo-pythia8/RunIISummer20UL16NanoAODv9-106X_mcRun2_asymptotic_v17-v1/NANOAODSIM	0.082
/TWZToLL_thad_Wlept_5f_DR_TuneCP5_13TeV-amcatnlo-pythia8/RunIISummer20UL16NanoAODAPVv9-106X_mcRun2_asymptotic_preVFP_v11-v1/NANOAODSIM	0.003004
/TWZToLL_tlept_Whad_5f_DR_TuneCP5_13TeV-amcatnlo-pythia8/RunIISummer20UL16NanoAODAPVv9-106X_mcRun2_asymptotic_preVFP_v11-v1/NANOAODSIM	0.003004
/TWZToLL_tlept_Wlept_5f_DR_TuneCP5_13TeV-amcatnlo-pythia8/RunIISummer20UL16NanoAODAPVv9-106X_mcRun2_asymptotic_preVFP_v11-v1/NANOAODSIM	0.0015
/WWTo2L2Nu_TuneCP5_13TeV-powheg-pythia8/RunIISummer20UL16NanoAODAPVv9-106X_mcRun2_asymptotic_preVFP_v11-v1/NANOAODSIM	12.178
/WWW_4F_TuneCP5_13TeV-amcatnlo-pythia8/RunIISummer20UL16NanoAODAPVv9-106X_mcRun2_asymptotic_preVFP_v11-[v1,ext1-v1]/NANOAODSIM	0.2086
/WWZ_4F_TuneCP5_13TeV-amcatnlo-pythia8/RunIISummer20UL16NanoAODAPVv9-106X_mcRun2_asymptotic_preVFP_v11-[v1,ext1-v1]/NANOAODSIM	0.1651
/WLLJJ_WToLNu_EWK_TuneCP5_13TeV_madgraph-madspin-pythia8/RunIISummer20UL16NanoAODAPVv9-106X_mcRun2_asymptotic_preVFP_v11-v2/NANOAODSIM	5.2843
/WZZ_TuneCP5_13TeV-amcatnlo-pythia8/RunIISummer20UL16NanoAODAPVv9-106X_mcRun2_asymptotic_preVFP_v11-[v1,ext1-v1]/NANOAODSIM	0.05565
/ZZTo4L_TuneCP5_13TeV_powheg_pythia8/RunIISummer20UL16NanoAODAPVv9-106X_mcRun2_asymptotic_preVFP_v11-v1/NANOAODSIM	1.256
/GluGluToContinToZZTo2e2mu_TuneCP5_13TeV-mcfm701-pythia8/RunIISummer20UL16NanoAODAPVv9-106X_mcRun2_asymptotic_preVFP_v11-v2/NANOAODSIM	0.00319
/GluGluToContinToZZTo2e2nu_TuneCP5_13TeV-mcfm701-pythia8/RunIISummer20UL16NanoAODAPVv9-106X_mcRun2_asymptotic_preVFP_v11-v3/NANOAODSIM	0.00319
/GluGluToContinToZZTo2e2tau_TuneCP5_13TeV-mcfm701-pythia8/RunIISummer20UL16NanoAODAPVv9-106X_mcRun2_asymptotic_preVFP_v11-v2/NANOAODSIM	0.00319
/GluGluToContinToZZTo2mu2tau_TuneCP5_13TeV-mcfm701-pythia8/RunIISummer20UL16NanoAODAPVv9-106X_mcRun2_asymptotic_preVFP_v11-v2/NANOAODSIM	0.00319
/GluGluToContinToZZTo4e_TuneCP5_13TeV-mcfm701-pythia8/RunIISummer20UL16NanoAODAPVv9-106X_mcRun2_asymptotic_preVFP_v11-v2/NANOAODSIM	0.00159
/GluGluToContinToZZTo4mu_TuneCP5_13TeV-mcfm701-pythia8/RunIISummer20UL16NanoAODAPVv9-106X_mcRun2_asymptotic_preVFP_v11-v2/NANOAODSIM	0.00159
/GluGluToContinToZZTo4tau_TuneCP5_13TeV-mcfm701-pythia8/RunIISummer20UL16NanoAODAPVv9-106X_mcRun2_asymptotic_preVFP_v11-v2/NANOAODSIM	0.00159
/ZZZ_TuneCP5_13TeV-amcatnlo-pythia8/RunIISummer20UL16NanoAODAPVv9-106X_mcRun2_asymptotic_preVFP_v11-[v1,ext1-v1]/NANOAODSIM	0.01398
/ZGToLLG_01J_5f_TuneCP5_13TeV-amcatnloFXFX-pythia8/RunIISummer20UL16NanoAODAPVv9-106X_mcRun2_asymptotic_preVFP_v11-v1/NANOAODSIM	55.78
/DYJetsToLL_M-10to50_TuneCP5_13TeV-madgraphMLM-pythia8/RunIISummer20UL16NanoAODAPVv9-106X_mcRun2_asymptotic_preVFP_v11-v1/NANOAODSIM	18610.0
/DYJetsToLL_M-50_TuneCP5_13TeV-amcatnloFXFX-pythia8/RunIISummer20UL16NanoAODAPVv9-106X_mcRun2_asymptotic_preVFP_v11-v1/NANOAODSIM	6025.2
/ST_s-channel_4f_leptonDecays_TuneCP5_13TeV-amcatnlo-pythia8/RunIISummer20UL16NanoAODAPVv9-106X_mcRun2_asymptotic_preVFP_v11-v1/NANOAODSIM	3.68
/ST_t-channel_top_4f_InclusiveDecays_TuneCP5_13TeV-powheg-madspin-pythia8/RunIISummer20UL16NanoAODAPVv9-106X_mcRun2_asymptotic_preVFP_v11-v1/NANOAODSIM	136.02
/ST_t-channel_antitop_4f_InclusiveDecays_TuneCP5_13TeV-powheg-madspin-pythia8/RunIISummer20UL16NanoAODAPVv9-106X_mcRun2_asymptotic_preVFP_v11-v1/NANOAODSIM	80.95
/ST_tW_antitop_5f_inclusiveDecays_TuneCP5_13TeV-powheg-pythia8/RunIISummer20UL16NanoAODAPVv9-106X_mcRun2_asymptotic_preVFP_v11-v1/NANOAODSIM	35.85
/ST_tW_top_5f_inclusiveDecays_TuneCP5_13TeV-powheg-pythia8/RunIISummer20UL16NanoAODAPVv9-106X_mcRun2_asymptotic_preVFP_v11-v1/NANOAODSIM	35.85
/TTJets_TuneCP5_13TeV-amcatnloFXFX-pythia8/RunIISummer20UL16NanoAODAPVv9-106X_mcRun2_asymptotic_preVFP_v11-v1/NANOAODSIM	831.76
/WJetsToLNu_TuneCP5_13TeV-amcatnloFXFX-pythia8/RunIISummer20UL16NanoAODAPVv9-106X_mcRun2_asymptotic_preVFP_v11-v2/NANOAODSIM	61526.7

TABLE A.12

LIST OF UL16 BACKGROUND SAMPLES.

UL16 Background Samples	Xsec (pb)
/TTGJets_TuneCP5_13TeV-amcatnloFXFX-madspin-pythia8/RunIISummer20UL16NanoAODv9-106X_mcRun2_asymptotic_v17-v1/NANOAODSIM	3.697
/TTTo2L2Nu_TuneCP5_13TeV-powheg-pythia8/RunIISummer20UL16NanoAODv9-106X_mcRun2_asymptotic_v17-v1/NANOAODSIM	87.31
/TTToSemiLeptonic_TuneCP5_13TeV-powheg-pythia8/RunIISummer20UL16NanoAODv9-106X_mcRun2_asymptotic_v17-v1/NANOAODSIM	364.351
/TTZToLL_M-1to10_TuneCP5_13TeV-amcatnlo-pythia8/RunIISummer20UL16NanoAODv9-106X_mcRun2_asymptotic_v17-v1/NANOAODSIM	0.082
/TWZToLL_thad_Wlept_5f_DR_TuneCP5_13TeV-amcatnlo-pythia8/RunIISummer20UL16NanoAODv9-106X_mcRun2_asymptotic_v17-v1/NANOAODSIM	0.003004
/TWZToLL_tlept_Whad_5f_DR_TuneCP5_13TeV-amcatnlo-pythia8/RunIISummer20UL16NanoAODv9-106X_mcRun2_asymptotic_v17-v1/NANOAODSIM	0.003004
/TWZToLL_tlept_Wlept_5f_DR_TuneCP5_13TeV-amcatnlo-pythia8/RunIISummer20UL16NanoAODv9-106X_mcRun2_asymptotic_v17-v1/NANOAODSIM	0.0015
/WWTo2L2Nu_TuneCP5_13TeV-powheg-pythia8/RunIISummer20UL16NanoAODv9-106X_mcRun2_asymptotic_v17-v1/NANOAODSIM	12.178
/WWW_4F_TuneCP5_13TeV-amcatnlo-pythia8/RunIISummer20UL16NanoAODv9-106X_mcRun2_asymptotic_v17-[v1,ext1-v1]/NANOAODSIM	0.2086
/WWZ_4F_TuneCP5_13TeV-amcatnlo-pythia8/RunIISummer20UL16NanoAODv9-106X_mcRun2_asymptotic_v17-[v1,ext1-v1]/NANOAODSIM	0.1651
/WLLJJ_WToLNu_EWK_TuneCP5_13TeV_madgraph-madspin-pythia8/RunIISummer20UL16NanoAODv9-106X_mcRun2_asymptotic_v17-v2/NANOAODSIM	5.2843
/WZTo3LNu_mllmin4p0_TuneCP5_13TeV-powheg-pythia8/RunIISummer20UL16NanoAODv9-106X_mcRun2_asymptotic_v17-v2	5.2843
/WLLJJ_WToLNu_EWK_TuneCP5_13TeV_madgraph-madspin-pythia8/RunIISummer20UL16NanoAODAPVv9-106X_mcRun2_asymptotic_preVFP_v11-v2/NANOAODSIM	0.2353
/WZZ_TuneCP5_13TeV-amcatnlo-pythia8/RunIISummer20UL16NanoAODv9-106X_mcRun2_asymptotic_v17-[v1,ext1-v1]/NANOAODSIM	0.05565
/ZZTo4L_TuneCP5_13TeV_powheg_pythia8/RunIISummer20UL16NanoAODv9-106X_mcRun2_asymptotic_v17-v1/NANOAODSIM	1.256
/GluGluToContinToZZTo2e2mu_TuneCP5_13TeV-mcfm701-pythia8/RunIISummer20UL16NanoAODv9-106X_mcRun2_asymptotic_v17-v1/NANOAODSIM	0.00319
/GluGluToContinToZZTo2e2nu_TuneCP5_13TeV-mcfm701-pythia8/RunIISummer20UL16NanoAODv9-106X_mcRun2_asymptotic_v17-v1/NANOAODSIM	0.00319
/GluGluToContinToZZTo2e2tau_TuneCP5_13TeV-mcfm701-pythia8/RunIISummer20UL16NanoAODv9-106X_mcRun2_asymptotic_v17-v1/NANOAODSIM	0.00319
/GluGluToContinToZZTo2mu2tau_TuneCP5_13TeV-mcfm701-pythia8/RunIISummer20UL16NanoAODv9-106X_mcRun2_asymptotic_v17-v1/NANOAODSIM	0.00319
/GluGluToContinToZZTo4e_TuneCP5_13TeV-mcfm701-pythia8/RunIISummer20UL16NanoAODv9-106X_mcRun2_asymptotic_v17-v2/NANOAODSIM	0.00159
/GluGluToContinToZZTo4mu_TuneCP5_13TeV-mcfm701-pythia8/RunIISummer20UL16NanoAODv9-106X_mcRun2_asymptotic_v17-v2/NANOAODSIM	0.00159
/GluGluToContinToZZTo4tau_TuneCP5_13TeV-mcfm701-pythia8/RunIISummer20UL16NanoAODv9-106X_mcRun2_asymptotic_v17-v1/NANOAODSIM	0.00159
/ZZZ_TuneCP5_13TeV-amcatnlo-pythia8/RunIISummer20UL16NanoAODv9-106X_mcRun2_asymptotic_v17-[v1,ext1-v1]/NANOAODSIM	0.01398
/ZGToLLG_01J_5f_TuneCP5_13TeV-amcatnloFXFX-pythia8/RunIISummer20UL16NanoAODv9-106X_mcRun2_asymptotic_v17-v1/NANOAODSIM	55.78
/DYJetsToLL_M-10to50_TuneCP5_13TeV-madgraphMLM-pythia8/RunIISummer20UL16NanoAODv9-106X_mcRun2_asymptotic_v17-v1/NANOAODSIM	18610.0
/DYJetsToLL_M-50_TuneCP5_13TeV-amcatnloFXFX-pythia8/RunIISummer20UL16NanoAODv9-106X_mcRun2_asymptotic_v17-v1/NANOAODSIM	6025.2
/ST_s-channel_4f_leptonDecays_TuneCP5_13TeV-amcatnlo-pythia8/RunIISummer20UL16NanoAODv9-106X_mcRun2_asymptotic_v17-v2/NANOAODSIM	3.68
/ST_t-channel_top_4f_InclusiveDecays_TuneCP5_13TeV-powheg-madspin-pythia8/RunIISummer20UL16NanoAODv9-106X_mcRun2_asymptotic_v17-v1/NANOAODSIM	136.02
/ST_t-channel_antitop_4f_InclusiveDecays_TuneCP5_13TeV-powheg-madspin-pythia8/RunIISummer20UL16NanoAODv9-106X_mcRun2_asymptotic_v17-v1/NANOAODSIM	80.95
/ST_tW_antitop_5f_inclusiveDecays_TuneCP5_13TeV-powheg-pythia8/RunIISummer20UL16NanoAODv9-106X_mcRun2_asymptotic_v17-v2/NANOAODSIM	35.85
/ST_tW_top_5f_inclusiveDecays_TuneCP5_13TeV-powheg-pythia8/RunIISummer20UL16NanoAODv9-106X_mcRun2_asymptotic_v17-v2/NANOAODSIM	35.85
/TTJets_TuneCP5_13TeV-amcatnloFXFX-pythia8/RunIISummer20UL16NanoAODv9-106X_mcRun2_asymptotic_v17-v1/NANOAODSIM	831.76
/WJetsToLNu_TuneCP5_13TeV-madgraphMLM-pythia8/RunIISummer20UL16NanoAODv9-106X_mcRun2_asymptotic_v17-v1/NANOAODSIM	61526.7

TABLE A.13

LIST OF UL17 BACKGROUND SAMPLES.

UL17 Background Samples	Xsec (pb)
/TTGJets_TuneCP5_13TeV-amcatnloFXFX-madspin-pythia8/RunIISummer20UL17NanoAODv9-106X_mc2017_realistic_v9-v1/NANOAODSIM	3.697
/TTTo2L2Nu_TuneCP5_13TeV-powheg-pythia8/RunIISummer20UL17NanoAODv9-106X_mc2017_realistic_v9-v1/NANOAODSIM	87.31
/TTToSemiLeptonic_TuneCP5_13TeV-powheg-pythia8/RunIISummer20UL17NanoAODv9-106X_mc2017_realistic_v9-v1/NANOAODSIM	364.351
/TTZToLL_M-1to10_TuneCP5_13TeV-amcatnlo-pythia8/RunIISummer20UL17NanoAODv9-106X_mc2017_realistic_v9-v1/NANOAODSIM	0.082
/TWZToLL_thad_Wlept_5f_DR_TuneCP5_13TeV-amcatnlo-pythia8/RunIISummer20UL17NanoAODv9-106X_mc2017_realistic_v9-v1/NANOAODSIM	0.003004
/TWZToLL_tlept_Whad_5f_DR_TuneCP5_13TeV-amcatnlo-pythia8/RunIISummer20UL17NanoAODv9-106X_mc2017_realistic_v9-v1/NANOAODSIM	0.003004
/TWZToLL_tlept_Wlept_5f_DR_TuneCP5_13TeV-amcatnlo-pythia8/RunIISummer20UL17NanoAODv9-106X_mc2017_realistic_v9-v1/NANOAODSIM	0.0015
/WWTo2L2Nu_TuneCP5_13TeV-powheg-pythia8/RunIISummer20UL17NanoAODv9-106X_mc2017_realistic_v9-v2/NANOAODSIM	12.178
/WWW_4F_TuneCP5_13TeV-amcatnlo-pythia8/RunIISummer20UL17NanoAODv9-106X_mc2017_realistic_v9-[v1,ext1-v2]/NANOAODSIM	0.2086
/WWZ_4F_TuneCP5_13TeV-amcatnlo-pythia8/RunIISummer20UL17NanoAODv9-106X_mc2017_realistic_v9-v1/NANOAODSIM	0.1651
/WZTo3LNu_mllmin4p0_TuneCP5_13TeV-powheg-pythia8/RunIISummer20UL17NanoAODv9-106X_mc2017_realistic_v9-v2/NANOAODSIM	5.2843
/WLLJJ_WToLNu_EWK_TuneCP5_13TeV_madgraph-madspin-pythia8/RunIISummer20UL16NanoAODv9-106X_mcRun2_asymptotic_v17-v2/NANOAODSIM	0.2353
/WZZ_TuneCP5_13TeV-amcatnlo-pythia8/RunIISummer20UL17NanoAODv9-106X_mc2017_realistic_v9-[v1,ext1-v2]/NANOAODSIM	0.05565
/ZZTo4L_TuneCP5_13TeV_powheg_pythia8/RunIISummer20UL17NanoAODv9-106X_mc2017_realistic_v9-v2/NANOAODSIM	1.256
/GluGluToContinToZZTo2e2mu_TuneCP5_13TeV-mcfm701-pythia8/RunIISummer20UL17NanoAODv9-106X_mc2017_realistic_v9-v2/NANOAODSIM	0.00319
/GluGluToContinToZZTo2e2nu_TuneCP5_13TeV-mcfm701-pythia8/RunIISummer20UL17NanoAODv9-106X_mc2017_realistic_v9-v2/NANOAODSIM	0.00319
/GluGluToContinToZZTo2e2tau_TuneCP5_13TeV-mcfm701-pythia8/RunIISummer20UL17NanoAODv9-106X_mc2017_realistic_v9-v2/NANOAODSIM	0.00319
/GluGluToContinToZZTo2mu2tau_TuneCP5_13TeV-mcfm701-pythia8/RunIISummer20UL17NanoAODv9-106X_mc2017_realistic_v9-v2/NANOAODSIM	0.00319
/GluGluToContinToZZTo4e_TuneCP5_13TeV-mcfm701-pythia8/RunIISummer20UL17NanoAODv9-106X_mc2017_realistic_v9-v2/NANOAODSIM	0.00159
/GluGluToContinToZZTo4mu_TuneCP5_13TeV-mcfm701-pythia8/RunIISummer20UL17NanoAODv9-106X_mc2017_realistic_v9-v2/NANOAODSIM	0.00159
/GluGluToContinToZZTo4tau_TuneCP5_13TeV-mcfm701-pythia8/RunIISummer20UL17NanoAODv9-106X_mc2017_realistic_v9-v2/NANOAODSIM	0.00159
/ZZZ_TuneCP5_13TeV-amcatnlo-pythia8/RunIISummer20UL17NanoAODv9-106X_mc2017_realistic_v9-v1/NANOAODSIM	0.01398
/ZGToLLG_01J_5f_TuneCP5_13TeV-amcatnloFXFX-pythia8/RunIISummer20UL17NanoAODv9-106X_mc2017_realistic_v9-v1/NANOAODSIM	55.78
/DYJetsToLL_M-10to50_TuneCP5_13TeV-madgraphMLM-pythia8/RunIISummer20UL17NanoAODv9-106X_mc2017_realistic_v9-v1/NANOAODSIM	18610.0
/DYJetsToLL_M-50_TuneCP5_13TeV-amcatnloFXFX-pythia8/RunIISummer20UL17NanoAODv9-106X_mc2017_realistic_v9-v2/NANOAODSIM	6025.2
/ST_s-channel_4f_leptonDecays_TuneCP5_13TeV-amcatnlo-pythia8/RunIISummer20UL17NanoAODv9-106X_mc2017_realistic_v9-v1/NANOAODSIM	3.68
/ST_t-channel_top_4f_InclusiveDecays_TuneCP5_13TeV-powheg-madspin-pythia8/RunIISummer20UL17NanoAODv9-106X_mc2017_realistic_v9-v1/NANOAODSIM	136.02
/ST_t-channel_antitop_4f_InclusiveDecays_TuneCP5_13TeV-powheg-madspin-pythia8/RunIISummer20UL17NanoAODv9-106X_mc2017_realistic_v9-v1/NANOAODSIM	80.95
/ST_tW_antitop_5f_inclusiveDecays_TuneCP5_13TeV-powheg-pythia8/RunIISummer20UL17NanoAODv9-106X_mc2017_realistic_v9-v2/NANOAODSIM	35.85
/ST_tW_top_5f_inclusiveDecays_TuneCP5_13TeV-powheg-pythia8/RunIISummer20UL17NanoAODv9-106X_mc2017_realistic_v9-v2/NANOAODSIM	35.85
/TTJets_TuneCP5_13TeV-amcatnloFXFX-pythia8/RunIISummer20UL17NanoAODv9-106X_mc2017_realistic_v9-v1/NANOAODSIM	831.76
/WJetsToLNu_TuneCP5_13TeV-madgraphMLM-pythia8/RunIISummer20UL17NanoAODv9-106X_mc2017_realistic_v9-v1/NANOAODSIM	61526.7

TABLE A.14

LIST OF UL18 BACKGROUND SAMPLES.

UL18 Background Samples	Xsec (pb)
/TTGJets_TuneCP5_13TeV-amcatnloFXFX-madspin-pythia8/RunIISummer20UL18NanoAODv9-106X_upgrade2018_realistic_v16_L1v1-[v1,ext1-v1]/NANOAODSIM	3.697
/TTTo2L2Nu_TuneCP5_13TeV-powheg-pythia8/RunIISummer20UL18NanoAODv9-106X_upgrade2018_realistic_v16_L1v1-v1/NANOAODSIM	87.31
/TTToSemiLeptonic_TuneCP5_13TeV-powheg-pythia8/RunIISummer20UL18NanoAODv9-106X_upgrade2018_realistic_v16_L1v1-v1/NANOAODSIM	364.351
/TTZToLL_M-1to10_TuneCP5_13TeV-amcatnlo-pythia8/RunIISummer20UL18NanoAODv9-106X_upgrade2018_realistic_v16_L1v1-v1/NANOAODSIM	0.082
/TWZToLL_thad_Wlept_5f_DR_TuneCP5_13TeV-amcatnlo-pythia8/RunIISummer20UL18NanoAODv9-106X_upgrade2018_realistic_v16_L1v1-v1/NANOAODSIM	0.003004
/TWZToLL_tlept_Whad_5f_DR_TuneCP5_13TeV-amcatnlo-pythia8/RunIISummer20UL18NanoAODv9-106X_upgrade2018_realistic_v16_L1v1-v1/NANOAODSIM	0.003004
/TWZToLL_tlept_Wlept_5f_DR_TuneCP5_13TeV-amcatnlo-pythia8/RunIISummer20UL18NanoAODv9-106X_upgrade2018_realistic_v16_L1v1-v1/NANOAODSIM	0.0015
/WWTo2L2NuWWTo2L2Nu_TuneCP5_13TeV-powheg-pythia8/RunIISummer20UL18NanoAODv9-106X_upgrade2018_realistic_v16_L1v1-v2/NANOAODSIM	12.178
/WWW_4F_TuneCP5_13TeV-amcatnlo-pythia8/RunIISummer20UL18NanoAODv9-106X_upgrade2018_realistic_v16_L1v1-v1/NANOAODSIM	0.2086
/WWZ_4F_TuneCP5_13TeV-amcatnlo-pythia8/RunIISummer20UL18NanoAODv9-106X_upgrade2018_realistic_v16_L1v1-v1/NANOAODSIM	0.1651
/WZTo3LNu_mllmin4p0_TuneCP5_13TeV-powheg-pythia8/RunIISummer20UL18NanoAODv9-106X_upgrade2018_realistic_v16_L1v1-v2/NANOAODSIM	5.2843
/WLLJJ_WToLNu_EWK_TuneCP5_13TeV_madgraph-madspin-pythia8/RunIISummer20UL18NanoAODv9-106X_upgrade2018_realistic_v16_L1v1-v2/NANOAODSIM	0.2353
/WZZ_TuneCP5_13TeV-amcatnlo-pythia8/RunIISummer20UL18NanoAODv9-106X_upgrade2018_realistic_v16_L1v1-[v1,ext1-v2]/NANOAODSIM	0.05565
/ZZTo4L_TuneCP5_13TeV_powheg_pythia8/RunIISummer20UL18NanoAODv9-106X_upgrade2018_realistic_v16_L1v1-v2/NANOAODSIM	1.256
/GluGluToContinToZZTo2e2mu_TuneCP5_13TeV-mcfm701-pythia8/RunIISummer20UL18NanoAODv9-106X_upgrade2018_realistic_v16_L1v1-v2/NANOAODSIM	0.00319
/GluGluToContinToZZTo2e2nu_TuneCP5_13TeV-mcfm701-pythia8/RunIISummer20UL18NanoAODv9-106X_upgrade2018_realistic_v16_L1v1-v2/NANOAODSIM	0.00319
/GluGluToContinToZZTo2e2tau_TuneCP5_13TeV-mcfm701-pythia8/RunIISummer20UL18NanoAODv9-106X_upgrade2018_realistic_v16_L1v1-v2/NANOAODSIM	0.00319
/GluGluToContinToZZTo2mu2tau_TuneCP5_13TeV-mcfm701-pythia8/RunIISummer20UL18NanoAODv9-106X_upgrade2018_realistic_v16_L1v1-v2/NANOAODSIM	0.00319
/GluGluToContinToZZTo4e_TuneCP5_13TeV-mcfm701-pythia8/RunIISummer20UL18NanoAODv9-106X_upgrade2018_realistic_v16_L1v1-v2/NANOAODSIM	0.00159
/GluGluToContinToZZTo4mu_TuneCP5_13TeV-mcfm701-pythia8/RunIISummer20UL18NanoAODv9-106X_upgrade2018_realistic_v16_L1v1-v2/NANOAODSIM	0.00159
/GluGluToContinToZZTo4tau_TuneCP5_13TeV-mcfm701-pythia8/RunIISummer20UL18NanoAODv9-106X_upgrade2018_realistic_v16_L1v1-v2/NANOAODSIM	0.00159
/ZZZ_TuneCP5_13TeV-amcatnlo-pythia8/RunIISummer20UL18NanoAODv9-106X_upgrade2018_realistic_v16_L1v1-v1/NANOAODSIM	0.01398
/ZGToLLG_01J_5f_TuneCP5_13TeV-amcatnloFXFX-pythia8/RunIISummer20UL18NanoAODv9-106X_upgrade2018_realistic_v16_L1v1-v1/NANOAODSIM	55.78
/DYJetsToLL_M-10to50_TuneCP5_13TeV-madgraphMLM-pythia8/RunIISummer20UL18NanoAODv9-106X_upgrade2018_realistic_v16_L1v1-v1/NANOAODSIM	18610.0
/DYJetsToLL_M-50_TuneCP5_13TeV-amcatnloFXFX-pythia8/RunIISummer20UL18NanoAODv9-106X_upgrade2018_realistic_v16_L1v1-v2/NANOAODSIM	6025.2
/ST_s-channel_4f_leptonDecays_TuneCP5_13TeV-amcatnlo-pythia8/RunIISummer20UL18NanoAODv9-106X_upgrade2018_realistic_v16_L1v1-v1/NANOAODSIM	3.68
/ST_t-channel_top_4f_InclusiveDecays_TuneCP5_13TeV-powheg-madspin-pythia8/RunIISummer20UL18NanoAODv9-106X_upgrade2018_realistic_v16_L1v1-v1/NANOAODSIM	136.02
/ST_t-channel_antitop_4f_InclusiveDecays_TuneCP5_13TeV-powheg-madspin-pythia8/RunIISummer20UL18NanoAODv9-106X_upgrade2018_realistic_v16_L1v1-v1/NANOAODSIM	80.95
/ST_tW_antitop_5f_inclusiveDecays_TuneCP5_13TeV-powheg-pythia8/RunIISummer20UL18NanoAODv9-106X_upgrade2018_realistic_v16_L1v1-v2/NANOAODSIM	35.85
/ST_tW_top_5f_inclusiveDecays_TuneCP5_13TeV-powheg-pythia8/RunIISummer20UL18NanoAODv9-106X_upgrade2018_realistic_v16_L1v1-v2/NANOAODSIM	35.85
/TTJets_TuneCP5_13TeV-amcatnloFXFX-pythia8/RunIISummer20UL18NanoAODv9-106X_upgrade2018_realistic_v16_L1v1-v1/NANOAODSIM	831.76
/WJetsToLNu_TuneCP5_13TeV-madgraphMLM-pythia8/RunIISummer20UL18NanoAODv9-106X_upgrade2018_realistic_v16_L1v1-v1/NANOAODSIM	61526.7

APPENDIX B

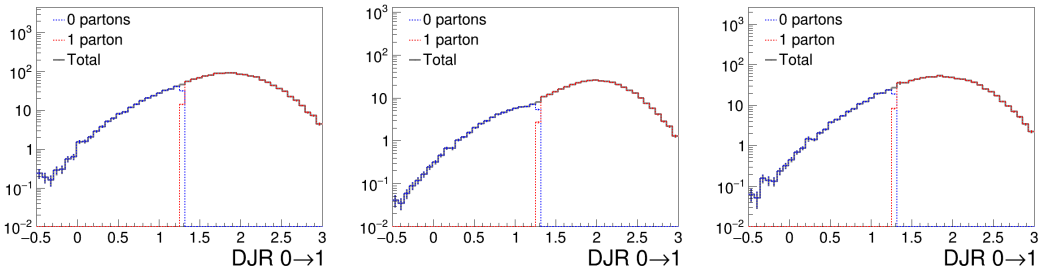
VALIDATION OF EFT SAMPLES

This appendix provides details for checks that were performed to validate the LO EFT samples produced for this analysis. In Section B.1, the validation of the LO matching procedure is discussed. In Section B.2, validation of the starting point (for the reweighting procedure) is discussed.

B.1 Validation of leading order matching procedure for EFT samples

When using LO matched samples, validation should be performed to ensure that the matching procedure is properly filling the overlapping phase space. As explained in Section 4.2.1, this validation is especially important for EFT samples, since the matching procedure can lead to complications when applied to EFT samples because EFT effects are included in the ME contribution, but not in the PS contribution. This appendix presents empirical validation checks for the matching procedure.

The differential jet rate (DJR) can be used as a method of validation for the LO matching procedure [84] [85]. For the k_T algorithm, the DJR histogram represent the distribution of k_T values for which an n jet event transitions to an $n+1$ jet event. A discontinuity in the transition between the n and $n+1$ curves would indicate that there is a mismatch in the overlapping regions of phase space, while a smooth transition is an indication that the ME generator and PS generator are working together to properly fill the phase space without any gaps or double counting. For the matched samples generated for this analysis, we observe smooth DJR plots, as shown in Figure B.1.



(a) (b) (c)

Figure B.1. DJR histograms for LO matched samples $t\bar{t}H$ (a), $t\bar{t}l\nu$ (b), and $t\bar{t}l\bar{l}$ (c). The distributions have been reweighted to a non-SM point.

In Figure B.1, all WCs have been set to non-zero values. The x axis shows the log base 10 of the scale at which an n jet event transitions into an $n+1$ jet event. The line labeled "0 partons" refers to the contribution from the parton shower, while the line labeled "1 parton" refers to the contribution from the matrix element. The line labeled "Total" is the sum of the two contributions. The smooth transition indicates that the matching scales have allowed the matrix element generator and parton shower to smoothly fill the overlapping phase space.

For the $t\bar{t}X$ samples generated for this analysis, the matching scales used with MadGraph and Pythia (i.e. the `xqcut` and `qQut`) were 10 and 20, respectively. A more detailed description of the meaning of these matching scales is provided in [10].

DJR plots should be studied at the SM and at non-SM points. It is also important to vary the matching scales around the nominal values and produce DJR plots for these variations as well. While it would not be unexpected for the DJR plots to look somewhat less smooth farther away from the nominal values (since the nominal values in principle should be chosen such that they produce smooth DJR plots), there should not be any large discontinuities. As an example of what would be considered to be a large discontinuity, Ref. [10] may be referenced (especially Figure 7).

In order to account for uncertainties associated with the matching scale, a systematic uncertainty should be evaluated. In this analysis, we evaluated this uncertainty by producing samples with the `qQut` varied around the nominal (to `qQut` = 15 and `qQut` = 25). As a quick check for large/unexpected differences caused by varying the scales, the one-dimensional inclusive cross section quadratic parameterizations for each sample (for the up, down, and nominal variations) can be compared for all WCs. Note that this check can be performed at gen level, and can be checked before producing large statistics samples for the systematic uncertainty evaluation (for this analysis, we performed this inclusive check even before producing the large statistics nominal `qQut` samples). For this analysis, the inclusive checks were consistent, so the N_{jet} yields in the signal region were studied for the up and down variations. The size of the variations were very small, and were determined to be not larger than statistical uncertainty of the sample. For this reason, a `qQut` uncertainty was not included for this analysis.

B.2 Validation of starting points for EFT samples

As discussed in Section 4.2.2, it is important to perform checks to ensure that the reweighting is working properly. This section will step through some examples of how this validation may be performed. Additionally, the slides in Ref. [86] provide an example of the types of validation checks performed for this analysis.

For this analysis, we considered several candidate starting points. We considered several similar starting point to the TOP-19-001 analysis (one that was identical to the TOP-19-001 analysis for the 16 WCs in common, and two that were based on the limits obtained in TOP-19-001), as well as starting points where the WCs scaled the inclusive cross section by a given amount (e.g. 10%, 30%, etc.). The samples can be reweighted to their own starting points, as well as to the starting points of the other candidate samples. To check for consistency, the reweighted cross section can

be compared to the cross section of the dedicated sample.

For example, let us consider sample A with a starting point of $\vec{a}$ and sample B with a starting point of $\vec{b}$. We can reweight A to $\vec{a}$, and compare the reweighted cross section against the original cross section at $\vec{a}$ (since $\vec{a}$ is the starting point of A, the original cross section at this point is included in the LHE file). Sample A can also be reweighted to $\vec{b}$, and this cross section can be compared to the original cross section at $\vec{b}$ (since this is the starting point of sample B). The cross section should also be checked at the SM. Since the SM should always included as one of the reweight points, the reweighted SM cross section can be compared against the original SM cross section for every sample.

It is expected that the agreement will be better for a sample's own starting point than at an arbitrary point in WC space. It is also expected that the agreement will be better at points that are more similar to a sample's starting point than at points that are further from the sample's starting point. However, it is difficult to define a universal threshold that determines whether a starting point is good or bad. Nevertheless, to provide some quantitive numbers for future reference, in general agreement much better than 1% was expected at the sample's own starting point (on the order of 0.001%). At reasonable points in the WC space (e.g. where all of the WCs scale the process by 30%), agreement at better than 1% was deemed to be good. At more extreme points in the WC space (e.g. points where all of the WCs scale the cross section by 50%), larger disagreement (a couple percent) was deemed to be acceptable. At very extreme points in the WC space (e.g. points where the WCs scale the cross section by 5 times the standard model cross section) even larger disagreement (e.g. 5-10%) was deemed to be acceptable, as these regions are unlikely to be explored in the analysis.

As mentioned in Section 4.2.2, the distributions of the weights at the candidate starting points should also be checked. Figures B.2 and B.3 show some example

weight distributions for the candidate starting point samples considered in this analysis. With this check, we are attempting to ensure that the chosen samples have good statistical power. We are also looking for single events with very large weights (which degrade the statistical power of the sample, and can also be an indication that the starting point was not good). To check the distribution of weights, we reweight the samples to the SM and to various points in EFT space, and plot a histogram of the values of the weights of the events in the sample. For example, in plot (c) of Figure B.2, the blue distribution ("run0") is relatively good (as its peak is very sharp and tall, indicating that most of the events have a similar weight), while the yellow distribution ("run 5") is relatively bad (as is very broad, so will not provide good statistical power).

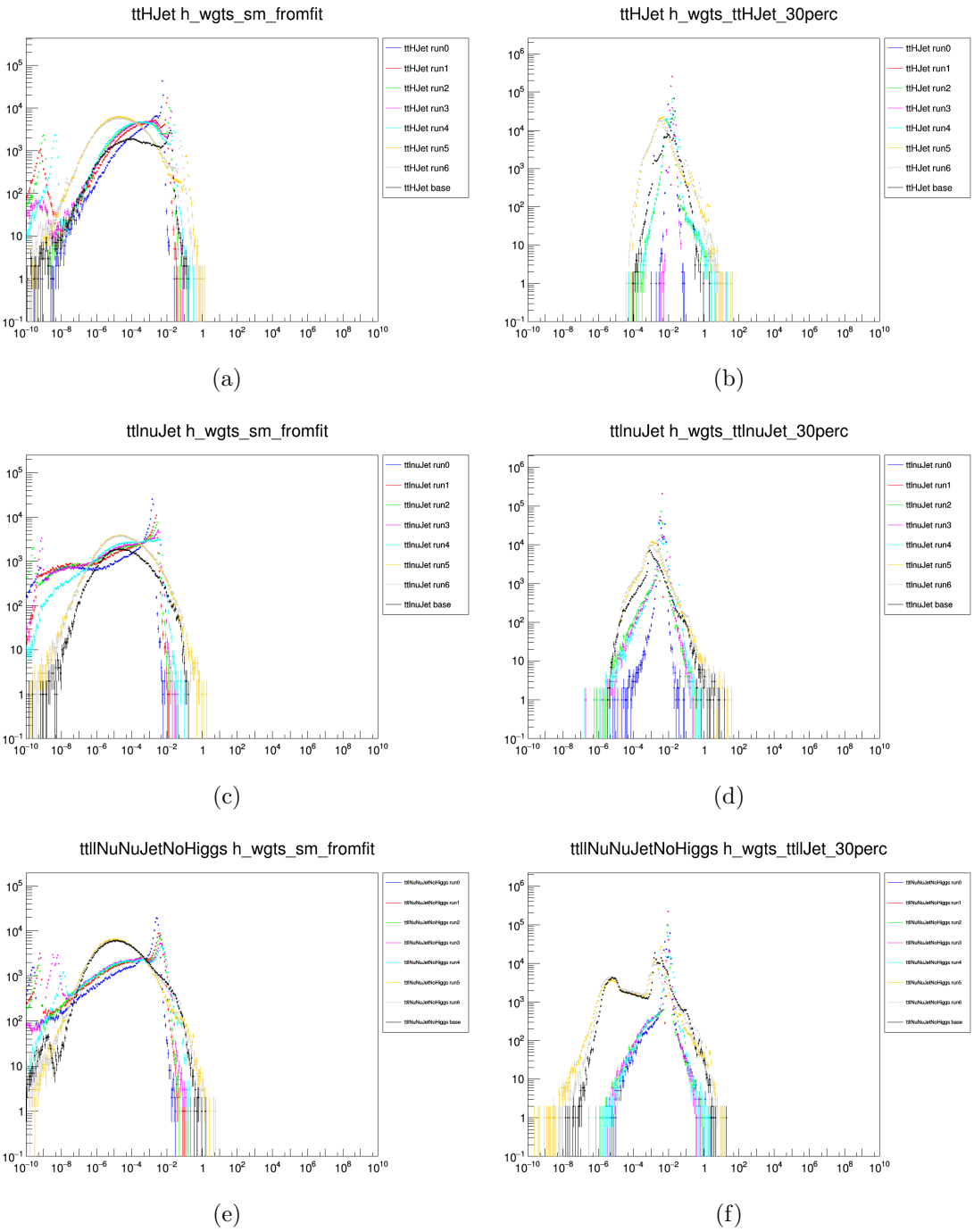


Figure B.2. Distribution of weights for $t\bar{t}H$ at the SM (a) and a non-SM point (b), $t\bar{t}l\nu$ at the SM (c) and a non-SM point (d), and $t\bar{t}l\bar{l}$ at the SM (e) and a non-SM point (f).

