

EFFECTIVE FIELD THEORY INTERPRETATION FOR MEASUREMENTS
OF TOP QUARK PAIR-PRODUCTION IN ASSOCIATION WITH A W OR Z BOSON

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

Anna Elizabeth Woodard

Kevin Lannon, Director

Graduate Program in Physics

Notre Dame, Indiana

April 2018



© Copyright by

Anna Elizabeth Woodard

2018

This work is licensed under a “CC BY-SA 4.0” license.



EFFECTIVE FIELD THEORY INTERPRETATION FOR MEASUREMENTS
OF TOP QUARK PAIR-PRODUCTION IN ASSOCIATION WITH A W OR Z BOSON

Abstract

by

Anna Elizabeth Woodard

Using effective field theory, the effects of new particles or interactions can be captured in a model-independent way by supplementing the standard model Lagrangian with higher-dimensional operators. A measurement of the cross section for top quarks produced in association with a W or Z boson, using 19.5 fb^{-1} of proton-proton collisions with a center-of-mass energy of 8 TeV, collected by the CMS experiment at the LHC, is extended within this framework to set constraints on the Wilson coefficients of five dimension-six operators. An additional measurement of the same processes using 35.9 fb^{-1} of proton-proton collisions with a center-of-mass energy of 13 TeV collected by the CMS experiment is used to perform a more sophisticated study of eight dimension-six operators and present bounds on their Wilson coefficients.

Dedicated to Gilbert,
despite a talk which began frantically after he sat on my keyboard and deleted my slides

CONTENTS

Figures	vi
Tables	ix
Acknowledgments	xi
Chapter 1: Introduction	1
Chapter 2: Theory	3
2.1 The Standard Model	3
2.1.1 The Lagrangian formulation and principle of stationary action	5
2.1.2 Symmetry and gauge invariance	7
2.1.3 The Standard Model Lagrangian	10
2.1.4 Quantum chromodynamics	12
2.1.4.1 Confinement and hadronization	13
2.1.5 Electroweak interactions	14
2.1.6 Spontaneous symmetry breaking	17
2.1.7 Perturbation theory	23
2.2 Physics beyond the Standard Model	24
2.2.1 Deficiencies in the Standard Model	25
2.2.2 New physics in the top sector	26
2.2.3 Effective field theory	29
Chapter 3: Experimental apparatus	33
3.1 The Large Hadron Collider	34
3.2 The CMS detector	40
3.2.1 Overview	40
3.2.2 Coordinate system and conventions	44
3.2.3 Solenoidal magnet	46
3.2.4 Inner tracker	47
3.2.4.1 Pixel detector	48
3.2.4.2 Silicon strip tracker	49
3.2.5 Calorimeters	50
3.2.5.1 Electromagnetic calorimeter	50
3.2.5.2 Hadronic calorimeter	52

3.2.6	Muon subsystem	52
3.2.7	Trigger and data acquisition	55
Chapter 4: Object reconstruction and selection		58
4.1	Particle flow	58
4.2	Object selection	63
4.2.1	Leptons	63
4.2.1.1	8 TeV	64
4.2.1.2	13 TeV	65
4.2.2	Jets and missing energy	67
Chapter 5: Observation of top quark pairs produced in association with a vector boson in proton–proton collisions at $\sqrt{s}=8$ TeV		69
5.1	Event selection	69
5.2	Event modeling	72
5.2.1	Prompt backgrounds and signal processes	73
5.2.1.1	Simulation corrections	73
5.2.2	Nonprompt backgrounds	75
5.2.3	Charge-misidentified backgrounds	77
5.3	Full event reconstruction	78
5.4	Signal extraction	80
5.5	Statistical procedure	83
5.6	Systematic uncertainties	92
5.7	Results	94
Chapter 6: Cross section measurement for top quark pair production in association with a W or Z boson in proton–proton collisions at $\sqrt{s}=13$ TeV		103
6.1	Samples	103
6.2	Event selection	104
6.3	Event modeling	107
6.3.1	Nonprompt backgrounds	107
6.3.2	Charge-misidentified backgrounds	108
6.3.3	WZ background	109
6.3.4	Rare backgrounds and $t(\bar{t})X$	109
6.4	Systematic uncertainties	110
6.5	Results	113
Chapter 7: Effective field theory interpretation for measurements of top quark pair- production in association with a W or Z boson		120
7.1	General strategy	120
7.2	8 TeV analysis	123
7.3	13 TeV analysis	129
7.3.1	Independent constraints	129
7.3.2	Preliminary simultaneous constraints	147

7.4	Future directions	163
7.4.1	Matrix element reweighting	165
7.4.2	Parameterized matrix element reweighting	166
Chapter 8: Conclusions		171
Appendix A: Reproducibility		173
Appendix B: Preliminary cross section scaling and likelihood scans for pairs of selected coefficients		176
Bibliography		201

FIGURES

2.1	Spontaneous symmetry breaking (classical example)	18
2.2	Potential of a complex scalar field	19
2.3	Representative Feynman diagrams for $t\bar{t}Z$ and $t\bar{t}W$	25
2.4	One-loop diagrams with top-Higgs or top-Z couplings	28
2.5	Example Feynman diagrams for Fermi theory	30
3.1	Overview of the LHC accelerator complex	34
3.2	Schematic layout of the LHC sectors	35
3.3	Luminosity and average pileup at CMS	41
3.4	Overview of the CMS detector	42
3.5	Transverse slice through the CMS detector	43
3.6	Longitudinal cross section of the CMS inner tracker	47
3.7	Layout of the CMS electromagnetic calorimeter	50
3.8	Cross section of one quadrant of the CMS muon system	53
5.1	Dijet mass, for matched versus all possible dijet combinations	82
5.2	Final discriminant for events in the $t\bar{t}W$ channel	100
5.3	Final discriminant for events in the $t\bar{t}Z$ channel	101
5.4	Profile likelihood for simultaneous fit as a function of $\sigma_{t\bar{t}Z}$ and $\sigma_{t\bar{t}W}$	102
6.1	Post-fit yields for each category in the SS $t\bar{t}W$ channel	115
6.2	Post-fit yields for each category in the 3ℓ $t\bar{t}Z$ channel	115
6.3	Simultaneous fit for the $t\bar{t}W$ and $t\bar{t}Z$ cross sections	119
7.1	Expected cross section scaling at 13 TeV due to $\bar{c}_{uG}$	122
7.2	Best fit values in the $\sigma_{t\bar{t}W}, \sigma_{t\bar{t}Z}$ plane (8 TeV)	125
7.3	Profile likelihood scans (8 TeV)	127
7.4	Profile likelihood, $\mu(c_1)$, and best fit c_1 for $\bar{c}_{uG}$ (13 TeV)	139
7.5	Profile likelihood, $\mu(c_1)$, and best fit c_1 for $\bar{c}_{uW}$ (13 TeV)	140
7.6	Profile likelihood, $\mu(c_1)$, and best fit c_1 for $\bar{c}_{3G}$ (13 TeV)	141

7.7	Profile likelihood, $\mu(c_1)$, and best fit c_1 for $\tilde{c}_{3G}$ (13 TeV)	142
7.8	Profile likelihood, $\mu(c_1)$, and best fit c_1 for $\tilde{c}_{2G}$ (13 TeV)	143
7.9	Profile likelihood, $\mu(c_1)$, and best fit c_1 for $\tilde{c}_{uB}$ (13 TeV)	144
7.10	Profile likelihood, $\mu(c_1)$, and best fit c_1 for $\tilde{c}_H$ (13 TeV)	145
7.11	Profile likelihood, $\mu(c_1)$, and best fit c_1 for $\tilde{c}_{Hu}$ (13 TeV)	146
7.12	Ratio of two and one-dimensional fit yields for $\tilde{c}_{uG}$ and $\tilde{c}_{uW}$ (13 TeV)	148
7.13	Ratio of two and one-dimensional fit yields for $\tilde{c}_{3G}$ and $\tilde{c}_{3G}$ (13 TeV)	149
7.14	Ratio of two and one-dimensional fit yields for $\tilde{c}_{2G}$ and $\tilde{c}_{uB}$ (13 TeV)	150
7.15	Ratio of two and one-dimensional fit yields for $\tilde{c}_H$ and $\tilde{c}_{Hu}$ (13 TeV)	151
7.16	Example code for multidimensional fitting	155
7.17	Percent error of cross section fit	157
7.18	Signal scaling fit quality as a function of included points	158
7.19	Signal scaling as a function of Wilson coefficient pairs	159
7.20	Signal scaling as a function of $\tilde{c}_{uB}$ and reference model	166
7.21	MADGRAPH event weights and quality of weight parameterization	169
7.22	Performance of the matrix element reweighting parameterization	170
B.1	Signal scaling and profile likelihood scan in the $\tilde{c}_H, \tilde{c}_{2G}$ plane	177
B.2	Signal scaling and profile likelihood scan in the $\tilde{c}_{Hu}, \tilde{c}_{2G}$ plane	178
B.3	Signal scaling and profile likelihood scan in the $\tilde{c}_{2G}, \tilde{c}_{uB}$ plane	179
B.4	Signal scaling and profile likelihood scan in the $\tilde{c}_{uG}, \tilde{c}_{2G}$ plane	180
B.5	Signal scaling and profile likelihood scan in the $\tilde{c}_{uW}, \tilde{c}_{2G}$ plane	181
B.6	Signal scaling and profile likelihood scan in the $\tilde{c}_{3G}, \tilde{c}_{2G}$ plane	182
B.7	Signal scaling and profile likelihood scan in the $\tilde{c}_{3G}, \tilde{c}_{2G}$ plane	183
B.8	Signal scaling and profile likelihood scan in the $\tilde{c}_H, \tilde{c}_{3G}$ plane	184
B.9	Signal scaling and profile likelihood scan in the $\tilde{c}_{Hu}, \tilde{c}_{3G}$ plane	185
B.10	Signal scaling and profile likelihood scan in the $\tilde{c}_{3G}, \tilde{c}_{uB}$ plane	186
B.11	Signal scaling and profile likelihood scan in the $\tilde{c}_{3G}, \tilde{c}_{uG}$ plane	187
B.12	Signal scaling and profile likelihood scan in the $\tilde{c}_{3G}, \tilde{c}_{uW}$ plane	188
B.13	Signal scaling and profile likelihood scan in the $\tilde{c}_{3G}, \tilde{c}_{3G}$ plane	189
B.14	Signal scaling and profile likelihood scan in the $\tilde{c}_{Hu}, \tilde{c}_H$ plane	190
B.15	Signal scaling and profile likelihood scan in the $\tilde{c}_H, \tilde{c}_{uB}$ plane	191
B.16	Signal scaling and profile likelihood scan in the $\tilde{c}_H, \tilde{c}_{uG}$ plane	192
B.17	Signal scaling and profile likelihood scan in the $\tilde{c}_H, \tilde{c}_{uW}$ plane	193

B.18	Signal scaling and profile likelihood scan in the $\bar{c}_H, \tilde{c}_{3G}$ plane	194
B.19	Signal scaling and profile likelihood scan in the $\bar{c}_{uB}, \bar{c}_{uG}$ plane	195
B.20	Signal scaling and profile likelihood scan in the $\bar{c}_{uB}, \bar{c}_{uW}$ plane	196
B.21	Signal scaling and profile likelihood scan in the $\bar{c}_{uB}, \tilde{c}_{3G}$ plane	197
B.22	Signal scaling and profile likelihood scan in the $\bar{c}_{uG}, \bar{c}_{uW}$ plane	198
B.23	Signal scaling and profile likelihood scan in the $\bar{c}_{uG}, \tilde{c}_{3G}$ plane	199
B.24	Signal scaling and profile likelihood scan in the $\bar{c}_{uW}, \tilde{c}_{3G}$ plane	200

TABLES

2.1	Fermions of The Standard Model	6
4.1	Lepton Selection Criteria (8 TeV)	65
4.2	Lepton Selection Criteria (13 TeV)	66
5.1	Targeted decay modes and branching fractions	69
5.2	Event selection criteria	72
5.3	Misidentification rates for muons and electrons	77
5.4	Input variables to the full event reconstruction	81
5.5	Input variable quality ranking: SS $t\bar{t}W$ versus $t\bar{t}$ BDT	84
5.6	Input variable quality ranking: 3ℓ $t\bar{t}W$ vs. $t\bar{t}$ BDT	85
5.7	Input variable quality ranking: 3ℓ $t\bar{t}Z$ vs. $t\bar{t}$ and WZ BDT	86
5.8	Input variable quality ranking: OS $t\bar{t}Z$ vs. $t\bar{t}$ BDT	87
5.9	Input variable quality ranking: OS $t\bar{t}Z$ vs. Z and $t\bar{t}$ BDT	88
5.10	Post-fit yields for the OS $t\bar{t}Z$ channel	96
5.11	Post-fit yields for the SS $t\bar{t}W$ channel	97
5.12	Post-fit yields for the 3ℓ $t\bar{t}W$, 3ℓ $t\bar{t}Z$, and 4ℓ $t\bar{t}Z$ channels	98
5.13	Measured cross sections and significances for $t\bar{t}W$	99
5.14	Measured cross sections and significances for $t\bar{t}Z$	99
6.1	Simulated datasets and cross sections	105
6.2	Summary of selection requirements for each channel	106
6.3	Systematic uncertainty impacts	112
6.4	Post-fit yields for the SS $t\bar{t}W$ channel ($D < 0$)	116
6.5	Post-fit yields for the SS $t\bar{t}W$ channel ($D > 0$)	117
6.6	Post-fit yields for the 3ℓ $t\bar{t}Z$ channel	118
6.7	Post-fit yields for the 4ℓ $t\bar{t}Z$ channel	118
6.8	Expected and observed significances	119
7.1	Signal scaling for selected processes (8 TeV)	124

7.2	Lagrangian terms for selected operators	126
7.3	Constraints on selected Wilson coefficients (8 TeV)	128
7.4	The μ_e due to operators which fail the $ \mu_e - 1 < 0.7$ requirement	131
7.5	Extreme signal strength due to eliminated operators (13 TeV)	132
7.6	Extreme signal scaling due to selected operators (13 TeV)	133
7.7	Yield difference due to eliminated operators (13 TeV)	134
7.8	Yield difference due selected operators (13 TeV)	135
7.9	Operator selection requirements (13 TeV)	136
7.10	Expected CL intervals (13 TeV)	147
7.11	Observed best fit values and CL intervals (13 TeV)	152
7.12	Scaling due to NP effects at the best-fit c_1 (13 TeV)	153
7.13	Structure constants for $t\bar{t}Z$	160
7.14	Structure constants for $t\bar{t}W$	161
7.15	Structure constants for $t\bar{t}H$	162
7.16	CL surface boundaries for 8D simultaneous fit	164

ACKNOWLEDGMENTS

I first thank the Large Hadron Collider accelerator team for their tireless work building and running this spectacular machine. I would also like to thank all of the members of the Compact Muon Solenoid collaboration. Without the dedication of these two groups of people, this dissertation would not have been possible.