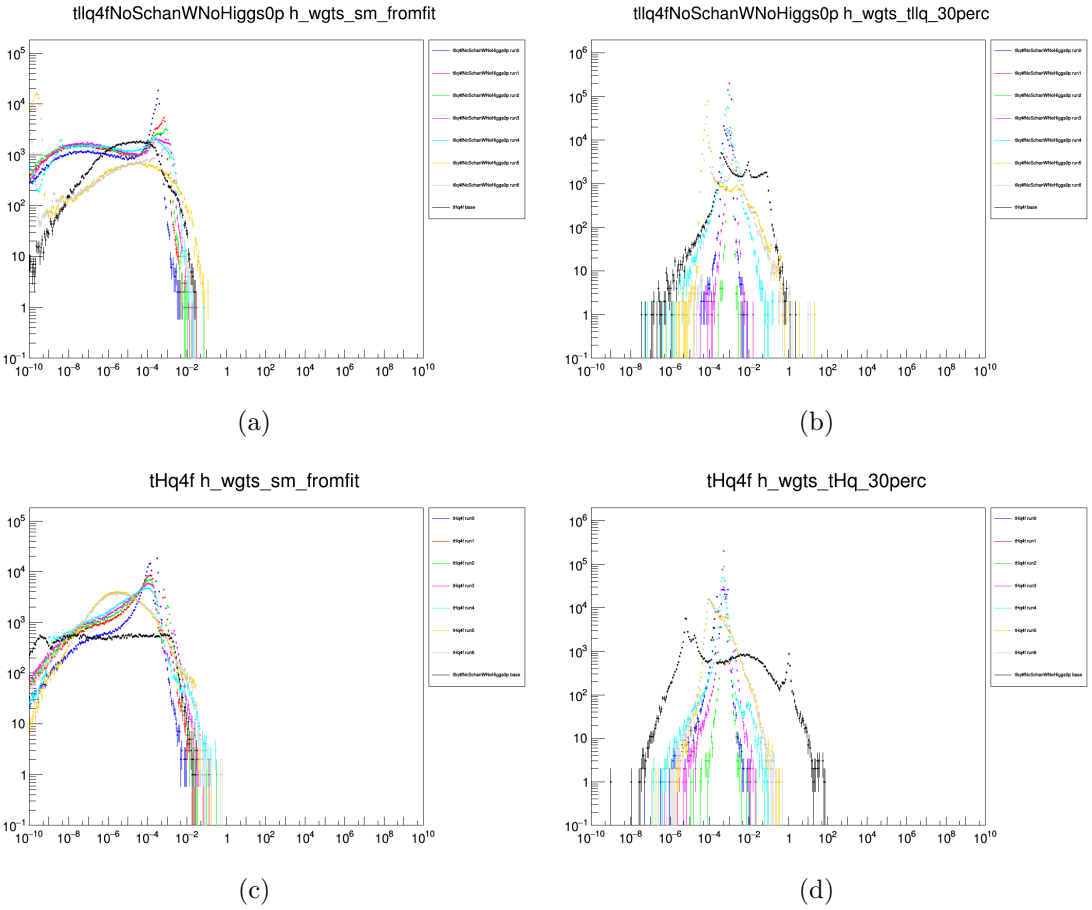


Figure B.3. Distribution of weights for $t\bar{l}lq$ at the SM (a) and a non-SM point (b) and tHq at the SM (c) and a non-SM point (d).

APPENDIX C

COMPARISON OF PRIVATELY GENERATED MC SAMPLES TO CENTRALLY GENERATED MC SAMPLES

This analysis uses privately generated LO MC samples for all signal processes, following the same approach as Ref. [4], which is referred to as TOP-19-001. This analysis uses the MG reweighting procedure to incorporate the effects of our 26 WCs into the samples, as explained in Chapter 4. The samples can be reweighted to any arbitrary point in the EFT space, including the SM point (i.e. the point where all WCs are equal to 0). When a sample is reweighted to the SM, it should equivalent (within uncertainties) to a sample that was generated at the SM. In order to verify that the samples produced for this analysis provide a good description at the SM, we can compare the predictions to centrally produced SM samples. Note that some differences are expected in these comparisons, since the central samples are NLO and the privately produced samples are LO. The central samples used for this comparison are listed in Table C.1.

The validation for each signal sample is discussed in more detail in the subsections of this appendix, and briefly summarized in the following bullets:

- $t\bar{t}H$, $t\bar{t}l\nu$, $t\bar{t}t\bar{t}$: These processes are discussed in sections C.1, C.2, and C.3 (respectively). For these processes, there is generally good agreement between privately produced LO samples and centrally produced NLO samples.
- $t\bar{t}l\bar{l}$: There is some tension for this process, as shown in the plots in section C.4. However, we do not believe this represents a problem for this analysis. The MC validation studies for TOP-19-001 (which involved all pre-UL samples) also showed this tension. While the studies for this analysis (which of course involves all UL samples) have found a somewhat larger disagreement, this seems

TABLE C.1

CENTRAL SAMPLES USED FOR COMPARISON AGAINST OUR PRIVATELY PRODUCED SAMPLES.

Year	Sample
2017	/TTZToLLNuNu_M-10_TuneCP5_PSweights_13TeV-amcatnlo-pythia8/RunIIFall17NanoAODv7-PU2017_12Apr2018_Nano02Apr2020_102X_mc2017_realistic_v8-v1/NANOAODSIM
UL16	/ttHJetToNonbb_M125_TuneCP5_13TeV_amcatnloFXFX_madspin_pythia8/RunIISummer20UL16NanoAODv9-106X_mcRun2_asymptotic_v17-v1/NANOAODSIM
UL16	/TTWJetsToLNu_TuneCP5_13TeV-amcatnloFXFX-madspin-pythia8/RunIISummer20UL16NanoAODv9-106X_mcRun2_asymptotic_v17-v1/NANOAODSIM
UL16	/TTZToLLNuNu_M-10_TuneCP5_13TeV-amcatnlo-pythia8/RunIISummer20UL16NanoAODv9-106X_mcRun2_asymptotic_v17-v1/NANOAODSIM
UL16	/tZq_ll_4f_ckm_NLO_TuneCP5_13TeV-amcatnlo-pythia8/RunIISummer20UL16NanoAODv9-106X_mcRun2_asymptotic_v17-v1/NANOAODSIM
UL16	/TTTT_TuneCP5_13TeV-amcatnlo-pythia8/RunIISummer20UL16NanoAODv9-106X_mcRun2_asymptotic_v17-v2/NANOAODSIM
UL16APV	/ttHJetToNonbb_M125_TuneCP5_13TeV_amcatnloFXFX_madspin_pythia8/RunIISummer20UL16NanoAODAPVv9-106X_mcRun2_asymptotic_preVFP_v11-v1/NANOAODSIM
UL16APV	/TTWJetsToLNu_TuneCP5_13TeV-amcatnloFXFX-madspin-pythia8/RunIISummer20UL16NanoAODAPVv9-106X_mcRun2_asymptotic_preVFP_v11-v2/NANOAODSIM
UL16APV	/TTZToLLNuNu_M-10_TuneCP5_13TeV-amcatnlo-pythia8/RunIISummer20UL16NanoAODAPVv9-106X_mcRun2_asymptotic_preVFP_v11-v1/NANOAODSIM
UL16APV	/tZq_ll_4f_ckm_NLO_TuneCP5_13TeV-amcatnlo-pythia8/RunIISummer20UL16NanoAODAPVv9-106X_mcRun2_asymptotic_preVFP_v11-v1/NANOAODSIM
UL16APV	/TTTT_TuneCP5_13TeV-amcatnlo-pythia8/RunIISummer20UL16NanoAODAPVv9-106X_mcRun2_asymptotic_preVFP_v11-v2/NANOAODSIM
UL17	/ttHJetToNonbb_M125_TuneCP5_13TeV_amcatnloFXFX_madspin_pythia8/RunIISummer20UL17NanoAODv9-106X_mc2017_realistic_v9-v1/NANOAODSIM
UL17	/TTWJetsToLNu_TuneCP5_13TeV-amcatnloFXFX-madspin-pythia8/RunIISummer20UL17NanoAODv9-106X_mc2017_realistic_v9-v1/NANOAODSIM
UL17	/TTZToLLNuNu_M-10_TuneCP5_13TeV-amcatnlo-pythia8/RunIISummer20UL17NanoAODv9-106X_mc2017_realistic_v9-v1/NANOAODSIM
UL17	/tZq_ll_4f_ckm_NLO_TuneCP5_13TeV-amcatnlo-pythia8/RunIISummer20UL17NanoAODv9-106X_mc2017_realistic_v9-v1/NANOAODSIM
UL17	/TTTT_TuneCP5_13TeV-amcatnlo-pythia8/RunIISummer20UL17NanoAODv9-106X_mc2017_realistic_v9-v2/NANOAODSIM
UL18	/ttHJetToNonbb_M125_TuneCP5_13TeV_amcatnloFXFX_madspin_pythia8/RunIISummer20UL18NanoAODv9-106X_upgrade2018_realistic_v16_L1v1-v1/NANOAODSIM
UL18	/TTWJetsToLNu_TuneCP5_13TeV-amcatnloFXFX-madspin-pythia8/RunIISummer20UL18NanoAODv9-106X_upgrade2018_realistic_v16_L1v1-v1/NANOAODSIM
UL18	/TTZToLLNuNu_M-10_TuneCP5_13TeV-amcatnlo-pythia8/RunIISummer20UL18NanoAODv9-106X_upgrade2018_realistic_v16_L1v1-v1/NANOAODSIM
UL18	/tZq_ll_4f_ckm_NLO_TuneCP5_13TeV-amcatnlo-pythia8/RunIISummer20UL18NanoAODv9-106X_upgrade2018_realistic_v16_L1v1-v1/NANOAODSIM
UL18	/TTTT_TuneCP5_13TeV-amcatnlo-pythia8/RunIISummer20UL18NanoAODv9-106X_upgrade2018_realistic_v16_L1v1-v2/NANOAODSIM

to be due to a change in the default shower starting scale for central UL samples. There does not seem to be any reason to believe that this change represents in improvement in the modeling. For these reasons, it is believed that the current modeling of this process is sufficient (as in TOP-19-001).

- $t\bar{l}\bar{l}q$: This process (along with tHq) is discussed in section C.5. The comparison for this sample should be handled carefully because of the fact that we cannot include an additional parton in the matrix element (as explained in Section 4.2.1). However, we already apply an additional uncertainty (derived in N_{jets}) to account for this (as discussed in Chapter 9). The uncertainty covers the discrepancy for the differential distributions used in the analysis, so the current modeling is believed to be sufficient.

C.1 Summary of comparisons for the $t\bar{t}H$ sample

For this process, there is good agreement between privately produced LO samples and centrally produced NLO samples.

C.2 Summary of comparisons for the $t\bar{t}l\nu$ sample

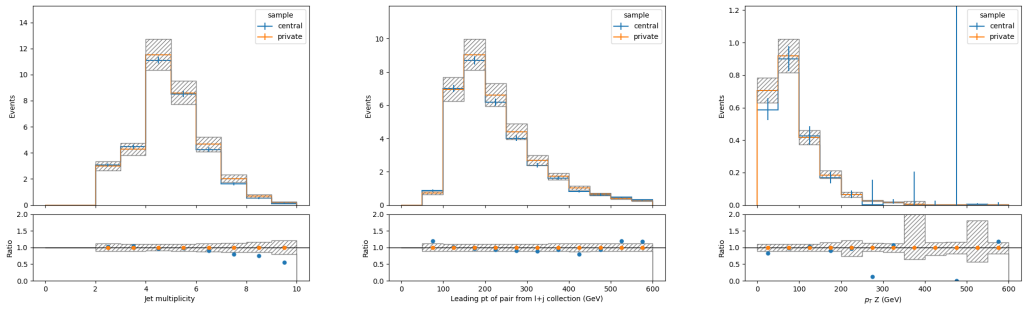
For this process, there is good agreement between privately produced LO samples and centrally produced NLO samples.

C.3 Summary of comparisons for the $t\bar{t}t\bar{t}$ sample

For this process, there is good agreement between privately produced LO samples and centrally produced NLO samples.

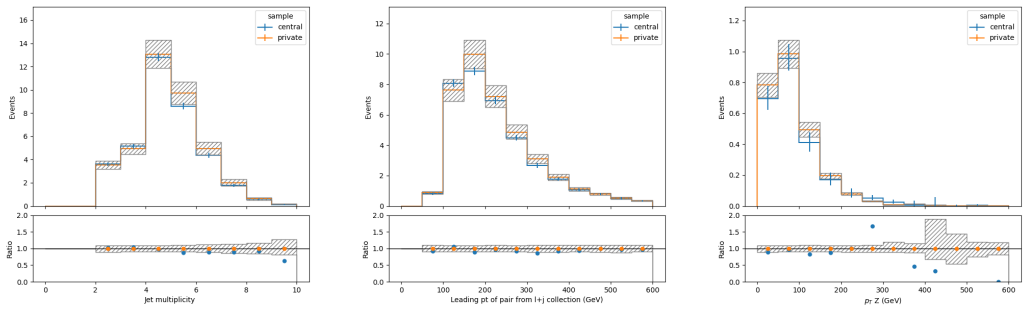
C.4 Summary of comparisons for the $t\bar{t}l\bar{l}$ sample

There is some tension between the private LO and central NLO samples for this process. However, in TOP-19-001, there was also tension between our private LO samples and the NLO central samples. Unlike the central $t\bar{t}H$ and $t\bar{t}l\nu$ samples, the central $t\bar{t}l\bar{l}$ sample does not explicitly include an extra parton in the matrix element, and we believe it is possible that some portion of the disagreement may be linked



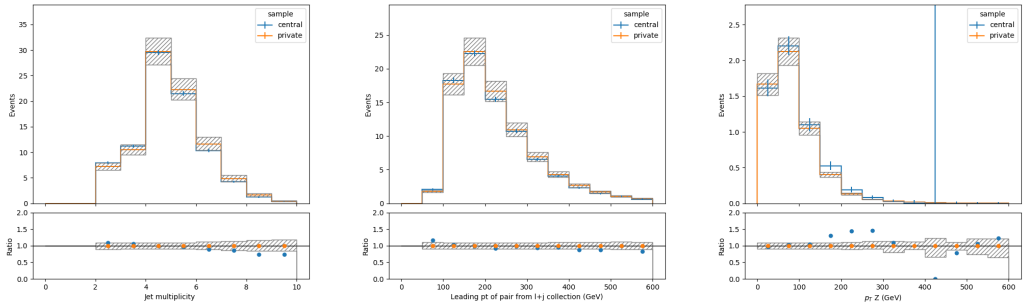
(a) (b) (c)

Figure C.1. RECO level comparison for UL16 $t\bar{t}H$. This plot shows the privately produced LO samples (reweighted to the SM) and the centrally produced NLO samples (datasets used for the central samples are listed in Table C.1). For this comparison, we have summed over all selection categories in the SR. The shaded band represents the systematic uncertainties for the private sample.



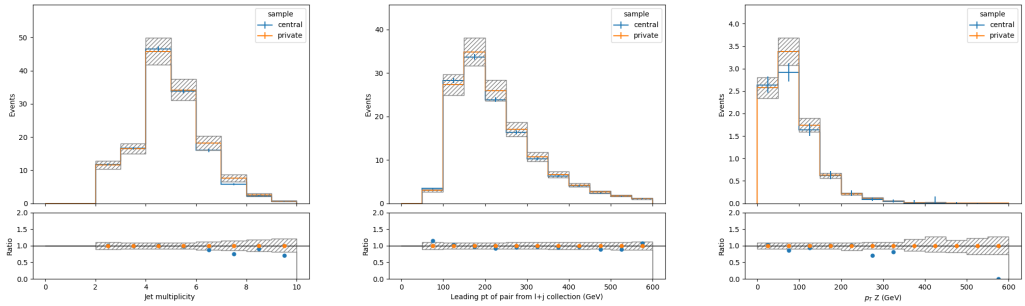
(a) (b) (c)

Figure C.2. RECO level comparison for UL16APV $t\bar{t}H$. All other relevant details are the same as described in Figure C.1.



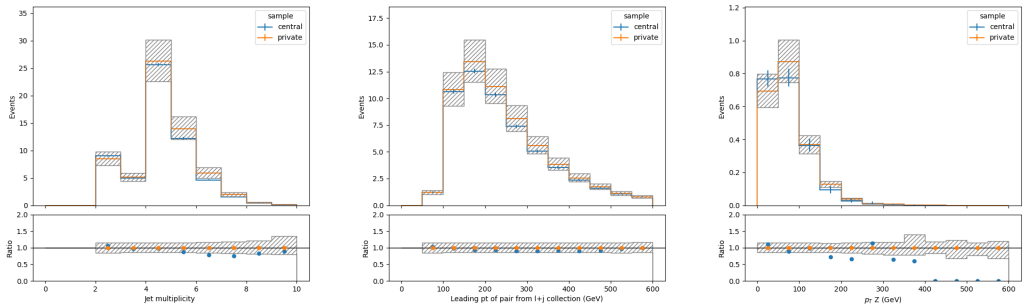
(a) (b) (c)

Figure C.3. RECO level comparison for UL17 $t\bar{t}H$. All other relevant details are the same as described in Figure C.1.



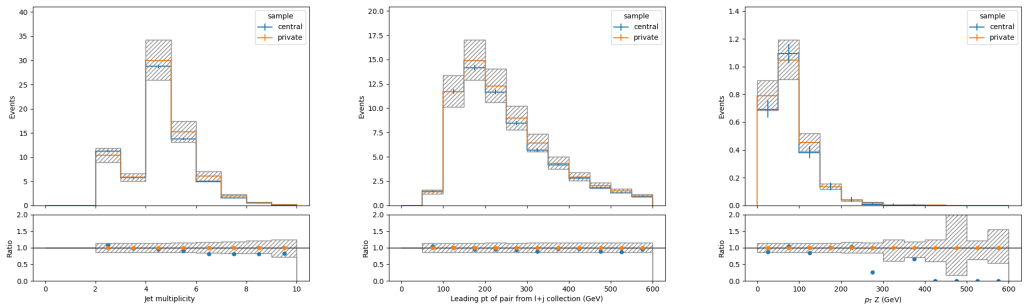
(a) (b) (c)

Figure C.4. RECO level comparison for UL18 $t\bar{t}H$. All other relevant details are the same as described in Figure C.1.



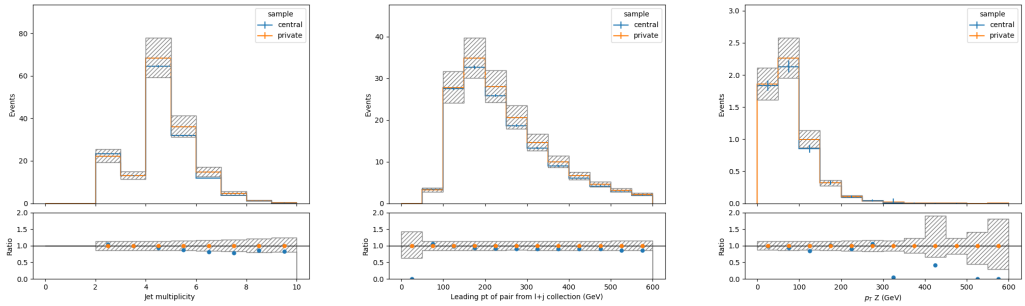
(a) (b) (c)

Figure C.5. RECO level comparison for UL16 $t\bar{t}l\nu$. All other relevant details are the same as described in Figure C.1.



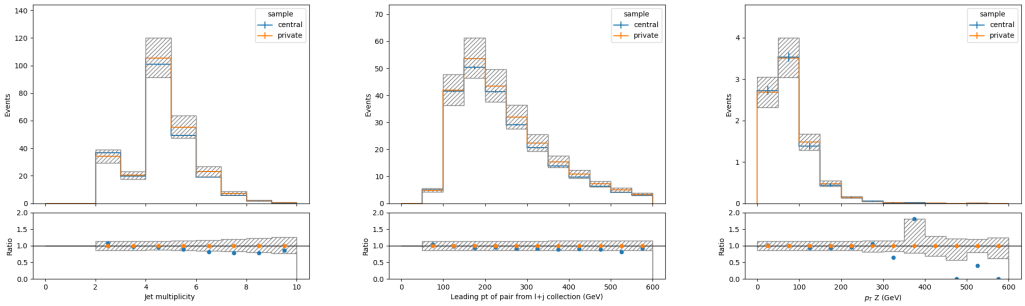
(a) (b) (c)

Figure C.6. RECO level comparison for UL16APV $t\bar{t}l\nu$. All other relevant details are the same as described in Figure C.1.



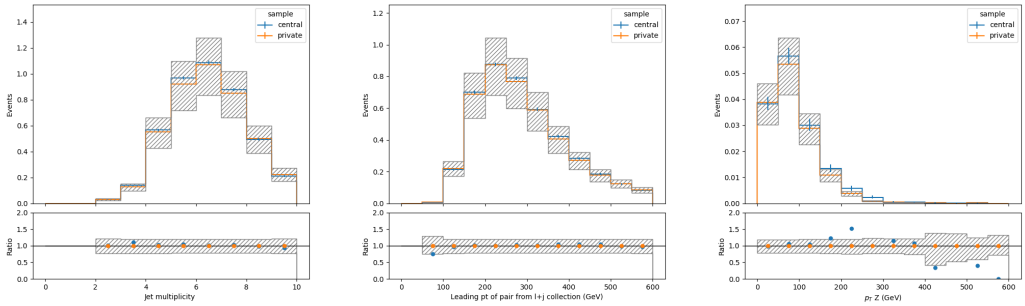
(a) (b) (c)

Figure C.7. RECO level comparison for UL17 $t\bar{t}l\nu$. All other relevant details are the same as described in Figure C.1.



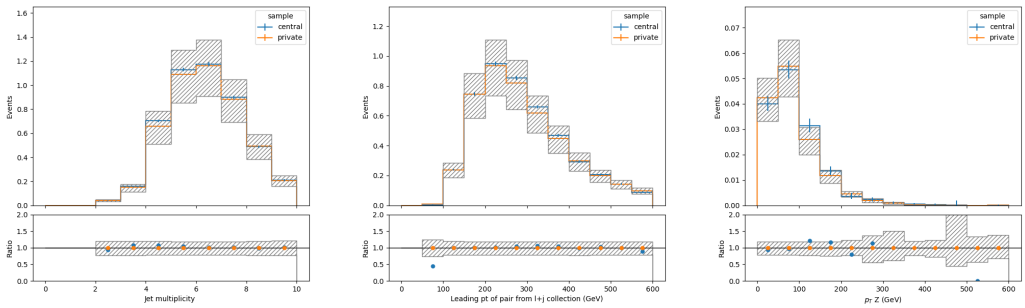
(a) (b) (c)

Figure C.8. RECO level comparison for UL18 $t\bar{t}l\nu$. All other relevant details are the same as described in Figure C.1.



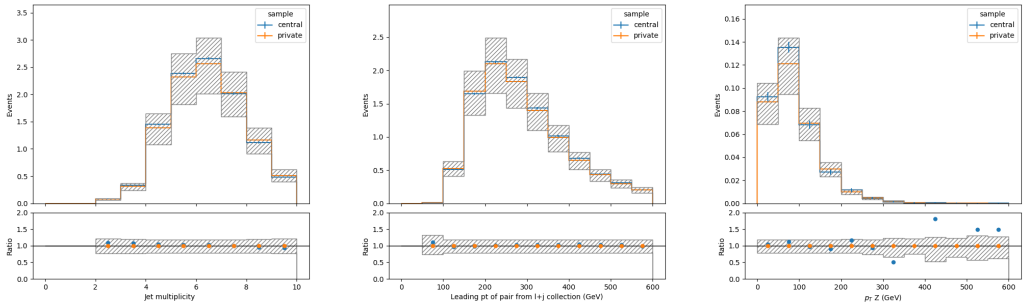
(a) (b) (c)

Figure C.9. RECO level comparison for UL16 $t\bar{t}t\bar{t}$. All other relevant details are the same as described in Figure C.1.



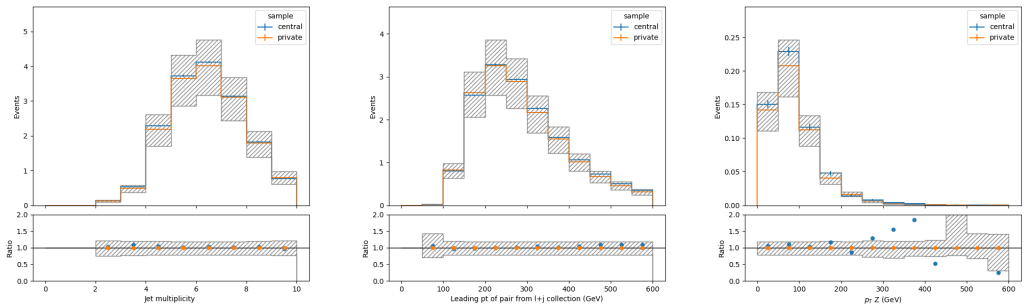
(a) (b) (c)

Figure C.10. RECO level comparison for UL16APV $t\bar{t}t\bar{t}$. All other relevant details are the same as described in Figure C.1.



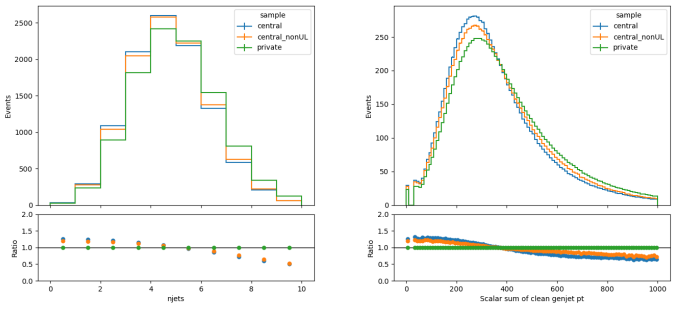
(a) (b) (c)

Figure C.11. RECO level comparison for UL17 $t\bar{t}t\bar{t}$. All other relevant details are the same as described in Figure C.1.



(a) (b) (c)

Figure C.12. RECO level comparison for UL18 $t\bar{t}t\bar{t}$. All other relevant details are the same as described in Figure C.1.



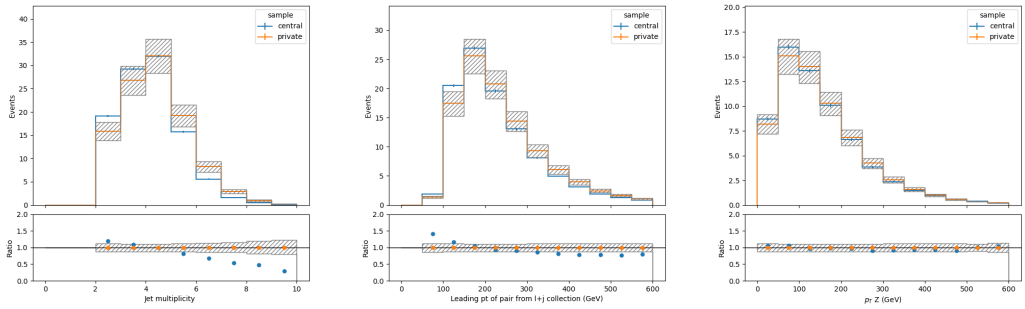
(a) (b)

Figure C.13. GEN level comparison for 2017 and UL17 $t\bar{t}l\bar{l}$ for the N_{jets} (a) and H_T (b). Some basic jet cleaning has been applied. As can be seen in the plots, the central UL sample has changed in comparison to the central pre-UL samples (and this change happens to make the tension with the private $t\bar{t}l\bar{l}$ sample somewhat worse). As discussed in the text, the change in the central sample seems to be due to a change in the default shower starting scale, and does not seem to represent an improvement in the modeling of the $t\bar{t}l\bar{l}$ process.

to this difference. The level of agreement between our private LO samples and the central NLO samples was deemed acceptable for TOP-19-001.

From TOP-19-001 (which used pre-UL samples) to this analysis (which uses all UL samples), our privately produced LO $t\bar{t}l\bar{l}$ samples have remained consistent. However, there has been a change in the central $t\bar{t}l\bar{l}$ sample from pre-UL to UL. The change is apparent primarily in jet-related variables, e.g. N_{jets} or H_T (where H_T is defined as the scalar sum of the p_T of all jets in the event). This change seems to be caused by a change in the MadGraph default shower starting scale that was implemented in MadGraph version 2.5.3. Based on discussions with the experts, it seems this change does not necessarily represent an improvement in the modeling of the $t\bar{t}l\bar{l}$ process. A GEN level comparison of the pre-UL and UL central $t\bar{t}l\bar{l}$ samples (along with the private UL $t\bar{t}l\bar{l}$ sample) is shown in Figure C.13.

This change in the central UL samples (which seems to have been caused by the



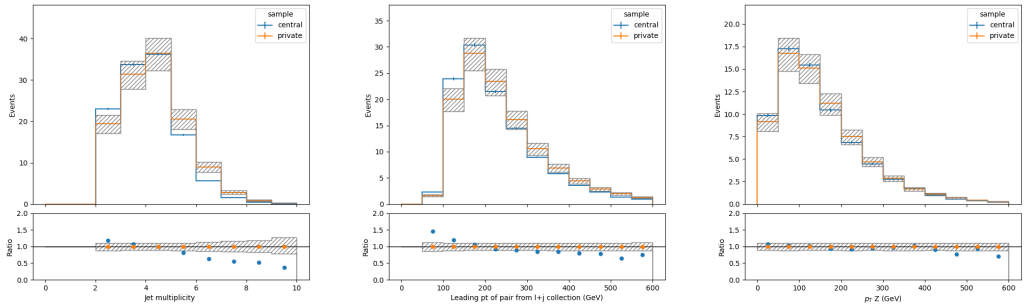
(a) (b) (c)

Figure C.14. RECO level comparison for UL16 $t\bar{t}l\bar{l}$. All other relevant details are the same as described in figure C.1.

change in the shower starting scale) has moved the central UL samples further from our private sample. This makes the tension between our private LO sample and the central NLO sample somewhat worse than was observed in TOP-19-001. However, since there is not reason to believe the change in the shower starting scale represents an improved modeling of the central $t\bar{t}l\bar{l}$ sample, it is not believed that this would imply that the private $t\bar{t}l\bar{l}$ sample is somehow less-well modeled in this analysis than in TOP-19-001. The $t\bar{t}l\bar{l}$ modeling in the private LO samples was already deemed to be acceptable for TOP-19-001. Thus, for these reasons, it is not believed any additional uncertainties would be required in order to account for the differences between the LO and NLO predictions for $t\bar{t}l\bar{l}$.

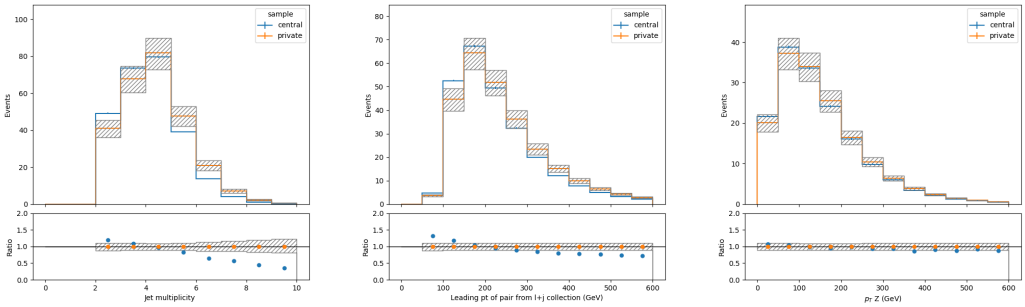
C.5 Summary of comparisons for the $tl\bar{l}q$ sample

As explained in Section 4.2.1, the LO $tl\bar{l}q$ sample (along with tHq) cannot be generated with an extra parton in the matrix element. For this reason, we may expect the modeling to be less accurate (especially at higher jet multiplicities). As in TOP-19-001, we apply an additional systematic uncertainty to this sample in



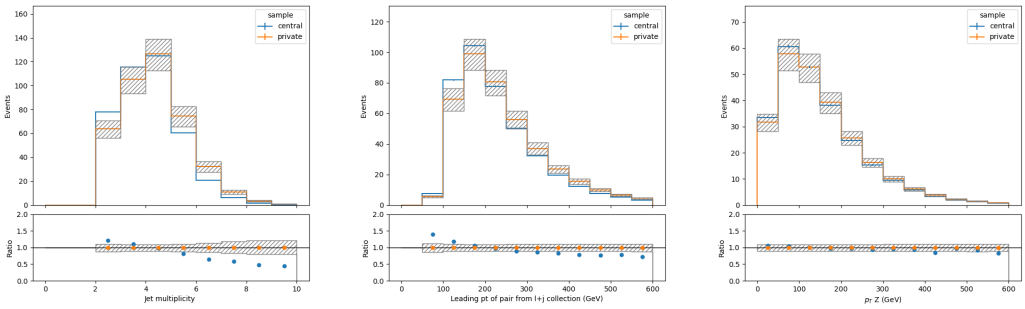
(a) (b) (c)

Figure C.15. RECO level comparison for UL16APV $t\bar{t}l\bar{l}$. All other relevant details are the same as described in figure C.1.



(a) (b) (c)

Figure C.16. RECO level comparison for UL17 $t\bar{t}l\bar{l}$. All other relevant details are the same as described in figure C.1.



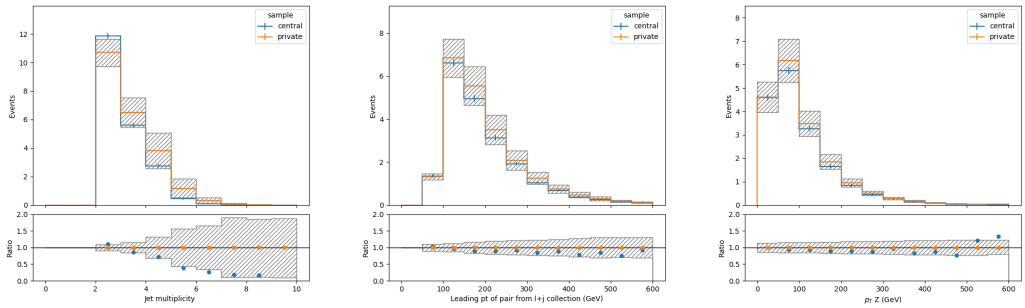
(a) (b) (c)

Figure C.17. RECO level comparison for UL18 $t\bar{t}l\bar{l}$. All other relevant details are the same as described in figure C.1.

order to account for these potential discrepancies. As described in Chapter 9, this uncertainty is derived based based on the comparison between the central NLO $tl\bar{l}q$ N_{jets} distribution, and the private LO $tl\bar{l}q$ N_{jets} distribution. This uncertainty is referred to as the "missing parton" uncertainty. For consistency, the missing parton uncertainty is also applied to the tHq sample (for which we are similarly unable to include an additional parton).

The missing parton systematic was derived and applied in a similar manner in TOP-19-001. However, this analysis takes a more differential approach than TOP-19-001, fitting differential distributions in each jet bin, so it is important to check that this systematic (which was derived in N_{jets}) also covers any discrepancies in the relevant differential distributions used in this analysis. As shown in figures C.18, C.19, C.20, and C.21, the systematic uncertainties (indicated by the shaded grey bands, which includes the "missing parton" uncertainty) generally cover the discrepancies between the samples for the $p_T(lj)_0$ and $p_T(Z)$ distributions (the kinematic distributions used in this analysis).

However, it should be mentioned that the missing parton systematic does not seem

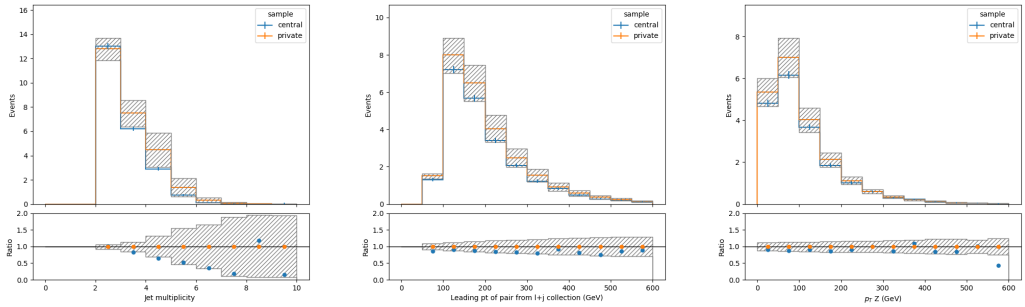


(a) (b) (c)

Figure C.18. RECO level comparison for UL16 $t\bar{t}lq$. All other relevant details are the same as described in figure C.1.

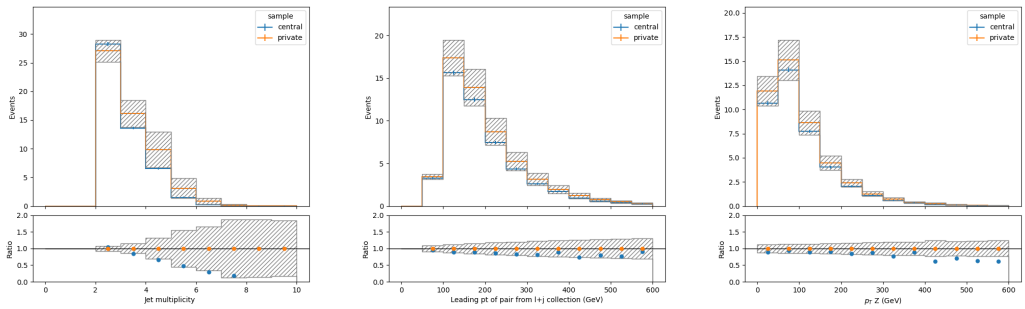
to fully cover the private vs central discrepancies for all kinematic distributions. For example, we have observed that there is tension between the central and private samples for the distribution of the invariant mass of the leading two leptons. The comparison for this distribution for the UL17 samples is shown in figure C.22; the tension is similar for other years as well. Upon additional investigation, we have found that the invariant mass of the samples agrees well when we restrict ourselves to lepton pairs coming from the Z/γ^*, so the discrepancy does not seem to be related to leptons from Z/γ^*.

Nevertheless, it should be emphasized that the invariant mass variable is not one of the variables that is directly used in the analysis. Thus, since the missing parton uncertainty covers the discrepancy for the relevant differential distributions used in the analysis ($p_T(lj)_0$ and $p_T(Z)$), we believe the modeling of this processes is sufficient, and there would not be a need to implement any additional uncertainties.



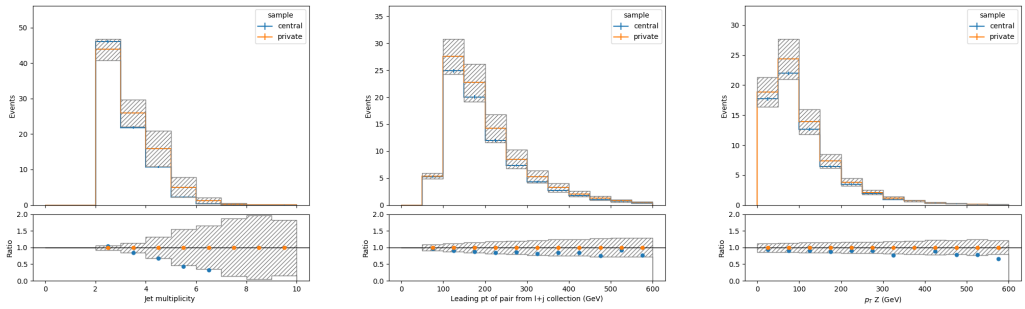
(a) (b) (c)

Figure C.19. RECO level comparison for UL16APV $t\bar{t}lq$. All other relevant details are the same as described in figure C.1.



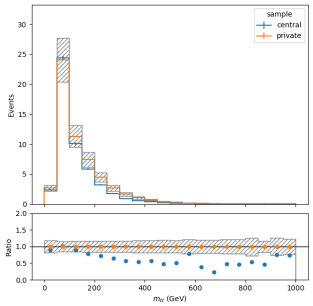
(a) (b) (c)

Figure C.20. RECO level comparison for UL17 $t\bar{t}lq$. All other relevant details are the same as described in figure C.1.



(a) (b) (c)

Figure C.21. RECO level comparison for UL18 $t\bar{t}llq$. All other relevant details are the same as described in figure C.1.



(a)

Figure C.22. RECO level comparison for UL17 $t\bar{t}llq$. All other relevant details are the same as described in figure C.18. As discussed in the text, there is tension between the central and private $t\bar{t}llq$ samples for this distribution (invariant mass of the leading two leptons). However, this distribution is not directly used in the analysis.

APPENDIX D

ONE-DIMENSIONAL QUADRATIC PARAMETERIZATION PLOTS

This appendix includes plots that show the one-dimensional quadratic dependence of the EFT samples on the 26 WCs studied in this analysis. The x axis range on the plots corresponds to the 2σ confidence intervals extracted from the profiled fits to data. In all cases, the one-dimensional quadratics have been normalized to the SM prediction (i.e. all curves pass through 1 at x=0). It is important to keep in mind that these plots only show one-dimensional slices of the full quadratic parameterization; interferences among WCs are thus not shown by these plots.

Section D.1 includes the $t\bar{t}H$ plots, Section D.2 includes the $t\bar{t}l\nu$ plots, Section D.3 includes the $t\bar{t}l\bar{l}$ plots, Section D.4 includes the $tl\bar{l}q$ plots, Section D.5 includes the tHq plots, and Section D.6 includes the $t\bar{t}t\bar{t}$ plots.

D.1 $t\bar{t}H$

This section includes plots that show the for the one-dimensional quadratic dependence of $t\bar{t}H$ on the 22 WCs included in the $t\bar{t}H$ EFT samples. The two-heavy-two-light WCs are shown in Figure D.1. The two-heavy-with-boson WCs are shown in Figure D.2. The two-heavy-two-lepton WCs are shown in Figure D.3.

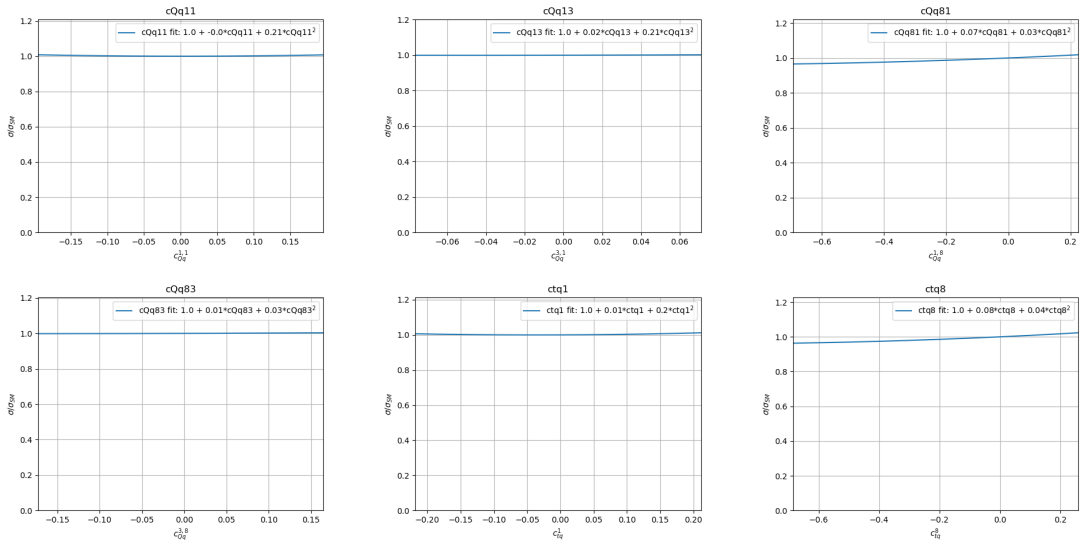


Figure D.1. The one-dimensional quadratic dependence of $t\bar{t}H$ on the two-heavy-two-light WCs.

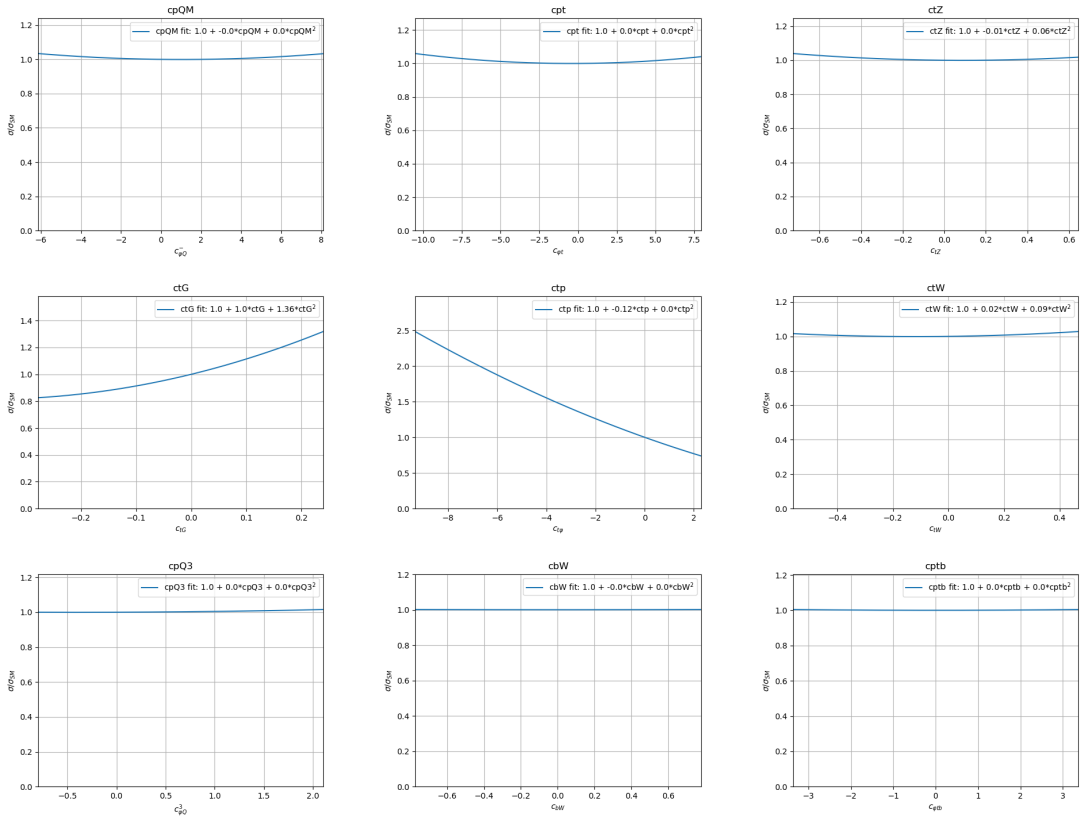


Figure D.2. The one-dimensional quadratic dependence of $t\bar{t}H$ on the two-heavy-with-boson WCs.

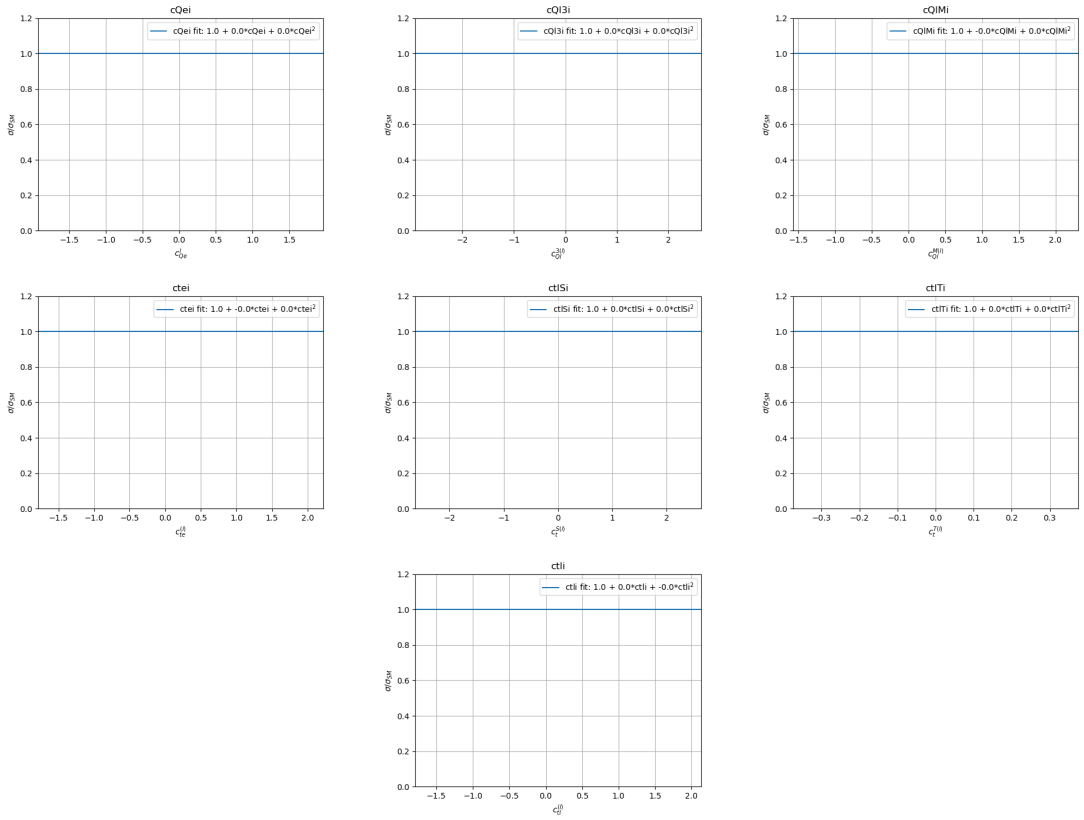


Figure D.3. The one-dimensional quadratic dependence of $t\bar{t}H$ on the two-heavy-two-lepton WCs.

D.2 $t\bar{t}l\nu$

This section includes plots that show the for the one-dimensional quadratic dependence of $t\bar{t}l\nu$ on the 22 WCs included in the $t\bar{t}l\nu$ EFT samples. The two-heavy-two-light WCs are shown in Figure D.4. The two-heavy-with-boson WCs are shown in Figure D.5. The two-heavy-two-lepton WCs are shown in Figure D.6.

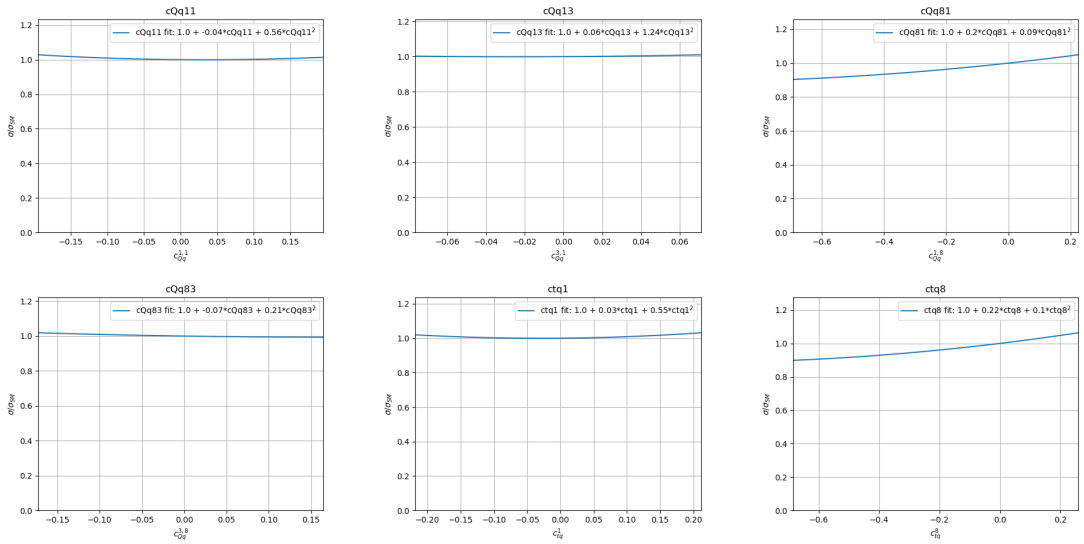


Figure D.4. The one-dimensional quadratic dependence of $t\bar{t}l\nu$ on the two-heavy-two-light WCs.

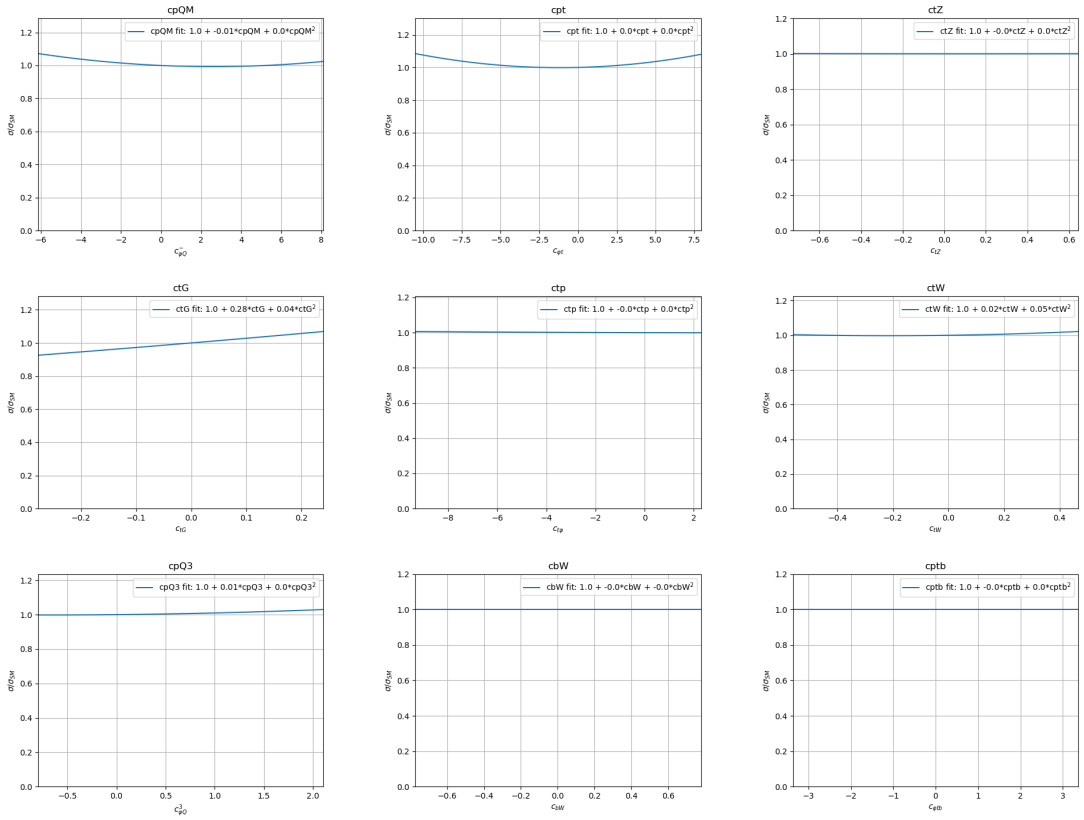


Figure D.5. The one-dimensional quadratic dependence of $t\bar{t}l\nu$ on the two-heavy-with-boson WCs.

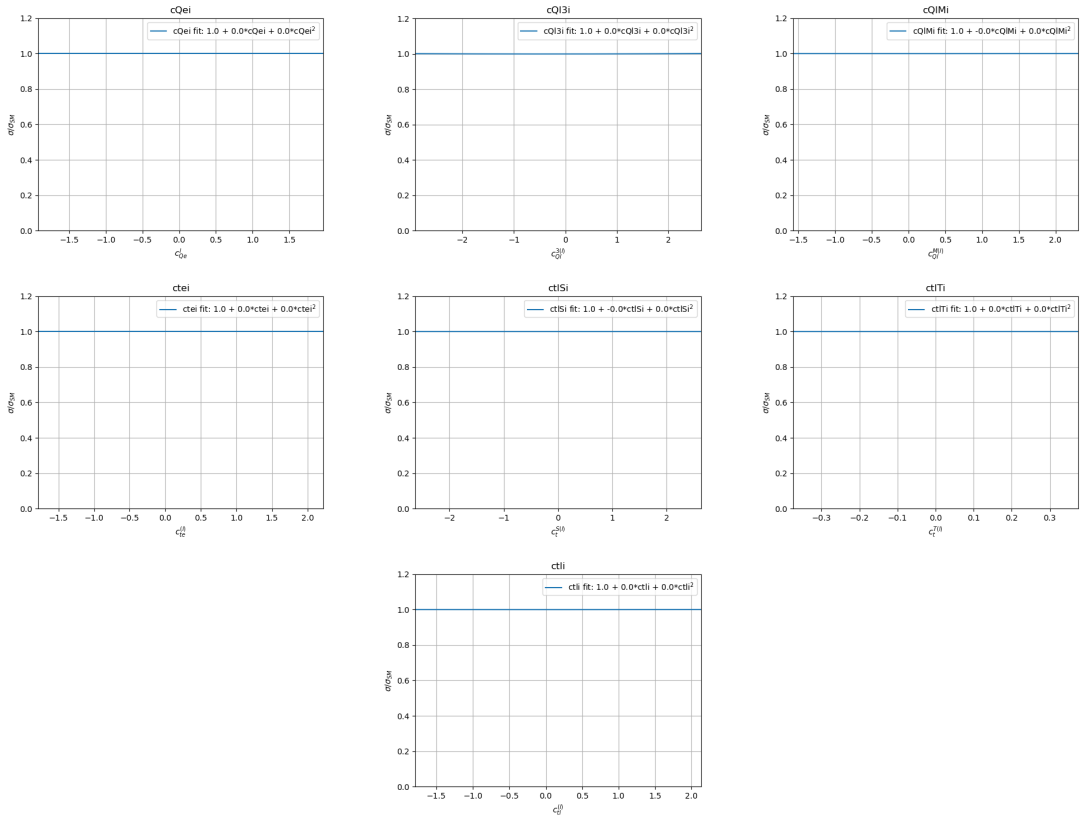


Figure D.6. The one-dimensional quadratic dependence of $t\bar{t}l\nu$ on the two-heavy-two-lepton WCs.

D.3 $t\bar{t}l\bar{l}$

This section includes plots that show the for the one-dimensional quadratic dependence of $t\bar{t}l\bar{l}$ on the 22 WCs included in the $t\bar{t}l\bar{l}$ EFT samples. The two-heavy-two-light WCs are shown in Figure D.7. The two-heavy-with-boson WCs are shown in Figure D.8. The two-heavy-two-lepton WCs are shown in Figure D.9.

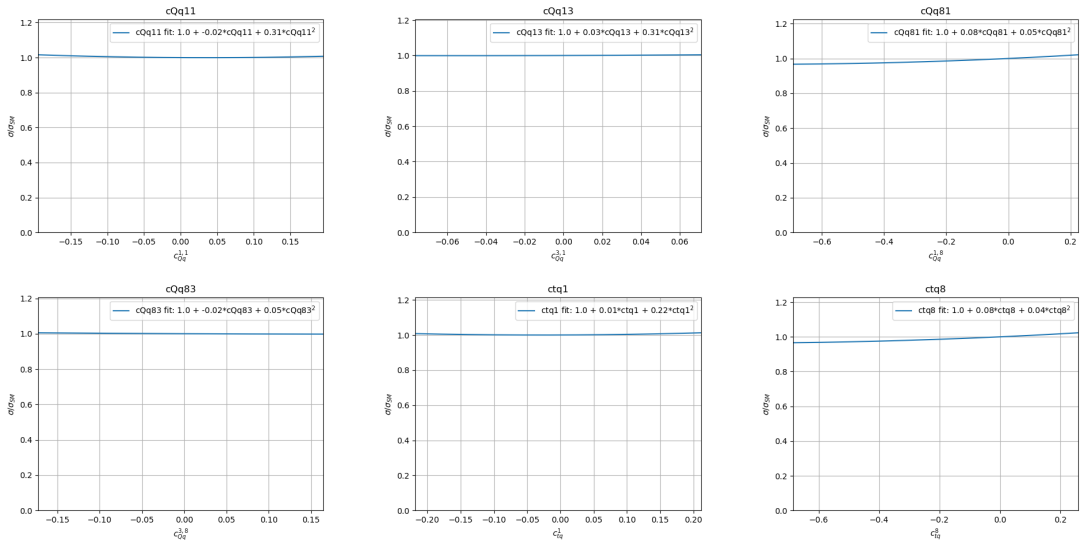


Figure D.7. The one-dimensional quadratic dependence of $t\bar{t}l\bar{l}$ on the two-heavy-two-light WCs.

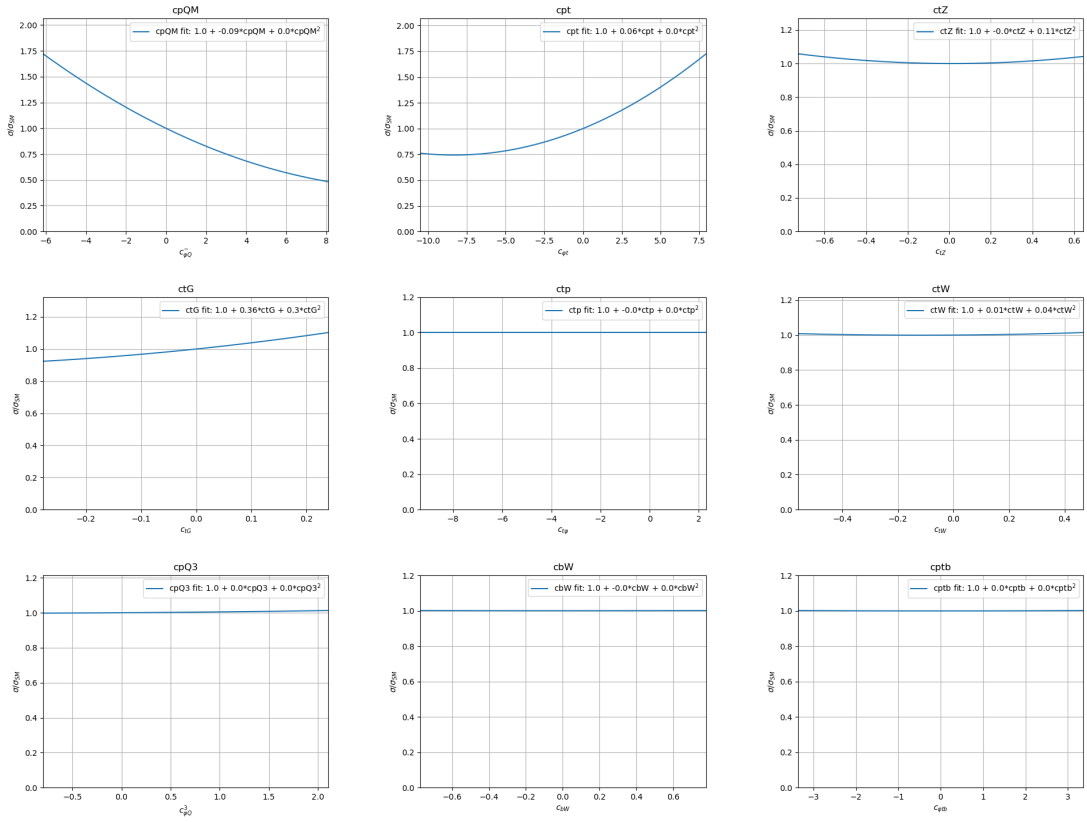


Figure D.8. The one-dimensional quadratic dependence of $t\bar{t}l\bar{l}$ on the two-heavy-with-boson WCs.

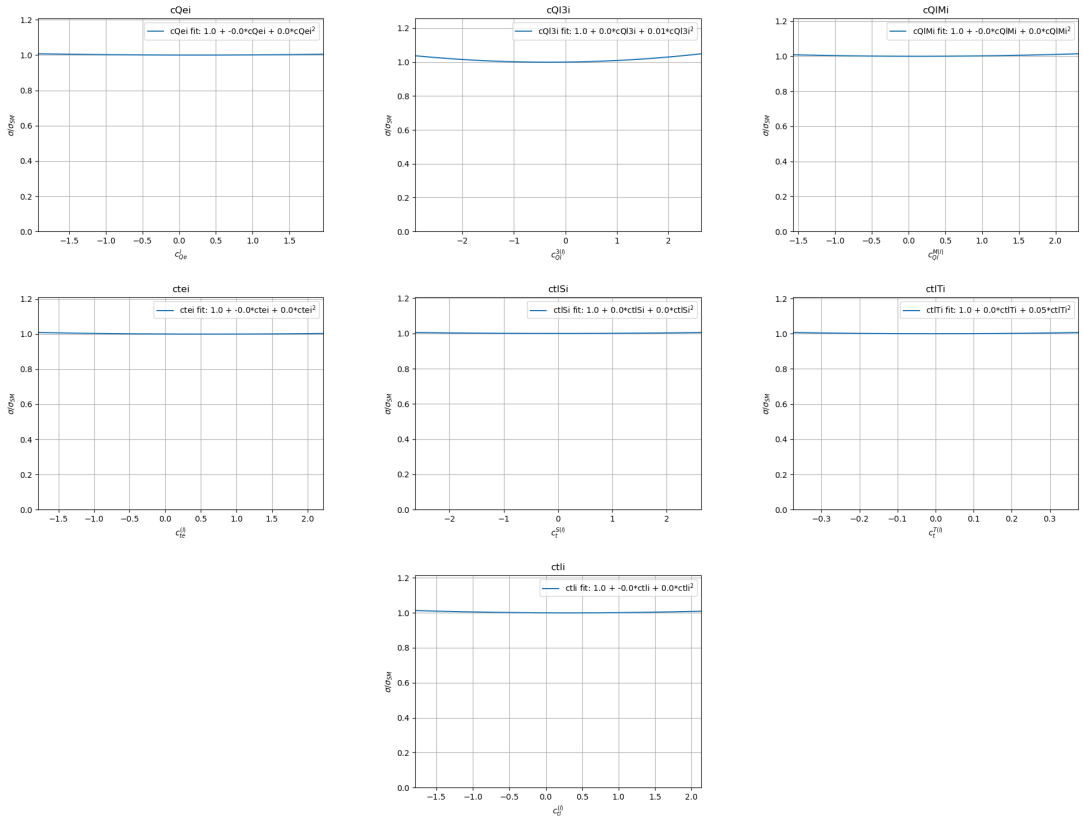


Figure D.9. The one-dimensional quadratic dependence of $t\bar{t}l\bar{l}$ on the two-heavy-two-lepton WCs.

D.4 $t\bar{t}\ell\bar{\ell}q$

This section includes plots that show the for the one-dimensional quadratic dependence of $t\bar{t}\ell\bar{\ell}q$ on the 22 WCs included in the $t\bar{t}\ell\bar{\ell}q$ EFT samples. The two-heavy-two-light WCs are shown in Figure D.10. The two-heavy-with-boson WCs are shown in Figure D.11. The two-heavy-two-lepton WCs are shown in Figure D.12.

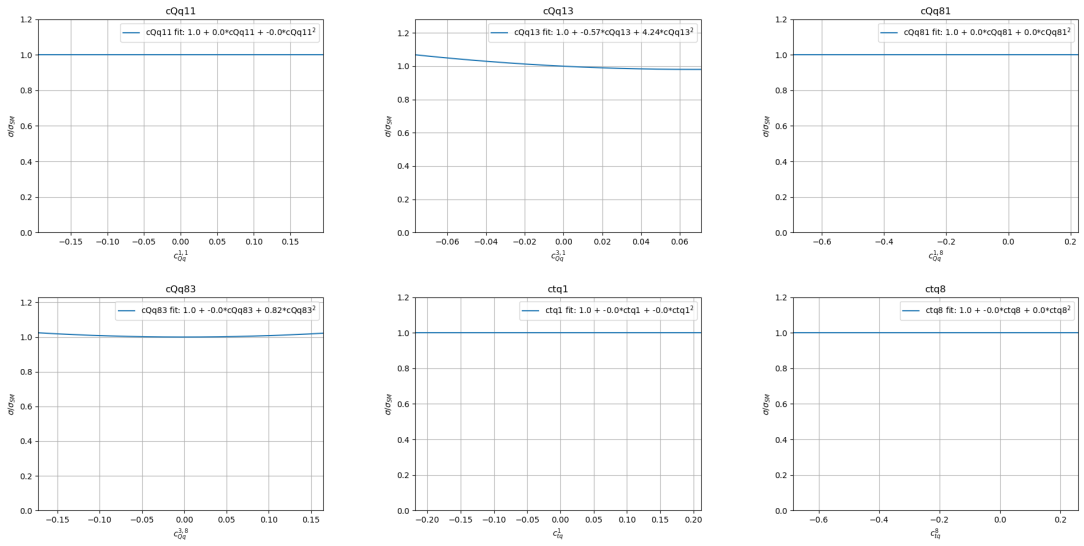


Figure D.10. The one-dimensional quadratic dependence of $t\bar{t}\ell\bar{\ell}q$ on the two-heavy-two-light WCs.

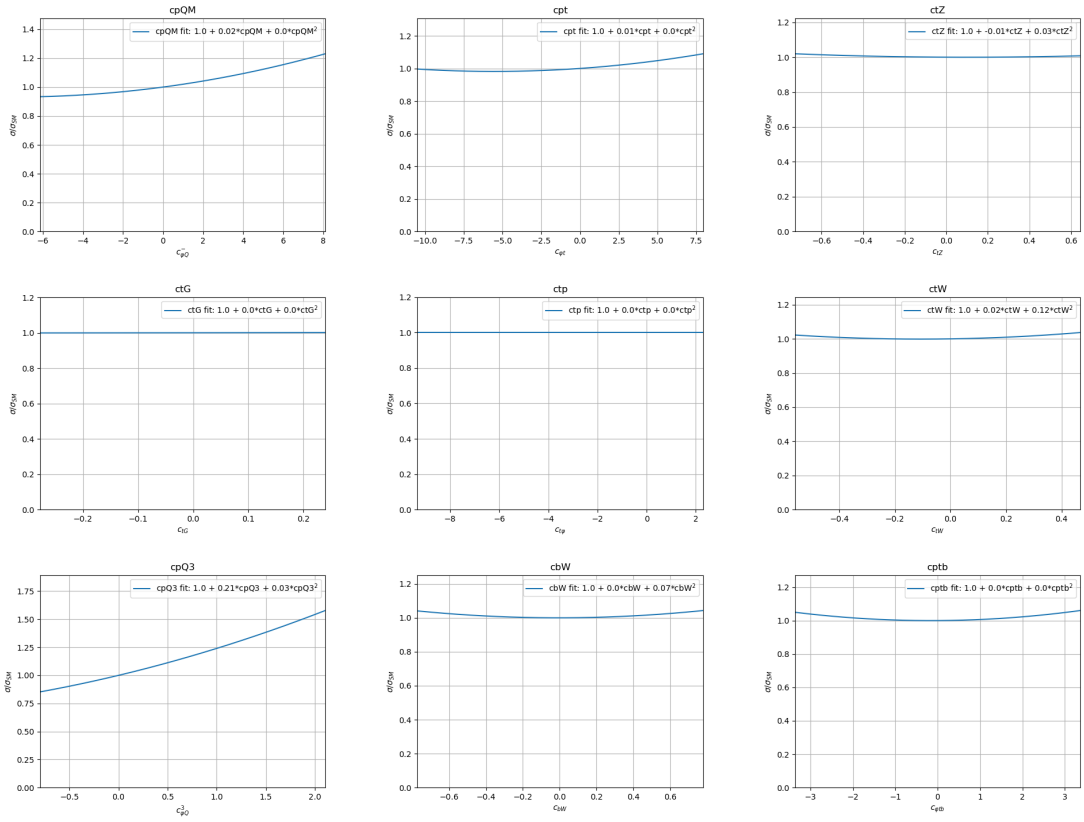


Figure D.11. The one-dimensional quadratic dependence of $\mathrm{tl\bar{l}q}$ on the two-heavy-with-boson WCs.

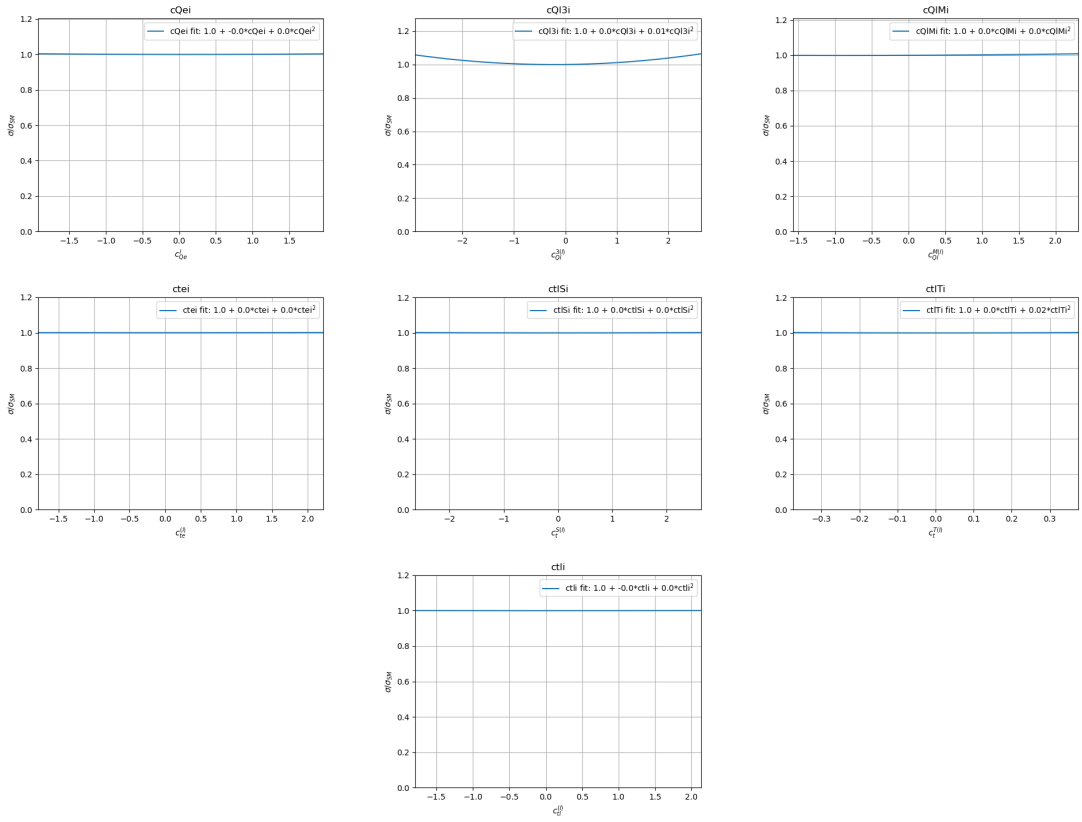


Figure D.12. The one-dimensional quadratic dependence of $t\bar{l}lq$ on the two-heavy-two-lepton WCs.

D.5 tHq

This section includes plots that show the for the one-dimensional quadratic dependence of tHq on the 22 WCs included in the tHq EFT samples. The two-heavy-two-light WCs are shown in Figure D.13. The two-heavy-with-boson WCs are shown in Figure D.14. The two-heavy-two-lepton WCs are shown in Figure D.15.

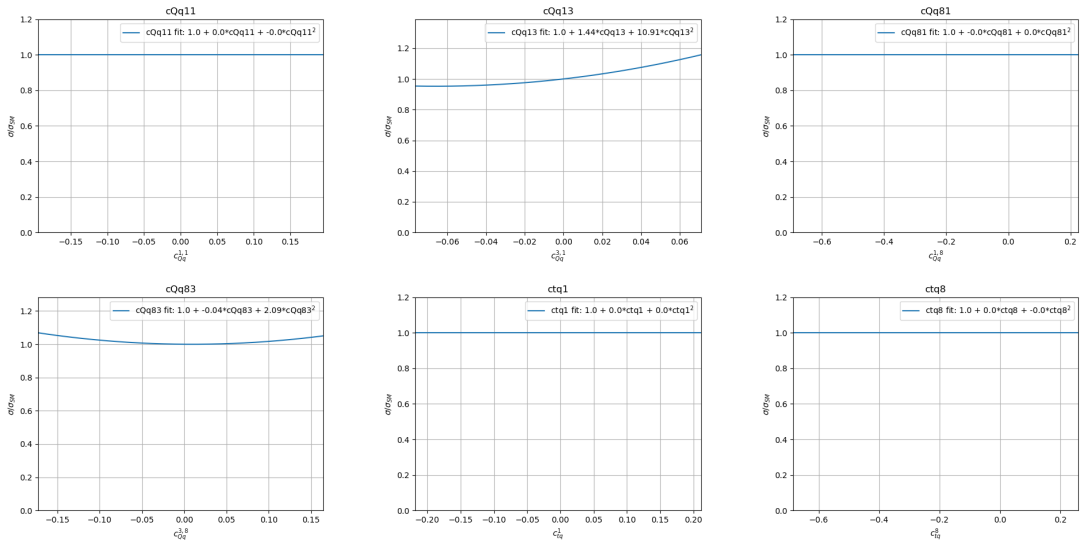


Figure D.13. The one-dimensional quadratic dependence of tHq on the two-heavy-two-light WCs.

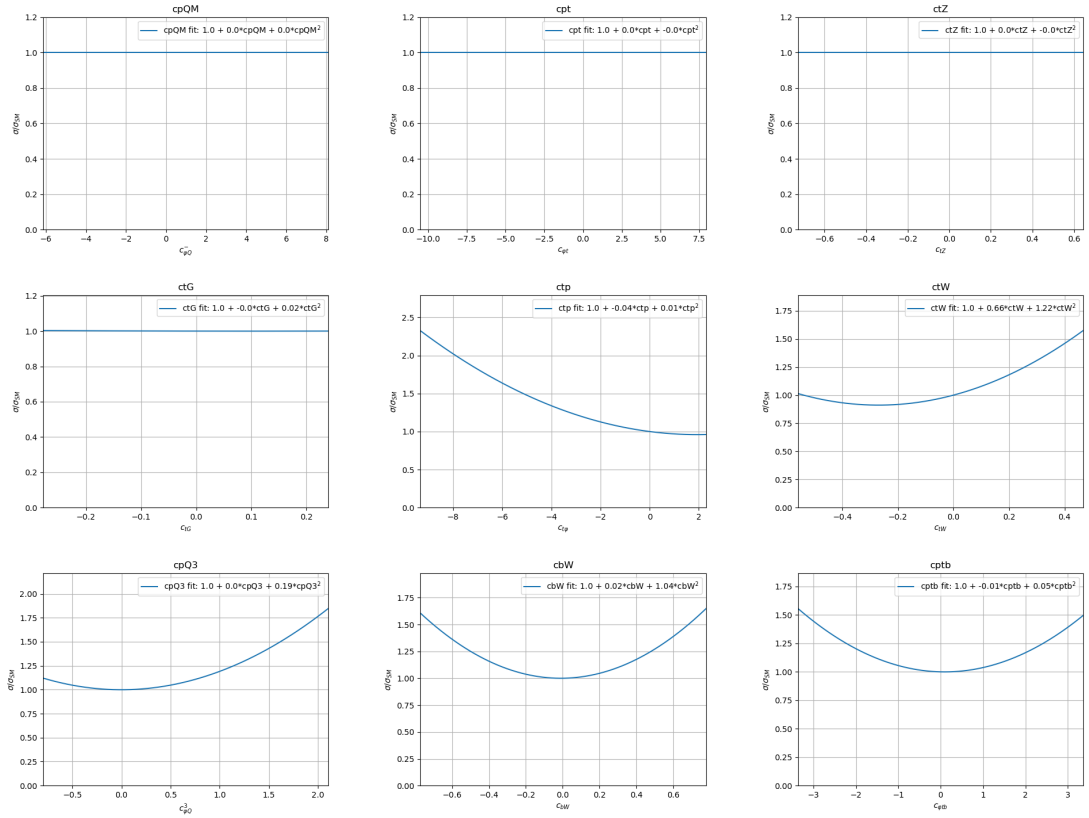


Figure D.14. The one-dimensional quadratic dependence of tHq on the two-heavy-with-boson WCs.