At Notre Dame, I have had the good fortune of working with and learning from many talented professors. Special thanks are due to my advisor Kevin Lannon. He has provided me with guidance, encouragement, and many opportunities to pursue interesting research. He has afforded me freedom to work on side projects which have significantly enriched my time in graduate school. I am grateful for his unwavering support, especially during a period when circumstances required I take some time away from school. I am also thankful to Kevin's postdoc Jason Slaunwhite and to my fellow student Andrew Brinkerhoff. Jason and Andrew were critical to my day-to-day training as a particle physicist. Notre Dame faculty members Michael Hildreth, Randy Ruchti, Mitch Wayne, and Colin Jessop deserve thanks for creating a great atmosphere for doing particle physics research at Notre Dame.

Before graduate school, there were a number of professors at Florida State that encouraged me. I would like to thank my undergraduate advisor Volker Crede in particular, who gave me my first opportunity to participate in physics research, and helped me get started down this path. He took a chance on me, despite having me in his introductory physics class, which was before I learned how to study.

My time at CERN and in South Bend have been enormously enriched by the friends I have made there. I would like to thank Matthias Wolf especially. We built Lobster together, a software tool which has been one of the most rewarding side projects I have had the

opportunity to work on. I thank him for being such a great friend throughout the years. I would also like to thank Ted Kolberg, Jamie Antonelli, Sean Lynch, Rachel Yohay, Dave and Kat Morse, Nil Valls, John Wood, Nathan Kellams, Michael Planer, Douglas Berry, Tessa Pearson, Allison Reinsvold, Fanbo Meng, Sarah Boutle, Chris Neu, Justin Griffiths, Nabarun Dev, Mayukh Raj Gangopadhyay, Charlie Mueller, Wuming Luo, Geoffrey Smith, Andrew Wightman, Wenzhao Li, Jeff Kolb, and countless others.

I owe a great debt to the many theorists who have helped me along the way. Thank you for providing guidance and patiently answering my questions. Special thanks go to Adam Martin, Fatimah Elahi, Landon Lehman, Ayan Paul, Joe Bramante, Brian Ostdiek, Fabio Maltoni and Eleni Vryonidou.

My analysis depended on Lobster, which could not have been developed without the tools and support provided by Douglas Thain's Cooperative Computing Laboratory at Notre Dame. I am grateful to Professor Thain for his insight and guidance in computing matters. I would also like to thank Benjamin Tovar, Patrick Donnelly, Peter Ivie, Haiyan Meng, and the rest of the CCL team. I am also thankful to Kevin's postdoc Kenyi Paolo Hurtado Anampa for his friendship and computing help. Likewise, my work would not have been possible without Serguei Federov, Paul Brenner, Scott Hampton, and the rest of the team at Notre Dame's Center for Research Computing. Thanks for putting up with me. Sorry about that time I accidentally hammered the cluster with misconfigured jobs writing two hundred gigabyte log files. I would also like to thank the CMS CRAB3 team and the HLT team. I am forever in the debt of the unnamed masses at Stack Overflow: this dissertation would have taken several more years without you.

I am grateful to the 13 TeV ttV team for taking me in: Didar Dobur, Illia Khvastunov, Mirena Paneva, and Deniz Poyraz.

I could never have written this dissertation without the love and support of my friends and family: my parents, Margaret and Daniel Woodard (who also deserves thanks for proof-reading this dissertation); my uncle Michael Barnes; my brothers and many other cousins,

aunts, and uncles; and Nora Jacobsen Ben Hammed, Mallory Smith, and Glenn Anderson. Finally, I would like to thank my best friend Amy Doran and my husband Orin Harris for their love and support; and to thank Orin for his patience, and Amy for her impatience.

CHAPTER 1

INTRODUCTION

The Standard Model (SM) is a spectacularly successful model of nature. It correctly predicted the charm and top quarks, the gluon, and the W, Z and Higgs bosons before their discoveries, and its numerical predictions agree with experimental results to a good (and in some cases remarkable) degree of precision. Nevertheless, the SM does not describe everything we observe: there is not enough charge parity violation to produce the imbalance between baryonic and antibaryonic matter in the visible universe, there is no good candidate for dark matter, and it does not address gravity. Furthermore, some believe the SM is not entirely satisfactory because it depends on many arbitrary parameters, and it does not explain, for example, the origins of electroweak symmetry breaking and the reasons underlying the existence of three generations of particles and the masses of those particles. The SM includes all of the currently known particles, but it is possible that more particles are yet to be discovered. Such particles may not have been detected because the probability that they would interact with currently known particles is small, or they may be so massive that the energy needed to produce them exceeds the amount that current experiments can provide. The SM may only be an effective low-energy theory, an approximation of reality that is valid at scale Λ . Through the use of effective field theory (EFT), potential deviations from the SM due to currently unknown particles or forces (new physics, or NP) can be parameterized in a model-independent way by extending the SM Lagrangian with higher-dimension operators.

NP involved in electroweak symmetry breaking would have a large coupling to the top quark because it is the heaviest particle. This dissertation describes two measurements of top-quark pairs produced in association with W or Z bosons ($t\bar{t}W$ and $t\bar{t}Z$). These mea-

measurements are interpreted within the framework of EFT in order to constrain the Wilson coefficients of selected dimension-six operators; these coefficients characterize the strength of the NP interaction.

This dissertation is organized as follows: the theoretical framework is introduced in Chapter 2. The Large Hadron Collider (LHC) and the Compact Muon Solenoid (CMS) experiment are described in Chapter 3. Chapter 4 explains how particle observables are reconstructed from the raw detector electrical signals and the quality requirements that are imposed on them. Chapters 5 and 6 present cross section measurements of $t\bar{t}W$ and $t\bar{t}Z$ production performed using LHC data from proton–proton collisions collected during Run 1 at $\sqrt{s}=8$ TeV and Run 2 at $\sqrt{s}=13$ TeV, respectively. Chapter 7 is the main subject of this dissertation. It describes the interpretation of the 8 TeV and 13 TeV $t\bar{t}W$ and $t\bar{t}Z$ measurements within the framework of EFT in order to constrain the Wilson coefficients of selected dimension-six operators that might signal the presence of NP contributions in $t\bar{t}W$ and $t\bar{t}Z$ production.

Chapters 2 and 3 and Section 4.1 are mainly reviews of existing literature. I made significant contributions to the analysis described in Chapter 5, but not to the analysis described in Chapter 6, and was the main analyst for the EFT interpretation described in Chapter 7. The 8 TeV measurements and their EFT interpretation have been published in the Journal of High Energy Physics [1]. The 13 TeV measurements and their EFT interpretation have been submitted to the Journal of High Energy Physics [2].

CHAPTER 2

THEORY

2.1 The Standard Model

The SM is currently considered the best available theoretical description of the fundamental particles as well as three of the four known fundamental forces of the universe, with gravity being the only exception. In the quantum mechanical description of the universe, classical point particles are replaced by wavefunctions that evolve according to the Schrödinger equation and describe the probability of a measurement outcome. In contrast, the SM is a Quantum Field Theory (QFT), which applies quantum mechanics to classical *fields*. These fields are continuous mathematical objects that specify a value or vector at each spacetime point. Quantum fields may have localized excitations, or ripples, that can be thought of as *particles*. The SM includes a set of fields whose excitations, or particles, are fundamental in that they are not expected to be approximations of alternative field descriptions at smaller distance scales. The SM is relativistic in the sense that its equations of motion are Lorentz invariant. In other words, they do not depend on the choice of inertial frame. If that were not the case, the laws of physics could depend on an arbitrary choice of reference frame.

Fundamental particles in the SM can be classified by how their joint wavefunctions behave under interchange of two identical particles: *bosons*, whose wavefunctions do not change under such exchange have integer values of the quantum number corresponding to intrinsic angular momentum (i.e., $spin^1$), while *fermions*, whose wavefunctions change sign under exchange of two identical particles have half-integer spin. A profound implication

¹Spin is a multiple of the reduced Planck constant $\hbar$, but throughout this work natural units will be used for which $c = \hbar = 1$.

of this difference is that no two fermions can occupy the same quantum state, while no such restriction exists for bosons. Fundamental particles can have one of two possible chiralities. This intrinsic property specifies the direction in which the quantum phase rotates in the complex plane (clockwise is referred to as left-handed, and counterclockwise as right-handed). Fundamental particles are further characterized by their mass and by quantum numbers called *charges* that determine which of the fundamental forces they participate in.

Fermions are divided into two groups: quarks and leptons. Quarks carry *color* charge, indicating that they participate in the strong interaction, while *leptons* do not. Quarks also carry an electric charge and therefore participate in the electromagnetic interaction. They can be of an *up-type*, with an electric charge of $+\frac{2}{3}|e|$, or a *down-type*, with an electric charge of $-\frac{1}{3}|e|$. Leptons likewise can be *electron-type* or *neutrino-type*. Electrons carry an electric charge of $-|e|$ while neutrinos have no electric charge.

There are three *generations*² of fermions, which are identical except for mass. Each generation consists of an up-type quark, a down-type quark, an electron-type lepton, and a neutrino-type lepton. The more massive generations are not stable, and ordinary matter is consequently made up of first-generation particles. For each fermion, a corresponding antifermion with the same mass but with opposite charge exists. Charged massive fermions are described by Dirac spinors, which have a component for each of the two chirality states. The uncharged fermions, the neutrinos, carry only left-handed chirality (anti-neutrinos carry right-handed chirality). Properties of the fundamental fermions are summarized in Table 2.1.

Because spin is conserved, a fermion with a spin projection of $+\frac{1}{2}$ may emit a boson with a spin of 1 and remain a fermion with a spin projection of $-\frac{1}{2}$. Similarly, a boson with integer spin can emit a boson and will remain a boson. However a fermion cannot emit another fermion because this would leave a particle with integer spin. Consequently, only bosons can be “exchanged”, that is, emitted or absorbed while leaving the emitter or absorber intact.

²Note that the number of generations is an experimentally driven feature of the SM; there is no theoretical justification for why there should be exactly three generations.

The fundamental forces result from couplings between fundamental fields. For example, the force between two excitations of a fermionic field is mediated by an excitation of a fundamental bosonic field to which it is coupled, which can be understood as an interaction in which a boson is exchanged between two fermions. It should be noted that a force is understood to be involved more generally in any interaction involving the excitation of a bosonic field. For example, radioactive beta decay is a weak force process because it necessarily involves the emission of a bosonic excitation of one of the fields associated with the weak force.

All massive particles interact gravitationally; however, gravity is such a weak force that it is irrelevant and undetectable in particle interactions at the LHC. The gravitational interaction is not described in the SM. The electromagnetic interaction, responsible for binding electrons in atoms, is mediated through the exchange of an excitation of the electromagnetic field, a *photon* (γ), which is electrically neutral and massless. Because the photon is massless, the range of the electromagnetic force is infinite. The strong force is responsible for holding quarks together to form *hadrons* and for binding protons and neutrons together to form atomic nuclei. It is mediated via the exchange of massless bosons called gluons. The weak force, responsible for the decay of heavier fermion generations into lighter ones, is mediated by the W^+ , W^- , and Z bosons. The W and Z bosons, discovered at CERN in 1983 [3, 4], play a special role in this thesis. The Higgs field is a scalar field with a non-zero expectation value that permeates all space. Excitations of this field (the Higgs boson) couple to the W and Z bosons and the fermions to give them mass.

2.1.1 The Lagrangian formulation and principle of stationary action

The *Lagrangian density* $\mathcal{L}$ (hereafter, simply the Lagrangian) is a functional that takes the configuration of one or more fields ϕ_i and their derivatives $\partial_\mu \phi_i$ and outputs a number. The Lagrangian encodes the system's dynamics. A sequence of configurations in spacetime comprise a *path*. The integral of the Lagrangian along a given path in spacetime is called that

TABLE 2.1

FERMIONS OF THE STANDARD MODEL

		Symbol	Mass / GeV	Electric charge / $ e $	T_3
Quarks	up	u	2.2×10^{-3}	$+\frac{2}{3}$	$+\frac{1}{2}$
	down	d	4.7×10^{-3}	$-\frac{1}{3}$	$-\frac{1}{2}$
	charm	c	1.28	$+\frac{2}{3}$	$+\frac{1}{2}$
	strange	s	0.096	$-\frac{1}{3}$	$-\frac{1}{2}$
	top	t	173.1	$+\frac{2}{3}$	$+\frac{1}{2}$
	bottom	b	4.18	$-\frac{1}{3}$	$-\frac{1}{2}$
Leptons	electron	e	0.51×10^{-3}	-1	$-\frac{1}{2}$
	muon	μ	0.106	-1	$-\frac{1}{2}$
	tau	τ	1.777	-1	$-\frac{1}{2}$
	neutrino	$\nu_{e,\mu,\tau}$	$< 20 \times 10^{-3}$	0	$+\frac{1}{2}$

path's *action*³ and is proportional to the change in complex phase accrued along that path:

$$S = \int \mathcal{L} d^4x$$

where $x^\mu = (t, x, y, z)$. One of Feynman's remarkable insights was that a system evolving from an initial to a final configuration can be considered as following all possible paths between them (even those paths that do not satisfy usual physical laws) and that each possible path is equally probable. When the paths are added together, those in which the action is *stationary* (i.e., in which considering small variations to arbitrary neighboring paths produces no change in the action) tend to add constructively, while the variance in

³See [5] chapter 18 for a good introductory discussion of this topic.

the other paths' actions are large and random such that those paths add incoherently. As a result, the most highly weighted path is the one in which the action is stationary. This is called the *principle of stationary action*, although in most cases the path of stationary action minimizes the action. A prominent historical example is the observation by Fermat that light follows the path of least time (in the case of light, time is proportional to action).

2.1.2 Symmetry and gauge invariance

Symmetry (i.e., invariance under transformation) seems to play a profound role with respect to the laws of nature. Symmetries can be either continuous or discrete. Rotation of a square is an example of a discrete transformation: only 90, 180, and 270 degree turns are symmetric. Rotation of a circle is an example of a continuous symmetry: after a rotation of any arbitrary angle, the circle looks the same as before. Continuous transformations can be represented by *Lie groups*, in which an infinitesimal transformation g , expressed in terms of infinitesimal parameters α^a , is an infinitesimal perturbation of the identity transformation: $g(\alpha) = 1 + i\alpha^a T^a + \mathcal{O}(\alpha^2)$, where the Hermitian operators T^a are the *generators* of the group. Finite transformations can be constructed from the repeated application of such infinitesimal transformations.

A Lagrangian exhibits a continuous symmetry if it is invariant under a transformation of fields $\phi_i(x) \rightarrow \phi'_i(x) = U_{ij}\phi_j(x)$. If the continuous transformation U_{ij} is constant at every point x in spacetime it is said to be *global*. Noether's first theorem states that a global continuous symmetry in the Lagrangian implies that an associated quantity is conserved. For example, invariance with respect to translations in time gives rise to the conservation of energy, invariance with respect to translations in space gives rise to the conservation of momentum, and invariance with respect to rotations in space gives rise to conservation of angular momentum. In electrodynamics, the Lagrangian is invariant under global phase transformations, which gives rise to the conserved electric charge.