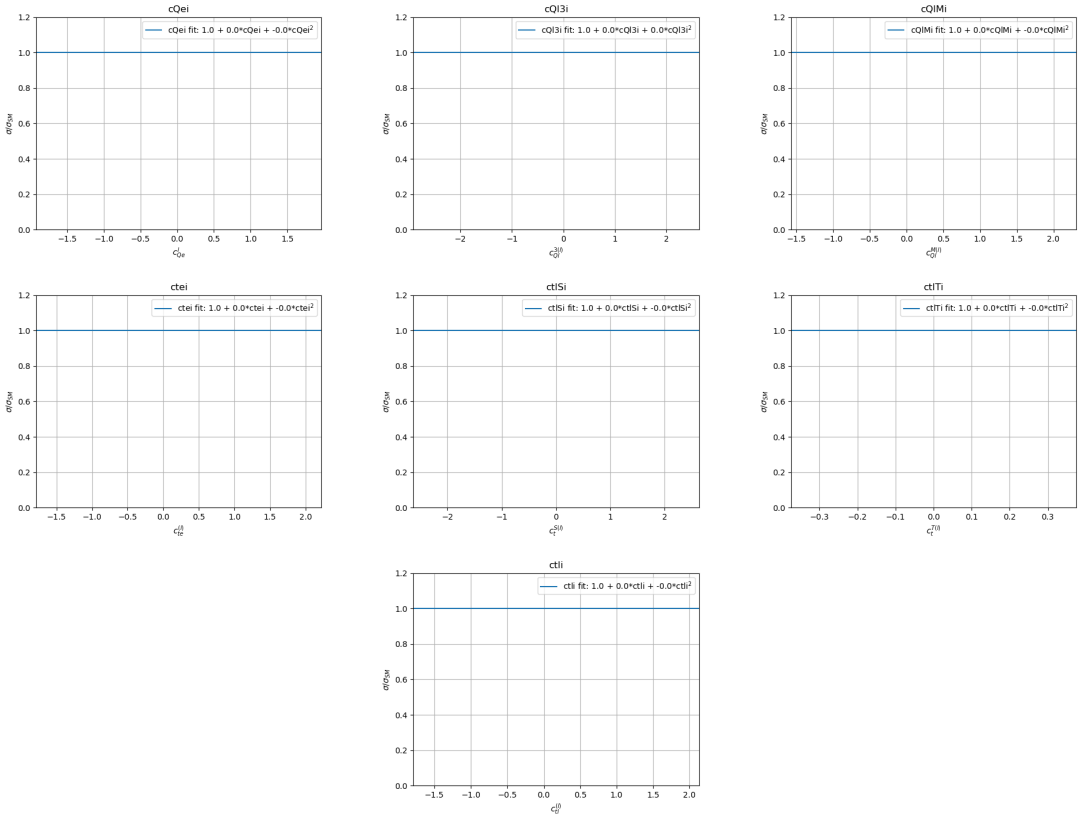


Figure D.15. The one-dimensional quadratic dependence of tHq on the two-heavy-two-lepton WCs.

D.6 $t\bar{t}t\bar{t}$

This section includes plots that show the for the one-dimensional quadratic dependence of $t\bar{t}t\bar{t}$ on the 26 WCs included in the $t\bar{t}t\bar{t}$ EFT samples. The two-heavy-two-light WCs are shown in Figure D.16. The two-heavy-with-boson WCs are shown in Figure D.17. The two-heavy-two-lepton WCs are shown in Figure D.18. The four-heavy WCs are shown in Figure D.19.

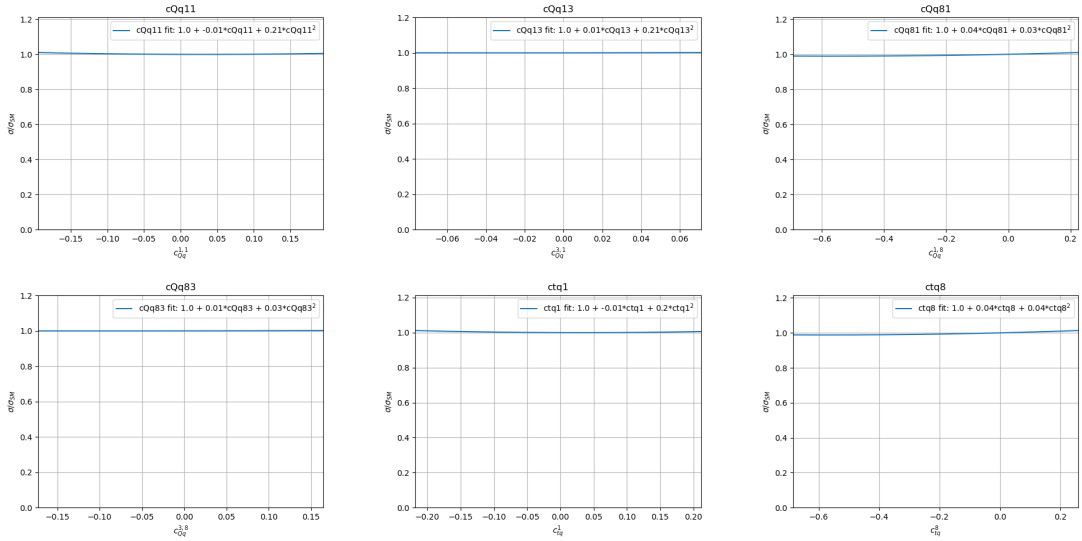


Figure D.16. The one-dimensional quadratic dependence of $t\bar{t}t\bar{t}$ on the two-heavy-two-light WCs.

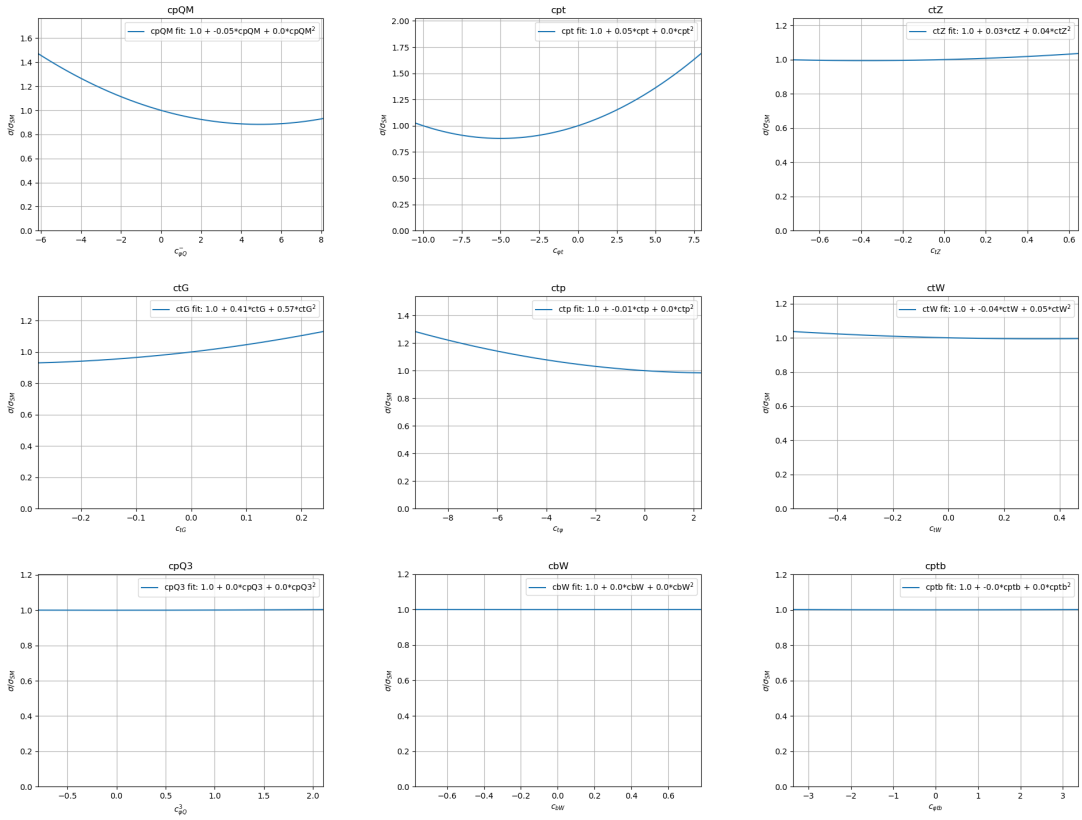


Figure D.17. The one-dimensional quadratic dependence of $t\bar{t}t\bar{t}$ on the two-heavy-with-boson WCs.

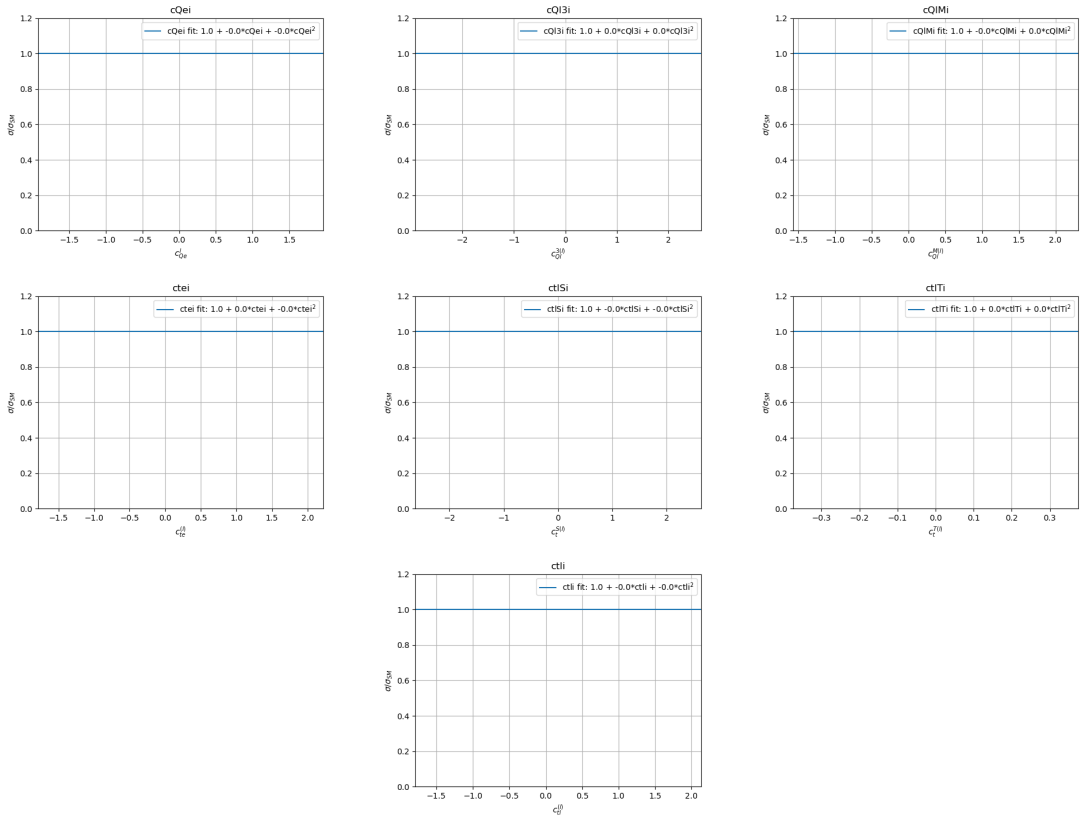


Figure D.18. The one-dimensional quadratic dependence of $t\bar{t}t\bar{t}$ on the two-heavy-two-lepton WCs.

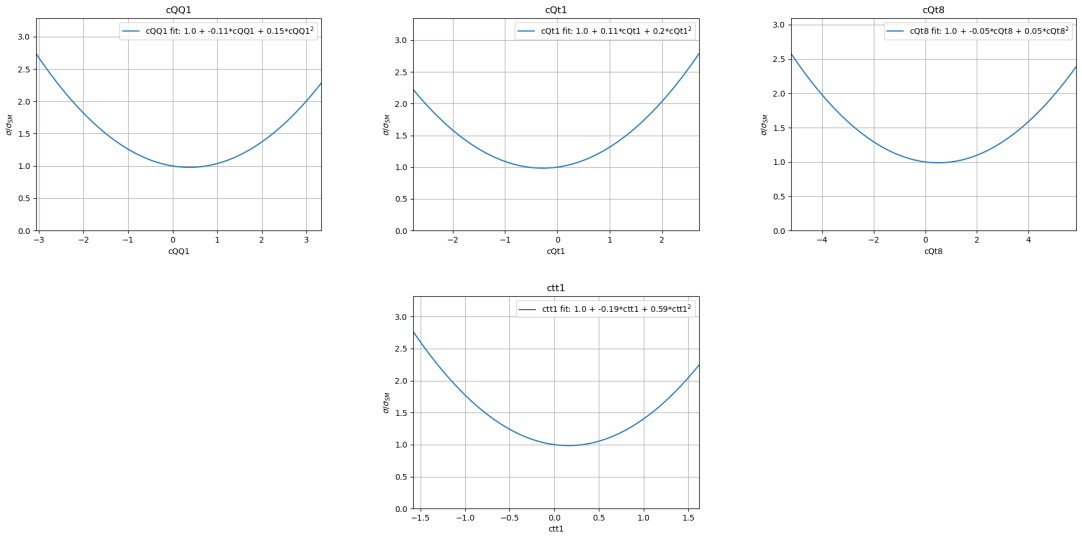


Figure D.19. The one-dimensional quadratic dependence of $t\bar{t}t\bar{t}$ on the four-heavy WCs.

APPENDIX E

CONTROL REGION PLOTS

This section includes the CR plots for the five CRs studied in this analysis. The selection criteria for the CRs are described below, and plots for each of the CRs for all of the UL years are shown in Figures E.1 through E.20.

- 2lss CR: For the 2lss CR, we require two same signed tight leptons, exactly 1 medium b tag, and exactly 1 or 2 jets. Note that for the jet multiplicity plots in this CR, the contribution from 3j is also included (as this makes any trends in the jet multiplicity easier to see). However, the signal samples contribute more significantly to the 3j bin than to the lower jet multiplicity bins; we would like to minimize the signal contamination in this CR, so we do not include the 3j contributions in the kinematic distributions for this CR.
- 3l CR: For the 3l CR, we require exactly three tight leptons, and exactly 0 medium b tagged jets. We do not apply a jet requirement.
- 2l os Z CR: For the 2l os Z CR, we require exactly two tight leptons of opposite sign and same flavor. We also require the invariant mass of the lepton pair to be within 10 GeV of the Z mass. Exactly 0 medium b tagged jets are required. No requirement on N_{jets} is specified.
- 2l os $t\bar{t}$ CR: For the 2l os $t\bar{t}$ CR, we require exactly two opposite sign, opposite flavor leptons. We require exactly 2 medium b tags, and exactly two jets. Note that we do not apply top p_T reweighting, so we do not expect exact agreement in the p_T distributions.
- 2l ss flip CR: For the 2lss charge flip CR, we require exactly two tight electrons with an invariant mass within 30 GeV of the Z peak. We place no requirement on b jets, but require fewer than 4 total jets.

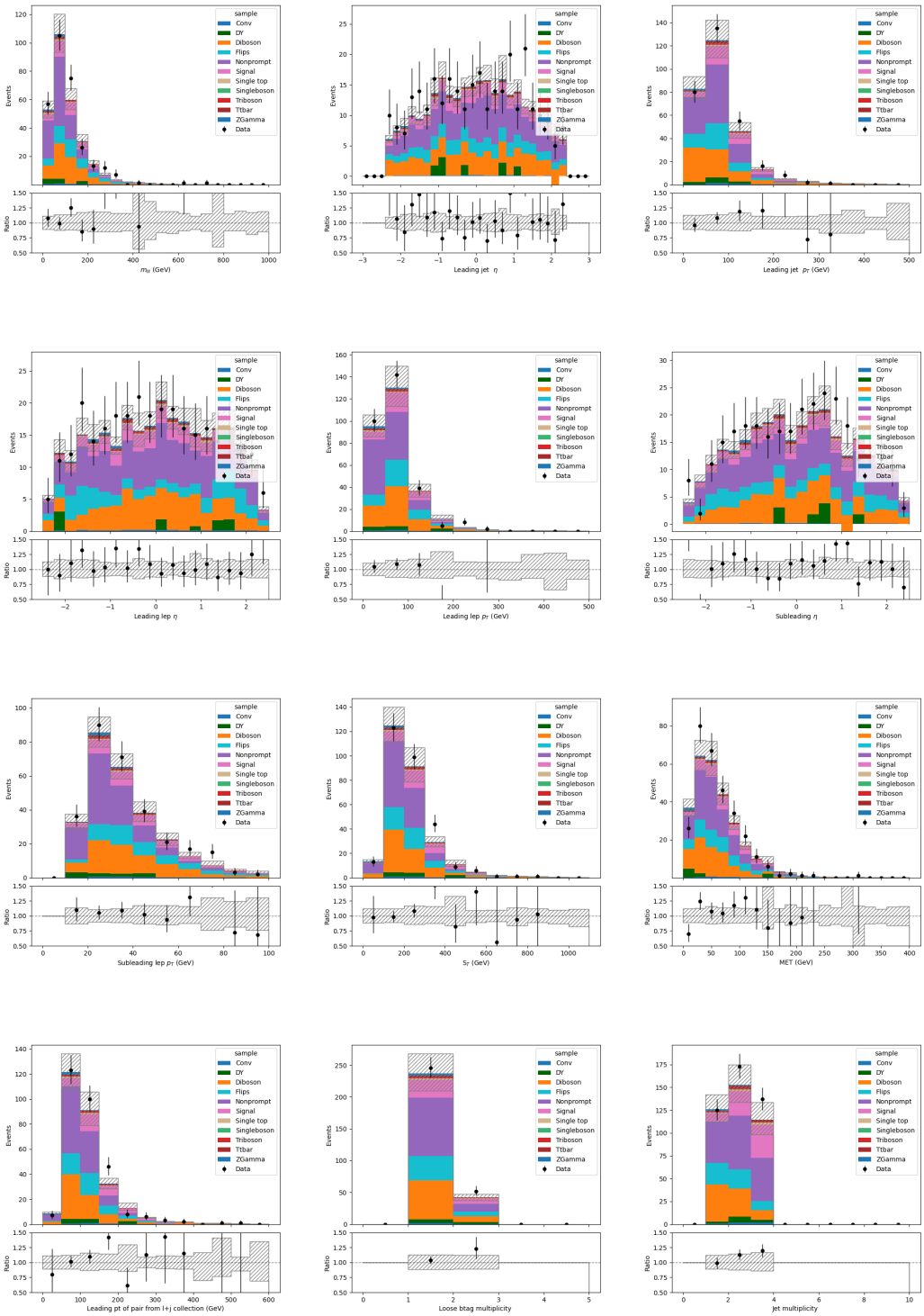


Figure E.1. Control region plots for the 2lss CR (for UL16 samples). The shaded gray band indicates the systematic uncertainty.

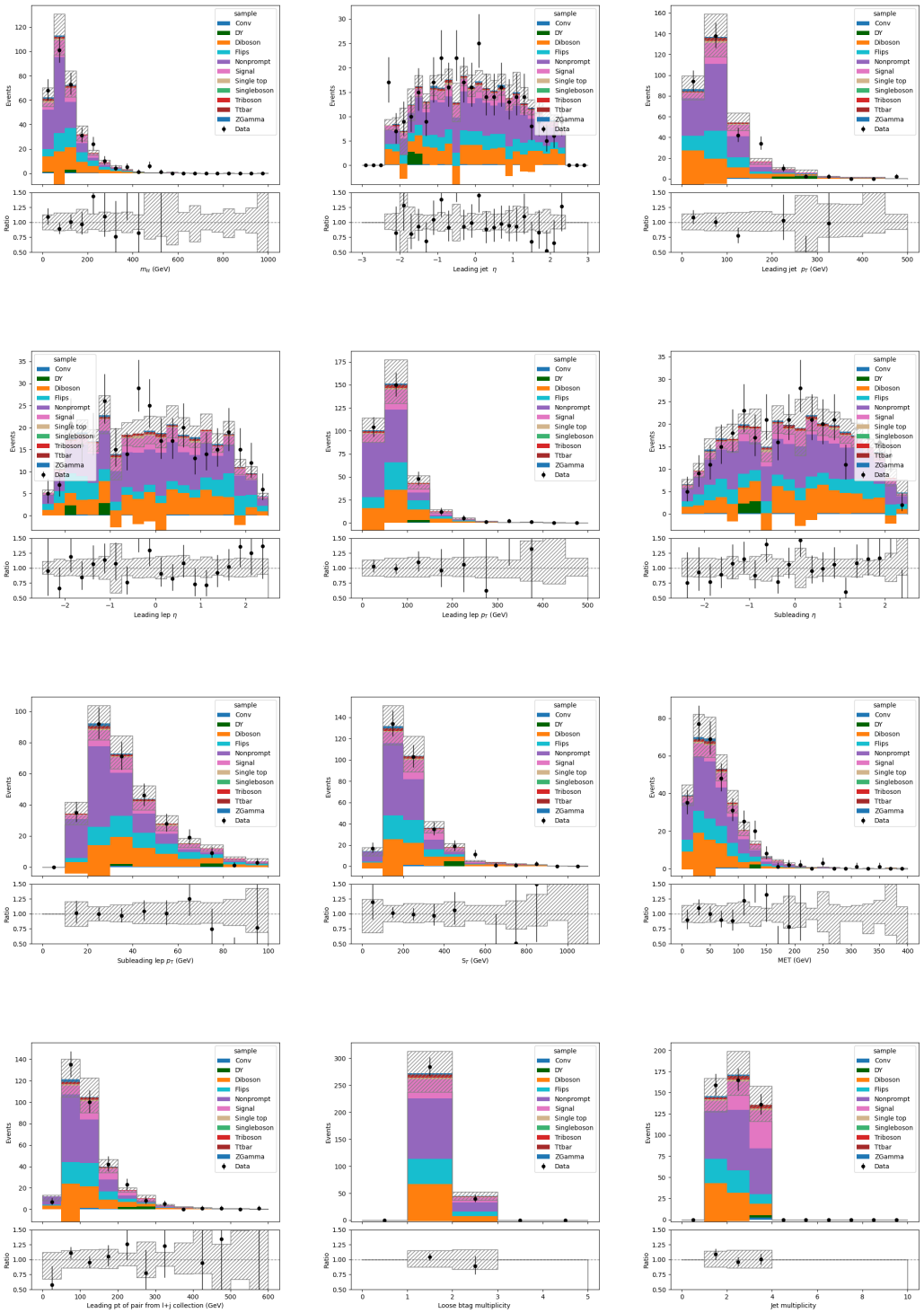


Figure E.2. Control region plots for the 2lss CR (for UL16APV samples). The shaded gray band indicates the systematic uncertainty.

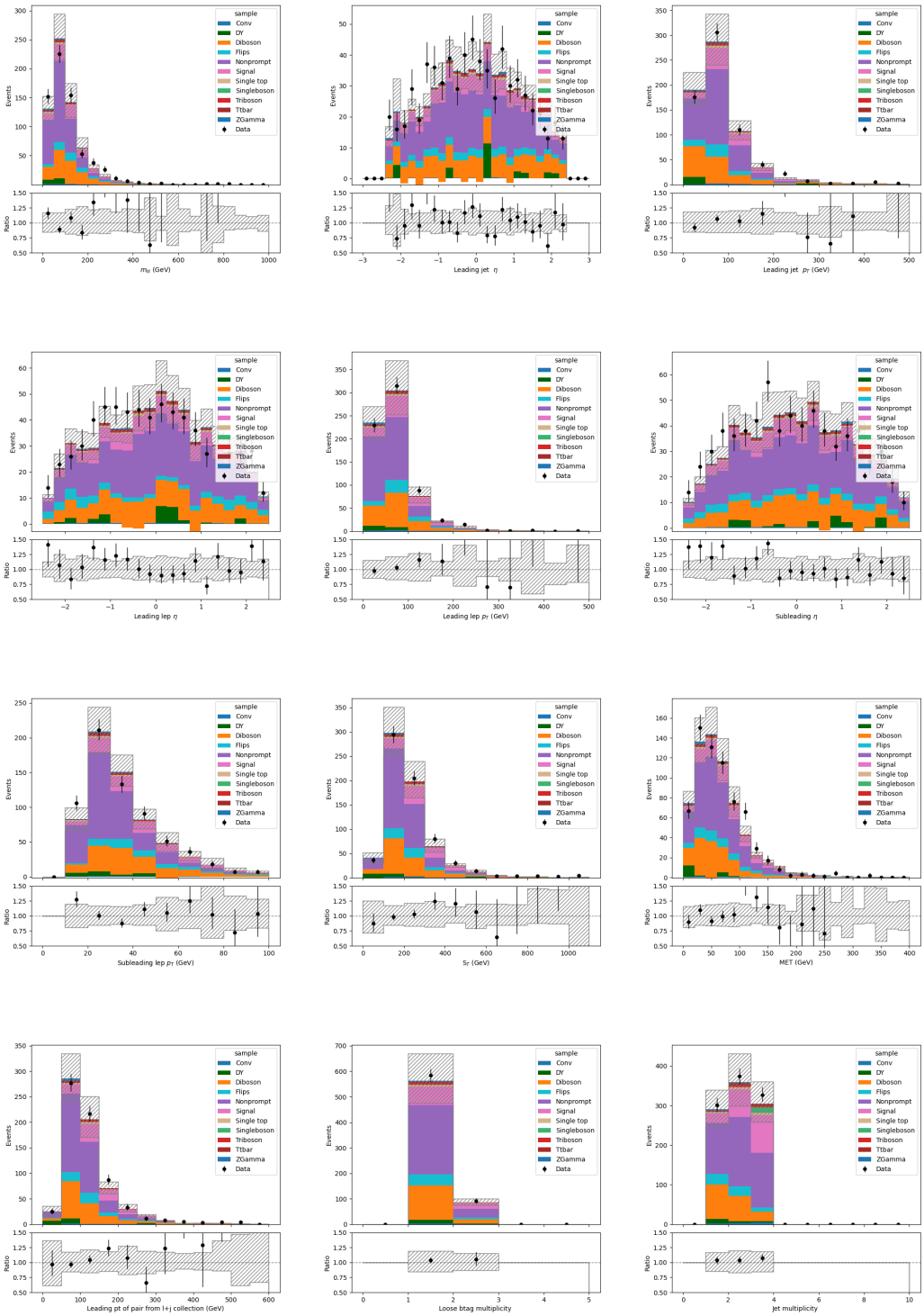


Figure E.3. Control region plots for the 2lss CR (for UL17 samples). The shaded gray band indicates the systematic uncertainty.

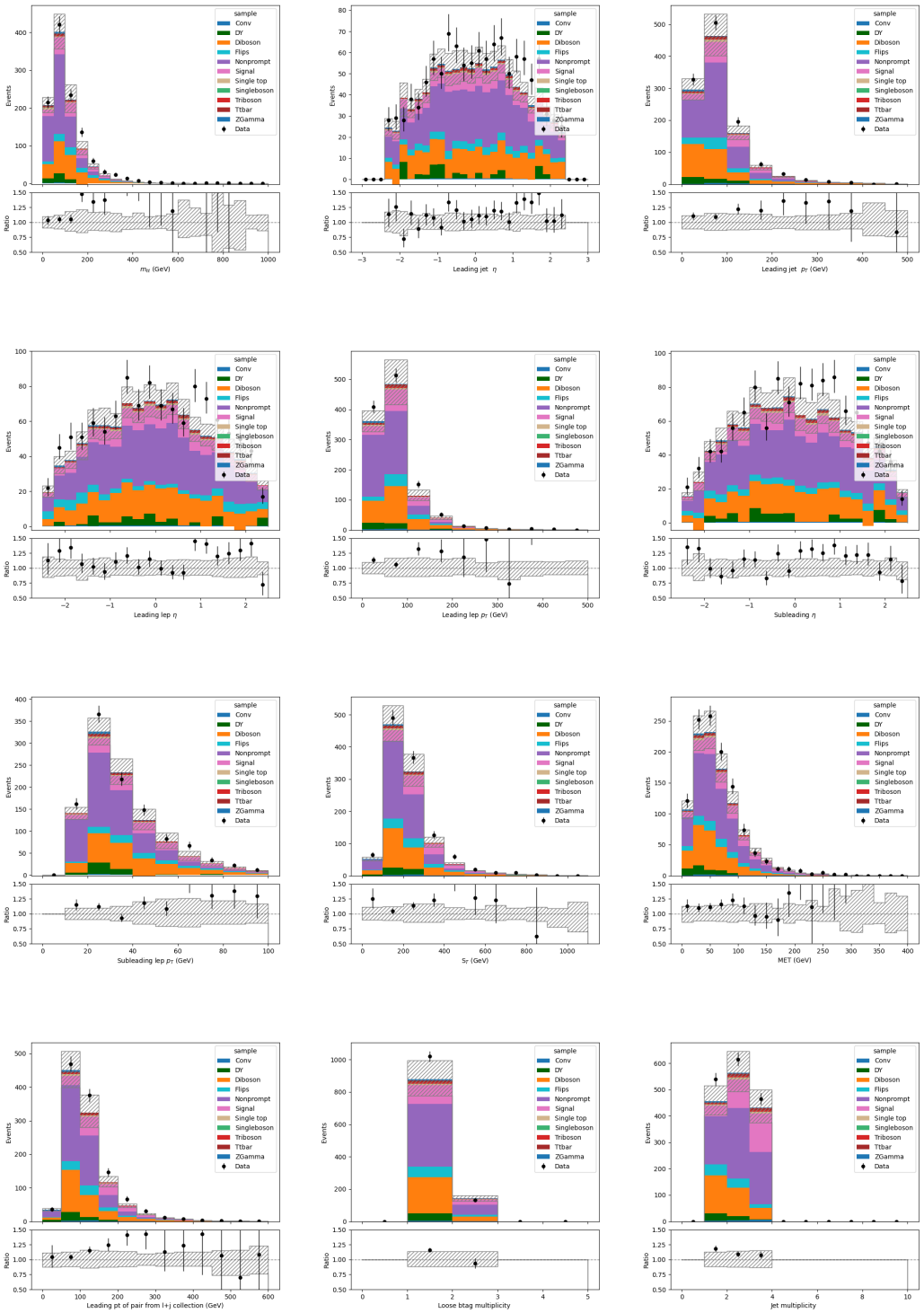


Figure E.4. Control region plots for the 2lss CR (for UL18 samples). The shaded gray band indicates the systematic uncertainty.

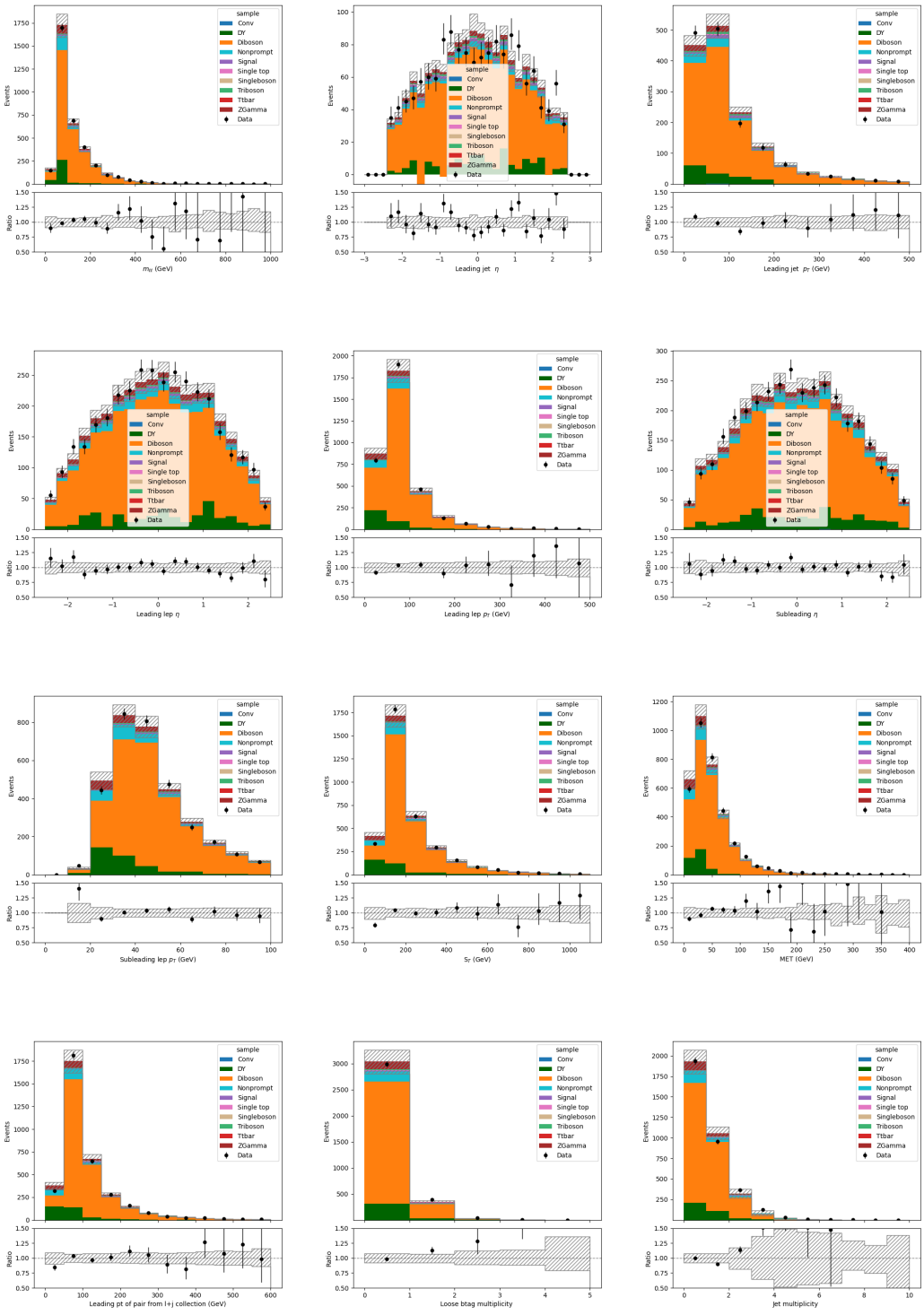


Figure E.5. Control region plots for the 3l CR (for UL16 samples). The shaded gray band indicates the systematic uncertainty.

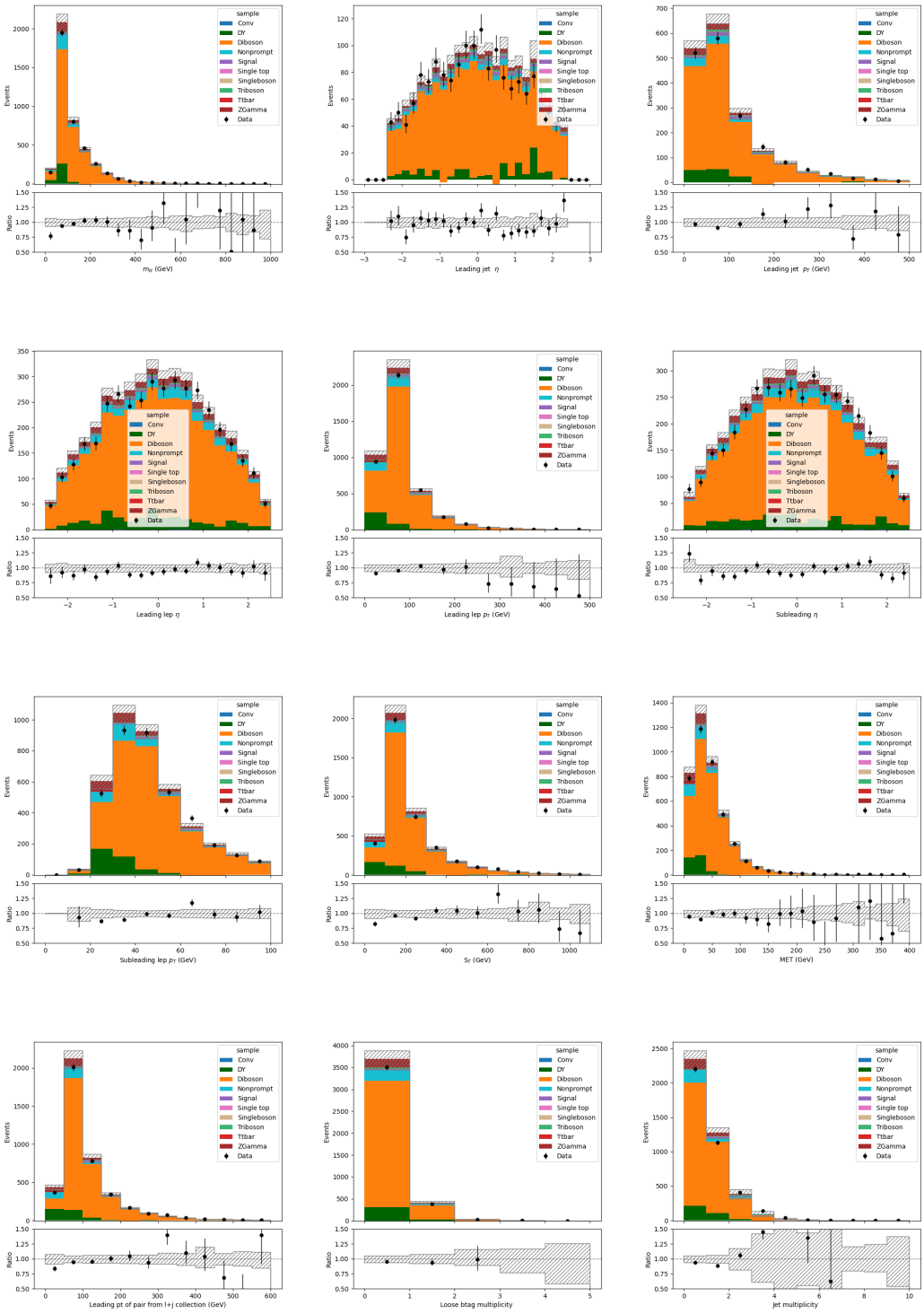


Figure E.6. Control region plots for the 3l CR (for UL16APV samples). The shaded gray band indicates the systematic uncertainty.

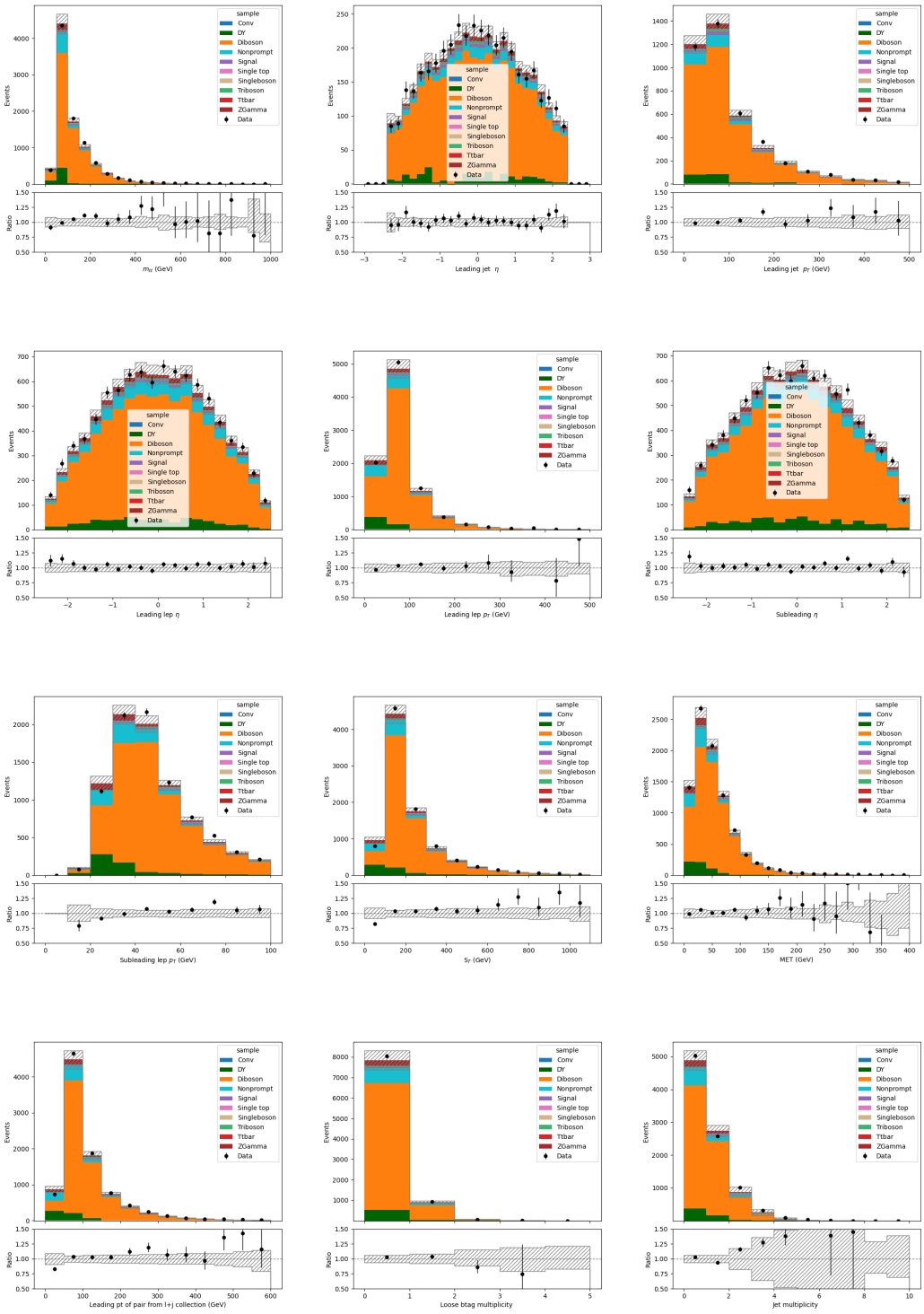


Figure E.7. Control region plots for the 3l CR (for UL17 samples). The shaded gray band indicates the systematic uncertainty.

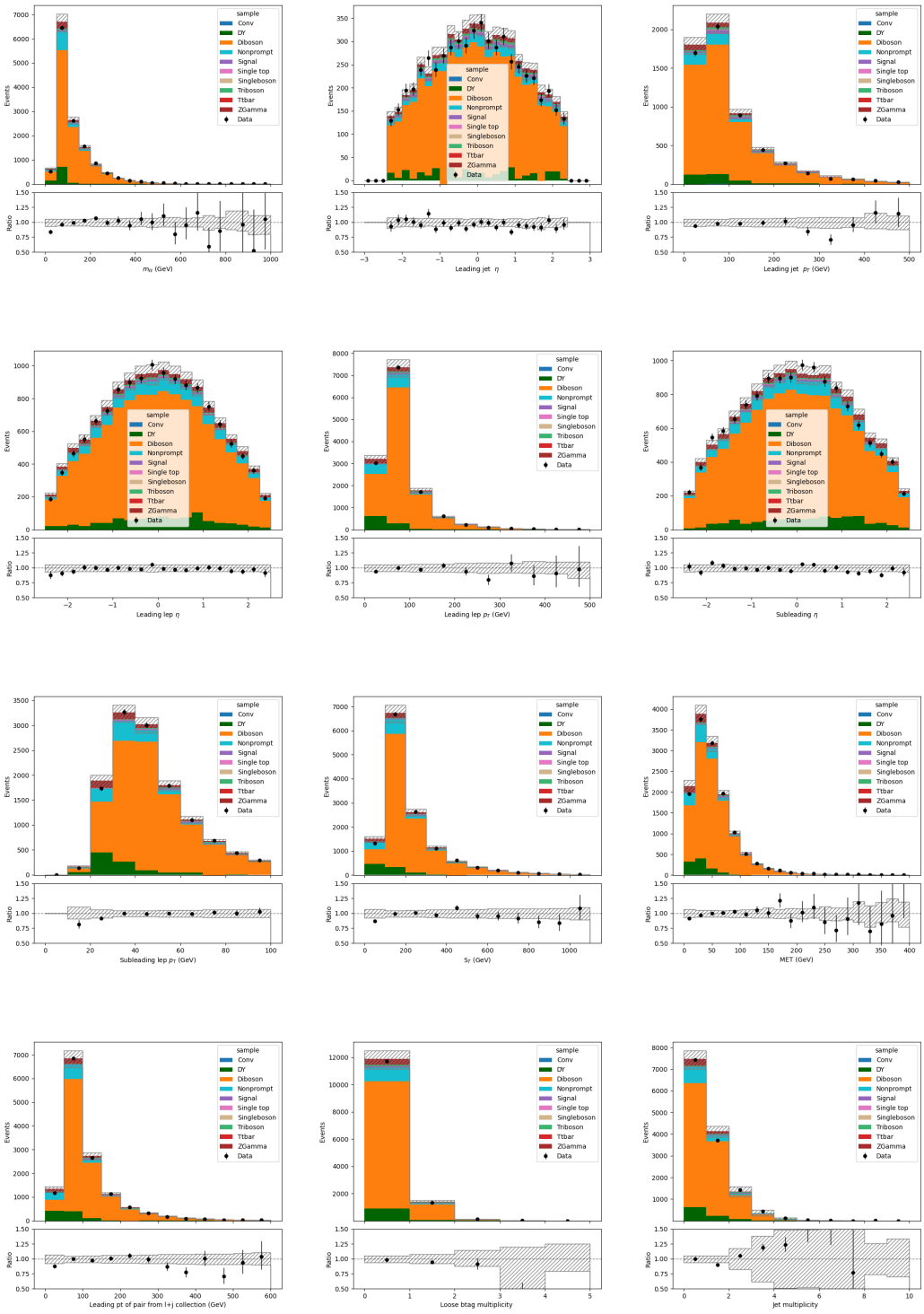


Figure E.8. Control region plots for the 3l CR (for UL18 samples). The shaded gray band indicates the systematic uncertainty.

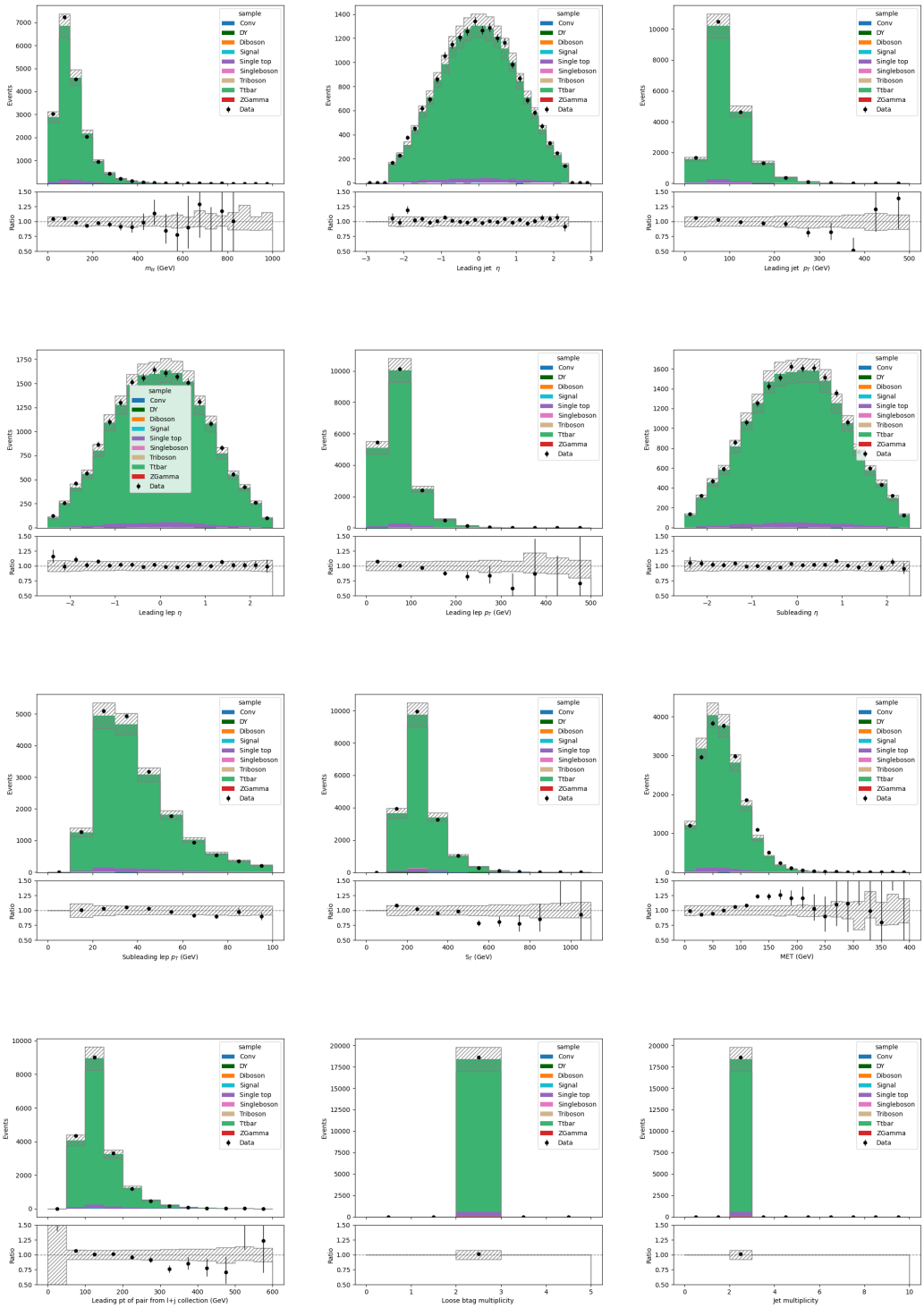


Figure E.9. Control region plots for the 2los $t\bar{t}$ CR (for UL16 samples). The shaded gray band indicates the systematic uncertainty.

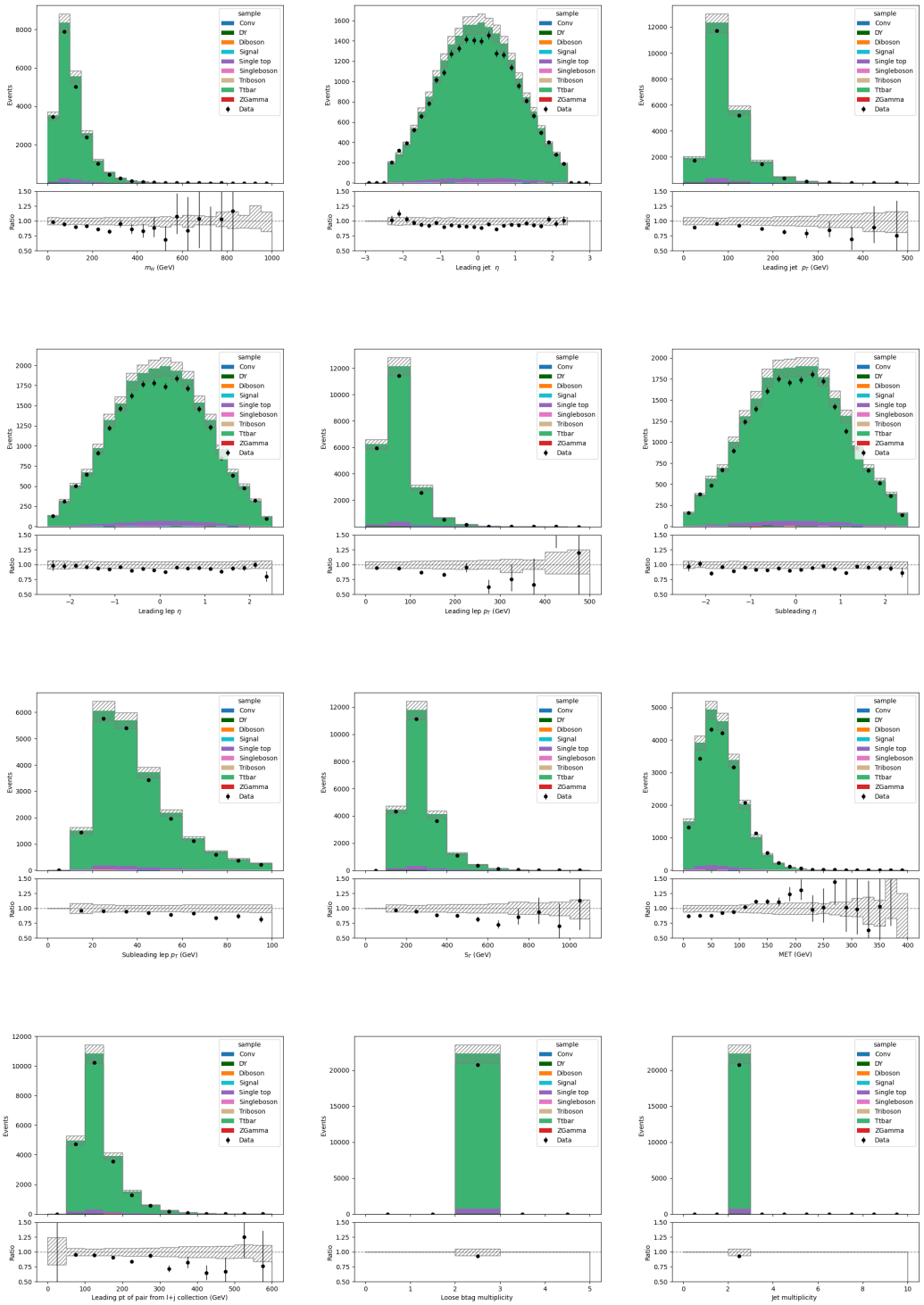


Figure E.10. Control region plots for the 2los $t\bar{t}$ CR (for UL16APV samples). The shaded gray band indicates the systematic uncertainty.

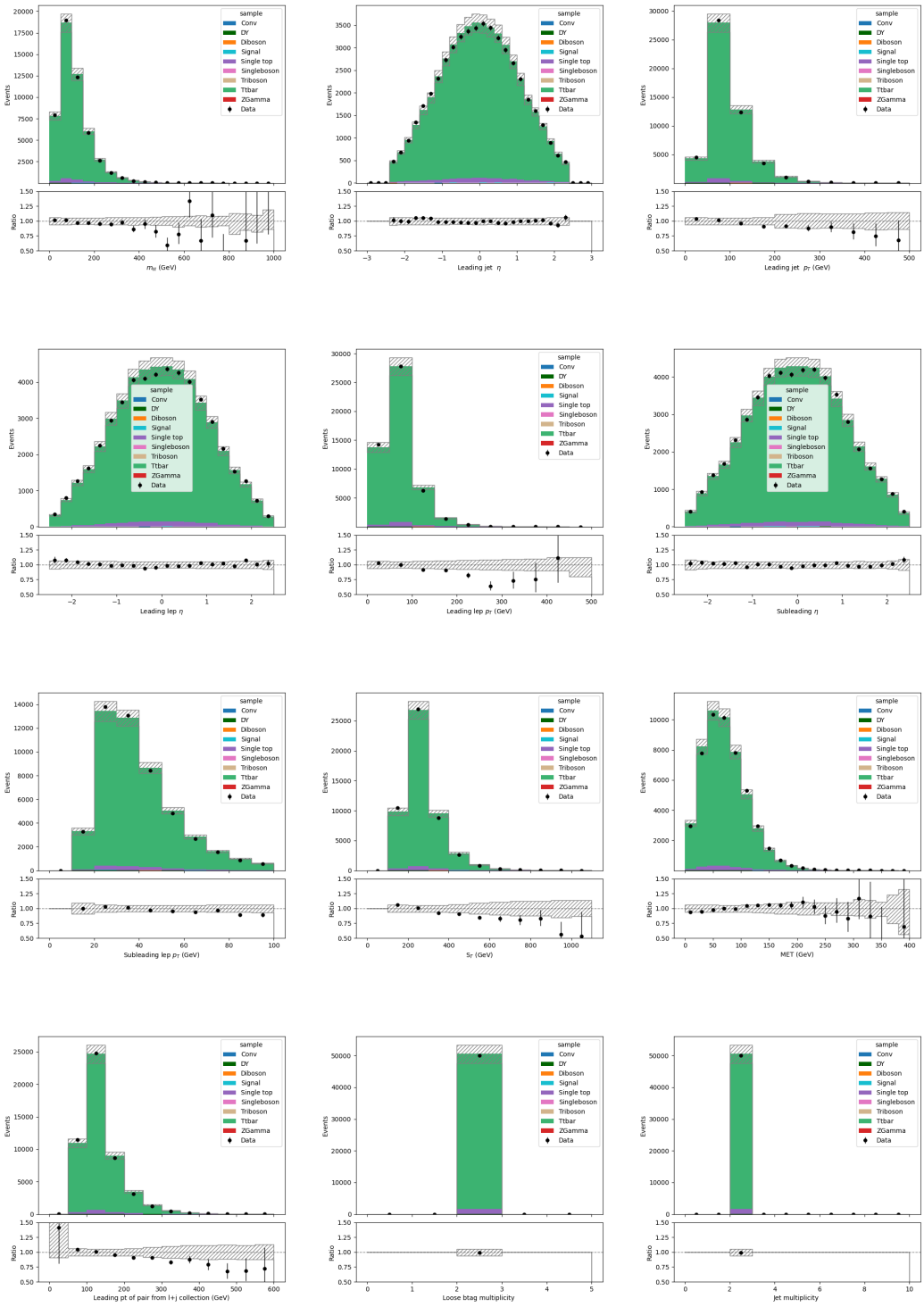


Figure E.11. Control region plots for the 2los $t\bar{t}$ CR (for UL17 samples). The shaded gray band indicates the systematic uncertainty.

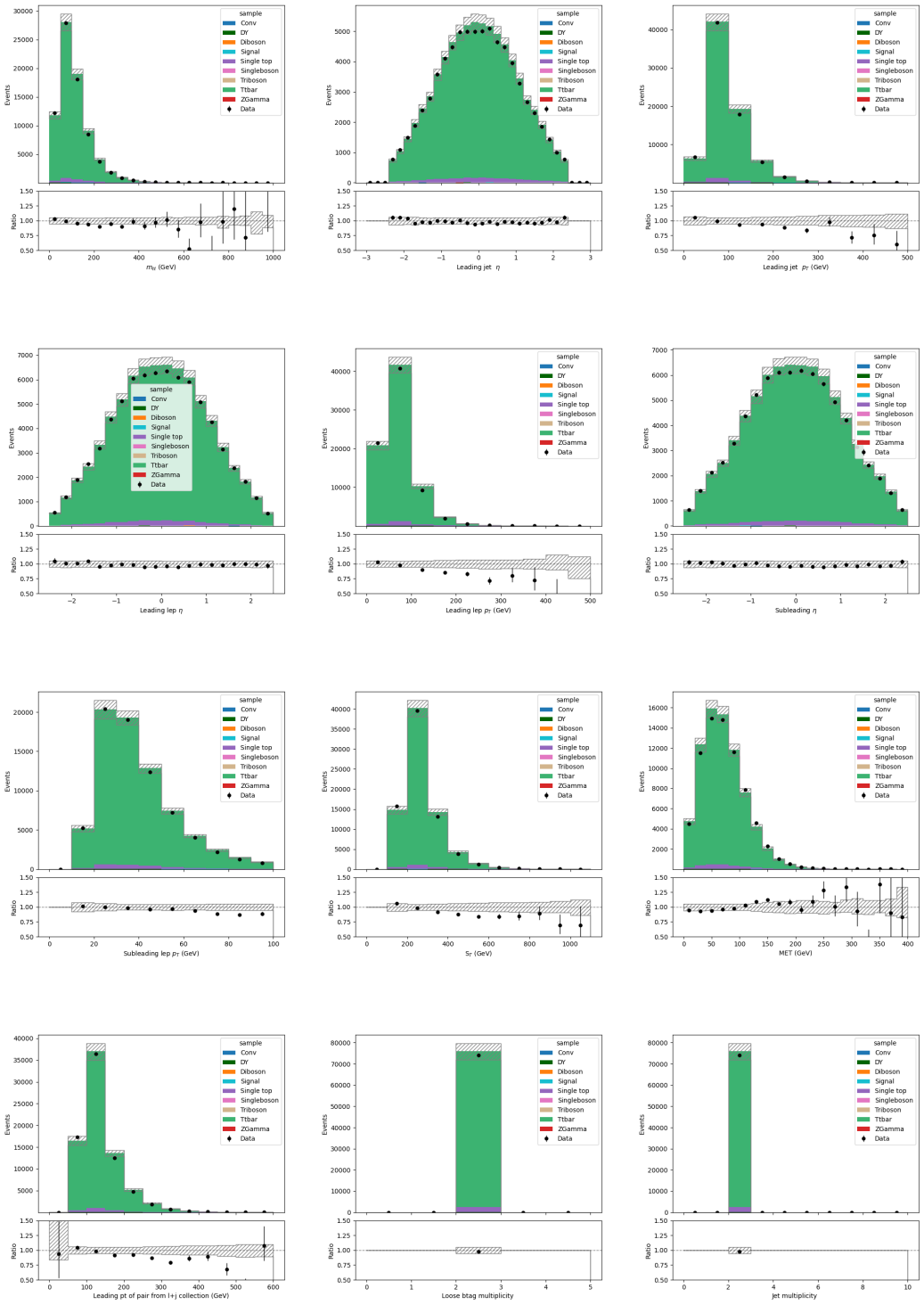


Figure E.12. Control region plots for the 2los $t\bar{t}$ CR (for UL18 samples). The shaded gray band indicates the systematic uncertainty.

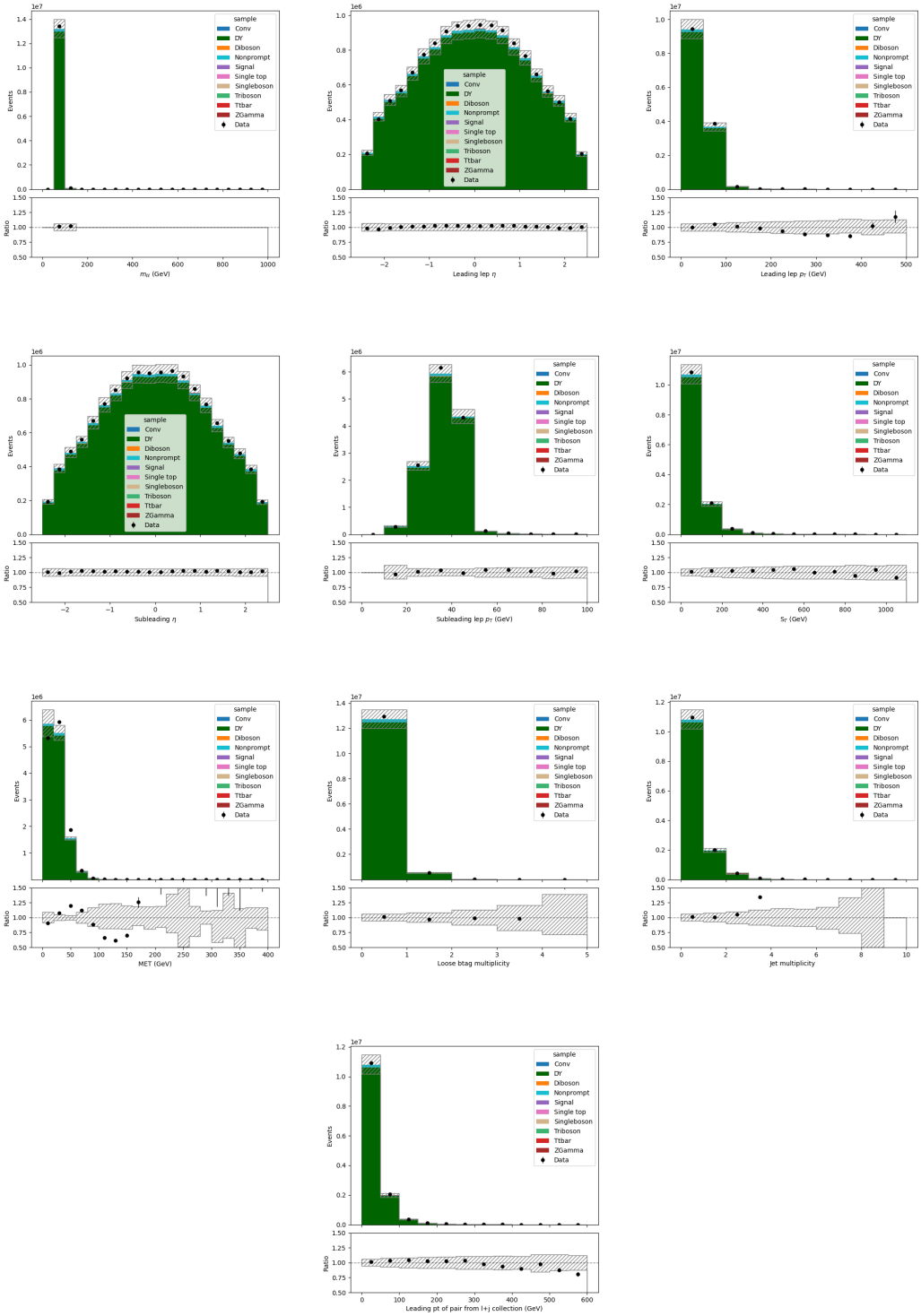


Figure E.13. Control region plots for the 2los Z CR (for UL16 smaples). The shaded gray band indicates the systematic uncertainty.

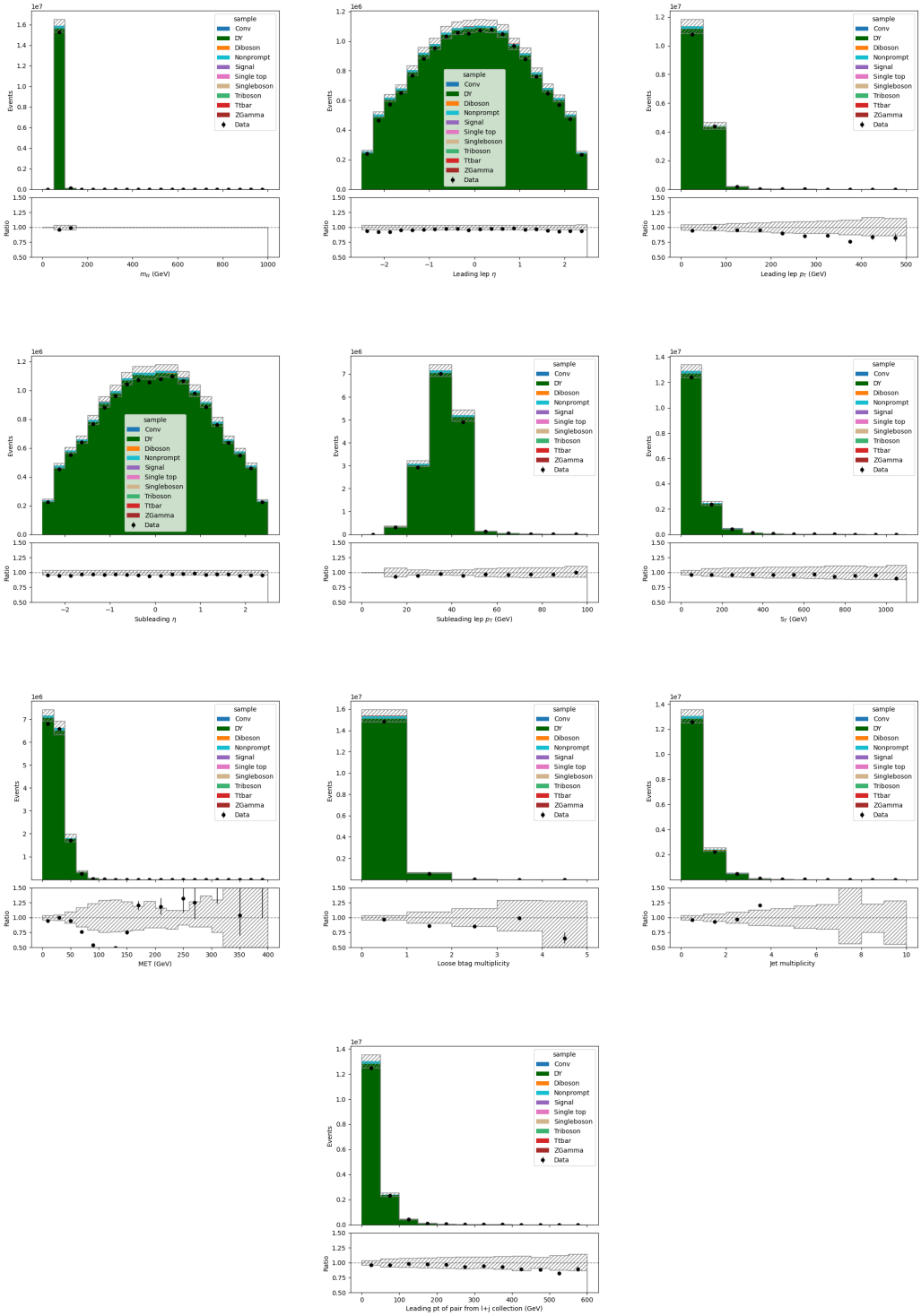


Figure E.14. Control region plots for the 2los Z CR (for UL16APV smaples). The shaded gray band indicates the systematic uncertainty.

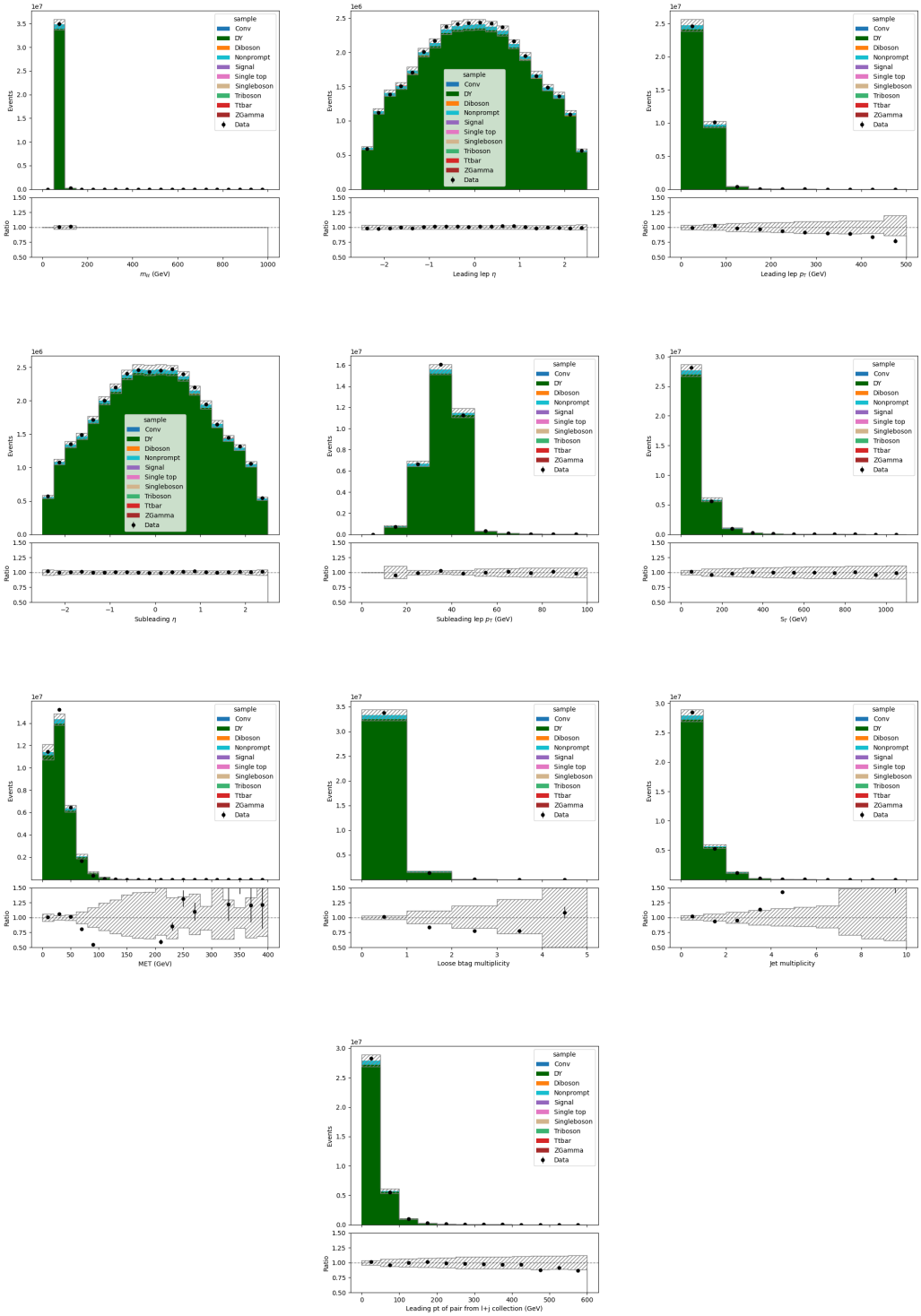


Figure E.15. Control region plots for the 2los Z CR (for UL17 smaples). The shaded gray band indicates the systematic uncertainty.

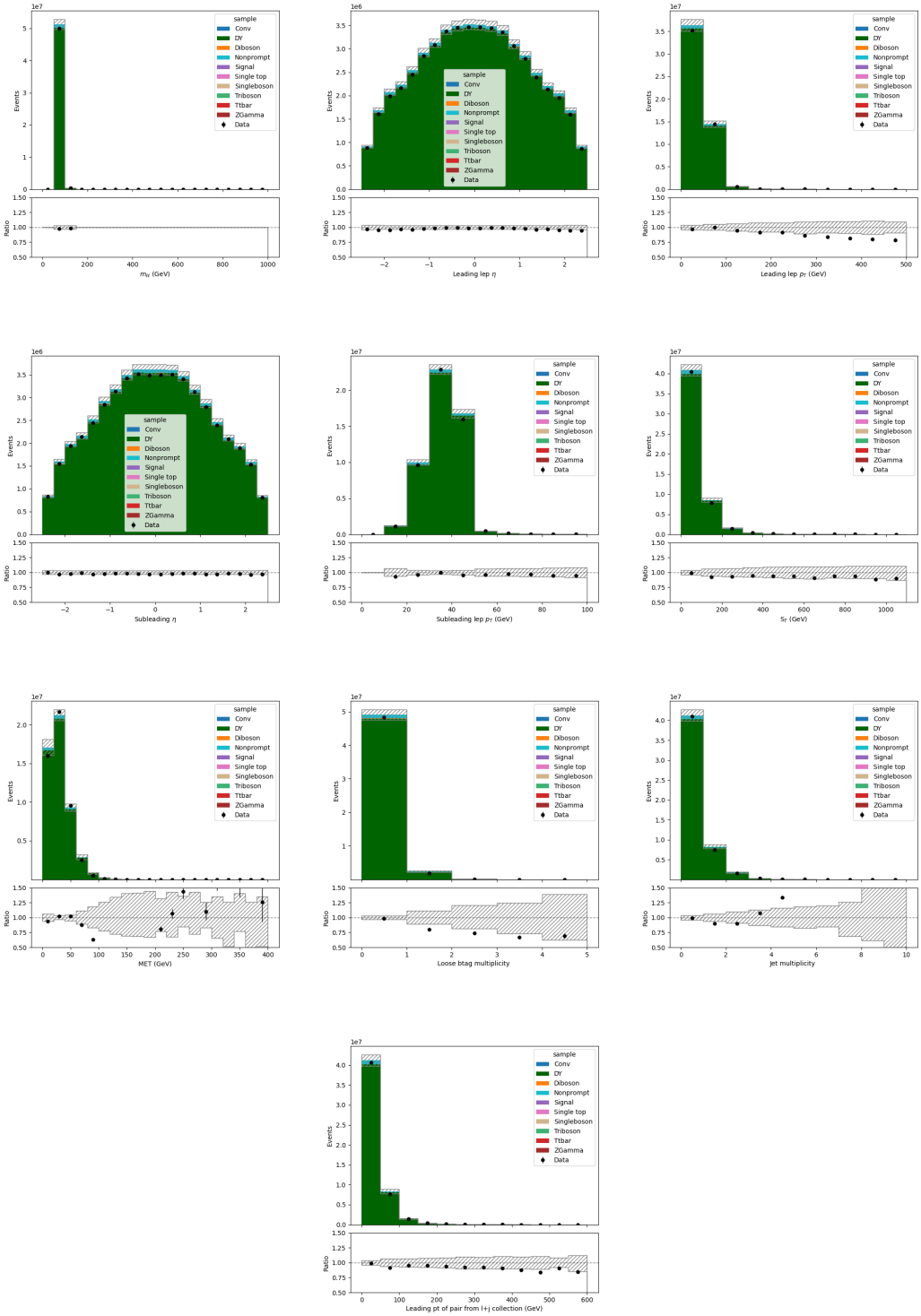


Figure E.16. Control region plots for the 2los Z CR (for UL18 smaples). The shaded gray band indicates the systematic uncertainty.