Ordinary transformations connect physically inequivalent states, leaving observables

unchanged. If a ball is thrown inside a train car, the calculation of its trajectory will be invariant with respect to the train car's velocity, even though different velocities correspond to different physical states. In contrast, a *gauge* transformation only changes our mathematical description and does not connect different physical states. Field theories describe fields that are not directly measured; instead, measurements are made of various observables such as energies, velocities, and charges. Different configurations of the underlying fields (i.e., different *gauges*) may possibly lead to the same set of observations; there is an inherent redundancy. If the observable quantities of a physical theory do not change when a gauge transformation is performed (i.e., one possible gauge configuration is transformed to another), then the physical theory is said to be gauge invariant.

The SM is based on the requirement that the theory be invariant under certain types of *local* continuous gauge transformations $U_{ij}(x)$, which unlike the global transformation U_{ij} , can vary from point to point. The program of insisting that the theory respect some set of local gauge symmetries is called *gauge theory*, and the procedure used to enforce the symmetry on the fermionic fields necessitates the existence of additional *gauge fields*, bosonic fields that are associated with the SM forces.

To see how this works, consider the Dirac Lagrangian for a fermion:

$$\mathcal{L} = \bar{\psi}(i\gamma^\mu\partial_\mu - m)\psi \quad (2.1)$$

where m is the mass associated with a spin-1/2 particle ψ , γ^μ are the Dirac matrices, and $\bar{\psi}$ is the Hermitian conjugate of ψ . This Lagrangian is invariant under the global transformation of the $U(1)$ group $\psi \rightarrow e^{iq\theta}\psi$, where q is the charge of the particle involved, which is conserved via Noether's theorem. If we consider a local transformation $\psi \rightarrow e^{iq\theta(x)}\psi$ which varies from point to point, we see that the Lagrangian is no longer invariant because of the

derivative in first term:

$$\partial_\mu(e^{iq\theta}\psi) = iq\partial_\mu\theta e^{iq\theta}\psi + e^{iq\theta}\partial_\mu\psi.$$

For the Lagrangian this means that

$$\begin{aligned}\mathcal{L} &\rightarrow -q\bar{\psi}\gamma^\mu\partial_\mu\theta\psi + \bar{\psi}i\gamma^\mu\partial_\mu\psi - m\bar{\psi}\psi \\ &\rightarrow \mathcal{L} - q\bar{\psi}\gamma^\mu\partial_\mu\theta\psi.\end{aligned}\tag{2.2}$$

Clearly the invariance has been spoiled, but a general procedure is known to recover the symmetry. The derivative in Equation (2.1) is replaced with the *covariant derivative*,

$$D_\mu = \partial_\mu - iqA_\mu,$$

which includes an additional term whose role is to produce a cancellation of the term in Equation (2.2) that spoiled the invariance. This procedure has resulted in the addition of a new vector field A_μ , transforming as $A_\mu \rightarrow A_\mu - \partial_\mu\theta$, which has been coupled to the fermionic field through cross terms in Equation (2.1). Because a new field has been added to the theory, a corresponding kinetic term must be added to describe the propagation of a vector field of mass m to the Lagrangian:

$$\mathcal{L}_{\text{free}} = -\frac{1}{4}(\partial^\mu A^\nu - \partial^\nu A^\mu)(\partial_\mu A_\nu - \partial_\nu A_\mu) + m^2 A^\mu A_\mu.\tag{2.3}$$

The factors in the first term transform as

$$\begin{aligned}\partial^\mu A^\nu - \partial^\nu A^\mu &\rightarrow \partial^\mu(A^\nu - \partial^\nu\theta) - \partial^\nu(A^\mu - \partial^\mu\theta) \\ &\rightarrow \partial^\mu A^\nu - \partial^\nu A^\mu\end{aligned}$$

and so are invariant under the gauge transformation. The second term transforms as $m^2 A^\mu A_\mu \rightarrow m^2 A^\mu A_\mu - 2m^2 A^\mu \partial_\mu \theta$. Therefore, to leave the Lagrangian invariant for arbitrary θ , we must have $m=0$. Thus, starting from the globally invariant Dirac Lagrangian for fermions and requiring that it be locally invariant⁴, a new massless vector field has arisen as a matter of necessity. Putting everything together and identifying the electromagnetic field tensor $F^{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu$ we recognize this result as the quantum electrodynamics (QED) Lagrangian:

$$\begin{aligned}\mathcal{L}_{QED} &= \bar{\psi}(i\gamma^\nu D_\nu - m)\psi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} \\ &= \bar{\psi}(i\gamma^\nu \partial_\nu - m)\psi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} + q\bar{\psi}\gamma^\mu\psi A_\mu\end{aligned}$$

where q is the electric charge, and the final term describes the interaction between the photon field and charged fermions. QED is an example of an *abelian* gauge theory because successive applications of the generator (i.e., successive changes in the phase) commute with one another: the order in which they are applied can be reversed without changing the result.

2.1.3 The Standard Model Lagrangian

QED has been remarkably successful, but to make further progress we must also consider non-abelian gauge theories for which the generators T^a do not commute:

$$[T^a, T^b] = if^{abc}T^c,$$

where f^{abc} are called the structure constants. In the following sections the electroweak and quantum chromodynamics (QCD) sectors are discussed separately, but it is helpful to first note the common elements that are required. In each case we need to determine 1) which gauge group specifies the vector fields, 2) which fields are present, 3) how the fields transform under the group, and 4) a locally gauge-invariant Lagrangian in terms

⁴Indeed, it seems unnatural *not* to require local invariance, as spacelike-separated points cannot communicate with each other; the alternative suggests action at a distance.

of the fields. For each group, a representation t^a is first chosen for the generators of the group. Next, we demand that the theory be invariant under a local $SU(N)$ transformation⁵, $\psi \rightarrow \psi' = e^{i\alpha^a(x)t^a} \psi$, where $\alpha^a(x)$ are arbitrary functions of x .

Similarly, as was introduced in Section 2.1.2, the covariant derivative $D_\mu = \partial_\mu - igA_\mu^a t^a$ introduces a new coupling constant g parameterizing the strength of the interaction and vector field A_μ^a for each generator of the local symmetry. The field tensor is defined in terms of the covariant derivative:

$$\begin{aligned}
-igF_{\mu\nu}^a t^a &= [D_\mu, D_\nu] \\
&= [\partial_\mu - igA_\mu^a t^a, \partial_\nu - igA_\nu^b t^b] \\
&= [\partial_\mu, \partial_\nu] - ig[A_\mu^a t^a, \partial_\nu] - ig[\partial_\mu, A_\nu^b t^b] + (ig)^2 [A_\mu^b t^b, A_\nu^c t^c] \\
&= ig[\partial_\nu A_\mu^a - \partial_\mu A_\nu^a + (ig)f^{abc} A_\mu^b A_\nu^c] t^a \\
&\rightarrow F_{\mu\nu}^a = \partial_\mu A_\nu^a - \partial_\nu A_\mu^a + gf^{abc} A_\mu^b A_\nu^c.
\end{aligned} \tag{2.4}$$

Note that the final term in the field tensor appears as a result of the non-commutative nature of the generators of $SU(N)$. It indicates that when the kinetic term is expressed as in Equation (2.3), terms arise with three and four factors of the vector field. These factors are the gauge boson self-interactions which do not appear in abelian theories such as QED. Note also that a gauge field has emerged corresponding to each generator of the group. The symmetry group used in the case of QED is $U(1)$, which has a single generator and a single gauge field. In addition, Noether's theorem results in a single conserved charge. $SU(N)$ has $N^2 - 1$ generators; therefore, an $SU(N)$ gauge theory has $N^2 - 1$ gauge fields, and application of Noether's theorem yields N conserved charges.

The SM is based on the product of two special unitary groups and one unitary group:

⁵A unitary group $U(N)$ is a group of $N \times N$ unitary matrices, where a unitary matrix is one for which the conjugate transpose equals the matrix inverse; these groups have a special importance in physics because they preserve norms. A special unitary group has the additional property that the determinant of each group element is 1. A unitary group has N^2 generators. For a special unitary group, the condition of unitarity reduces the number of independent elements by one, so a special unitary group has $N^2 - 1$ generators.

$SU(3) \times SU(2) \times U(1)$. The $SU(3)$ group describes the strong interactions, while the $SU(2) \times U(1)$ group describes both the weak and electromagnetic interactions.

2.1.4 Quantum chromodynamics

QCD, which provides a successful description of the strong interaction, emerges when the Lagrangian is required be invariant under local transformations of the non-abelian gauge group $SU(3)_C$.

Because it is $SU(N)$ with $N = 3$, there are three conserved charges, referred to as *color* charges *red*, *green*, and *blue*. Similarly, there are $N^2 - 1 = 8$ generators of the group corresponding to the eight gauge fields, called *gluons*. Leptons, which carry no color charge, do not interact via the strong force. Each quark flavor can be expressed as a color triplet:

$$q = \begin{pmatrix} q_r \\ q_g \\ q_b \end{pmatrix},$$

where the q_x are Dirac spinors and the subscripts label the color states. The QCD Lagrangian is

$$\mathcal{L}_{QCD} = \mathcal{L}_{gluon} + \mathcal{L}_{quark}, \text{ with}$$

$$\mathcal{L}_{gluon} = -\frac{1}{4} \sum_{a=1}^8 G_{\mu\nu}^a G_a^{\mu\nu},$$

where the field strength tensor is defined similarly to Equation (2.4):

$$-ig_s G_{\mu\nu}^a t^a = [D_\mu, D_\nu], \text{ with}$$

$$D_\mu^{ij} = \partial_\mu \delta_{ij} - ig_s t_{ij}^a G_\mu^a$$

$$\rightarrow G_{\mu\nu}^a = \partial_\mu G_\nu^a - \partial_\nu G_\mu^a - g_s f^{abc} G_\mu^b G_\nu^c, \quad (2.5)$$

where t^a are the generators of $SU(3)$, the i s and j s are color indices running from 1 to 3,

the sum over a runs from 1 to 8, G_μ^a are the gluon fields, g_s is the strong coupling constant, and f^{abc} are the structure constants of SU(3) defined by the commutation relation between the generators of SU(3). As carriers of color charge, gluons can couple to one another; this “self-interaction” arises from the last term in Equation (2.5).

The quark Lagrangian is

$$\mathcal{L}_{quark} = \sum_f \bar{q}_i^{(f)} (i\gamma_\mu D_{ij}^\mu - m_f \delta_{ij}) q_j^{(f)}$$

where the sum over f corresponds to the six quark flavors and i and j again refer to the three colors.

2.1.4.1 Confinement and hadronization

Unlike in QED, the force between quarks is large at large separations and does not become weaker with increasing separation. Conversely, quarks separated by short distances (e.g., during high energy scattering) are only weakly coupled. This striking feature of QCD is called *asymptotic freedom*. To understand this concept, it is helpful to draw an analogy to QED. When a photon propagates through the vacuum, it can induce the production of virtual pairs of electrons and positrons (called vacuum fluctuations) that have a *screening* effect, making the interaction coupling weaker at larger distances (similar to how the force between two electrons is effectively reduced when submerged in a dielectric medium). QCD has the same screening effect due to the production of quark-antiquark pairs, which tends to make the interaction coupling decrease with increasing distance. Unlike the electrically neutral photons in QED, however, gluons carry color charge. This charge leads to additional *anti-screening* effects from nonlinear self-interactions between the gluons, making the interaction coupling larger at large distances. In his Nobel prize lecture [6], Wilczek uses the analogy of a thundercloud, where a small charge induces a cloud of virtual particles in a self-reinforcing process. The cloud gets bigger as one moves away from the source,

enhancing its power. This *running coupling constant* means that QCD can only be calculated perturbatively at large momentum transfers with $q^2 \gg \Lambda_{\text{QCD}}^2$, where $\Lambda_{\text{QCD}} \sim 250 \text{ MeV}$.

A consequence of asymptotic freedom at low energies is *color confinement*: quarks and gluons do not exist in isolation at distances larger than $1/\Lambda_{\text{QCD}}$. As an example, consider trying to pull apart a stable, color-neutral meson (quark-antiquark pair). As the two quarks are separated, the energy in the gluon field between them grows until it is energetically favorable to produce a quark-antiquark pair that combines with the original quark and antiquark to form two separate mesons. The collision energies at the LHC are high enough (corresponding to short interaction distances) that the collisions are between quarks and gluons rather than protons. After high-energy quarks or gluons emerge from such a collision, they undergo two separate processes. First, a *shower* of quarks is produced due to the tendency of quarks to radiate gluons and gluons to in turn produce quark-antiquark pairs. Second, due to the process of color confinement, the shower of isolated quarks recombine or combine with quark-antiquark pairs, produced from the energy in the above-mentioned gluon field between separated quarks, to ultimately result in a collimated “spray” of hadrons called a *jet*. The total energy of the jet reflects the energy of the quark or gluon that initiated it. Because jets are showers of hadrons that do not unambiguously reflect the nature of the quark or gluon that initiated them, and because jets are produced so commonly in high energy proton collisions, they result in challenging backgrounds for many LHC analyses.

2.1.5 Electroweak interactions

A unified description of the weak and electromagnetic forces was proposed by Glashow [7], Salam [8], and Weinberg [9], and it results from the assumption of invariance under an $SU(2) \times U(1)$ gauge symmetry. The $SU(2)$ symmetry gives rise to the conserved weak isospin charge with three corresponding massless bosons (W_1 , W_2 , and W_3), and the $U(1)$ symmetry gives rise to the conserved weak hypercharge and its associated massless B boson. Section 2.1.6 covers how these bosons relate to the familiar bosons of the electromagnetic

and weak interactions.

A major difference between the weak force and the other fundamental forces is that the weak force is a *chiral* theory: it does not conserve *parity* (i.e., inversion of spatial coordinates). Consequently, to specify the field content it is necessary to split it into left- and right-handed components, which can be accomplished via the *projection operators*:

$$\begin{aligned}\psi_L &= \frac{1}{2}(1 - \gamma^5)\psi \\ \psi_R &= \frac{1}{2}(1 + \gamma^5)\psi,\end{aligned}$$

where the right-handed fermions do not participate in the weak interaction. For this reason, the weak isospin group is referred to as $SU(2)_L$. Mathematically, this is encoded in the transformation properties of the two types of fermions: right-handed fermions transform as isospin singlets ($I = 0$) and left-handed fermions transform as isospin doublets ($I = \frac{1}{2}$):

$$\begin{aligned}Q_L^i &= \begin{pmatrix} u_L^i \\ d_L^i \end{pmatrix}, \quad E_L^i = \begin{pmatrix} \nu_L^i \\ l_L^i \end{pmatrix}, \\ f_R^i &= e_R^i, u_R^i, d_R^i\end{aligned}$$

where i runs over the three fermion generations. Note that in this formulation, the fundamental fermions are all massless because a mass term would not be gauge invariant. The mass term is

$$\begin{aligned}-m_f \bar{\psi}\psi &= -m_f(\bar{\psi}_R + \bar{\psi}_L)(\psi_R + \psi_L) \\ &= -m_f(\bar{\psi}_R\psi_L + \bar{\psi}_L\psi_R).\end{aligned}\tag{2.6}$$

This expression is not gauge invariant because ψ_L changes under isospin rotations while ψ_R does not. The mechanism by which the fermion masses arise is explained in the following

section. The covariant derivative D_μ is

$$D_\mu = \partial_\mu - ig \frac{1}{2} W_\mu^a \sigma^a - i \frac{1}{2} g' B_\mu \quad (2.7)$$

where the a index runs from 1 to 3 over the generators of $SU(2)$ and

σ^a = Pauli matrices

B_ν = $U(1)$ gauge field

W_μ^a = 3 $SU(2)$ gauge fields

g' = $U(1)$ coupling constant

g = $SU(2)$ coupling constant.

A three-dimensional *weak isospin* space is defined by the three generators σ^a ; global $SU(2)$ transformations correspond to rotations of the W_μ^a vectors in this space. The Lagrangian describing the massless gauge fields is given by the following:

$$\mathcal{L}_{gauge} = -\frac{1}{4} W_{\mu\nu}^a W_a^{\mu\nu} - \frac{1}{4} B_{\mu\nu} B^{\mu\nu}$$

where the field strength tensors $W_{\mu\nu}^a$ and $B^{\mu\nu}$ are defined as

$$B_{\mu\nu} = \partial_\mu B_\nu - \partial_\nu B_\mu,$$

$$W_{\mu\nu}^a = \partial_\mu W_\nu^a - \partial_\nu W_\mu^a + g f^{abc} W_\mu^b W_\nu^c.$$

To complete the picture, we add a term to describe how matter interacts with the B_ν and W_μ^a fields:

$$\mathcal{L}_{int} = \sum_{i,f} \bar{f}^i (i \gamma^\mu D_\mu) f^i,$$

where $i = 1, 2, 3$ refers to the three fermion generations, and $f = E_L^i, Q_L^i, e_R^i, u_R^i, d_R^i$. The elec-

troweak sector is then described by $\mathcal{L}_{EW} = \mathcal{L}_{gauge} + \mathcal{L}_{int}$. This completes the description of an electroweak sector that is flawed in several respects. First, all four gauge bosons are massless, while the observed gauge bosons associated with the weak force are massive. In addition, there are four conserved charges, while the observed electroweak sector conserves only electric charge and the third component of weak isospin. Further, as we saw in Equation (2.6), the chiral nature of the weak interaction implies that none of the SM fermions that participate in the weak interaction can have mass. In the next section, these seemingly grave problems are addressed.

2.1.6 Spontaneous symmetry breaking

Within the SM, as long as the $SU(3) \times SU(2) \times U(1)$ symmetry is exact, all gauge bosons should be massless as shown in Section 2.1.2 and the fermions should be massless as shown in Section 2.1.5. However, these outcomes are in conflict with the experimental evidence: the $SU(2)$ bosons and the fermions have been observed to have mass.

In order for the fundamental fermions and gauge bosons to acquire mass, the gauge symmetry must somehow be broken, which occurs through a mechanism called *spontaneous symmetry breaking* (SSB). SSB is “spontaneous” in the sense that no outside cause of the symmetry breaking is present. For a skydiver falling to the ground, gravity breaks the symmetry, making up and down different from left and right. In SSB, the state of the system is not symmetric, but the laws of physics remain so. As a classical example of SSB, consider a ball placed at the top of the surface in Figure 2.1 subject to gravitational force pointing downward along the axis of symmetry. The Lagrangian of the system is symmetric with respect to rotations around the central axis, but the ball is unstable due to quantum fluctuations (i.e., Heisenberg’s uncertainty relations require that the wave function of the ball cannot be indefinitely localized). To reach the stable ground state, the system must spontaneously break the rotational symmetry, with the ball rolling to the bottom. In the ground state, the ball and the surface individually still possess rotational symmetry, but the system as a whole does not.

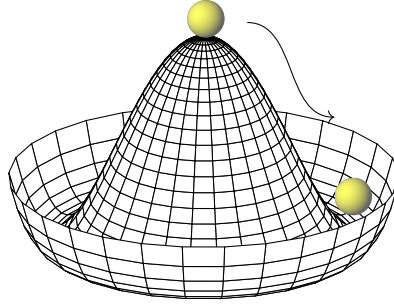


Figure 2.1. A classical example of an SSB is obtained by considering the motion of a ball placed at the top of this surface.

Another example is the orbit of the moon around the Earth. Newton's laws are rotationally invariant, but the solution of these laws that describe the moon's motion is not; it depends on initial conditions, so the rotational symmetry of the underlying law is not manifest.

The fact that the fermion and vector bosons have mass indicates that SM gauge invariance must be spontaneously broken. However, the electric charge is observed to be conserved and the photon is observed to be massless, so whatever SSB the SM observes should preserve $SU(3)_c \times U(1)_{EM}$.

To see how this can be accomplished⁶, it is important to note that the description of field fluctuations assumes fluctuations with respect to some ground state, that is, a configuration that has minimum energy. In the previous discussion the implicit assumption is that the lowest-energy field configuration is $\phi = 0$, so that trivially the vacuum expectation value is invariant under the symmetry. SSB will arise if $\phi = 0$ is not the ground state.

To see how this works, consider the case of a complex scalar field $\phi = \phi_1 + i\phi_2$ with the

⁶This section follows the derivations in [10, 11, 12].

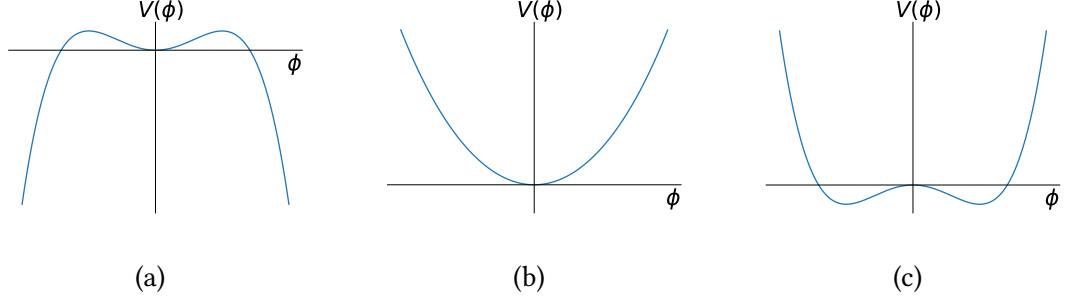


Figure 2.2. The potential for a complex scalar field given $\lambda < 0$ (a), $\mu^2 > 0$ (b), and $\mu^2 < 0$ (c).

Lagrangian

$$\mathcal{L} = (\partial_\mu \phi)^\dagger (\partial^\mu \phi) - U(\phi^\dagger \phi), \quad U(\phi^\dagger \phi) = \frac{1}{2} \mu^2 \phi^\dagger \phi + \frac{1}{4} \lambda (\phi^\dagger \phi)^2. \quad (2.8)$$

This Lagrangian is invariant under the global $U(1)$ phase transformation $\phi \rightarrow e^{i\theta} \phi$. What are the constraints on μ^2 and λ ? If λ is negative, the potential will not be bounded from below (see Figure 2.2 (a)); such a system would be unstable and have no defined ground state. There is no such constraint on μ^2 ; as shown in Figure 2.2 (b), there is one minimum in the case that $\mu^2 > 0$, and there are degenerate minima along a circle of radius $\mu/\sqrt{\lambda}$ in the case that $\mu^2 < 0$ (Figure 2.2 (c)):

$$\begin{aligned} \frac{dU}{d(\phi^\dagger \phi)} &= -\frac{1}{2} \mu^2 + \frac{1}{2} \lambda (\phi^\dagger \phi) = -\frac{1}{2} \mu^2 + \frac{1}{2} \lambda (\phi_1^2 + \phi_2^2) = 0 \\ &\rightarrow \phi_1^2 + \phi_2^2 = \mu^2 / \lambda. \end{aligned}$$

In the case that $\mu^2 < 0$, the system spontaneously breaks the $U(1)$ symmetry by choosing an arbitrary point along the circle of minima. The particular direction in which ϕ is real can be chosen without any loss of generality by choosing the ground state to be $\phi_{\text{ground}} = (\mu/\lambda) \phi_1$ (with $\phi_2 = 0$). Fluctuations of this field are now with respect to this vacuum state rather than $\phi = 0$, so it is possible to expand around it by introducing two coupled scalar fields with

$\eta \equiv \phi_1 - \mu/\lambda$ and $\xi \equiv \phi_2$, in terms of which the Lagrangian is

$$\begin{aligned}\mathcal{L} &= \frac{1}{2} \partial_\mu \eta \partial^\mu \eta + \frac{1}{2} \partial_\mu \xi \partial^\mu \xi - \frac{\lambda^2}{4} \left(\phi^\dagger \phi - \frac{\mu^2}{\lambda^2} \right)^2 + \frac{\mu^4}{4\lambda^2} \\ &= \frac{1}{2} \partial_\mu \eta \partial^\mu \eta + \frac{1}{2} \partial_\mu \xi \partial^\mu \xi - \frac{\lambda^2}{4} \left(\eta^2 + \frac{2\mu\eta}{\lambda} + \xi^2 \right)^2 + \frac{\mu^4}{4\lambda^2}.\end{aligned}$$

The new Lagrangian no longer appears symmetric with respect to $\phi \rightarrow e^{i\theta} \phi$; the underlying symmetry is not manifest after redefinition of the fields with respect to the vacuum state.

The Lagrangian can be broken up as follows:

$$\mathcal{L} = \mathcal{L}_{\text{free}} + \mathcal{L}_{\text{int}} + \frac{\mu^4}{4\lambda^2}$$

where

$$\mathcal{L}_{\text{free}} = \frac{1}{2} \partial_\mu \xi \partial^\mu \xi + \frac{1}{2} \partial_\mu \eta \partial^\mu \eta - \mu^2 \eta^2. \quad (2.9)$$

Here $\mathcal{L}_{\text{free}}$ represents the free particle fields and $\mathcal{L}_{\text{int}}$ represents higher-order interactions between them.⁷ The final term in Equation (2.9) corresponds⁸ to a scalar spin-zero particle of mass $\sqrt{2}\mu$. The ξ field is massless; Goldstone's theorem [13] proves that SSB of a continuous global symmetry will always give rise to a massless scalar boson for each broken generator. No such “Goldstone bosons” have been observed.

While the above procedure exemplifies how a gauge symmetry can be spontaneously broken, it is not quite sufficient for the purpose of breaking $SU(2) \times U(1)$ in the SM. The

⁷The $\mu^4/4\lambda^2$ term is, like all constant terms in a Lagrangian, irrelevant because it will not affect the Euler-Lagrange equations and hence the equations of motion.

⁸Compare with the Klein-Gordon Lagrangian for a free scalar field: $\mathcal{L} = \frac{1}{2} [\partial_\mu \phi \partial^\mu \phi - m^2 \phi^2]$.

correct procedure requires ϕ to be an isospin doublet of complex scalar fields:

$$\phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} \phi_1 + i\phi_2 \\ \phi_3 + i\phi_4 \end{pmatrix}$$

where the + and 0 refer to electrically charged and electrically neutral, and the requirement of $SU(2) \times U(1)$ local gauge invariance $\phi \rightarrow e^{i\alpha^a \sigma^a/2} e^{i\beta/2} \phi$. The gauge-invariant Lagrangian for ϕ is analogous⁹ to Equation (2.8) except that, as introduced in Section 2.1.2, the derivatives are replaced by covariant derivatives in order to maintain the local symmetry:

$$\mathcal{L}_H = (D^\mu \phi)^\dagger (D_\mu \phi) - U(\phi) \quad (2.10)$$

$$U(\phi) = -\mu^2 \phi^\dagger \phi + \lambda (\phi^\dagger \phi)^2,$$

where D_μ is the covariant derivative for $SU(2) \times U(1)$ as defined in Equation (2.7). Similar to the globally invariant example above, the vacuum state for this potential is degenerate in the four-dimensional space of scalar fields. Corresponding to the three Goldstone bosons, there are three real parameters that specify an element of $SU(2)$; we can use two of these conditions to select the gauge where $\phi^+ = 0$ and an additional condition to fix ϕ^0 to be real, which results in the absorption of the Goldstone bosons into the longitudinal components of the three massive bosons, while the remaining degree of freedom corresponds to the massive scalar Higgs boson. Specifically, after SSB the ground state becomes

$$\phi_{\text{ground}} = \begin{pmatrix} 0 \\ v \end{pmatrix}$$

where $v = \mu/\sqrt{\lambda}$, with excited states defined with respect to this ground state in terms of a

⁹The form of $U(\phi)$ may seem arbitrary; it turns out to be the most general renormalizable form [12].

real field $h(x)$:

$$\phi = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + h(x) \end{pmatrix}. \quad (2.11)$$

After expressing Equation (2.10) in terms of Equation (2.11), a term emerges associated with the massive neutral scalar boson field $h(x)$ corresponding to $m_H = \sqrt{2}\mu = \sqrt{2\lambda}v$. Each of the three generators of $SU(2)$ correspond to a massless scalar Goldstone boson, which is absorbed into the longitudinal components of the final massive bosons that emerge from evaluating the $(D^\mu \phi)^\dagger (D_\mu \phi)$ term. The $W^\pm$ bosons are defined as $W_\mu^\pm \equiv 1/\sqrt{2}(W_\mu^1 \mp iW_\mu^2)$ with $m_W = gv/2$. The Z boson is defined as $Z_\mu = \frac{1}{\sqrt{g'^2 + g^2}}(gA_\mu^3 - g'B_\mu)$ with $m_Z = \sqrt{g'^2 + g^2}v/2$. The photon is defined as being orthogonal to Z_μ^0 because $A_\mu = \frac{1}{\sqrt{g'^2 + g^2}}(gA_\mu^3 + g'B_\mu)$. The A_μ vector field has $M_A = 0$ and is a reflection of a remaining unbroken $U(1)$ symmetry. Note that the $U(1)$ symmetry that is respected is not the same $U(1)$ symmetry that was present before SSB because the A_μ field is a mixture of fields from the spontaneously broken $SU(2) \times U(1)$ symmetry. We say that $SU(2) \times U(1)_Y$ has been broken to $U(1)_{EM}$, where Y refers to the conserved charge corresponding to the generator of the $U(1)$ gauge field before SSB (called *weak hypercharge*), and EM refers to the familiar conserved charge of electromagnetism, corresponding to the combination of generators $Q = Y/2 + T_3$, where T_3 is the third component of the conserved charge corresponding to the generator of the $SU(2)$ gauge field before SSB (called *weak isospin*).