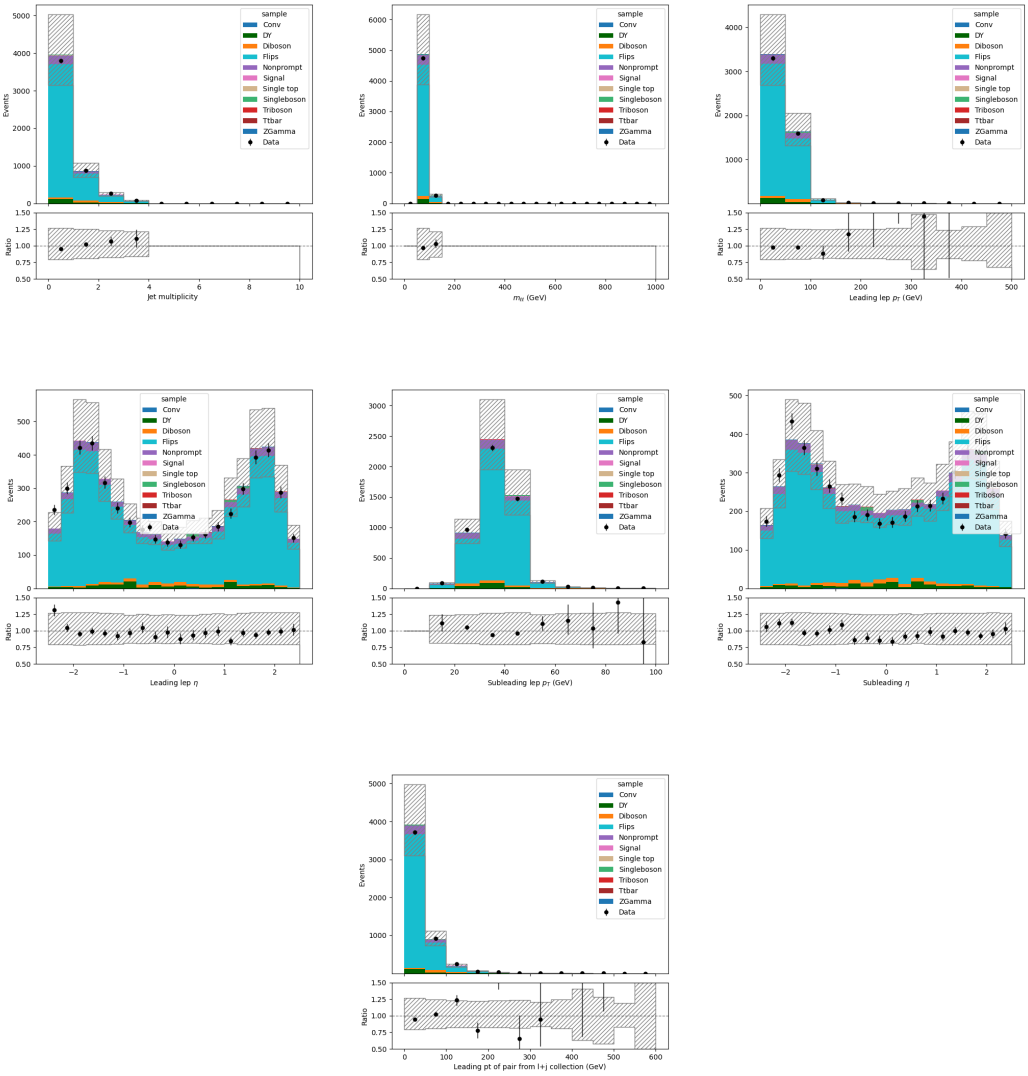


Figure E.17. Charge flip control regions for the UL16 samples after applying the scaling factors listed in 8.1. The shaded gray band indicates the systematic uncertainty.

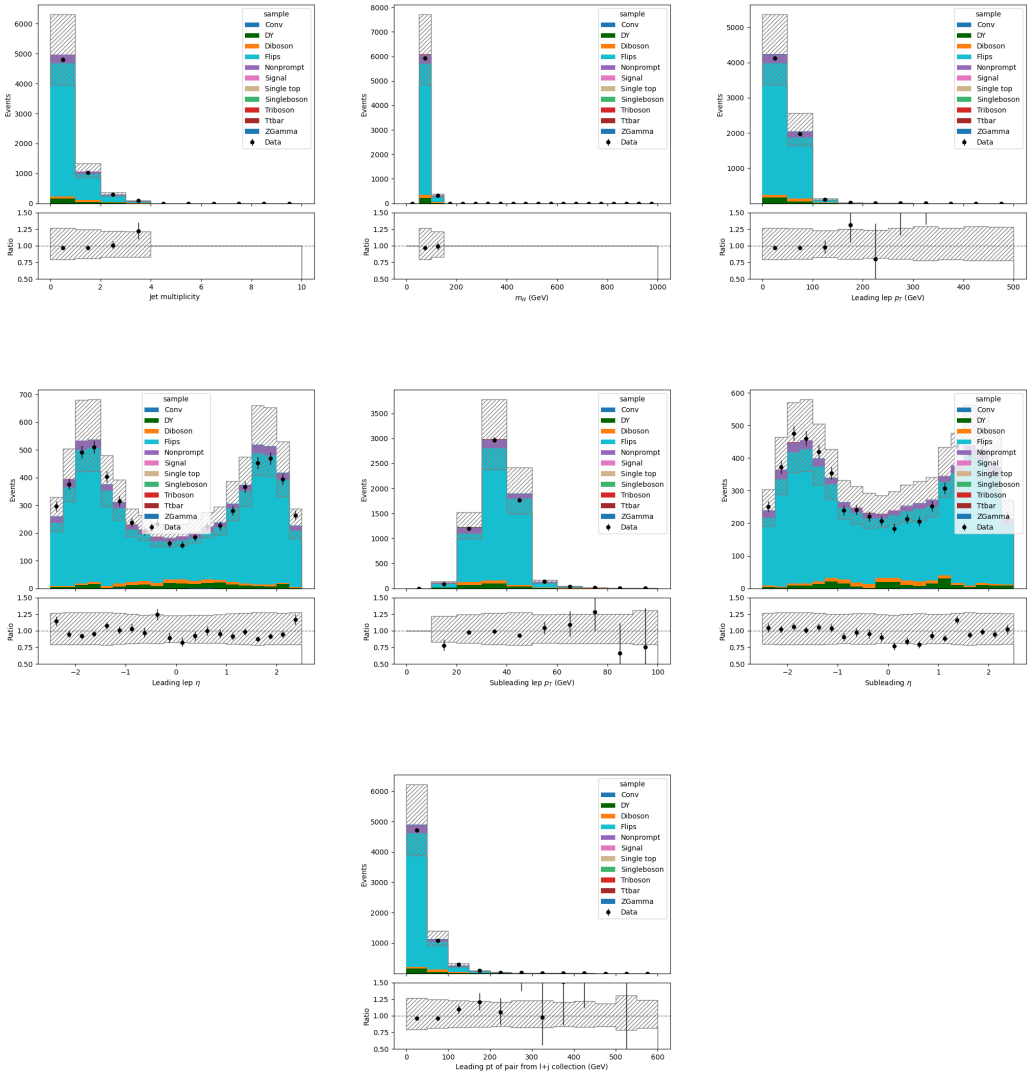


Figure E.18. Charge flip control regions for the UL16APV samples after applying the scaling factors listed in 8.1. The shaded gray band indicates the systematic uncertainty.

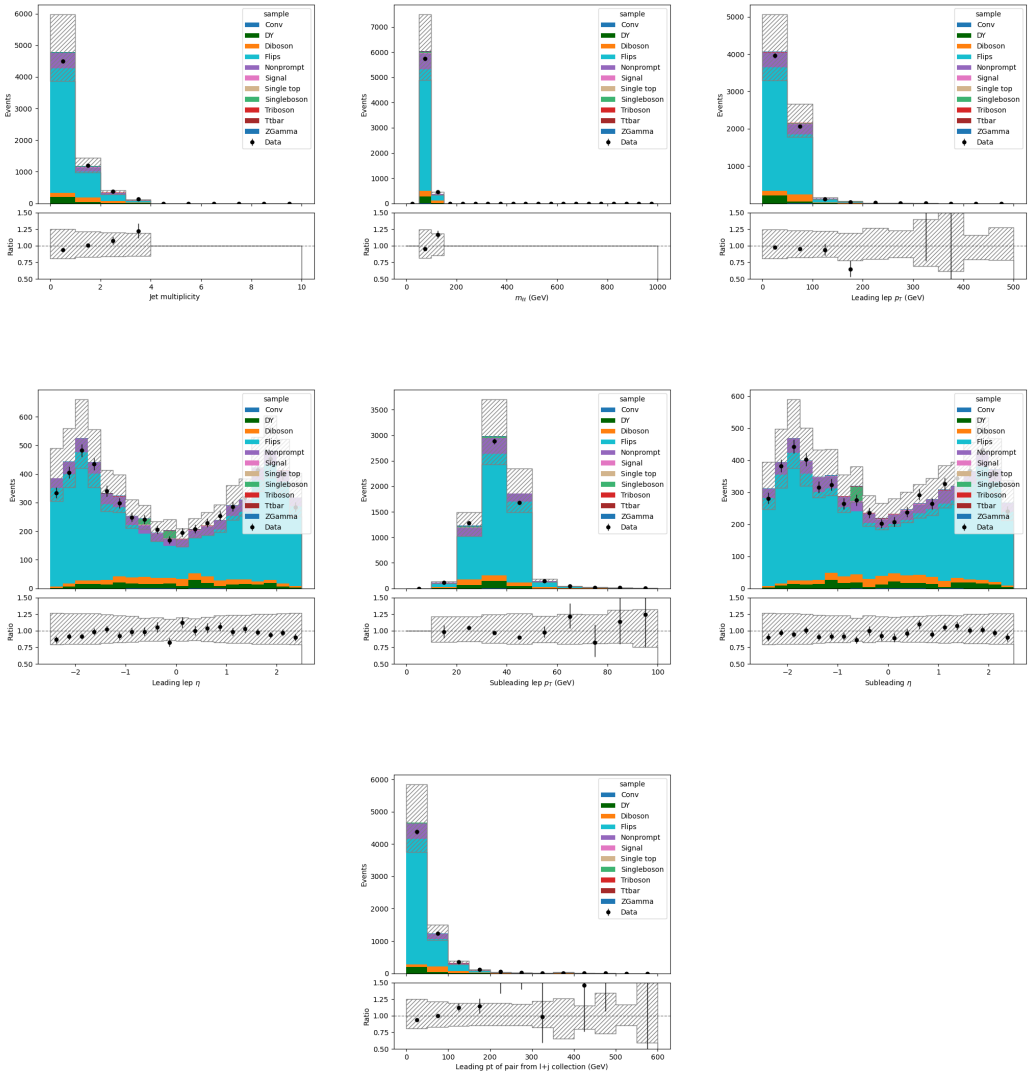


Figure E.19. Charge flip control regions for the UL17 samples after applying the scaling factors listed in 8.1. The shaded gray band indicates the systematic uncertainty.

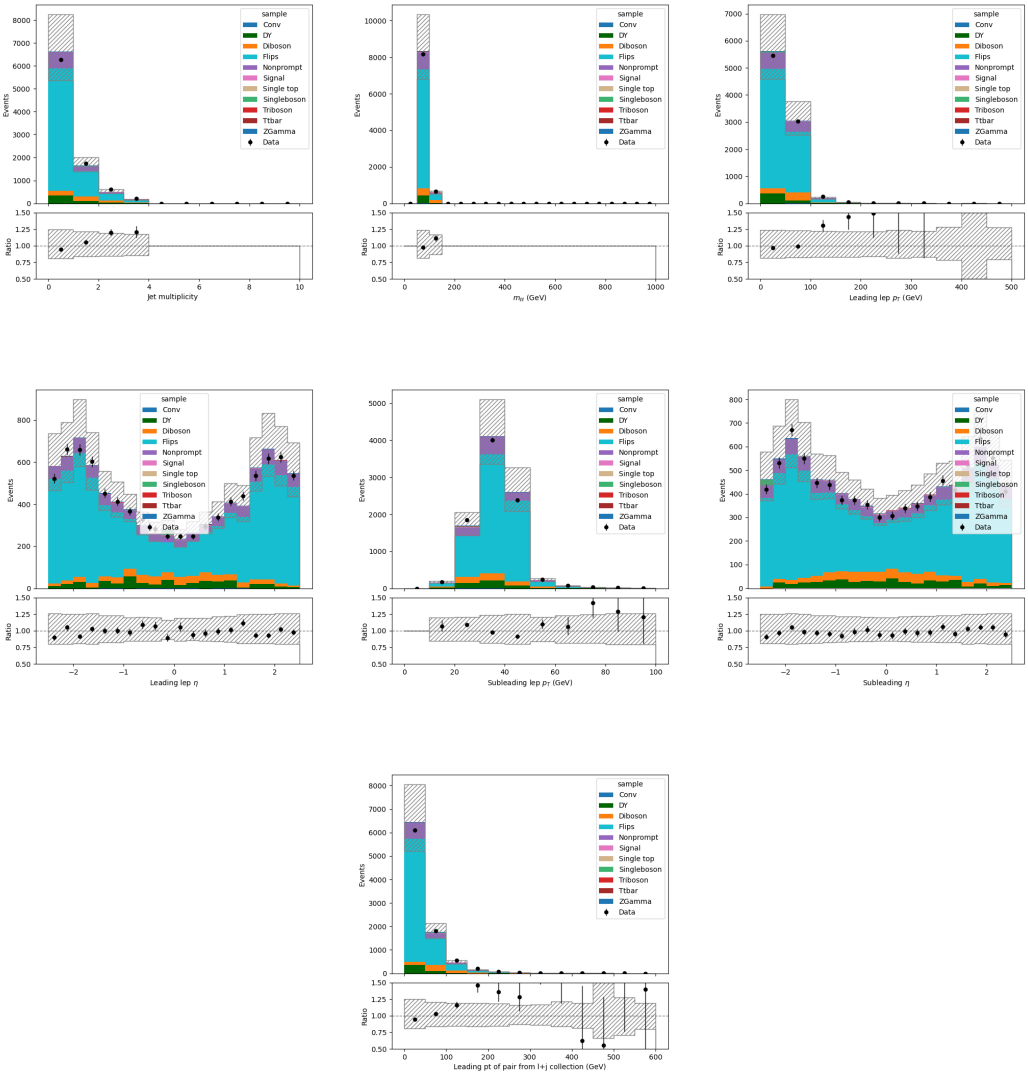


Figure E.20. Charge flip control regions for the UL18 samples after applying the scaling factors listed in 8.1. The shaded gray band indicates the systematic uncertainty.

APPENDIX F

DISCUSSION OF THE EFT DEPENDENCE OF THE SYSTEMATICS

The discussion in this appendix is based on the presentation linked in Ref. [87]. Before delving into the details of how the dependence on the WCs is being handled, there are two important points to keep in mind. The first important point is that TOP-19-001 (Ref. [4], the predecessor to the analysis presented in this thesis) and TOP-22-006 (Ref. [62], the analysis presented in this thesis) handle the quadratic dependence of the systematics on the WCs differently. The second important point is that neither approach is fully "correct".

We will start by describing the correct approach in Section F.1. Next, we will step through the implementation used in TOP-19-001 in Section F.2. Finally, we'll describe the implementation of TOP-22-006 in Section F.3.

F.1 The "correct" approach

Let us begin with the think about the "correct". approach For now, let us just consider the theory systematics that we obtain from the MC generation. The up/down weights for these systematics are calculated by MadGraph (MG) at the starting point. In principle, there is no reason to assume the up/down weights at any other point in the phase space would be the same as they are at the staring point.

Thus, if we really wanted the correct up/down variations, we would need to generate a dedicated sample at each point in the phase space we are interested in (or implement some modification to MG to calculate the up/down weights at every reweight point, but even if there was a way to do this, the number of points required

for e.g. PDFs would make the approach potentially infeasible). However, for now, let us set aside the fact that this approach is not feasible at scale, and simply try to explore the approach conceptually. A visualization of this approach is provided in Figure F.1.

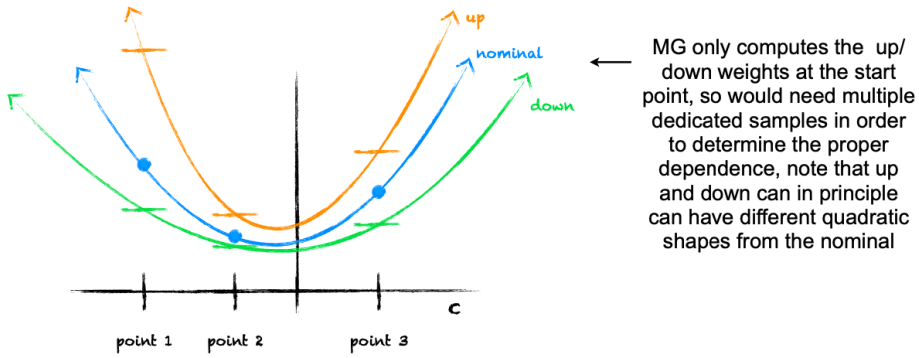


Figure F.1. Visualization of the "correct" quadratic dependence of the systematics on the WCs.

For 1 WC, we would need to produce 3 dedicated samples at 3 points in WC space, letting MG compute the correct up/down fluctuations at each point. With these weights, we could then fit a quadratic to the up, nominal, and down points. This should in principle provide the correct dependence, but this is of course not a feasible approach with many WCs.

In Ref. [10], we attempted a small-scale check of the "correct" implementation. For a few WCs, we produced 1d samples at multiple dedicated points, and we fit a nominal, an up, and a down quadratic. We then scaled all three quadratics to the SM so that we could directly compare the shapes without being distracted by

normalization differences. As shown in Figure 11 of Ref. [10], we found very minimal change to the shape. However, one caveat to keep in mind is that this study was done with inclusive cross sections (so any shape differences that might be more apparent differentially could be smeared out here).

F.2 The TOP-19-001 approach

TOP-19-001 applied the up/down weights (calculated at the starting point) to the reweighted SM point (doing this event by event) in order to estimate the up/down at the SM, then used this value as a rate uncertainty for each category. For example, let us consider a single event, visualized in Figure F.2.

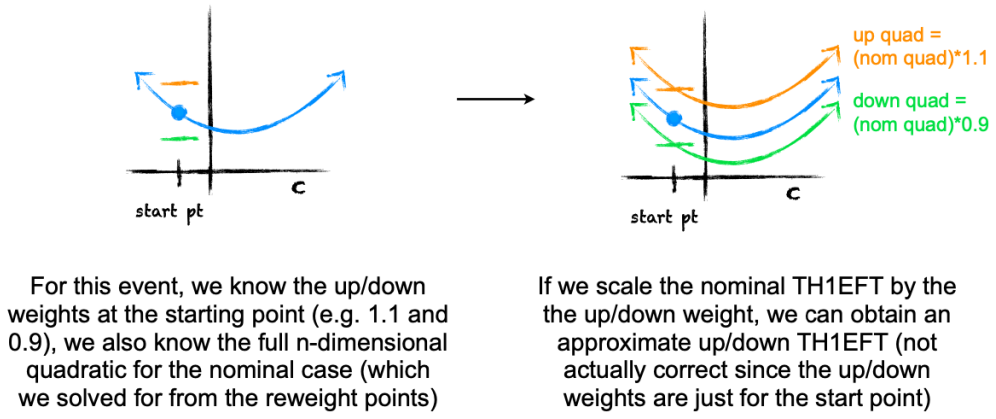


Figure F.2. Visualization of one event in the TOP-19-001 approach.

Now that we have walked through a single event, let us next think about a bin that has two events. In this example, let us assume the first event is the one visualized in Figure F.2, and the second event is one that has larger up/down weights. This is visualized in Figure F.3.

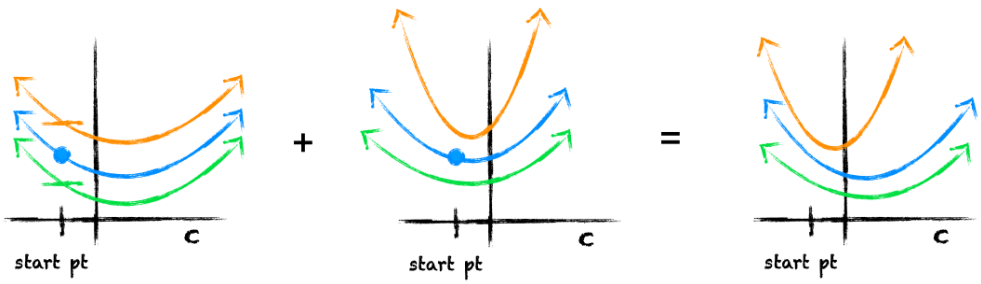


Figure F.3. Visualization of two events in the TOP-19-001 approach.

After scaling by the up/down weights and summing, the resulting up/down and nominal TH1EFTs in principle could have completely different shapes. However, it should again be emphasized that these differences in shape are not "correct" since they are just based on the starting point (and it is not known if this is an underestimation or overestimation of the shape differences).

Regardless, TOP-19-001 did not make use of this shape information in the fit. Rather, the next step in TOP-19-001's implementation was to evaluate the up/down TH1EFTs at the SM, and use that weight as a rate uncertainty to hand to Combine. As this uncertainty is implemented as a rate uncertainty in Combine, it will not change as a function of the WC (it is fixed at this value for all values of the WCs). While we should again emphasize that this is not correct, it may be a valid approximation. This final step in the implementation is visualized in Figure F.4.

F.3 The TOP-22-006 approach

The TOP-22-001 approach starts out in the same way as TOP-19-001. We have a HistEFT for the up/down and nominal histograms, which are calculated in the same way as TOP-19-001. In other words, each event's fit is scaled by the up/down weight (where the up/down weights were calculated for the start point) and we sum

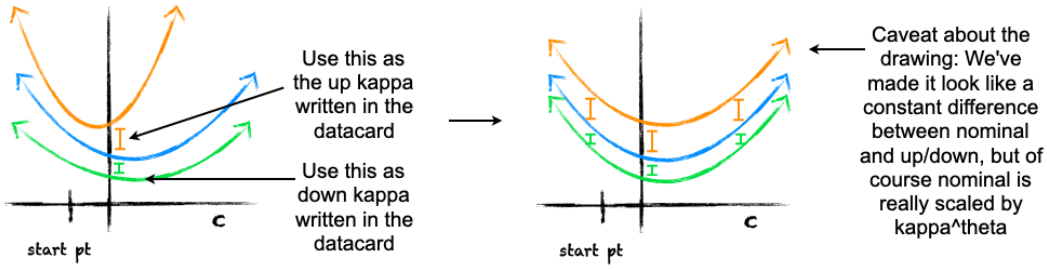


Figure F.4. Visualization of the evaluation of the systematics at the SM in the TOP-19-001 approach.

the nominal fits together, sum the up fits together, and sum the down fits together. The visualization for this has been shown in Figure F.3.

However, the next steps are different from TOP-19-001. Instead of evaluating the up/down fits at the SM and using that number as a rate uncertainty, we decompose the quadratic fits (for nominal/up/down) into template histograms that encode the quadratic parametrization (i.e. the method from CMS AN-20-204). The templates carrying the decomposed quadratic dependance are then what we pass to `text2workspace`.

Thus, with the TOP-22-006 approach, for every point in WC space that the fit might want to consider, the nominal/up/down quadratics can be evaluated to find the contribution at that particular point in the WC space. In other words, the shape systematics depend on the WCs. Please note that here "shape systematics" refer to systematics that we implement in Combine with templates (as opposed to systematics implemented via a number in the text datacard).

The following bullets summarize the main points to take away from this discussion of TOP-22-001's implementation of the quadratic dependence of the systematics on the WCs:

- In TOP-22-006, the systematics depend on the WCs, and the shapes of the quadratic dependences can in principle be different from the nominal.
- The TOP-22-006 approach is different from TOP-19-001's approach (since in TOP-19-001 the systematics did not depend on the WCs).
- The TOP-22-006 approach is also different from the "correct" approach (since we are estimating the shape of the up/down parametrization based on the up/down weights at the starting point).

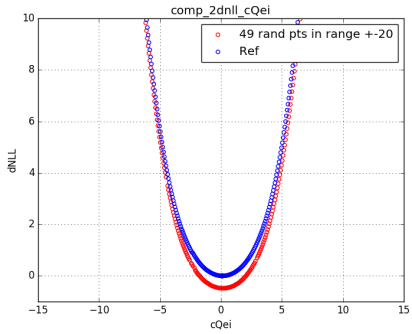
Even though this discussion has focused on theory uncertainties, it is worth mentioning that the same points apply to the rest of the shape systematics as well. For the other systematics we still scale each event's fit by the up/down weights, so we still get an up/down parametrization that is different from the nominal, and still pass this quadratic information to Combine. Thus, even experimental systematics (like lepSF or PreFire) will depend on the WCs.

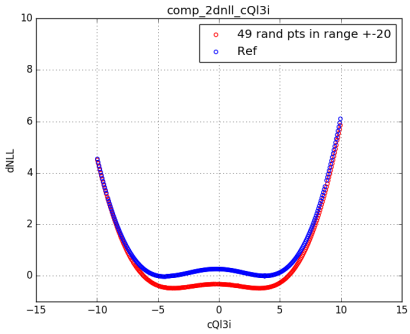
APPENDIX G

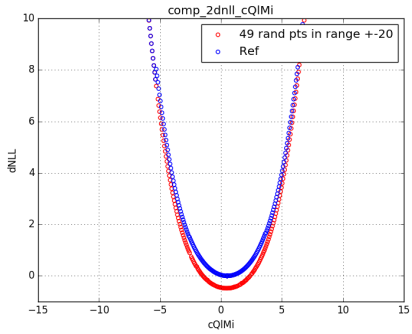
EXAMPLE USAGE OF RANDOM STARTING POINT APPROACH FOR NAVIGATING FALSE MINIMA

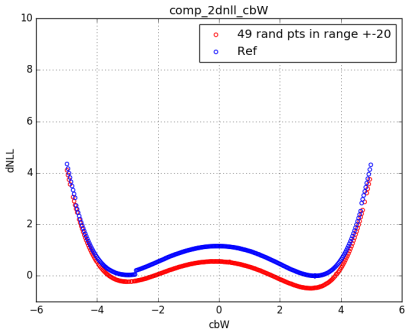
This appendix illustrates how the random starting point method can help to navigate local minima, using the Ref. [4] analysis as an example. The 1d profiled likelihood fits for each of the 16 WCs from [4] are first run without any random starting points; the results of these fits are shown by the blue points in Figure G.1. The blue scan points show several discontinuities, indicating that the fit had become "stuck" in a local minimum. We then performed the scan with 49 random starting points for the profiled parameters, shown in red in Figure G.1. The red scan points show continuous NLL values, and in many cases the random starting point scan was able to identify deeper global minima than had been identified in the original scan.

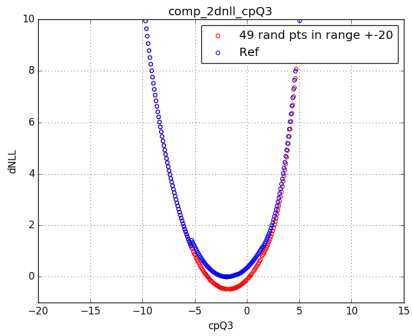
In Figure G.1, the y axis has been scaled such that $y = 0$ corresponds to the NLL obtained from the pre-fit fit that `Combine` performs prior to the scan. In this pre-fit fit, all 16 WCs are profiled. Ideally, this fit should identify the true global minimum. If the pre-fit fit correctly identifies the true global minimum, it would not be possible for a scan point to find a better likelihood value than the one obtained by the pre-fit fit, so there would never be a point on any scan that is below zero. However, in Figure G.1 we see many scanpoints with NLL values below zero (i.e. better than the NLL from the pre-fit fit). This is an indication that the pre-fit for [4] was also challenged by local minima. As discussed in Section 10.3, the [4] analysis worked around this issue by making use of the information in 2d scans. However, the random starting point approach is more general workaround to the issue.

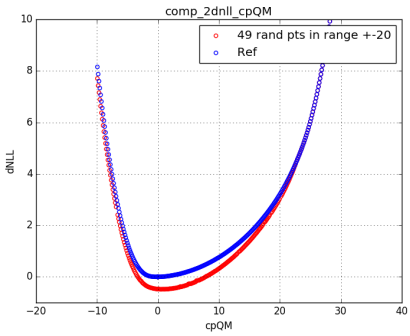

comp_2dnll_cQei
49 rand pts in range +-20
Ref
dNLL
cQei

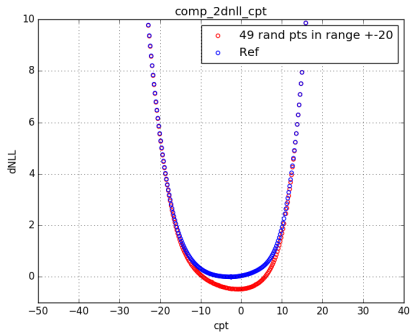

comp_2dnll_cQl3i
49 rand pts in range +-20
Ref
dNLL
cQl3i

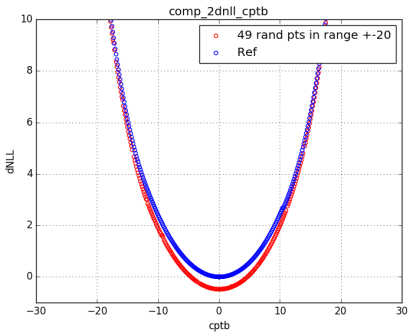

comp_2dnll_cQlMi
49 rand pts in range +-20
Ref
dNLL
cQlMi


comp_2dnll_cbW
49 rand pts in range +-20
Ref
dNLL
cbW


comp_2dnll_cpQ3
49 rand pts in range +-20
Ref
dNLL
cpQ3


comp_2dnll_cpQM
49 rand pts in range +-20
Ref
dNLL
cpQM


comp_2dnll_cpt
49 rand pts in range +-20
Ref
dNLL
cpt


comp_2dnll_cptb
49 rand pts in range +-20
Ref
dNLL
cptb

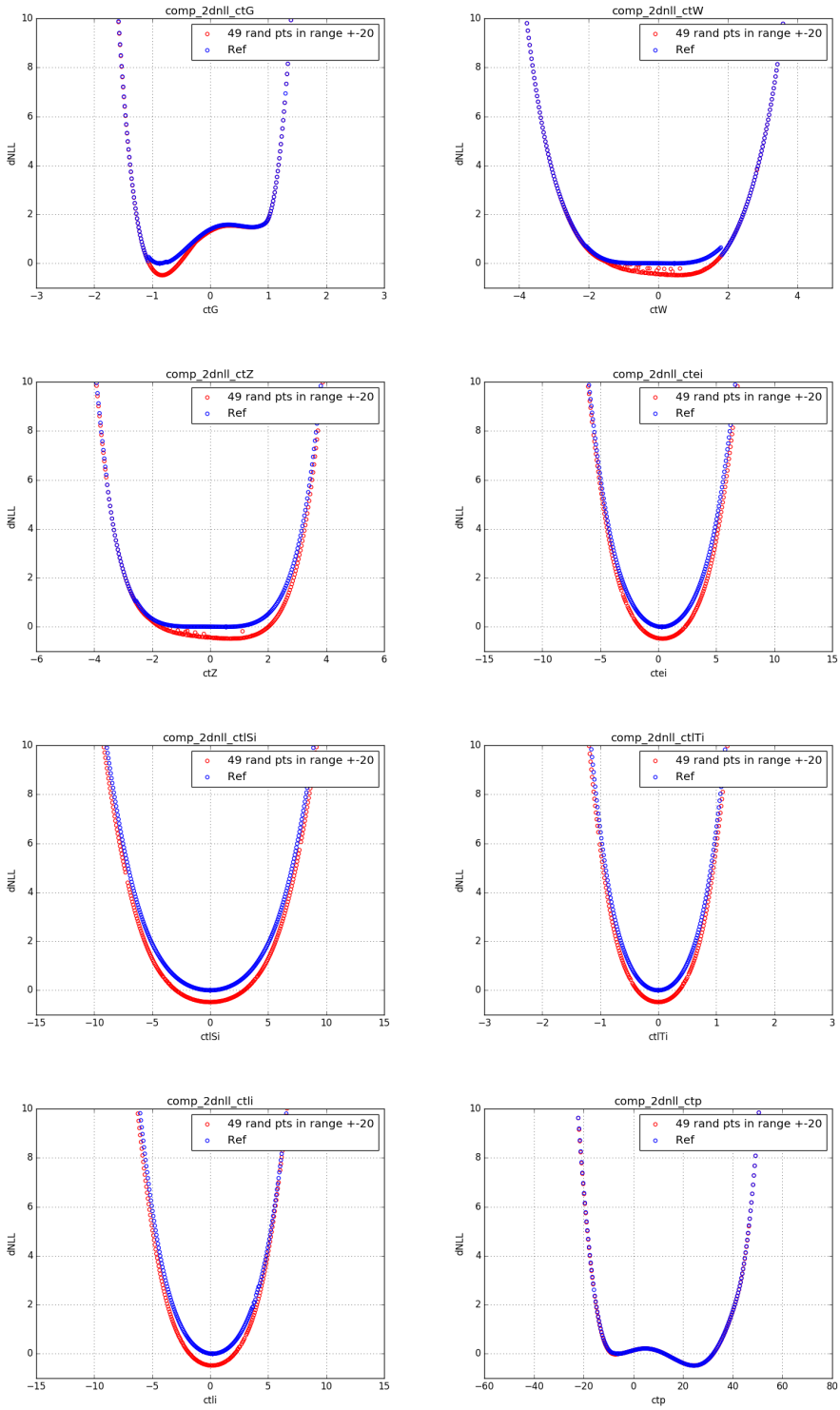


Figure G.1. Profile fits from the [4] analysis with (red) and without (blue) random starting points for the profiled POIs.

APPENDIX H

SENSITIVITY INTERPRETATION FIGURE OF MERIT PLOTS

This section will describe the details of the "bottom-up" approach to the sensitivity interpretation introduced in Section 12.3. The goal of this approach is to obtain a quantitative measure each bin's contribution to the setting of the 2σ limit. To this end, we define a figure of merit (FOM) that takes into account the difference between the prediction and the observation in each bin, scaling this difference by the total uncertainty on the prediction. We can thus write the FOM as follows:

$$\text{FOM} = \frac{|N_{\text{pred}} - N_{\text{obs}}|}{\sigma_{\text{tot}}}, \tag{H.1}$$

Where N_{pred} is the predicted number of events in the bin, N_{obs} is the observed number of events in the bin, and σ_{tot} is the total uncertainty on the prediction (which is the quadrature sum of the uncertainty from the fit, and the statistical uncertainty $\sqrt{N_{\text{pred}}}$).

The FOM provides a quantitative measure of how well the prediction agrees with the observation in a given bin. A FOM of 0 would imply the prediction is exactly equal to the observation, while a FOM of 1 would imply a roughly 1 standard deviation disagreement. We can use the FOM to understand which of the 178 bins are the most relevant for setting the 2σ limits by evaluating the FOM at the best fit point and at the $+2\sigma$ and -2σ limits. Bins where the FOM is much larger at the 2σ limit than at the best fit point would indicate that the bin is relevant to the setting of the 2σ limit. To summarize, we know that the reason why the fit hits the 2σ limit at a

particular point in the scan is because the agreement between the prediction and the observation (in one or more bins) becomes too large, and the FOM study is designed to help identify these bins.

Since we are interested in the profiled limits, it is necessary to run a profiled fit at the 2σ limits for all WCs (i.e. with the scanned WC fixed to its 2σ limit with all other floating). For 26 WCs, there are thus 52 total fits. Once these fits have been run, the ingredients for the FOM (as defined in Eq. H.1) can be read from the Combine fit result, and the FOM can be calculated for all 178 bins for each of the 52 fits.

Since the study of the FOMs results in a lot of information (two FOMs for each of the 178 bins for all 26 WCs at both the $+2\sigma$ and -2σ limits), it is important to visualize the information in a way that highlights the most relevant aspects of the results. We recall that we are interested in cases where the FOM is significantly worse (i.e. larger) at the 2σ limit than at the best fit point. It is thus useful to plot the FOM values such that these interesting cases stand out, as shown in Figure H.1. In this example, the FOM at the -2σ limit for c_{QQ}^{1} are shown in dark blue, and the FOM at the best fit point is shown in light blue. A line is drawn connecting each point, with the color of the line determined by whichever value is larger (e.g. if the FOM at the 2σ limit is larger than the FOM at the best fit, the line is dark blue). Bins that would be considered important are thus bins that exhibit long blue lines (as these long blue lines indicate cases where the FOM is much worse i.e. larger at the 2σ limit than the FOM at the best fit point).

We can make further adjustments to the visualization to help us to identify important bins. Since we are essentially interested in the length of the dark blue lines shown in Figure H.1 (as opposed to their overall y axis location), we can rescale each FOM relative to the best fit FOM. In other words, we can plot $\text{FOM} - \text{FOM}_{\text{bestfit}}$ on the y axis. Performing this scaling, we obtain Figure H.2 (still for the -2σ limit for

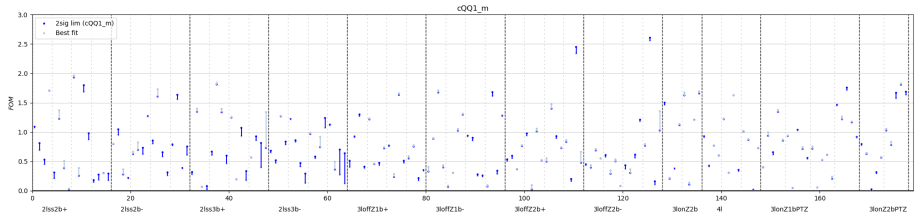


Figure H.1. An example FOM plot for the -2σ limit for c_{QQ}^{1}.

c_{QQ}^{1}). It now becomes much more clear that the important bins for this WC are the 2ℓss bins (especially in the bins that require at least 3 b tags). This result is consistent with our expectations, as this WC is from the four-heavy group of operators, which significantly impacts the $t\bar{t}t\bar{t}$ process.

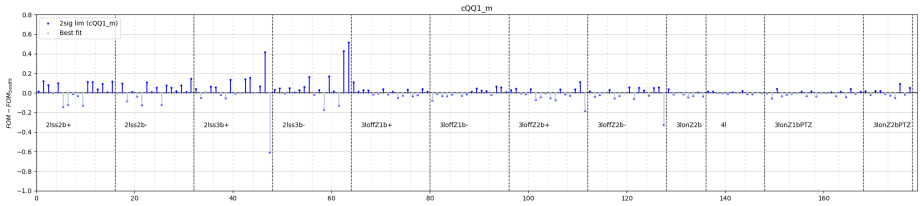
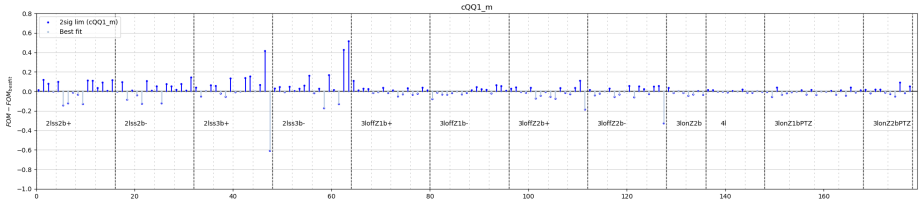
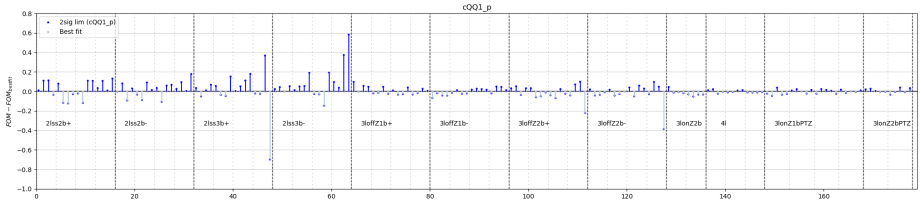
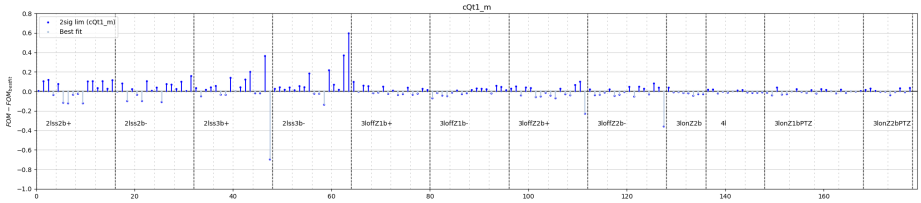


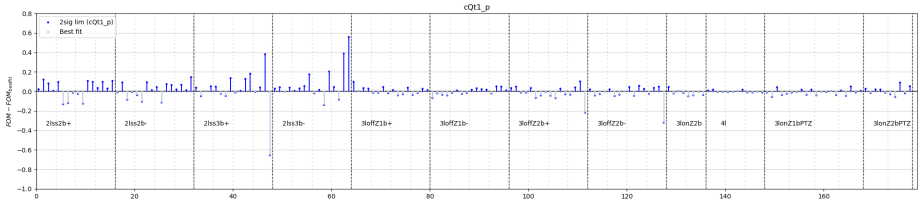
Figure H.2. An example FOM plot for the -2σ limit for c_{QQ}^{1}, scaled to the FOM at the best fit point.