The angle that describes the mixing between the (A_3, B) bosons that gives rise to the physical Z and γ bosons is called the weak mixing angle:

$$\begin{pmatrix} Z^0 \\ A \end{pmatrix} = \begin{pmatrix} \cos \theta_W & -\sin \theta_W \\ \sin \theta_W & \cos \theta_W \end{pmatrix} \begin{pmatrix} A^3 \\ B \end{pmatrix}$$

As discussed in Section 2.1.5, the observed fermion masses cannot arise through terms such as $-m_f \bar{\psi}_R \psi_L$ because such terms are not gauge invariant. However, coupling right- and

left-handed fermions is possible through the scalar Higgs doublet ϕ :

$$\mathcal{L}_{f1} = -\lambda_e \bar{E}_L \phi e_R - \lambda_d \bar{Q}_L \phi d_R - \lambda_u \epsilon^{ab} \bar{Q}_{La} \phi_b^\dagger u_R + \text{h.c.}$$

After the Higgs acquires its vacuum expectation value,

$$\mathcal{L}_{f1} = -\frac{\lambda_e}{\sqrt{2}} v \bar{e}_L e_R - \frac{\lambda_d}{\sqrt{2}} v \bar{d}_L d_R - \frac{\lambda_u}{\sqrt{2}} \bar{u}_L u_R + \text{h.c.},$$

corresponding to $m_X = \frac{\lambda_X}{\sqrt{2}} v$ for $X = d, u, e$. The procedure follows similarly for the second and third generations. The λ s are the Yukawa couplings, which determine each fermion's mass and must be determined experimentally.

Transitions between quark generations occur through flavor-changing interactions and are due to the mismatch between the mass and weak eigenstates. The unitary matrix describing the strength of these flavor-changing weak decays is called the Cabibbo-Kobayashi-Maskawa (CKM) matrix, which is approximately diagonal. The diagonal terms describe the quark transitions between up- and down-type quarks of the same generation, and the off-diagonal terms describe the transitions between generations. The CKM matrix must also be determined experimentally.

The Higgs boson was discovered at CERN in 2012 and so far appears to be consistent with SM predictions [14].

2.1.7 Perturbation theory

The SM includes several exactly conserved quantities: linear and angular momenta, energy, electric charge, weak isospin, and color. Charge-parity-time reversal is also a symmetry of the theory. If the particle type and initial momenta of a decaying particle are known, whether a given decay is allowed can be determined by conservation laws. If the couplings between initial, intermediate, and daughter particles are known, the rate for a given decay process can also in principle be calculated. Similar considerations apply to

particles interacting during a collision; the relative frequency of observing an allowed final state depends on the coupling between initial, intermediate, and final state particles, and in principle it can be calculated given the couplings described by the Lagrangian.

The Klein-Gordon and Dirac equations describe the motion of free fields and can be solved exactly; we can find their exact eigenvalues and eigenvectors. Unfortunately, when theories with interactions are considered, the picture becomes more complicated. To date, no exact solutions are known for quantum field theories with interactions in more than two spacetime dimensions [15]. Consequently the current approach is to treat the interaction Lagrangian as a perturbation of the free fields and calculate some finite number of terms in the expansion; if the coupling constant is sufficiently small, the approximation will be accurate enough to be useful.

Of interest are the interactions that take place during a high energy collision, involving the evolution of some initial state (e.g., a quark from each proton in a proton-proton collision) to some final state of measured outgoing particles. The initial and final state particles are assumed to be free fields, and the perturbative calculation describes how the coupling between those fields results in a probability that a given initial state transforms into a given final state. The observable that is calculated which describes the probability of such an occurrence is called the *cross section*. Given two colliding beams of given particle density and speed, the calculated cross section allows the expected frequency of occurrences of a given final state to be determined and thus allows a comparison of the theory to the experiment.

2.2 Physics beyond the Standard Model

Evidence for NP, which would extend the SM, would be found if some couplings were observed (via their effect on cross sections and relative frequencies of observing different final state particles) to deviate from SM expectations. These interactions could be mediated by new particles that have not yet been observed.

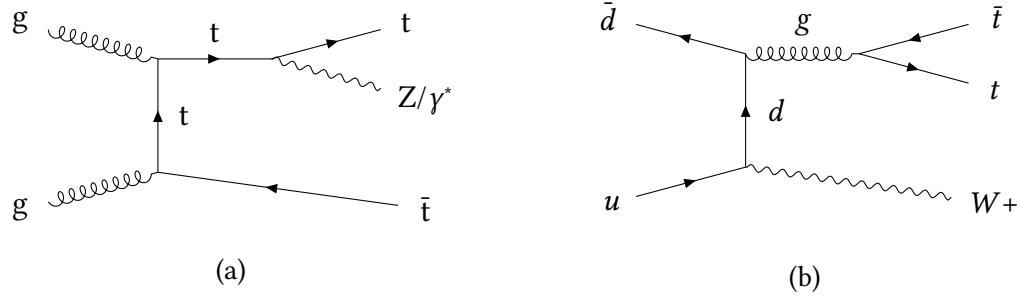


Figure 2.3. Dominant leading order Feynman diagrams for the (a) $t\bar{t}Z$ and (b) $t\bar{t}W$ processes at the LHC. For (b), the charge conjugate process is implied.

2.2.1 Deficiencies in the Standard Model

The SM has been very successful. It predicted the existence of the Higgs boson, W and Z bosons, gluon, and the top and charm quarks before they were confirmed experimentally to exist. A precision test of the SM involves evaluating the agreement between independent measurements of the fine structure constant α . The most precise measurement to date, with a precision of more than one part per billion, was made using a Penning trap [16]. This measurement, along with atomic-recoil measurements [17], has validated the SM to within 10 parts per billion. Despite its great success as a theory, the SM fails to answer various open questions.

Neutrino masses The SM was formulated assuming massless neutrinos, but observed neutrino oscillations imply that they are massive.

Hierarchy problem The SM has no explanation for what on its face is a surprising discrepancy: the weak force is more than 20 orders of magnitude stronger than gravity. Alternatively, the SM has no explanation for why the Higgs boson mass is so small in the face of QM corrections at much higher energy scales.

Matter-antimatter asymmetry The universe was extremely hot right after the Big Bang, with a lot of energy for producing particle-antiparticle pairs. We would naively expect that matter and antimatter would be produced in roughly the same proportion and then annihilate, leaving a radiation-dominated universe, but the universe instead appears to be composed almost entirely of matter rather than antimatter. It turns out that slightly more matter was produced than antimatter. The SM only accounts for only a small part of this asymmetry.

Dark energy Current cosmological estimates conclude that ordinary matter of the type

described by the SM makes up only $\sim 5\%$ of the total mass-energy content of the universe, while $\sim 68\%$ is made up of dark energy [18]. Dark energy is an unknown type of energy associated with the vacuum of space, and it is responsible for the observed acceleration of the expansion of the universe. Attempts to use the SM to calculate the vacuum energy fail spectacularly, because the predicted vacuum energy is over 100 orders of magnitude too large.

Dark matter About $\sim 27\%$ of the mass-energy content of the universe is composed of dark matter. Dark matter is a hypothetical type of matter that interacts with ordinary matter so weakly that it has not yet been observed directly, despite its gravitational influence having been indirectly observed in a wide range of astronomical and cosmological data. The SM neutrinos are similar to dark matter, but they are not massive enough to constitute a significant fraction of dark matter.

Gravity The unification of the electric and magnetic forces, and more recently the weak force, within a single theoretical construct has been a major accomplishment in the evolution of physics, but it remains incomplete. Although gravity is sometimes considered too weak to be relevant in experimental physics, it governs much of the state of the universe. Gravity is well described under relativity; however, it is not predicted by the SM, and a unified theory that encompasses both quantum phenomena and gravity remains elusive.

Clearly there must be NP. Since such phenomena have not been observed experimentally, they are presumably beyond the energy scale of experiments to date. In the next section, a strategy is introduced that may enable the search for such physics despite the limitation of currently available energy scales.

2.2.2 New physics in the top sector

This thesis presents NP interpretations for two SM analyses, each cross section measurements for a top quark pair produced in association with a W or Z boson ($t\bar{t}W$ and $t\bar{t}Z$, see Figure 2.3). The top quark was discovered [19, 20] at Fermilab in 1995 by the CDF and DØ collaborations, and it is the heaviest fundamental particle discovered to date. Its discovery was expected since the discovery of the bottom quark, because it would complete the third generation of fermions. At the LHC, top quarks are produced predominantly through gluon fusion, $gg \rightarrow t\bar{t}$, and quark-antiquark annihilation, $q\bar{q} \rightarrow t\bar{t}$. At $\sqrt{s} = 8$ TeV, about 85 % of $t\bar{t}$ production is via gluon fusion.

Various theoretical motivations justify paying special attention to processes involving top quarks. Because of its large mass, the top quark has a very short lifetime of around 5×10^{-25} s. Consequently, it decays before it has time to interact with any other particles, and it does not form bound states as other quarks do; this is our only opportunity to study a “bare” quark.

Again due to its mass, the top quark may play an important role in the electroweak sector. The Yukawa coupling of the Higgs to the top quark is ≈ 1 , which is 40 times larger than the next largest Yukawa coupling. This suggests that the top quark may have a special relationship to the Higgs and hints at NP. One of the most promising strategies to detect NP is by measuring the top quark coupling to gauge bosons and looking for deviations from the SM predictions. Moreover, by studying the interaction between the top quark and the Higgs, we may learn more about possible underlying principles that dictate the pattern of particle masses. If new particles couple to the Higgs field (or fields), the top quark would likely be the most affected by the new interactions because, having the largest mass, it has the largest coupling to the Higgs. Measuring the top-Higgs coupling via $H \rightarrow t\bar{t}$ is not possible because the Higgs is lighter than a pair of top quarks, so it cannot decay to them. The top-Higgs coupling can be constrained indirectly through measurements involving gluon fusion production [21] or by study of the decay of the Higgs to photons. Such processes involve fermion loops, which are dominated by contributions from the top quark (see Figure 2.4). While many extensions of the SM involve new particles that could contribute to fermion loops, indirect measurement of the top-Higgs coupling relies on the assumption of a strictly SM scenario. In other words, if the measured process contains loops, and if NP particles exist, they could contribute to loop diagrams with the same initial and final states: it is impossible for such measurements to distinguish between the true top couplings and possible NP contributions. In contrast, the rate at which the Higgs is produced in association with a top-quark pair ($t\bar{t}H$) provides a direct measurement of the top-Higgs couplings because it is a tree-level process. Similarly, the top-Z coupling has been measured [22] indirectly via the decay $Z \rightarrow b\bar{b}$, which has

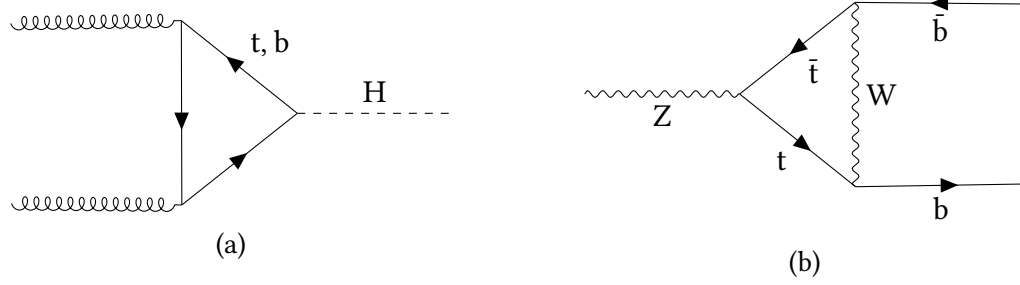


Figure 2.4. The gluon fusion production mode of the Higgs boson, which depends on the top-Higgs coupling, is shown in (a). The decay $Z \rightarrow b\bar{b}$, which depends on the top-Z coupling, is shown in (b).

higher-order corrections involving a top loop, while the rate of $t\bar{t}Z$ production provides a direct measurement of coupling between the top quark and the Z boson.

I contributed to a search for $t\bar{t}H$ [23], which studied events with three different signatures for the Higgs boson decay: $H \rightarrow \text{hadrons}$, $H \rightarrow \text{leptons}$, and $H \rightarrow \text{photons}$ using 5.1 fb^{-1} of 7 TeV and 19.7 fb^{-1} of 8 TeV proton-proton collision data recorded at CMS. The combination of these analyses measured the $t\bar{t}H$ cross section to be 2.8 ± 1.0 times the expected SM value. The $t\bar{t}W$ and $t\bar{t}Z$ measurements described in Chapter 5 build on many of the background estimation approaches developed for the $t\bar{t}H$ analysis. More recently, a search for $t\bar{t}H$ in final states with electrons, muons, or hadronically decaying τ leptons was conducted using 35.9 fb^{-1} of 13 TeV proton-proton collision data recorded at CMS. The cross section was measured to be $1.23^{+0.45}_{-0.43}$ times the expected SM value, with an observed (expected) significance of 3.2σ (2.8σ) [24, 25].

A search for $t\bar{t}Z$ was carried out by the CMS collaboration at $\sqrt{s} = 7 \text{ TeV}$, which measured the cross section with a precision of $\approx 50 \%$ [26]. With 8 TeV data, CMS [27] and ATLAS [28] measured the $t\bar{t}W$ and $t\bar{t}Z$ processes with 3σ significance.

2.2.3 Effective field theory

Various NP theories have been proposed that involve extra dimensions, supersymmetry, and other hypothesized phenomena. Each of these would have different experimental signatures, and dedicated searches are needed for each possibility. Moreover, no assurance exists that any of the NP theories proposed so far are correct, and there is no indication that any specific theory should be strongly preferred. Striking experimental signatures of NP in the form of heavy resonances are looking increasingly less likely as the LHC fails to observe “bumps” in invariant mass distributions. Given an overwhelming landscape of less-striking experimental signatures associated with a host of possible NP models, a more general approach for searching for NP without testing each model one by one is desirable. One approach is to use Effective Field Theory (EFT), which allows for the expression of NP in a model-independent way.

To understand EFT, it is helpful to first consider effective theories in general. An effective theory is a theory that describes a certain set of observations, without claiming that the mechanism of the theory is the actual cause of the observed effects. In other words, a more fundamental underlying theory is expected to exist. When we interpret things around us as made of solid matter, we are using an effective theory. Ordinary matter is mostly empty space, but if we are building a table, it is not necessary to take that into account as we swing a hammer. Newtonian mechanics works extremely well, but it is an approximation for low speeds and large objects. Relativity and quantum mechanics are more fundamental underlying theories, but using relativity and quantum mechanics to calculate the trajectory of a cannon ball would be wasted effort. Effective theories take advantage of scale separation; they are useful tools as long as the limitations (i.e., the conditions under which the theory can no longer provide an accurate description) are understood.

An EFT is a low-energy approximation of a higher energy QFT. Particles with masses beyond what is currently experimentally accessible cannot be seen directly, but with EFT their indirect effects can be quantified. Consider a search for a particle that has a mass greater

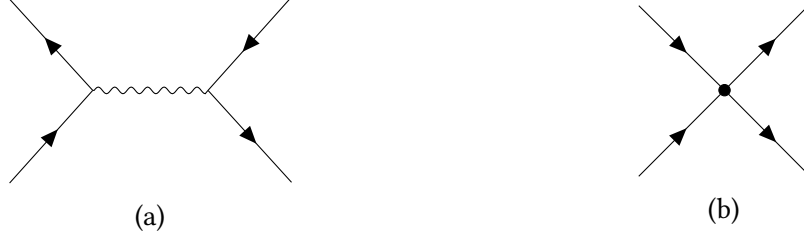


Figure 2.5. Representative Feynman diagrams: in the low-energy limit, the exchange of a massive vector boson (a) reduces to a four-point fermion interaction (b).

than what current collider energies can produce. From the uncertainty principle, we know that when the mass of a particle mediating a given force is larger, the range of that force will be shorter. In EFT, the mass of a given mediator is taken to be so large that the propagator connecting separate interaction vertices is reduced to a point, resulting in a simplified interaction vertex called an *effective operator*. Another effect is the reduction of an almost infinite class of possible theories down to a much smaller set of EFT operators since many different theories whose Feynman diagrams have different internal propagator arrangements can be reduced to the same point interaction vertex when the previously described procedure is performed.

An example of this procedure is the Fermi theory of electroweak interactions. In the SM, muon decay is mediated by a W boson, where the decay amplitude is [29]:

$$i\mathcal{M} = \left(-i\frac{g}{\sqrt{2}}\right)^2 \left[\bar{u}(v_u)\gamma^u \frac{1-\gamma^5}{2} u(\mu) \right] \frac{-i}{q^2 - m_W^2 + i\epsilon} \left(\eta_{\mu\nu} - \frac{q_\mu q_\nu}{m_W^2} \right) \left[\bar{u}(v_u)\gamma^u \frac{1-\gamma^5}{2} u(\mu) \right]$$

If we consider the low-energy limit where $q \ll m_W$, then the propagator can be expanded as

$$\frac{-i}{q^2 - m_W^2 + i\epsilon} \left(\eta_{\mu\nu} - \frac{q_\mu q_\nu}{m_W^2} \right) = i \frac{\eta_{\mu\nu}}{m_W^2} + \mathcal{O}\left(\frac{1}{m_W^4}\right)$$

This expansion corresponds to a four-fermion contact interaction with a strength that is suppressed by the W boson mass (see Figure 2.5), which is valid for energies below the W boson mass.

In this example, we know the exact theory, but obtaining a simpler theory is useful when considering low-energy interactions. In searching for NP, where the exact theory is unknown, we must go the other way by starting with the low-energy theory. All unique products of fields (i.e., *operators*) in the SM are dimension four¹⁰. To extend the SM, we add higher-dimension operators by combining the SM fields in all possible combinations that have the correct dimensionality and do not break any of the symmetries we believe must be respected. The most general effective Lagrangian can be written as:

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

where $\mathcal{L}_{\text{SM}}$ is the dimension-four SM Lagrangian, Λ is mass scale of the NP, and $\mathcal{O}_i^{(d)}$ are the i effective operators of dimension d . There is always a finite number of operators at a fixed dimension. The Wilson coefficients c_i parameterize the strength of the NP interaction. This infinite sum over operators with arbitrarily high dimension may seem unsettling, but note that at energies $E \ll \Lambda$ an operator's contribution will scale inversely with dimension, as higher-dimensional operators are suppressed by powers of Λ . It is therefore reasonable to truncate the series at the desired level of accuracy. There is only one gauge-invariant dimension five operator. It is a possible model for the generation of neutrino masses [30], but it is irrelevant in the present work. This work is restricted to dimension-six operators. There are 59 independent dimension-six operators [31]. This work follows the basis choice and notation from [32] and only examines the smaller set of 39 operators presented there, which are chosen because they have an effect on Higgs physics.

EFT has a number of desirable features [33]:

1. EFT respects unitarity (the sum of probabilities of all possible event outcomes should equal 1).
2. EFT respects SM symmetries.

¹⁰Here dimensionality refers to the mass dimension. Because natural units are being used in this work, $[m] = [E] = [p] = [x^{-1}] = [t^{-1}]$. The Lagrangian has dimension four, so $[\phi] = [A_\mu] = 1$, $[\psi] = 3/2$, and $[F_{\mu\nu}] = 2$.

3. EFT reduces to the SM in the low-energy limit.
4. EFT should be general enough to describe any NP.
5. EFT allows for the calculation of radiative corrections at any order in both the SM and NP interactions.

Note that EFT is a useful approach if the NP we are searching for has not been found yet because it is too heavy to be produced directly at current experimental energies. Conversely, EFT will not usefully describe NP which is too weakly coupled to detect.

CHAPTER 3

EXPERIMENTAL APPARATUS

The central goal of high energy physics is the study of new particles and interactions. Because of Einstein's mass-energy relation, heavy particles can be created through interactions with sufficient energy. If a projectile strikes a fixed target, the center-of-mass frame of the collision will fly down the beamline relative to the lab frame; much of the energy of the projectile is transferred to kinetic energy rather than into mass-energy. In a particle collider, two beams serve as target and projectile simultaneously. Because the momentum of the colliding particles is equal and opposite, for colliding point particles no energy is lost in the movement of the center of mass. Therefore, particle colliders maximize the potential for converting kinetic energy into mass-energy in particle interactions.

A proton is a composite particle made up of partons (valence quarks, virtual or *sea* quarks, and gluons), each carrying some fraction of the proton momentum. High-energy proton-proton (pp) collisions result in collisions between partons. While the momentum of each proton in the collision is equal and opposite, the momentum of each parton is not, with the result that some energy is carried off in the movement of the center of mass. While it is unfortunate that less than 100 % of each proton's energy can be used for making mass, particle colliders afford the best chance of producing heavy new particles and more generally studying particle interactions at the highest possible energies.

The LHC uses this approach to produce the most energetic collisions currently available, 13 TeV for pp collisions and 1 PeV (corresponding to an average of about 5 TeV per nucleon) for lead ion collisions. The collision energy in pp mode corresponds to a velocity about 3.1 m s^{-1} slower than the speed of light. The LHC delivers collisions to four main detectors lo-

cated along its circumference. ALICE examines collisions involving heavy ions, which generate extremely high temperatures at which matter transitions to a quark-gluon plasma, while the other detectors primarily study interactions between protons. LHCb examines heavy flavor physics and CP violation. ATLAS and CMS, which collect both proton-proton and heavy ion data, were designed to accomplish a broad physics program including the search for the Higgs boson, precision measurements to test the SM, and the search for NP beyond the SM.

The datasets used in the analyses presented in this thesis were collected using proton collisions generated by the LHC and recorded by CMS. An overview of these systems is presented in this chapter.

3.1 The Large Hadron Collider

The following description of the LHC is based on [36]. The LHC was completed in 2008, at a cost of about \$4 billion (2008 USD), by the European Organization for Nuclear Research,

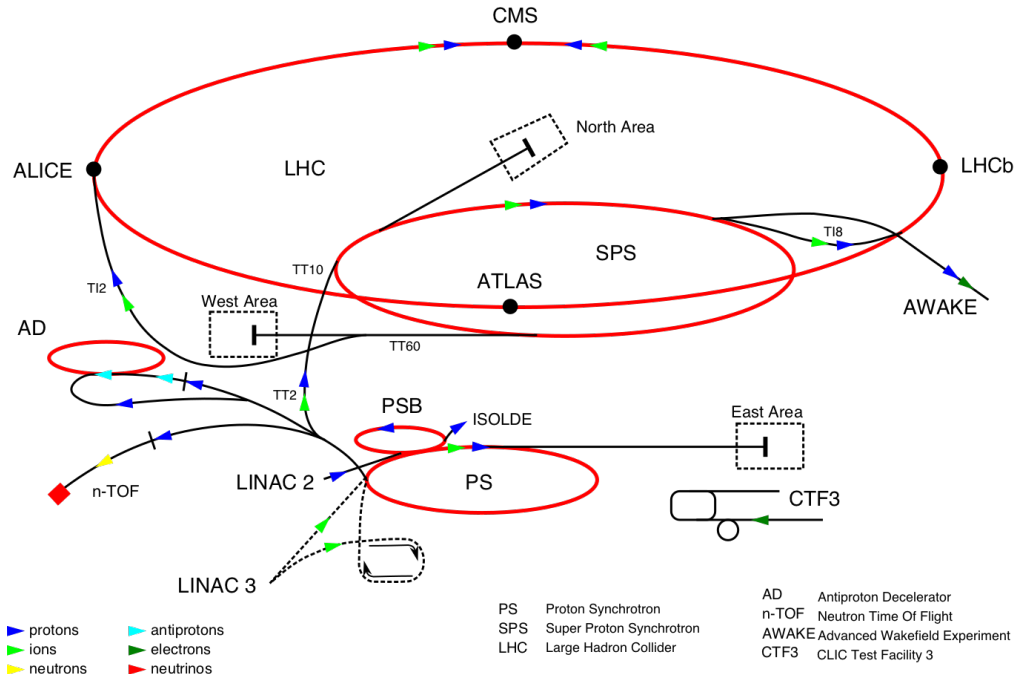


Figure 3.1. Overview of the LHC accelerator complex [34].

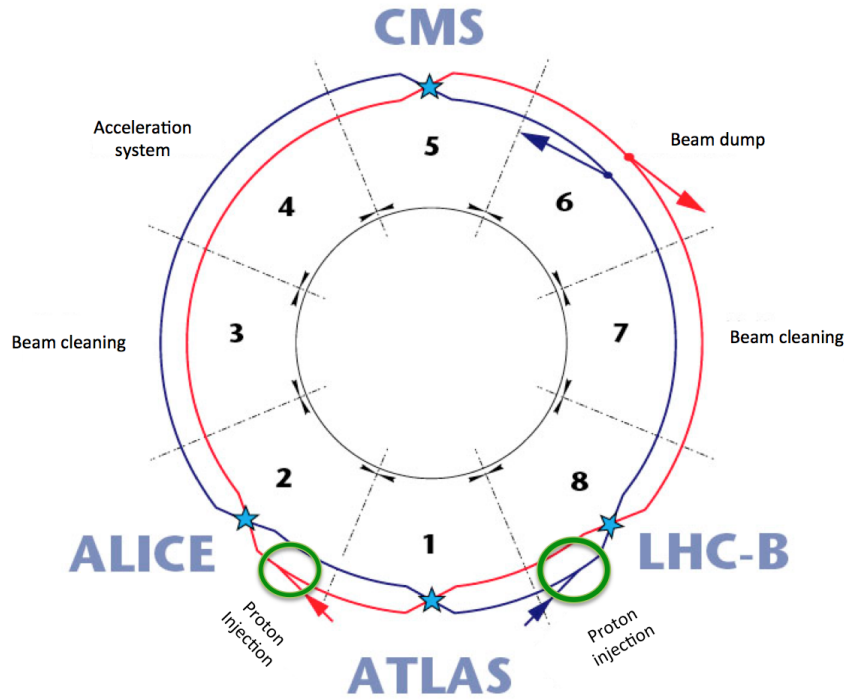


Figure 3.2. Schematic layout of the LHC sectors [35].

which is known by the acronym CERN (based on its original name, Conseil Européen pour la Recherche Nucléaire).

An overview of the CERN accelerator complex is provided in Figure 3.2. The LHC was constructed in a 26.7 km tunnel used in the 1980s for the Large Electron Positron (LEP) experiment. This tunnel is roughly circular, composed of eight arcs and eight straight sections between 45–170 m below ground level. The particle beams intersect at four interaction points (IPs) located in four of the straight sections. The protons are maintained in their trajectory around the ring via 1 232 dipole magnets, each 14.3 m long with a field strength of 8.3 T. The superconducting windings are copper-clad niobium-titanium and carry a current of 10 980 A [37]. There are also 392 quadrupole magnets (between 5–7 m) that focus the beam and keep it stable, while 6 400 smaller magnets perform higher-order corrections. The

total energy stored in all of the magnets is about 10 GJ. The magnets are cooled with 96 t of helium II, which is helium at a temperature and pressure such that it displays superfluid properties, namely zero viscosity. The low temperature, only 1.9 K, allows the conductors to remain superconducting at higher current densities and magnetic fields than would be possible with helium I, and the superfluid state of helium II allows it to remove heat about 10 times as efficiently [38].

LEP was a particle-antiparticle (electron-positron) collider, in which the particles move in opposite directions but also had opposite charge; consequently, both beams could share the same ring of magnets. The LHC is a particle-particle (proton-proton) collider, however, and it requires two closely spaced rings with opposite magnetic dipole fields. A twin-bore design was adopted for the magnets to allow the two rings to fit within the existing 3.7 m diameter tunnel.

Protons are injected into the LHC using a series of accelerators recycled from previous CERN experiments. The proton beam begins as hydrogen gas, which is ionized by passing it through a charged screen and then accelerated to 50 MeV in the LINAC2 linear accelerator. The protons then enter the Proton Synchrotron Booster (PSB), a set of four rings of 25 m radius where they are accelerated to 1.4 GeV and grouped into bunches. The four beams from the PSB are merged and enter the 100 m-radius Proton Synchrotron (PS). The PS was CERN's first synchrotron, brought online in 1959. The PS accelerates the protons to an energy of 25 GeV before they are injected into the Super Proton Synchrotron (SPS)¹. The SPS has a radius of over a kilometer and accelerates protons to 450 GeV. Protons are diverted from the SPS into two separate 2.5 km transfer tunnels that inject protons from opposite directions into the clockwise and counterclockwise rings of the LHC. When both beams are filled, the injection stops and the beams are pumped up to their peak energy. A beam *squeeze* is then performed in which the beams are focused, and finally, the beam paths are brought together

¹The SPS began operation in 1979, and the discovery of the W and Z bosons and CP violation were all made there.

at the IP and collisions begin.

Running at 6.5 TeV per proton with an intensity of 2.5×10^{14} protons per beam, the total energy² in the beam is about 260 MJ. The destructive power of the beam necessitates a beam dump system³. This system safely deposits this energy in the case of any of a wide range of possible failures in the LHC (such as a magnet *quench* in which a magnet becomes non-superconducting) or at the end of a physics run, after the luminosity of the beam has degraded below usable levels. Each of the two LHC beams has a dump system that includes a set of fast kicker magnets to deflect the beam and a septum magnet where the beam dump channel branches from the circulating channel. The dumped beam is then directed into an 8 m-long cylinder of graphite composite that is encased in concrete. A dilution magnet spreads the beam to minimize overheating; even under normal operation, parts of the graphite can reach 700 °C.