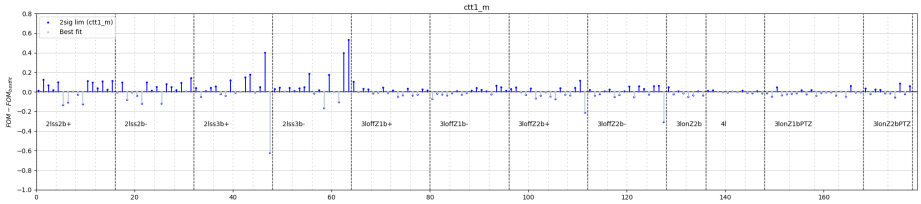
Equipped with these improvements in the visualization, we can proceed to make all 52 of the FOM plots (shown in Figure H.3). The next step is the carefully examine the 52 plots and look for trends. The trends and resulting conclusions obtained from this study (in combination with insight obtained from the "top-down" approach) is described in Section 12.3.

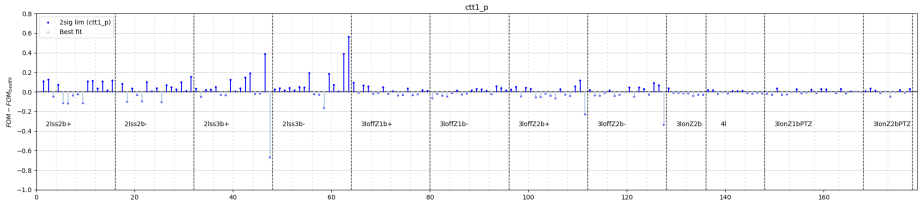

cQQ1_m
2sig lim (cQQ1_m)
Best fit
$FOM - FOM_{nom}$
2lss2b+
2lss2b-
2lss3b+
2lss3b-
3loffZ1b+
3loffZ1b-
3loffZ2b+
3loffZ2b-
3lonZ2b
4l
3lonZ1bPTZ
3lonZ2bPTZ

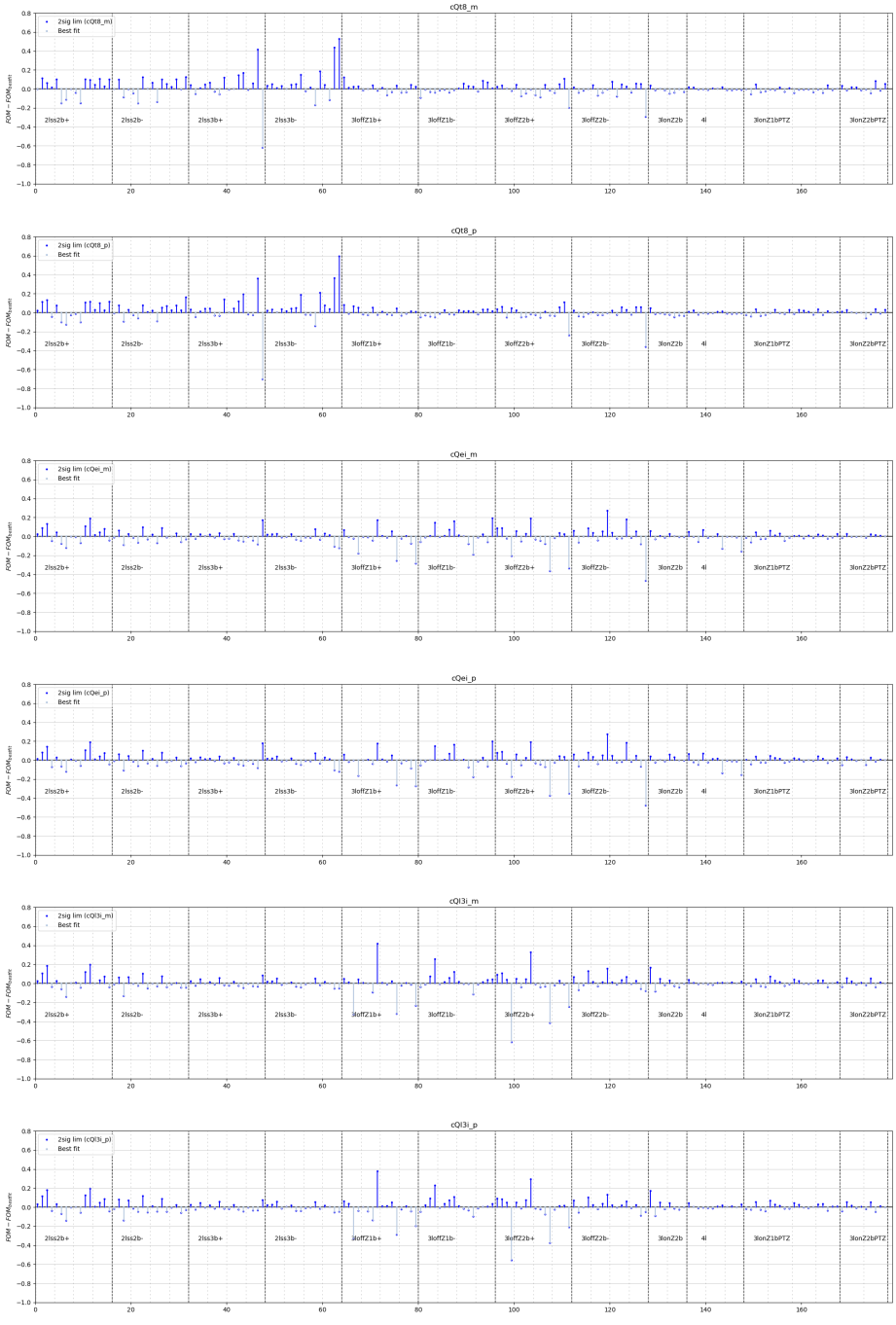

cQQ1_p
2sig lim (cQQ1_p)
Best fit
$FOM - FOM_{nom}$
2lss2b+
2lss2b-
2lss3b+
2lss3b-
3loffZ1b+
3loffZ1b-
3loffZ2b+
3loffZ2b-
3lonZ2b
4l
3lonZ1bPTZ
3lonZ2bPTZ

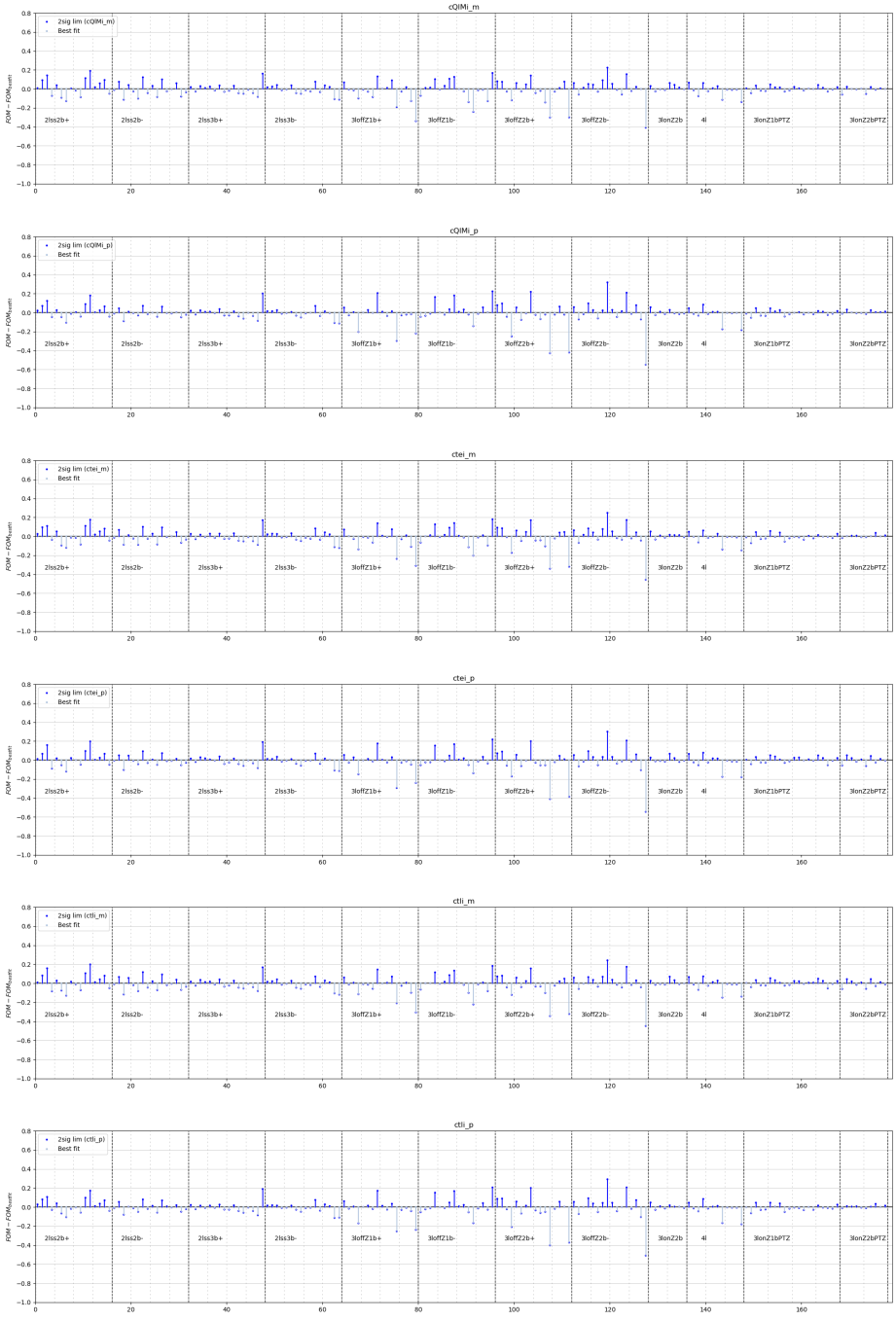

cQt1_m
2sig lim (cQt1_m)
Best fit
$FOM - FOM_{nom}$
2lss2b+
2lss2b-
2lss3b+
2lss3b-
3loffZ1b+
3loffZ1b-
3loffZ2b+
3loffZ2b-
3lonZ2b
4l
3lonZ1bPTZ
3lonZ2bPTZ

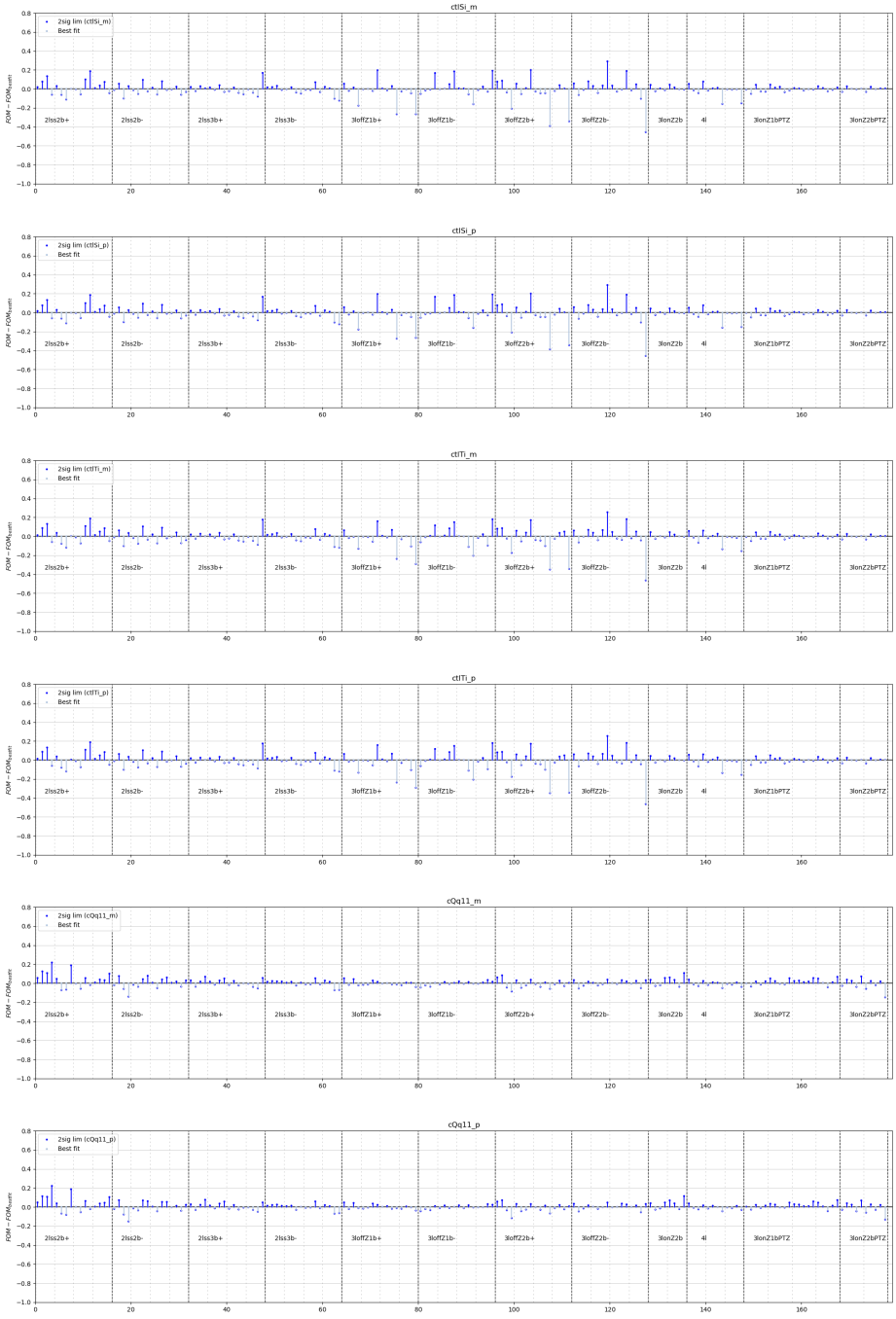

cQt1_p
2sig lim (cQt1_p)
Best fit
$FOM - FOM_{nom}$
2lss2b+
2lss2b-
2lss3b+
2lss3b-
3loffZ1b+
3loffZ1b-
3loffZ2b+
3loffZ2b-
3lonZ2b
4l
3lonZ1bPTZ
3lonZ2bPTZ

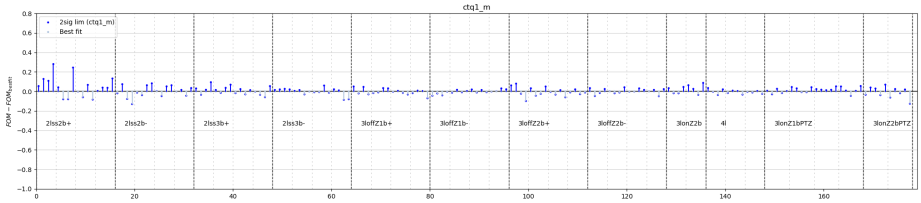

ctt1_m
2sig lim (ctt1_m)
Best fit
$FOM - FOM_{nom}$
2lss2b+
2lss2b-
2lss3b+
2lss3b-
3loffZ1b+
3loffZ1b-
3loffZ2b+
3loffZ2b-
3lonZ2b
4l
3lonZ1bPTZ
3lonZ2bPTZ

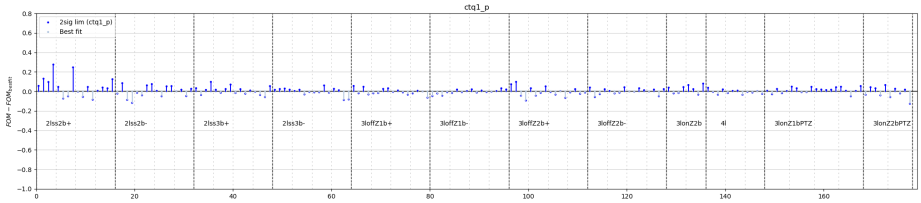

ctt1_p
2sig lim (ctt1_p)
Best fit
$FOM - FOM_{nom}$
2lss2b+
2lss2b-
2lss3b+
2lss3b-
3loffZ1b+
3loffZ1b-
3loffZ2b+
3loffZ2b-
3lonZ2b
4l
3lonZ1bPTZ
3lonZ2bPTZ

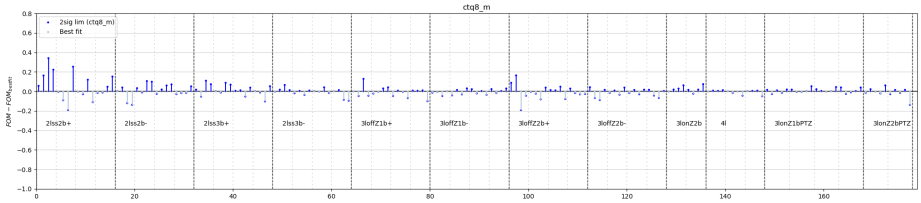

cQt8_m
2sig lim (cQt8_m)
Best fit
$FOM - FOM_{nom}$
2lss2b+
2lss2b-
2lss3b+
2lss3b-
3loffZ1b+
3loffZ1b-
3loffZ2b+
3loffZ2b-
3lonZ2b
4l
3lonZ1bPTZ
3lonZ2bPTZ
cQt8_p
2sig lim (cQt8_p)
Best fit
cQei_m
2sig lim (cQei_m)
Best fit
cQei_p
2sig lim (cQei_p)
Best fit
cQl3i_m
2sig lim (cQl3i_m)
Best fit
cQl3i_p
2sig lim (cQl3i_p)
Best fit

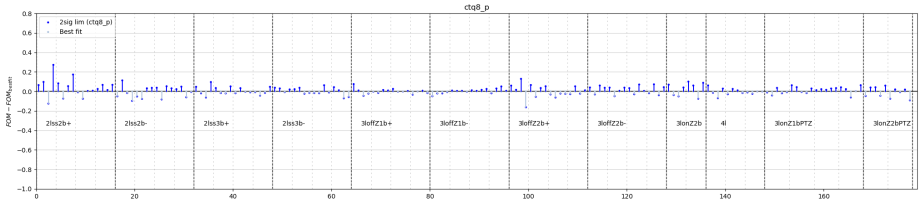

cQIMi_m
cQIMi_p
ctei_m
ctei_p
ctii_m
ctii_p
2sig lim (cQIMi_m)
2sig lim (cQIMi_p)
2sig lim (ctei_m)
2sig lim (ctei_p)
2sig lim (ctii_m)
2sig lim (ctii_p)
Best fit
FOM − FOM_worst
2lss2b+
2lss2b-
2lss3b+
2lss3b-
3loffZ1b+
3loffZ1b-
3loffZ2b+
3loffZ2b-
3lonZ2b
4l
3lonZ1bPTZ
3lonZ2bPTZ

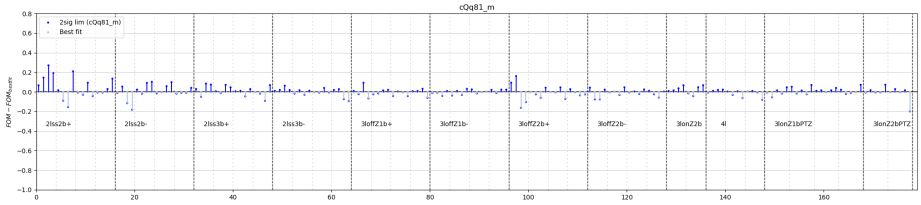

ctlSi_m
ctlSi_p
ctlTi_m
ctlTi_p
cQq11_m
cQq11_p
2sig lim
Best fit
2lss2b+
2lss2b-
2lss3b+
2lss3b-
3loffZ1b+
3loffZ1b-
3loffZ2b+
3loffZ2b-
3lonZ2b
4l
3lonZ1bPTZ
3lonZ2bPTZ

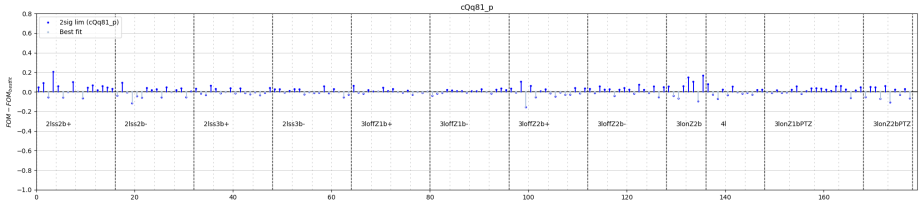

ctq1_m
2sig lim (ctq1_m)
Best fit
FOM − FOM_nominal
2lss2b+
2lss2b-
2lss3b+
2lss3b-
3loffZ1b+
3loffZ1b-
3loffZ2b+
3loffZ2b-
3lonZ2b
4l
3lonZ1bPTZ
3lonZ2bPTZ

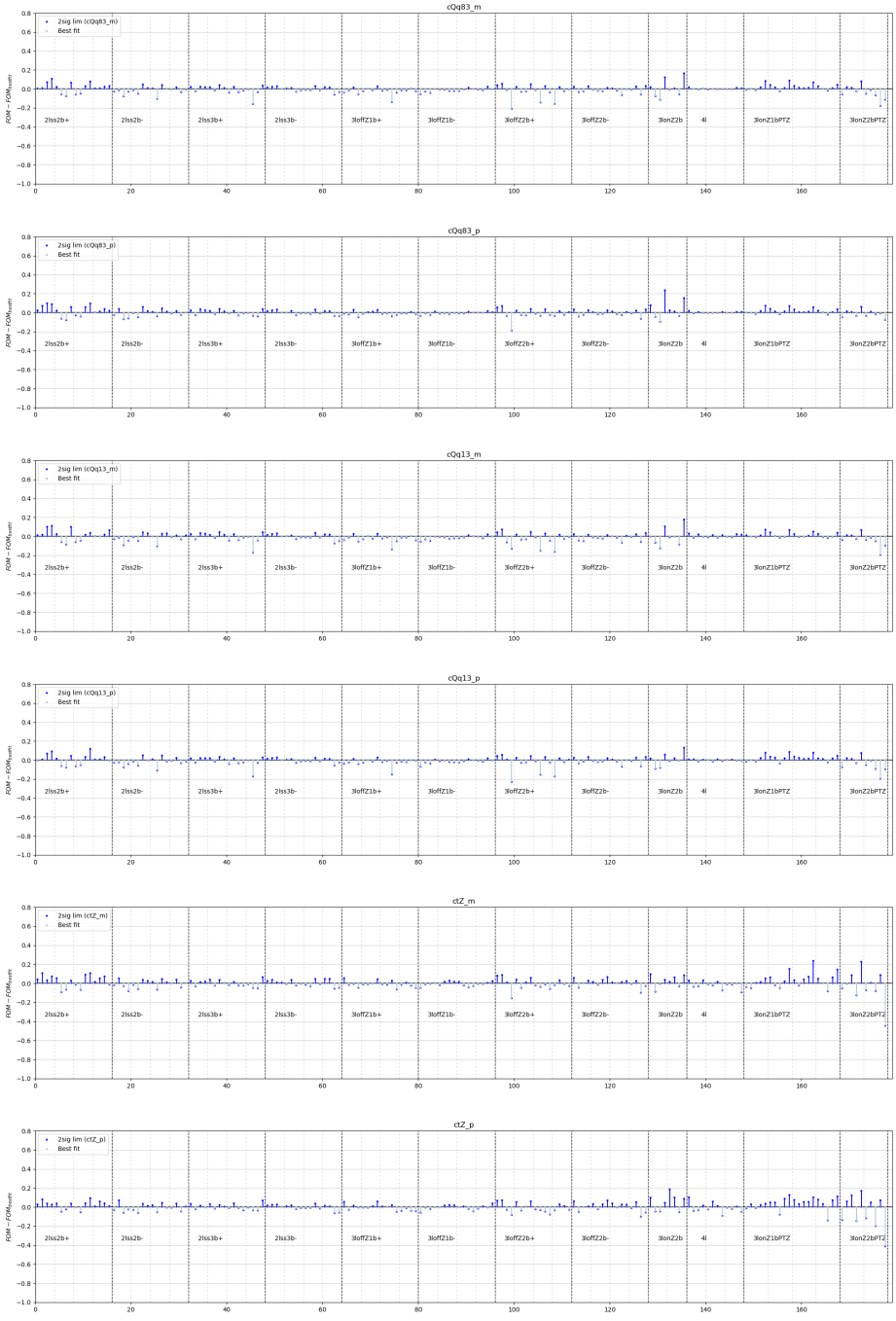

ctq1_p
2sig lim (ctq1_p)
Best fit
FOM − FOM_nominal
2lss2b+
2lss2b-
2lss3b+
2lss3b-
3loffZ1b+
3loffZ1b-
3loffZ2b+
3loffZ2b-
3lonZ2b
4l
3lonZ1bPTZ
3lonZ2bPTZ

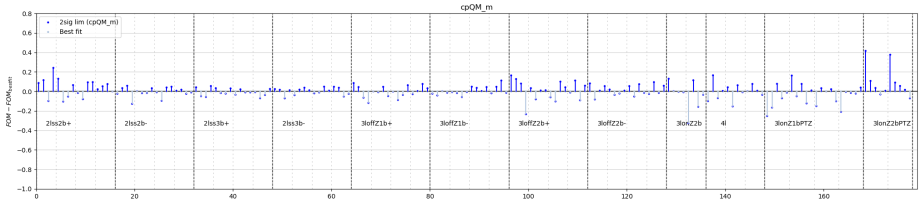

ctq8_m
2sig lim (ctq8_m)
Best fit
FOM − FOM_nominal
2lss2b+
2lss2b-
2lss3b+
2lss3b-
3loffZ1b+
3loffZ1b-
3loffZ2b+
3loffZ2b-
3lonZ2b
4l
3lonZ1bPTZ
3lonZ2bPTZ

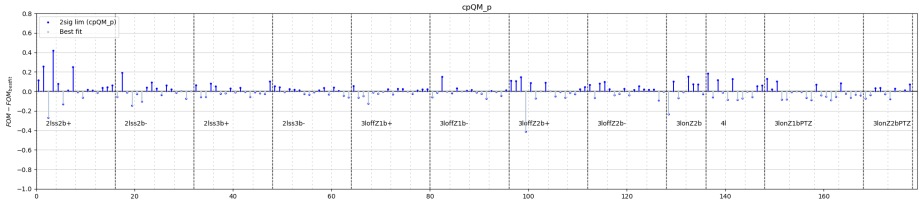

ctq8_p
2sig lim (ctq8_p)
Best fit
FOM − FOM_nominal
2lss2b+
2lss2b-
2lss3b+
2lss3b-
3loffZ1b+
3loffZ1b-
3loffZ2b+
3loffZ2b-
3lonZ2b
4l
3lonZ1bPTZ
3lonZ2bPTZ

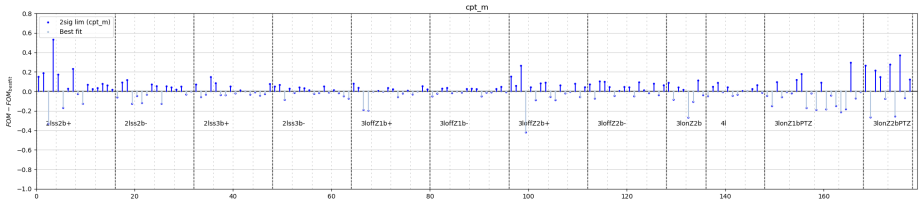

cQq81_m
2sig lim (cQq81_m)
Best fit
FOM − FOM_nominal
2lss2b+
2lss2b-
2lss3b+
2lss3b-
3loffZ1b+
3loffZ1b-
3loffZ2b+
3loffZ2b-
3lonZ2b
4l
3lonZ1bPTZ
3lonZ2bPTZ

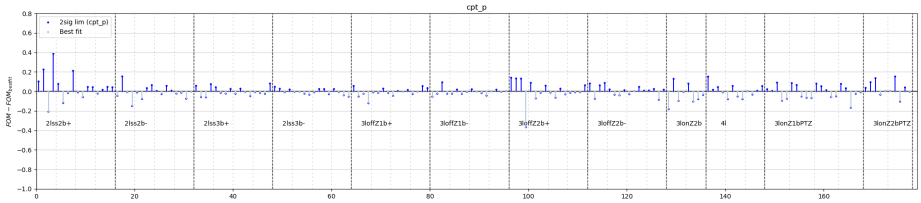

cQq81_p
2sig lim (cQq81_p)
Best fit
FOM − FOM_nominal
2lss2b+
2lss2b-
2lss3b+
2lss3b-
3loffZ1b+
3loffZ1b-
3loffZ2b+
3loffZ2b-
3lonZ2b
4l
3lonZ1bPTZ
3lonZ2bPTZ

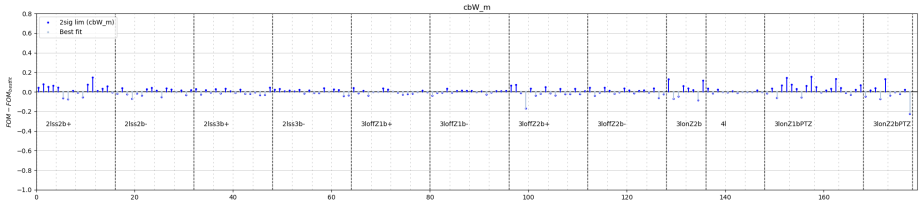

cQq83_m
cQq83_p
cQq13_m
cQq13_p
ctZ_m
ctZ_p
2sig lim
Best fit
FOM − FOM_nominal
2lss2b+
2lss2b-
2lss3b+
2lss3b-
3loffZ1b+
3loffZ1b-
3loffZ2b+
3loffZ2b-
3lonZ2b
4l
3lonZ1bPTZ
3lonZ2bPTZ

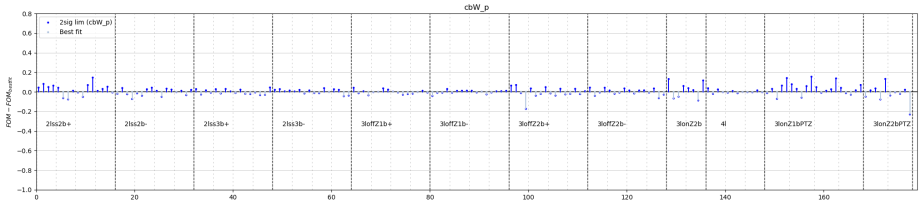

cpQM_m
2sig lim (cpQM_m)
Best fit
$FOM - FOM_{nom}$
2lss2b+
2lss2b-
2lss3b+
2lss3b-
3loffZ1b+
3loffZ1b-
3loffZ2b+
3loffZ2b-
3lonZ2b
4l
3lonZ1bPTZ
3lonZ2bPTZ

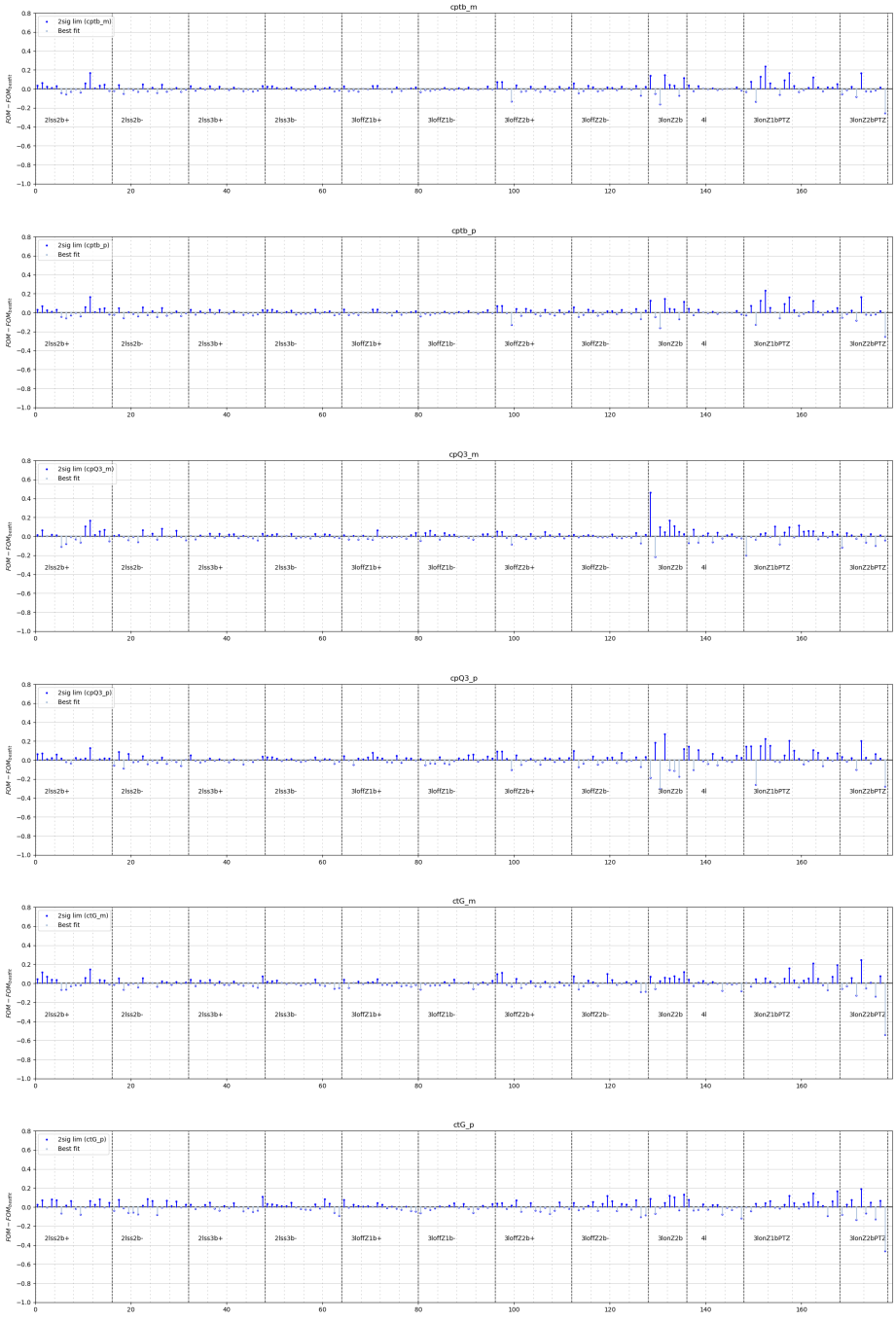

cpQM_p
2sig lim (cpQM_p)
Best fit
$FOM - FOM_{nom}$
2lss2b+
2lss2b-
2lss3b+
2lss3b-
3loffZ1b+
3loffZ1b-
3loffZ2b+
3loffZ2b-
3lonZ2b
4l
3lonZ1bPTZ
3lonZ2bPTZ


cpt_m
2sig lim (cpt_m)
Best fit
$FOM - FOM_{nom}$
2lss2b+
2lss2b-
2lss3b+
2lss3b-
3loffZ1b+
3loffZ1b-
3loffZ2b+
3loffZ2b-
3lonZ2b
4l
3lonZ1bPTZ
3lonZ2bPTZ


cpt_p
2sig lim (cpt_p)
Best fit
$FOM - FOM_{nom}$
2lss2b+
2lss2b-
2lss3b+
2lss3b-
3loffZ1b+
3loffZ1b-
3loffZ2b+
3loffZ2b-
3lonZ2b
4l
3lonZ1bPTZ
3lonZ2bPTZ


cbW_m
2sig lim (cbW_m)
Best fit
$FOM - FOM_{nom}$
2lss2b+
2lss2b-
2lss3b+
2lss3b-
3loffZ1b+
3loffZ1b-
3loffZ2b+
3loffZ2b-
3lonZ2b
4l
3lonZ1bPTZ
3lonZ2bPTZ


cbW_p
2sig lim (cbW_p)
Best fit
$FOM - FOM_{nom}$
2lss2b+
2lss2b-
2lss3b+
2lss3b-
3loffZ1b+
3loffZ1b-
3loffZ2b+
3loffZ2b-
3lonZ2b
4l
3lonZ1bPTZ
3lonZ2bPTZ


cptb_m
cptb_p
cpQ3_m
cpQ3_p
ctG_m
ctG_p
2sig lim (cptb_m)
2sig lim (cptb_p)
2sig lim (cpQ3_m)
2sig lim (cpQ3_p)
2sig lim (ctG_m)
2sig lim (ctG_p)
Best fit
FOM − FOM_multi
2lss2b+
2lss2b-
2lss3b+
2lss3b-
3loffZ1b+
3loffZ1b-
3loffZ2b+
3loffZ2b-
3lonZ2b
4l
3lonZ1bPTZ
3lonZ2bPTZ

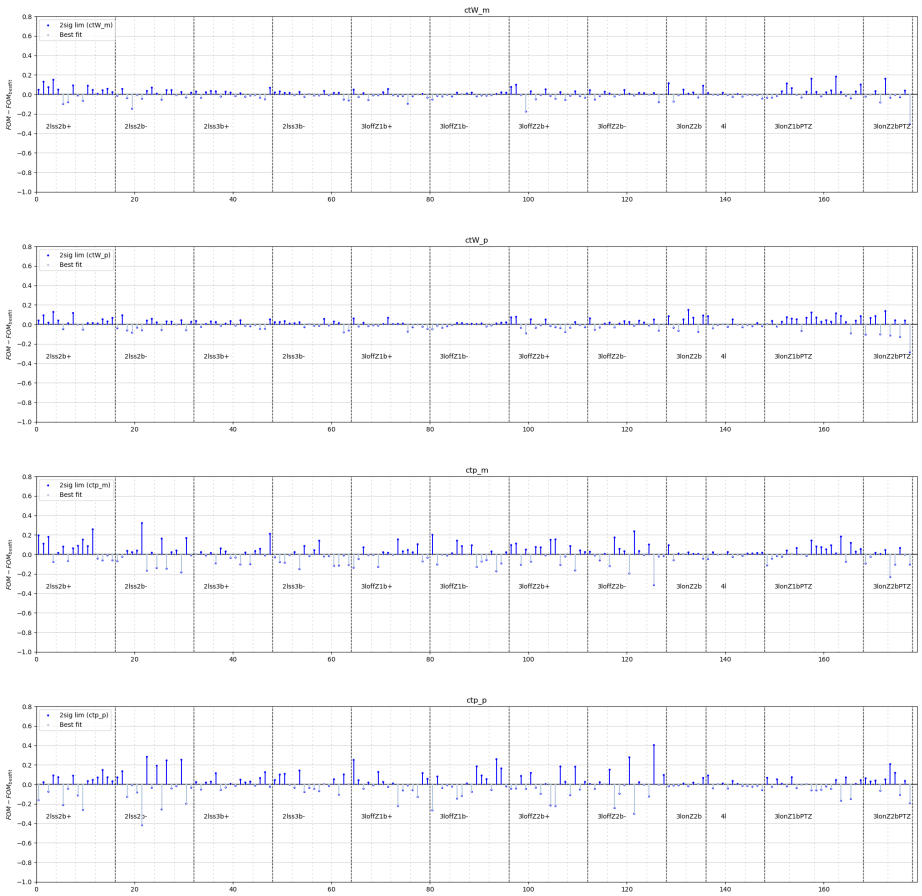


Figure H.3. The full set of 52 FOM plots (relative to the best fit point) for all 26 WCs at the $+2\sigma$ and -2σ limits.