Protons circulating in the LHC are divided into a train of regularly spaced bunches, each containing $\mathcal{O}(10^{11})$ protons. The particles are accelerated as they pass through RF cavities; an electric field is induced inside each cavity, which oscillates with a frequency that is an integer multiple of the particle revolution frequency and is timed to apply an accelerating force on each bunch of particles. During ramp up, as particles travel around the ring, they receive an energy kick each time they pass through a resonator, until they reach the target energy. At the LHC, the RF cavities operate at 400 MHz which corresponds to a ratio⁴ between the RF and particle revolution frequencies of 35 570. This value is the number of points, or *buckets*, circulating around each beam line, each of which can carry a bunch of particles.

Not all the buckets contain bunches, however. The kicker magnets of the beam dump

²This is approximately equivalent to the kinetic energy of 750 Honda Civics each moving at 55 MPH, or the energy in 319 chocolate bars.

³Some researchers have proposed using the beam dump system itself to do physics such as dark matter searches by placing a detector near the beam dump; see for example [39].

⁴The protons are moving at approximately c , and the LHC has a circumference of 26 659 m, so $f_{\text{RF}}/f_{\text{rev}} \approx 400 \times 10^6 / (c/26\,659\text{ m}) = 35\,570$.

system require $3\text{ }\mu\text{s}$ to reach full strength, and a partially diverted beam would damage sensitive downstream accelerator components. Consequently the beam contains a $3\text{ }\mu\text{s}$ *abort gap* of unfilled buckets; in the event of an abort the kicker magnets are initiated at the start of the abort gap and reach full strength before the next bunch of protons arrives.

The LHC was first brought online in 2008. Soon afterward, an electrical arc damaged the cooling system, which resulted in a large helium leak [40]. Over 50 magnets and additional components of the cryogenic system were damaged. Initial operation was delayed by a year for repairs as well as to carry out a detailed investigation and implement preventative measures. In 2010 the LHC began operations at a record-breaking center-of-mass energy ($\sqrt{s}$) of 7 TeV, exceeding the previous record of $\sqrt{s}=1.96\text{ TeV}$ set by the Tevatron, which ceased operations in 2011. In April 2012 the beam energy was increased to 8 TeV and remained at that level for the duration of 2012. This period is referred to as Run 1. After Run 1, the LHC began the first “long shutdown” (LS1) to make necessary upgrades for raising the center-of-mass energy. In June 2015, physics data collection resumed at $\sqrt{s}=13\text{ TeV}$. This point was the beginning of the data-taking period referred to as Run 2, which is scheduled to conclude in 2018. This thesis presents analyses based on both datasets.

Charged particles undergoing acceleration emit bremsstrahlung radiation. For circular motion, the radiated power due to centripetal acceleration is proportional to $1/m^4$. A significant advantage of proton-proton colliders over electron-positron colliders like LEP is the fact that, because of the higher proton mass, less energy is lost to bremsstrahlung radiation. However, in contrast to an electron-positron collider like LEP, where all of the energy in a collision is available to produce final state particles, the momentum in a proton is divided between its constituent partons. The momentum carried by a given parton follows a statistical distribution. So while less energy is lost to bremsstrahlung radiation in a proton-proton collider, allowing it to produce higher energy projectiles, only some collisions are produced near the proton-proton center-of-mass energy provided by the accelerator.

In a search for rare events, a critical metric is the rate at which the interaction of interest

occurs. This rate is proportional to the luminosity, which provides a measure of the rate at which particles pass through the interaction volume, and the cross section, which is the probability that each particle will experience the interaction of interest. The dimensions of luminosity are $(\text{area})^{-2}\text{s}^{-1}$. If Gaussian beams collide head-on, the instantaneous luminosity is [41]:

$$\mathcal{L} = \frac{N_b^2 n_b f \gamma}{4\pi \sigma_{xy}} \quad (3.1)$$

where N_b is the number of particles per bunch, n_b is the number of bunches per beam, f is the revolution frequency, γ is the relativistic gamma factor, and σ_{xy} is the transverse beam size.

Real machines such as the LHC have additional complications. Collisions are not head-on, but occur at a small but non-zero crossing angle. This circumstance introduces a factor S in the luminosity equation, which for small crossing angles and bunch length $\sigma_z \gg \sigma_{xy}$ is approximated as

$$S = \frac{1}{\sqrt{1 + \left(\frac{\theta_z \phi}{\theta_x 2}\right)^2}}.$$

Additionally Equation (3.1) assumes that the entire collision region has a constant beam size. In fact the transverse beam size is a function of the distance along the beam trajectory. To maximize the number of expected collisions, the focusing fields are designed to minimize beam size at the IP. The beams spread out in both directions, creating an hourglass shape that is characterized by the amplitude function $\beta(z)$ and the beam emittance ϵ , where

$$\theta_{xy}(z) = \sqrt{\epsilon \cdot \beta(z)}.$$

An important metric describing the extent to which the beam is “squeezed” is the β^* value, which gives the distance from the IP at which the transverse beam size doubles. For Runs 1 and 2, β^* was typically 40–60 cm.

Although multiple proton-proton collisions (*events*) occur per crossing, there are far fewer interesting high- p_T scattering events. The majority of the collisions simply produce

sprays of hadrons along the beamline that complicate reconstruction of interesting events. The occurrence of multiple collisions per bunch crossing is referred to as *pileup*; mitigating the effects of pileup by filtering out particles from uninteresting collisions during the reconstruction process is essential to the detection of rare events.

In both runs, the number of protons per bunch is $\mathcal{O}(10^{11})$. In Run 1, beam stability issues limited the number of bunches per beam to 1 380 with a bunch crossing every 50 ns and an average of 21 proton-proton collisions per bunch crossing. During Run 2, most of the data were obtained with a bunch crossing every 25 ns and an average of 27 collisions per bunch crossing. The maximum instantaneous luminosity reached during Run 1 was 7.7 Hz/nb. In total, about 30 fb^{-1} of proton-proton data were delivered during Run 1. Integrated and instantaneous luminosity along with pileup are presented in Figure 3.3 for 2012 (representative of Run 1) versus 2016 (representative of Run 2).

3.2 The CMS detector

3.2.1 Overview

CMS (see Figure 3.4 and Figure 3.5) is located at IP 5 near the French town of Cessy, in a cavern about 100 m underground. The name Compact Muon Solenoid reflects the special attention given to muon detection and the solenoid coil that surrounds the detector elements with a strong magnetic field. It is roughly cylindrical, measuring about 21.6 m in length and 15 m in diameter and weighing about 14 000 t. CMS is less than half the length of ATLAS and a little over half its diameter, while being significantly more massive. This status may be the source of the term “compact”⁵. The CMS collaboration consists of approximately 3 800 people at about 199 institutions in 43 countries. The total cost for design and construction of the CMS was about \$510 million (2008 USD).

CMS is a general-purpose detector, designed to characterize a wide range of proton-

⁵Both detectors are “compact” in comparison to other larger detectors, like the Laser Interferometer Gravitational Wave Observatory, which has two observatories each with 4 km interferometer arms.

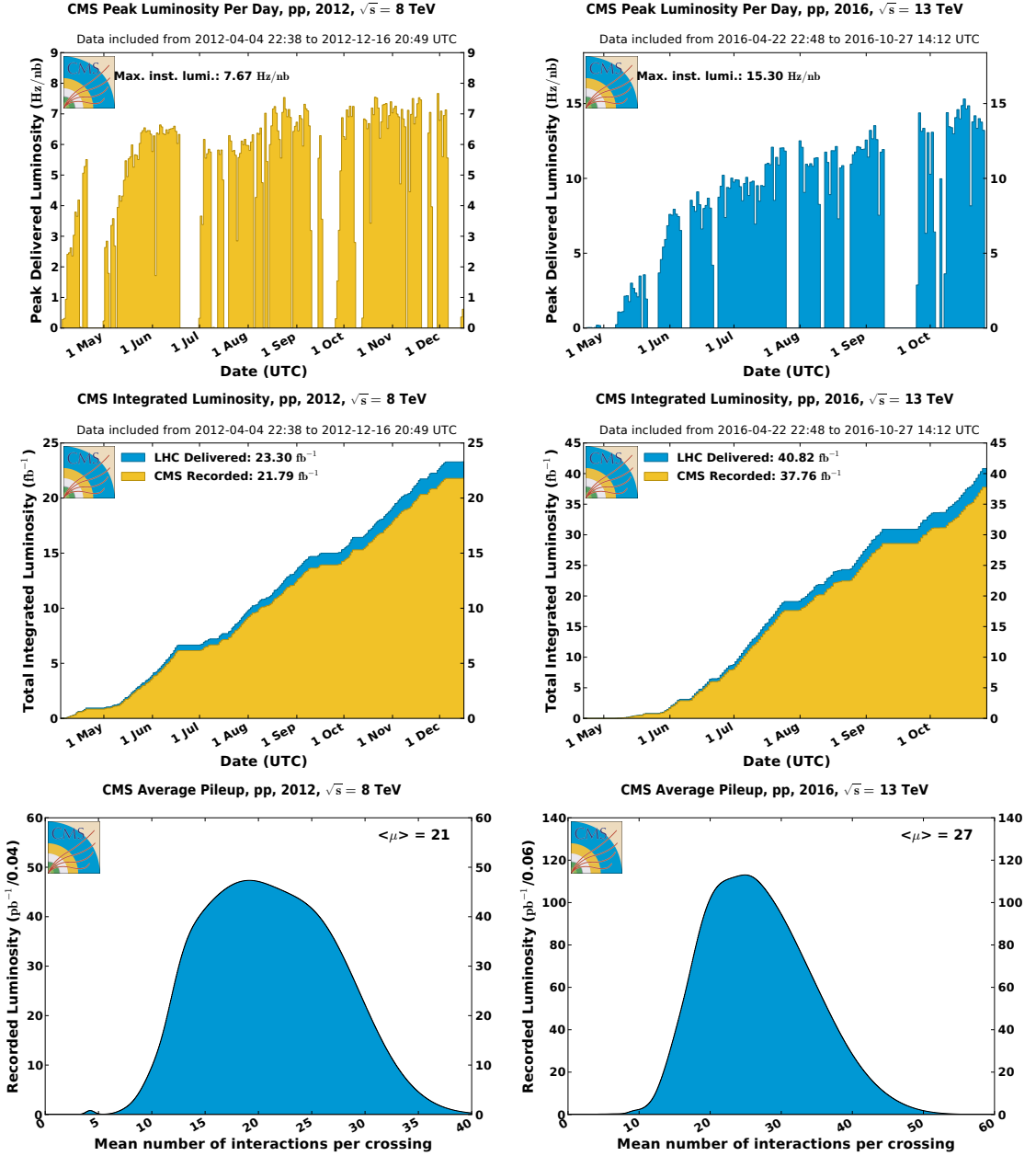


Figure 3.3. Maximum instantaneous luminosity (top), integrated luminosity (center), and mean interactions per crossing (bottom) at CMS for 2012 (left) and 2016 (right) [42].

proton scattering events. Consequently, the design had to be hermetic, in order to record the energy, charge, and momentum of as many scattering products as possible. The detector is made up of a cylindrical barrel and two endcaps with its longitudinal axis along the beamline.

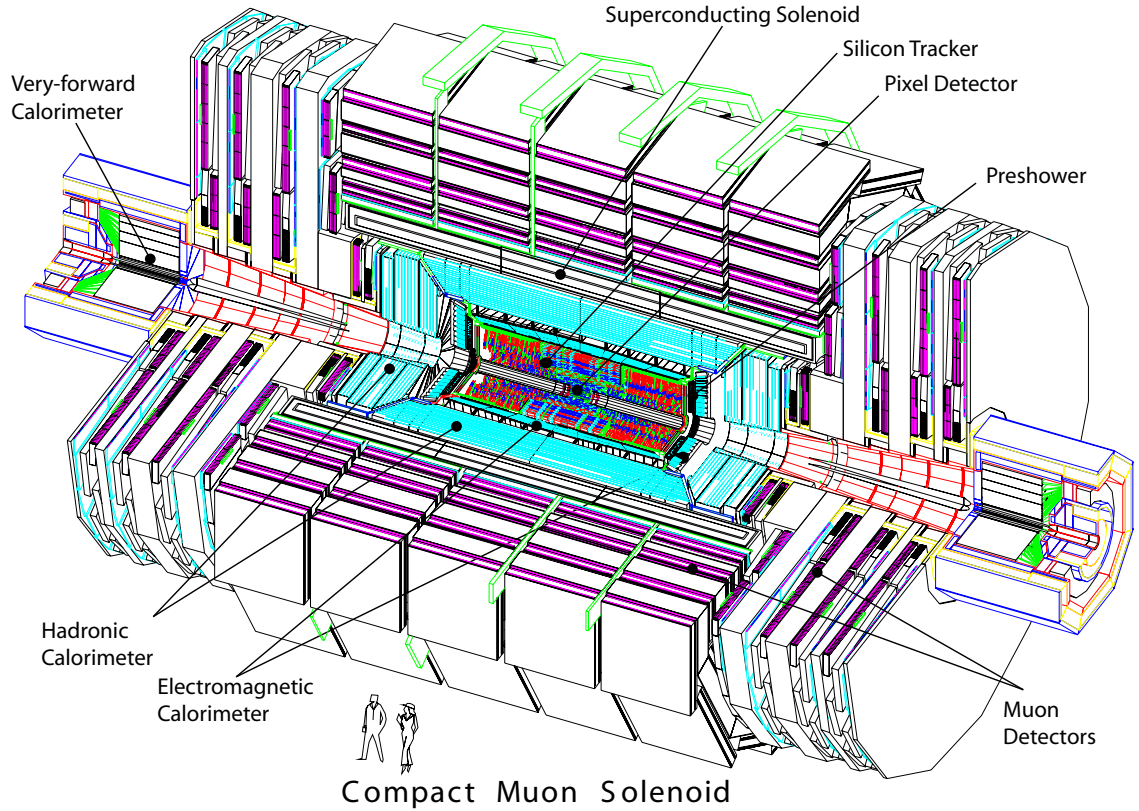


Figure 3.4. Overview of the CMS detector and its subcomponents [43].