APPENDIX I

IMPACT PLOTS FOR SELECTED WCS

This appendix includes the impact plots for the WCs where the systematic uncertainties are larger than the statistical uncertainties (as identified in Section 12.4.1). The impact for each systematic is defined as the change in the best fit value for the given WC when then given systematic is set to its plus or minus one sigma value. The fits are performed with Asimov data. The top 30 systematics (as ordered by impact) are shown in each plot. The full set of impact plots for all WCs (for unblind and Asimov fits) can be found in Ref. [46].

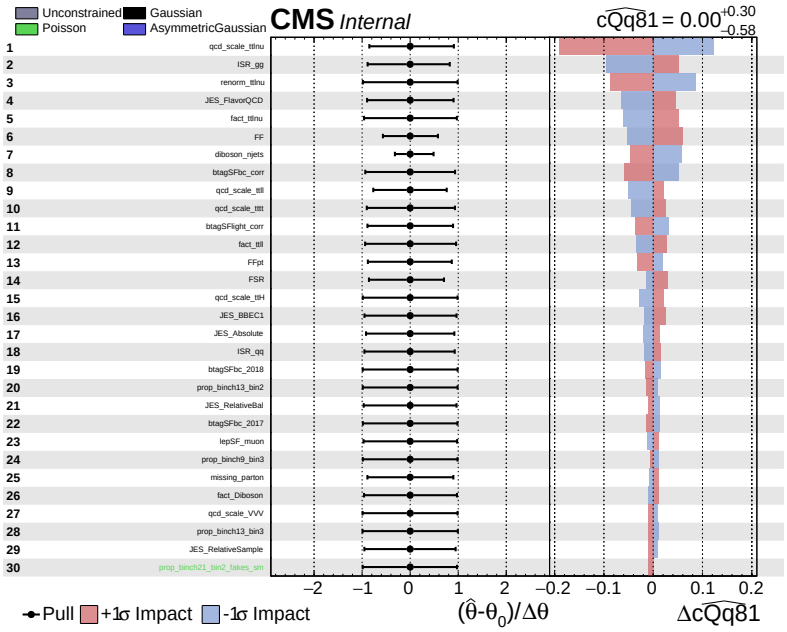


Figure I.1. Impact plot for c_{Qq}^{18} with Asimov data. The rest of the WCs are frozen to their SM values of 0. [46]

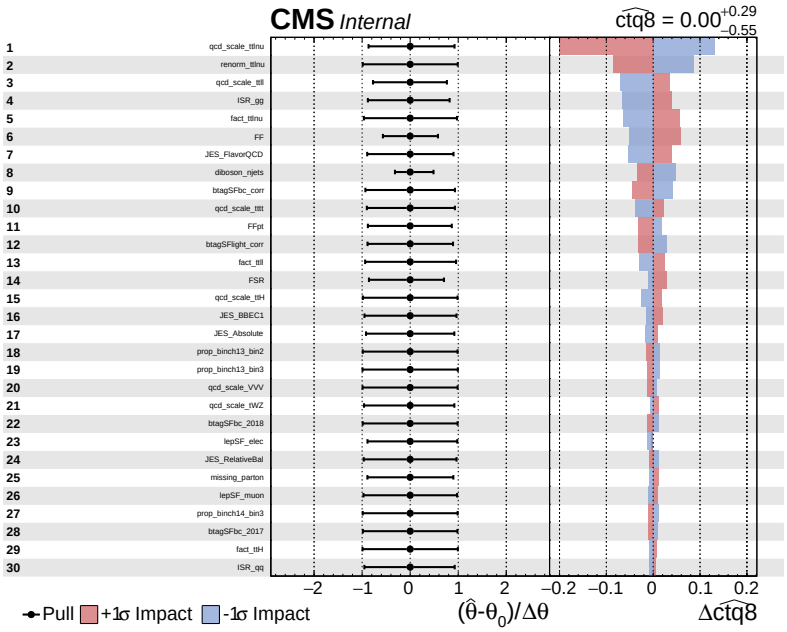


Figure I.2. Impact plot for c_{tq}^{8} with Asimov data. The rest of the WCs are frozen to their SM values of 0. [46]

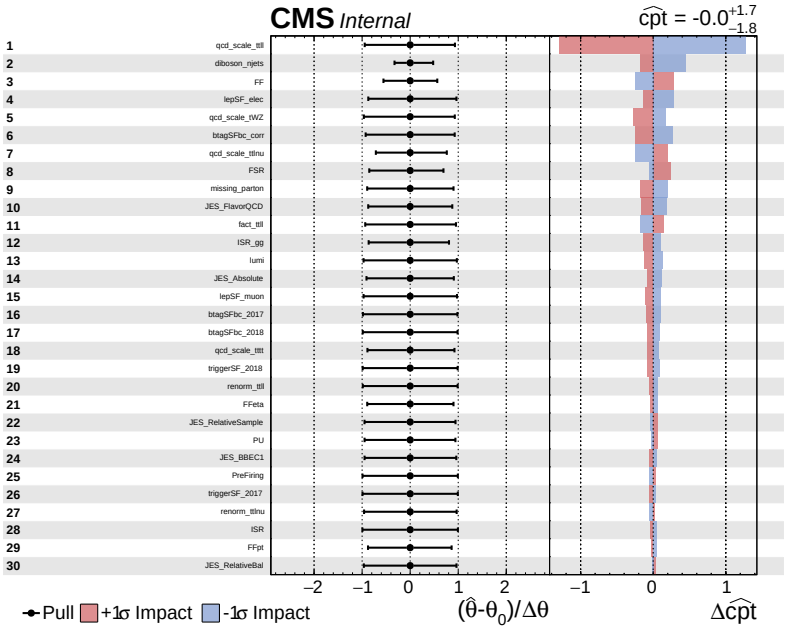


Figure I.3. Impact plot for $c_{\varphi t}$ with Asimov data. The rest of the WCs are frozen to their SM values of 0. [46]

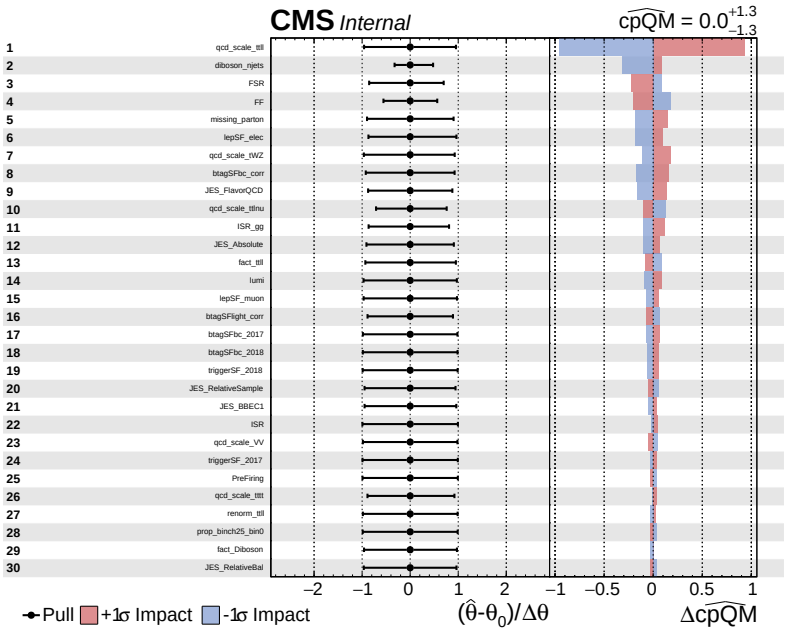


Figure I.4. Impact plot for $c^-_{\varphi Q}$ with Asimov data. The rest of the WCs are frozen to their SM values of 0. [46]

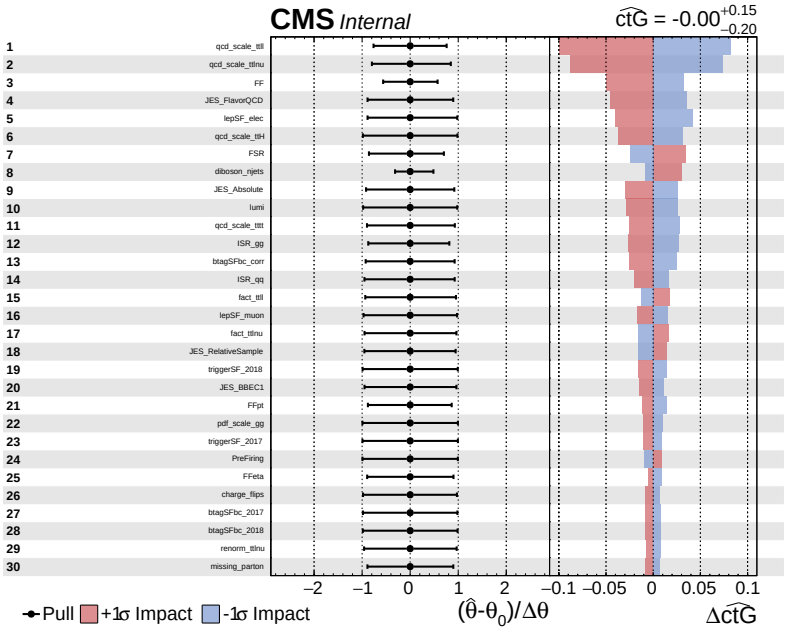


Figure I.5. Impact plot for c_{tG} with Asimov data. The rest of the WCs are frozen to their SM values of 0. [46]

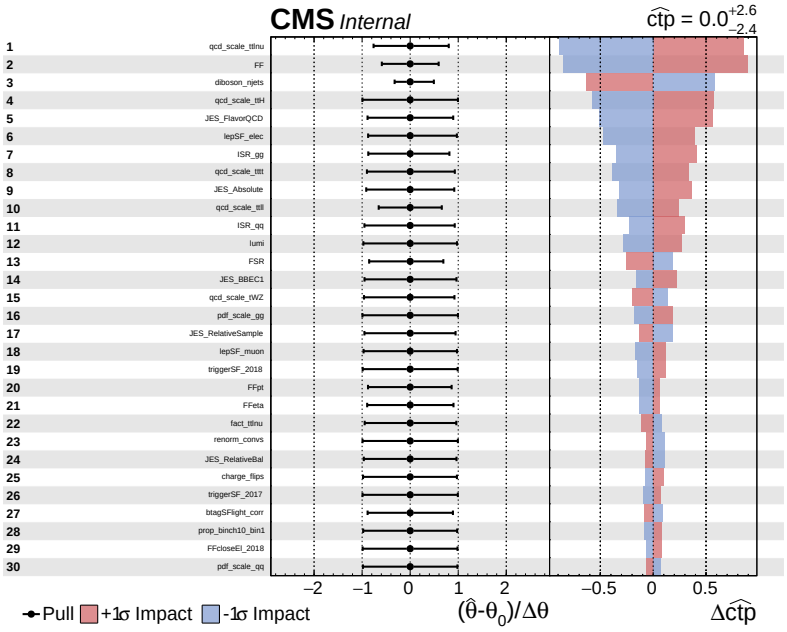


Figure I.6. Impact plot for $c_{t\varphi}$ with Asimov data. The rest of the WCs are frozen to their SM values of 0. [46]

APPENDIX J

LIMIT COMPARISONS TO OTHER ANALYSES

This appendix will compare the results of the analysis presented in this thesis (referred to as TOP-22-006) against two other CMS analysis. Section J.1 will provide a comparison to TOP-19-001 [4], the direct predecessor to TOP-22-006. Section J.2 will provide a comparison to TOP-21-003 [76].

J.1 TOP-19-001 comparison

First, we will compare TOP-22-006 against the results obtained by the predecessor to this analysis (Ref. [4]), which used 41.5 fb^{-1} of data collected by CMS in 2017. Figure J.1 shows this comparison of the TOP-22-006 results to the TOP-19-001 results for fits in which the other WCs are profiled and fits in which the other WCs are fixed to zero.

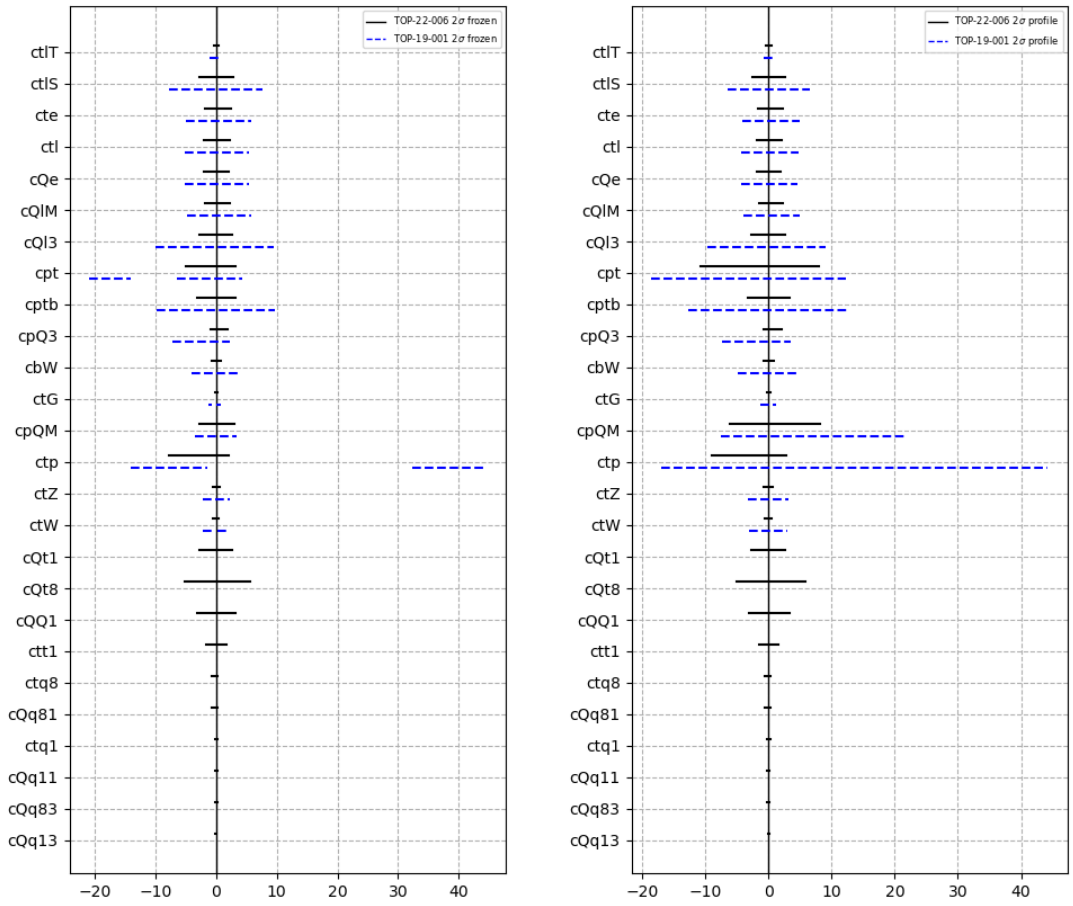
For all WCs, the limits presented in TOP-22-006 are consistently tighter than the limits obtained by Ref. [4], generally by more than a factor of two. In the case where the other WCs are fixed to zero (Figure J.1 (a)), the improvement is between approximately 20 to 80%, depending on the WC. It should be noted that Ref. [4] only studied a subset of the WCs included in this analysis, so the the profiled fits are not directly comparable. However, despite of the fact that TOP-22-006 includes 10 more free parameters than were included in Ref. [4], the profiled limits obtained by TOP-22-006 are consistently tighter. As shown in Figure J.1 (b), the improvement is between approximately 40 to 80%, depending on the WC.

The improved sensitivity is due multiple factors, including an increase in statistics (TOP-22-006 makes use of more than three times more data) and improved analysis technique (especially the use of kinematic distributions in the fit). The improvement varies by WC, but overall the statistics account for an approximately 25% improvement, while the differential binning accounts for improvements of approximately 10-50%.

J.2 TOP-21-003 comparison

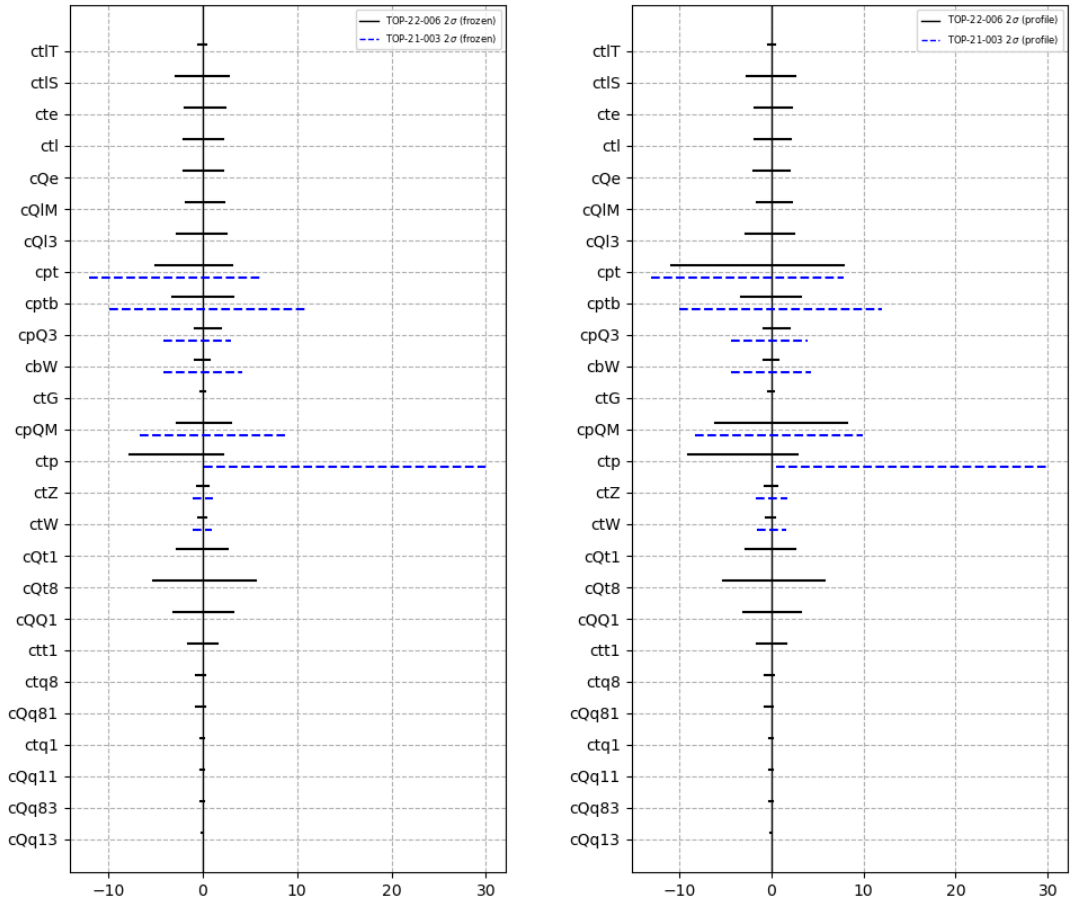
In this section, the results of TOP-22-006 will be compared against an analysis that studied top pairs produced in association with a boosted Z or Higgs boson, referred to as TOP-21-003 [76]. This boosted analysis used 138 fb^{-1} collected by CMS during 2016-2018 and studied eight dimension-six WCs (i.e. all WCs from the two-heavy-with-bosons category, except for c_{tG}). The EFT approach used in TOP-21-003 is similar to the approach used in TOP-22-006 (i.e. the prediction is parametrized in terms of WCs in order to study the EFT effects directly at detector level). The limit comparisons are presented in Figure J.2.

In most cases, the limits obtained by TOP-22-006 are tighter than the limits obtained by Ref. [76] by about a factor of two to a factor of five. However, there are two exceptions; the limits on $c_{\varphi t}$ and $c_{\varphi Q}^{-}$ in the profiled case are approximately equal in both analyses. In TOP-22-006, these WCs have significant correlations in the likelihood fit (as discussed in Section 12.2), potentially contributing to their relatively wide confidence intervals. It should also be noted that in the TOP-22-006 limits shown in Figure J.2 (b), all 26 WCs are profiled, so the limits are not directly comparable to the profiled results of TOP-21-003 (which only studies 8 WCs). However, as most of the correlations relevant to these 8 WCs are among each other (i.e. among other two-heavy-with-boson WCs) the TOP-22-006 profiled limits do not change significantly when profiling only these 8 WCs instead of all 26 WCs.



(a) (b)

Figure J.1. The 2σ profiled limits of this analysis (indicated with the solid black lines labeled TOP-22-001) compared against the limits obtained by the direct predecessor of this analysis (TOP-19-001 [4], indicated with the dashed blue lines). Figure (a) shows the results of the likelihood fits in which a single WC is fit with all other WCs fixed to their SM values of zero. Figure (b) shows the results of the likelihood fits in which the other WCs are profiled. In these plots, $\Lambda = 1$ TeV.



(a) (b)

Figure J.2. The 2σ profiled limits of this analysis (indicated with the solid black lines labeled TOP-22-001) compared against the limits obtained by Ref. [76] (TOP-21-003), indicated with the dashed blue lines). Figure (a) shows the results of the likelihood fits in which a single WC is fit with all other WCs fixed to their SM values of zero. Figure (b) shows the results of the likelihood fits in which the other WCs are profiled. In these plots, $\Lambda = 1$ TeV.

BIBLIOGRAPHY

1. J. L. Feng. Dark matter candidates from particle physics and methods of detection. *Ann. Rev. Astron. Astrophys.*, 48:495, 2010, doi:10.1146/annurev-astro-082708-101659, arXiv:1003.0904.

2. T. A. Porter, R. P. Johnson, and P. W. Graham. Dark matter searches with astroparticle data. *Ann. Rev. Astron. Astrophys.*, 49:155, 2011, doi:10.1146/annurev-astro-081710-102528, arXiv:1104.2836.

3. P. J. E. Peebles and B. Ratra. The cosmological constant and dark energy. *Rev. Mod. Phys.*, 75:559, 2003, doi:10.1103/RevModPhys.75.559, arXiv:astro-ph/0207347.

4. 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*, 03:095, 2021, doi:10.1007/JHEP03(2021)095, arXiv:2012.04120.

5. C. Degrande, N. Greiner, W. Kilian, O. Mattelaer, H. Mebane, T. Stelzer, S. Willenbrock, and C. Zhang. Effective Field Theory: A Modern Approach to Anomalous Couplings. *Annals Phys.*, 335:21–32, 2013, doi:10.1016/j.aop.2013.04.016, arXiv:1205.4231.

6. B. Grzadkowski, M. Iskrzynski, M. Misiak, and J. Rosiek. Dimension-Six Terms in the Standard Model Lagrangian. *JHEP*, 10:085, 2010, doi:10.1007/JHEP10(2010)085, arXiv:1008.4884.

7. 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*, 04:159, 2014, doi:10.1007/JHEP04(2014)159, arXiv:1312.2014.

8. D. Barducci et al. Interpreting top-quark LHC measurements in the standard-model effective field theory. 2 2018, arXiv:1802.07237.

9. J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, O. Mattelaer, H. S. Shao, T. Stelzer, P. Torrielli, and M. Zaro. The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations. *JHEP*, 07:079, 2014, doi:10.1007/JHEP07(2014)079, arXiv:1405.0301.

10. R. Goldouzian, J. H. Kim, K. Lannon, A. Martin, K. Mohrman, and A. Wightman. Matching in $pp \to t\bar{t}W/Z/h+$ jet SMEFT studies. *JHEP*, 06:151, 2021, doi:10.1007/JHEP06(2021)151, arXiv:2012.06872.

11. LHC Machine. *JINST*, 3:S08001, 2008, doi:10.1088/1748-0221/3/08/S08001.

12. S. Chatrchyan et al. The CMS Experiment at the CERN LHC. *JINST*, 3:S08004, 2008, doi:10.1088/1748-0221/3/08/S08004.

13. Radiofrequency cavities. 2012. URL https://cds.cern.ch/record/1997424.

14. B. Salvachua. Overview of Proton-Proton Physics during Run 2. In *9th LHC Operations Evian Workshop*, pages 7–14, 2019.

15. CMS. Public cms luminosity information, 2022. URL https://twiki.cern.ch/twiki/bin/view/CMSPublic/LumiPublicResults.

16. I. Neutelings. Cms coordinate system, 2021. URL https://tikz.net/axis3d_cms/. Licensed under a Creative Commons Attribution-ShareAlike 4.0 International License with © Copyright 2021 – TikZ.net. https://creativecommons.org/licenses/by-sa/4.0/.

17. A. M. Sirunyan et al. Particle-flow reconstruction and global event description with the CMS detector. *JINST*, 12(10):P10003, 2017, doi:10.1088/1748-0221/12/10/P10003, arXiv:1706.04965.

18. V. Veszpremi. Operation and performance of the CMS tracker. *JINST*, 9:C03005, 2014, doi:10.1088/1748-0221/9/03/C03005, arXiv:1402.0675.

19. S. Chatrchyan et al. Alignment of the CMS tracker with LHC and cosmic ray data. *JINST*, 9:P06009, 2014, doi:10.1088/1748-0221/9/06/P06009, arXiv:1403.2286.

20. C. Biino. The CMS Electromagnetic Calorimeter: overview, lessons learned during Run 1 and future projections. *J. Phys. Conf. Ser.*, 587(1):012001, 2015, doi:10.1088/1742-6596/587/1/012001.

21. A. M. Sirunyan et al. Calibration of the CMS hadron calorimeters using proton-proton collision data at $\sqrt{s} = 13$ TeV. *JINST*, 15(05):P05002, 2020, doi:10.1088/1748-0221/15/05/P05002, arXiv:1910.00079.

22. A. M. Sirunyan et al. Performance of the CMS muon detector and muon reconstruction with proton-proton collisions at $\sqrt{s} = 13$ TeV. *JINST*, 13(06):P06015, 2018, doi:10.1088/1748-0221/13/06/P06015, arXiv:1804.04528.

23. A. M. Sirunyan et al. Performance of the CMS Level-1 trigger in proton-proton collisions at $\sqrt{s} = 13$ TeV. *JINST*, 15(10):P10017, 2020, doi:10.1088/1748-0221/15/10/P10017, arXiv:2006.10165.

24. R. Frazier, S. Fayer, G. Hall, C. Hunt, G. Iles, D. Newbold, and A. Rose. A demonstration of a time multiplexed trigger for the CMS experiment. *JINST*, 7: C01060, 2012, doi:10.1088/1748-0221/7/01/C01060.

25. A. Triossi. The CMS Barrel Muon Trigger Upgrade. Technical report, CERN, Geneva, 2017. URL https://cds.cern.ch/record/2252301.

26. CMS SW Guide. The cms nanoaod data tier, 2022. URL https://twiki.cern.ch/twiki/bin/view/CMSPublic/WorkBookNanoAOD.

27. CMS MCM. Info for mc production for ultra legacy campaigns 2016, 2017, 2018, 2022. URL https://cms-pdmv.gitbook.io/project/mccontact/info-for-mc-production-for-ultra-legacy-campaigns-2016-2017-2018.

28. CMS Lumi POG. Luminosity physics object group, 2022. URL https://twiki.cern.ch/twiki/bin/viewauth/CMS/TWikiLUM.

29. T. Sjostrand, S. Mrenna, and P. Z. Skands. A brief introduction to PYTHIA 8.1. *Comput. Phys. Commun.*, 178:852, 2008, doi:10.1016/j.cpc.2008.01.036, arXiv:0710.3820.

30. J. Alwall et al. Comparative study of various algorithms for the merging of parton showers and matrix elements in hadronic collisions. *Eur. Phys. J.*, C53:473–500, 2008, doi:10.1140/epjc/s10052-007-0490-5, arXiv:0706.2569.

31. O. Mattelaer. On the maximal use of Monte Carlo samples: re-weighting events at NLO accuracy. *Eur. Phys. J. C*, 76(12):674, 2016, doi:10.1140/epjc/s10052-016-4533-7, arXiv:1607.00763.

32. M. Cacciari, G. P. Salam, and G. Soyez. The anti-k_t jet clustering algorithm. *JHEP*, 04:063, 2008, doi:10.1088/1126-6708/2008/04/063, arXiv:0802.1189.

33. M. Cacciari, G. P. Salam. Dispelling the N^3 myth for the k_t jet-finder. *Phys. Lett. B*, 641:57, 2006, doi:10.1016/j.physletb.2006.08.037, arXiv:hep-ph/0512210.

34. A. M. Sirunyan et al. Pileup mitigation at CMS in 13 TeV data. *JINST*, 15(09): P09018, 2020, doi:10.1088/1748-0221/15/09/P09018, arXiv:2003.00503.

35. A. M. Sirunyan et al. Identification of heavy-flavour jets with the CMS detector in pp collisions at 13 TeV. *JINST*, 13(05):P05011, 2018, doi:10.1088/1748-0221/13/05/P05011, arXiv:1712.07158.

36. CMS BTV POG. Heavy flavour tagging for 13 tev data in 2016prevfp ultra-legacy reprocessing and 10_6_x mc, 2022. URL https://twiki.cern.ch/twiki/bin/view/CMS/BtagRecommendation106XUL16preVFP.

37. CMS BTV POG. Heavy flavour tagging for 13 tev data in 2016postvfp ultra-legacy reprocessing and 10_6_x mc, 2022. URL https://twiki.cern.ch/twiki/bin/view/CMS/BtagRecommendation106XUL16postVFP.

38. CMS BTV POG. Heavy flavour tagging for 13 tev data in 2017 ultra-legacy reprocessing and 10_6_x mc, 2022. URL https://twiki.cern.ch/twiki/bin/view/CMS/BtagRecommendation106XUL17.

39. CMS BTV POG. Heavy flavour tagging for 13 tev data in 2018 ultra-legacy reprocessing and 10_6_x mc, 2022. URL https://twiki.cern.ch/twiki/bin/view/CMS/BtagRecommendation106XUL18.

40. ttH multilepton group. Differential ttH production in the multilepton final state. *CMS Analysis Note*, 2020/040, 2020.

41. Measurement of the ratio B(t to Wb)/B(t to Wq). Technical report, CERN, Geneva, 2013. URL https://cds.cern.ch/record/1520879.

42. V. Khachatryan et al. Performance of electron reconstruction and selection with the CMS detector in pp collisions at $\sqrt{s} = 8$ TeV. *JINST*, 10:P06005, 2015, doi:10.1088/1748-0221/10/06/P06005, arXiv:1502.02701.

43. CMS EGamma POG. Multivariate electron identification for run2, 2022. URL https://twiki.cern.ch/twiki/bin/view/CMS/MultivariateElectronIdentificationRun2#Recommended_MVA_Recipe_for_regul.

44. A. Hocker et al. TMVA - toolkit for multivariate data analysis. *PoS*, ACAT:040, 2007, arXiv:physics/0703039.

45. CMS Muon POG. Baseline muon selections for run-ii, 2022. URL https://twiki.cern.ch/twiki/bin/view/CMS/SWGuideMuonIdRun2#Medium_Muon.

46. Basnet, A and others. Search for new physics in top quark production with additional leptons in proton-proton collisions at center-of-mass energy of 13TeV using effective field theory in Run II. *CMS Analysis Note*, 2021/051, 2022.

47. ttH multilepton group. Measurement of the Higgs boson production in association with top quarks in final states with electrons, muons and hadronic taus using the full Run 2 dataset. *CMS Analysis Note*, 2019/111, 2019.

48. CMS SW Guide. Met analysis, 2019. URL https://twiki.cern.ch/twiki/bin/view/CMSPublic/WorkBookMetAnalysis.

49. A. M. Sirunyan et al. Electron and photon reconstruction and identification with the CMS experiment at the CERN LHC. *JINST*, 16(05):P05014, 2021, doi:10.1088/1748-0221/16/05/P05014, arXiv:2012.06888.

50. CMS Physics Validation Group. Utilities for accessing pileup information for data, 2022. URL https://twiki.cern.ch/twiki/bin/viewauth/CMS/PileupJSONFileforData.

51. CMS. nanoAOD-tools, 2021. URL https://github.com/cms-nanoAOD/nanoAOD-tools/tree/master/python/postprocessing/data/pileup.

52. CMS EGamma POG. Egamma run ii recommendations, 2020. URL https://twiki.cern.ch/twiki/bin/view/CMS/EgammaRunIIRecommendations.

53. CMS Muon POG. Muon physics object group, 2022. URL https://twiki.cern.ch/twiki/bin/view/CMS/MuonPOG.

54. CMS. Tag and probe, 2014. URL https://twiki.cern.ch/twiki/bin/view/CMSPublic/TagAndProbe.

55. CMS BTV POG. Methods to apply b-tagging efficiency scale factors, 2022. URL https://twiki.cern.ch/twiki/bin/view/CMS/BTagSFMethods.

56. An, S. and others. Particle-level differential cross section measurement of ttbb production in the lepton+jets channel. *CMS Analysis Note*, 2021/040, 2018.

57. CMS L1 Trigger Group. Reweighting recipe to emulate level 1 ecal and muon prefiring, 2021. URL https://twiki.cern.ch/twiki/bin/viewauth/CMS/L1PrefiringWeightRecipe.

58. V. Khachatryan et al. Jet energy scale and resolution in the CMS experiment in pp collisions at 8 TeV. *JINST*, 12(02):P02014, 2017, doi:10.1088/1748-0221/12/02/P02014, arXiv:1607.03663.

59. CMS JERC Group. Jet energy resolution, 2022. URL https://twiki.cern.ch/twiki/bin/viewauth/CMS/JetResolution.

60. A. Khukhunaishvili and A. Bodek. Rochester corrections for muon momentum for legacy run2 samples, 2020. URL https://indico.cern.ch/event/981770/contributions/4135530/attachments/2157765/3639739/roccor.pdf.

61. CMS MUON POG. Rochester corrections, 2021. URL https://twiki.cern.ch/twiki/bin/viewauth/CMS/RochcorMuon.

62. CMS. Search for new physics in top quark production with additional leptons in CMS Run II proton-proton collisions at $\sqrt{s} = 13$ TeV in the context of effective field theory. *In review by CMS collaboration in preparation for submission to JHEP*, 2023, arXiv:XXXX.XXXXX.

63. David, P and others. The cross-section for the production of a top quark pair in association with a W boson are measured in proton-proton collisions at a centre-of-mass energy of 13TeV, using the data sample recorded with the CMS detector during 2016, 2017, and 2018. *CMS Analysis Note*, 2019/127, 2019.

64. A. M. Sirunyan et al. Precision luminosity measurement in proton-proton collisions at $\sqrt{s} = 13$ TeV in 2015 and 2016 at CMS. *Eur. Phys. J. C*, 81:800, 2021, doi:10.1140/epjc/s10052-021-09538-2, arXiv:2104.01927.

65. CMS Collaboration. CMS luminosity measurement for the 2017 data-taking period at $\sqrt{s} = 13$ TeV. CMS Physics Analysis Summary CMS-PAS-LUM-17-004, 2018. URL https://cds.cern.ch/record/2621960/.

66. CMS Collaboration. CMS luminosity measurement for the 2018 data-taking period at $\sqrt{s} = 13$ TeV. CMS Physics Analysis Summary CMS-PAS-LUM-18-002, 2019. URL https://cds.cern.ch/record/2676164/.

67. CMS JERC Group. Jet energy scale uncertainty sources, 2022. URL https://twiki.cern.ch/twiki/bin/view/CMS/JECUncertaintySources#Run2_JEC_uncertainty_correlation.

68. CMS HiggsCombine. Automatic statistical uncertainties, 2021. URL https://cms-analysis.github.io/HiggsAnalysis-CombinedLimit/part2/bin-wise-stats/.

69. S. S. Wilks. The Large-Sample Distribution of the Likelihood Ratio for Testing Composite Hypotheses. *The Annals of Mathematical Statistics*, 9(1):60 – 62, 1938, doi:10.1214/aoms/1177732360. URL https://doi.org/10.1214/aoms/1177732360.

70. R. L. Workman and Others. Review of Particle Physics. *PTEP*, 2022:083C01, 2022, doi:10.1093/ptep/ptac097.

71. H.-C. project. "HiggsAnalysis-CombineLimit" [software], commit a1dc956452f5b8e055de43bfb093e5243a1e99d9. 2018. URL https://github.com/cms-analysis/HiggsAnalysis-CombinedLimit.

72. W. Verkerke and D. P. Kirkby. The RooFit toolkit for data modeling. *eConf*, C0303241:MOLT007, 2003, arXiv:physics/0306116.

73. F. James. MINUIT Function Minimization and Error Analysis: Reference Manual Version 94.1. 1994.

74. A. Massironi and G. Boldrini. EFT model for SMP measurements. *CMS Analysis Note*, 2020/204, 2020.

75. CMS. Top quark physics summary figures summary of eft measurements with top quark events, 2021. URL https://twiki.cern.ch/twiki/bin/view/CMSPublic/PhysicsResultsTOPSummaryFigures#Summary_of_EFT_measurements_with.

76. Search for new physics using effective field theory in 13 TeV pp collision events that contain a top quark pair and a boosted Z or Higgs boson. 8 2022, arXiv:2208.12837.

77. 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*, 03:056, 2020, doi:10.1007/JHEP03(2020)056, arXiv:1907.11270.

78. K. Lee et al. Probing effective field theory operators in the associated production of top quarks with a Z boson in multilepton final states at $\sqrt{s} = 13$ TeV. *JHEP*, 12:083, 2021, doi:10.1007/JHEP12(2021)083, arXiv:2107.13896.

79. A. Tumasyan et al. Measurement of the inclusive and differential $\mathrm{t\bar{t}}\gamma$ cross sections in the single-lepton channel and EFT interpretation at $\sqrt{s} = 13$ TeV. *JHEP*, 12: 180, 2021, doi:10.1007/JHEP12(2021)180, arXiv:2107.01508.

80. A. M. Sirunyan et al. Measurement of the top quark polarization and $\mathrm{t\bar{t}}$ spin correlations using dilepton final states in proton-proton collisions at $\sqrt{s} = 13$ TeV. *Phys. Rev. D*, 100(7):072002, 2019, doi:10.1103/PhysRevD.100.072002, arXiv:1907.03729.

81. A. M. Sirunyan et al. Search for the production of four top quarks in the single-lepton and opposite-sign dilepton final states in proton-proton collisions at $\sqrt{s} = 13$ TeV. *JHEP*, 11:082, 2019, doi:10.1007/JHEP11(2019)082, arXiv:1906.02805.

82. J. J. Ethier, G. Magni, F. Maltoni, L. Mantani, E. R. Nocera, J. Rojo, E. Slade, E. Vryonidou, and C. Zhang. Combined SMEFT interpretation of Higgs, diboson, and top quark data from the LHC. *JHEP*, 11:089, 2021, doi:10.1007/JHEP11(2021)089, arXiv:2105.00006.

83. 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*, 04:100, 2019, doi:10.1007/JHEP04(2019)100, arXiv:1901.05965.

84. J. Alwall, S. de Visscher, and F. Maltoni. QCD radiation in the production of heavy colored particles at the LHC. *JHEP*, 02:017, 2009, doi:10.1088/1126-6708/2009/02/017, arXiv:0810.5350.

85. P. Lenzi and J. M. Butterworth. A Study on Matrix Element corrections in inclusive Z/ gamma* production at LHC as implemented in PYTHIA, HERWIG, ALPGEN and SHERPA. 3 2009, arXiv:0903.3918.

86. K. Mohrman. Mc validation checks, 2021. URL https://indico.cern.ch/event/1023102/contributions/4313340/attachments/2223243/3765493/MC_starting_point_validation_and_matching_checks_April09_2021.pdf.

87. K. Mohrman. Discussion of the eft dependence of the systematics, 2022. URL https://indico.cern.ch/event/1123567/contributions/4787852/attachments/2410630/4124888/systematics_eft_dependence_march18_2022.pdf.