Events produce a variety of final state particles that participate in different sets of fundamental interactions. For this reason, CMS is composed of multiple subdetectors, each optimized for measurements of specific sets of particles. A central feature of CMS is its solenoid magnet, which produces a strong and roughly constant magnetic field around the IP. Charged particles moving in a magnetic field will follow a curved path with a radius that is proportional to the momentum. A *tracker* detects the ionization left by the passage of a charged particle. If at least three points along the particle's path can be determined, it is possible to reconstruct the radius of its trajectory and calculate its momentum to charge ratio. Since leptons and stable hadrons have integer units of elementary charge, particle momentum and the sign of its assumed unit of charge (reflected in the sign of the track

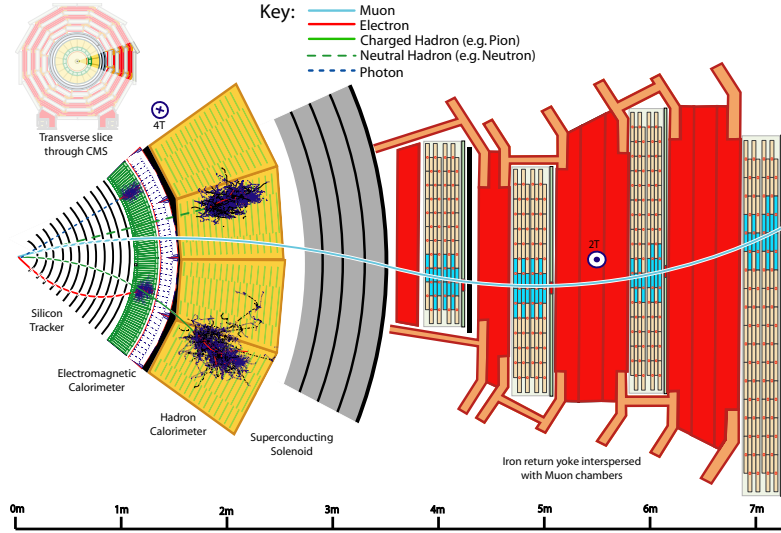


Figure 3.5. Transverse slice through the CMS detector. Particle traces are depicted to illustrate the function of the components [44].

curvature) can be reconstructed in practice.

A tracker alone is not sufficient to fully characterize events: neutral particles will pass through undetected, and particles with different masses can have the same momentum. To fully characterize a particle, only one unknown is permitted in the equation $E^2 = p^2 + m^2$. For these reasons, CMS includes *calorimeters* that measure particle energy. Because the calorimeters stop particles, the tracker must be the innermost detector; conversely, energy loss in the tracker must be minimized to allow particle energy to be measured accurately in the calorimeter.

Charged particles moving through matter are accelerated in the coulomb field of atomic nuclei and emit bremsstrahlung radiation. Power loss for muons is suppressed compared with electrons because of their higher mass, and as a result, muons are not easily stopped in the calorimeters. A dedicated muon tracking system is placed outside the calorimeters. Because other particles are stopped in the calorimeters, a sign-of-charge (sign of track curvature) and momentum measurement (radius of track curvature) are sufficient to fully

specify muons. One final type of particle, the neutrino, only interacts weakly and thus is invisible to both trackers and calorimeters. By making the detector as hermetic as possible, the presence of neutrinos or exotic neutral particles can be inferred using conservation of momentum (see Section 3.2.2).

The following sections describe the CMS components in more detail and are based on [45].

3.2.2 Coordinate system and conventions

The standard CMS coordinate system is used throughout this work, in which the origin is defined as the nominal interaction point. The x-axis points towards the center of the LHC ring, the y-axis points upwards towards the sky, and the z-axis points along the beamline such that the system is right-handed. Alternatively, polar coordinates are defined such that the distance from the beamline in the plane perpendicular to the z-axis is called r . The azimuthal angle ϕ is defined such that $\phi = 0$ points along the x-axis. The polar angle is defined such that $\theta = 0$ points along the z-axis.

Conservation of momentum requires that the total outgoing particle momenta equals the total incoming parton momenta. Because the momentum vector of the incoming protons is entirely along the beamline, the momentum in the plane transverse to the beamline must sum to zero, while the longitudinal component depends on the unknown fraction of the total proton momentum that each parton carried. Thus, longitudinal boosts vary from event to event in a distribution that depends on the parton distribution functions (PDFs). For this reason, we are usually only concerned with the transverse component of momentum p_T , from which the transverse energy E_T is defined for a particle of mass m :

$$p_T = |\vec{p}| \sin(\theta) = \sqrt{p_x^2 + p_y^2},$$

$$E_T = \sqrt{m^2 + p_T^2}.$$

Any deviation of the total transverse momentum from zero must have been carried away by

weakly interacting particles such as neutrinos. This quantity is called the *missing transverse momentum*: $p_T^{\text{miss}} = |-\sum_i \vec{p}_{T,i}|$, where the sum is over i visible particles.

The rapidity y and the pseudorapidity η are defined in terms of θ :

$$\begin{aligned} y &= \frac{1}{2} \ln \left(\frac{E + p_z}{E - p_z} \right) \\ \eta &= \frac{1}{2} \ln \left(\frac{|\vec{p}| + p_z}{|\vec{p}| - p_z} \right) = -\ln \tan \frac{\theta}{2}. \end{aligned} \quad (3.2)$$

The rapidity has the desirable feature that differences in rapidity are invariant under boosts along the beam axis. Characterizing angles between outgoing particles in terms of such a quantity allows for more convenient comparison of events with different boosts. In other words, a histogram showing the rapidity separation between two particles is meaningful despite being populated by events whose center-of-mass frames have arbitrary longitudinal boosts. The rapidity depends on both energy and momentum, which is difficult to measure at high rapidities for highly relativistic particles. For this reason, it is more common to work with the pseudorapidity, which is approximately equal to the rapidity in the limit $|\vec{p}| \gg m$ because $E^2 = |\vec{p}|^2 + m^2 \implies E \approx |\vec{p}|$. It follows from Equation (3.2) that $\eta = 0$ corresponds to the vector perpendicular to the beamline, and $\eta = \infty$ ($\eta = -\infty$) corresponds to a vector that is parallel (anti-parallel) to the beamline. As η gets closer to the beamline, equally spaced angular separations represent larger and larger changes in η . This reflects the intuitive fact that due to events being randomly boosted in the forward directions (along the beamline), the average flux of particles is much higher in the forward directions (at angles close to the beamline).

The angular separation in ϕ - η space that describes solid angles is called ΔR , and it is also approximately Lorentz invariant. ΔR is defined as

$$\Delta R = \sqrt{\Delta \eta^2 + \Delta \phi^2}.$$

3.2.3 Solenoidal magnet

The CMS magnet is a solenoid 6 m in diameter and 12.5 m in length, capable of producing a highly uniform magnetic field in its inner bore. It is large enough to accommodate the tracker and calorimeters, in order to minimize energy losses when traversing the magnet, which would interfere with their measurements. The standard operating field of 3.8 T requires a current of 18 160 A, corresponding to a stored energy of 2.3 GJ, which circulates in four layers of superconducting niobium-titanium windings. A 10 000 t steel return yoke, composed of five barrel wheels and six endcap disks, guides the lines of magnetic field around the outside of the solenoid and back into the other end. The return yoke, because of its ferromagnetism and high magnetic permeability, becomes magnetized by the field and thus passively adds to the field strength, focusing the field in the region outside the solenoid where the muon system is located.

Charged particles moving inside the solenoid are deflected by the Lorentz force, $\vec{F} = q(\vec{v} \times \vec{B})$. This force causes the component of the particle velocity that is perpendicular⁶ to the magnetic field v_T to produce circular motion while the component of velocity parallel to the magnetic field is unaffected, which results in a helical trajectory. Particle momentum and charge can be determined from the radius and direction of curvature of the particle's trajectory.

The windings of the solenoid are superconducting, which requires the magnet to be cooled with liquid helium. One component in the cryogenic system is the cold box, which contains the liquid helium heat exchangers. Run 1 went mainly smoothly for the CMS magnet. Unfortunately, the cold box became contaminated by oil in some components early in Run 2, causing erratic performance. As a result only about 75 % of 2015 proton-proton collision data were taken at full magnetic field strength. During the year-end technical stop the affected components were thoroughly cleaned and the oil recovery system was upgraded. Magnet operation has subsequently run smoothly.

⁶At CMS, the magnetic field points along the z axis, so the component perpendicular to it is the transverse component described in Section 3.2.2.

3.2.4 Inner tracker

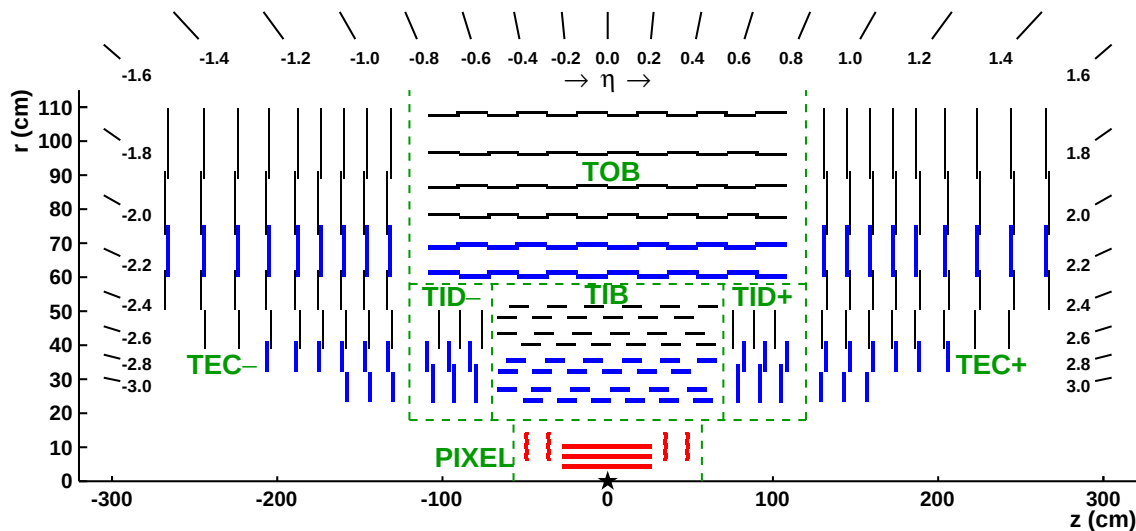


Figure 3.6. Longitudinal cross section of the CMS inner tracker [46].

The CMS inner tracker [47, 48] (pictured in Figure 3.6) is a cylindrical subdetector that is closest to the IP, consisting of an inner silicon pixel detector and an outer silicon strip detector. It is designed to measure the location and time that an electrically charged particle passes (a *hit*) while inducing minimal energy loss. Minimizing energy loss requires the tracker to be capable of taking such accurate hit measurements that the particle trajectory (or *track*) can be reconstructed with just a few hits. Being the innermost subdetector, it receives the highest flux of particles and therefore radiation hardness is very important. The inner tracker covers $|\eta| < 2.5$ and has about 200 m^2 of active silicon area, which is approximately the size of a tennis court. In the central region, the inner tracker has a typical momentum resolution of 0.7% at 1 GeV and 5.0% at 1000 GeV [49] and typical impact-resolution for high momentum tracks of $10 \mu\text{m}$. Unfortunately, the material making up the tracker corresponds

to up to 0.5 interaction lengths at some pseudorapidities, which means that in the worst case up to an 85 % chance exists for a photon to convert or an electron to emit bremsstrahlung radiation, and hadrons have a 20 % probability of a nuclear reaction before entering the calorimeters [50]. These interactions complicate event reconstruction.

Silicon detectors are semiconductors, with electrical properties that are determined by introducing small amounts of impurities that add mobile charge carriers. N-type silicon has electron charge carriers, while p-type silicon has lattice positions from which an electron is missing, resulting in mobile *holes*. The detector contains a junction between n-type and p-type regions. An externally applied voltage provides a positive charge to the n-type region, drawing electrons away from the junction, while a negative charge applied to the p-type region draws away holes. This produces a region with no mobile charge carriers near the junction, called the depletion zone, and no current flows. The detector is a diode under reverse bias.

When a charged particle passes through the detector, it ionizes silicon atoms near the junction, producing free electrons and holes. The electrons move to the positive terminal and the holes to the negative terminal, generating a small electric current that is amplified and recorded.

The tracker is subject to accumulated radiation damage, which creates extraneous electron acceptor sites, depleting the n-type silicon and increasing the bias voltage needed to maintain the depletion zone. This increases the *leakage current* that flows through the junction even when no ionization is present [51] and can cause breakdown of the junction and *thermal runaway*. To reduce the rate of radiation damage, the tracker is cooled to -10°C .

3.2.4.1 Pixel detector

The detector closest to the IP requires the highest spatial resolution. Referred to as the *pixel detector*, it contains millions of very small detector elements called pixels, each $100\text{ }\mu\text{m} \times 150\text{ }\mu\text{m}$ in area and $320\text{ }\mu\text{m}$ in thickness. The high-luminosity LHC environment

produces an effective p-doping over time, which requires a continually increasing bias voltage. For this reason, high dose n-implants are introduced onto an n-type substrate with p-type rings to isolate individual pixels and p-type silicon on the back side, which allows the sensor to operate while partially depleted. Because of the presence of electric and magnetic fields, the direction that the liberated electron-hole pairs drift is affected by the Lorentz force, leading to sharing of charge between neighboring pixels. This effect is exploited to improve the pixel resolution. During Run 1, the pixel detector had three concentric cylindrical barrel pixel layers (BPix) at radial distances between 4.4–10.2 cm from the IP, and two pixel layers (FPix) in each of the two endcaps which extend out to roughly $z = \pm 50$ cm. Partially into Run 2 during the year-end technical shutdown in 2016, the pixel detector was upgraded [52] to have four barrel layers and three endcap layers, in addition to faster readout chips and an improved cooling system. The upgrade resulted in an increase from 66 million total pixels to 124 million.

3.2.4.2 Silicon strip tracker

The pixel detector is surrounded by the silicon strip tracker (SST), which fills a volume between 20–116 cm radially and out to 118 cm in z , covering $|\eta| \leq 2.5$. Because the particle flux at this distance is smaller compared with the pixel region, a lower granularity design is sufficient. The SST has approximately 9.3 million strips composed of p-type implants on an n-type substrate with a solid n-type back. The strip system is divided into one barrel and two endcap regions. The barrel consists of the tracker inner barrel (TIB, four layers) and the tracker outer barrel (TOB, six layers). Each endcap consists of a tracker endcap (TEC, nine layers of disks, each with up to seven rings of radial-strip detectors) and a tracker inner disk (TID, nine layers). Similar to the pixel detector, charged particles liberate conduction-band electrons which drift towards readout sensors. The strip detectors are inclined with respect to the z -axis to be approximately perpendicular to the paths of particles radiating from the IP.

3.2.5 Calorimeters

Calorimeters measure the energy of particles by absorbing them; in other words, they induce particle showers within their volume, which dissipate kinetic energy. The active material of a calorimeter is a dense medium. Incident particles interact with the medium through electromagnetic or strong processes and produce secondary particles. These secondary particles also interact, generating a subsequent shower of particles with progressively smaller energies, which deposit energy that is proportional to the energy of the incident particle.

CMS has two calorimeters. The electromagnetic calorimeter (ECAL) absorbs mainly photons and electrons. Charged hadrons also deposit energy, but may not be absorbed. The hadronic calorimeter (HCAL) measures the energy of hadrons.

3.2.5.1 Electromagnetic calorimeter

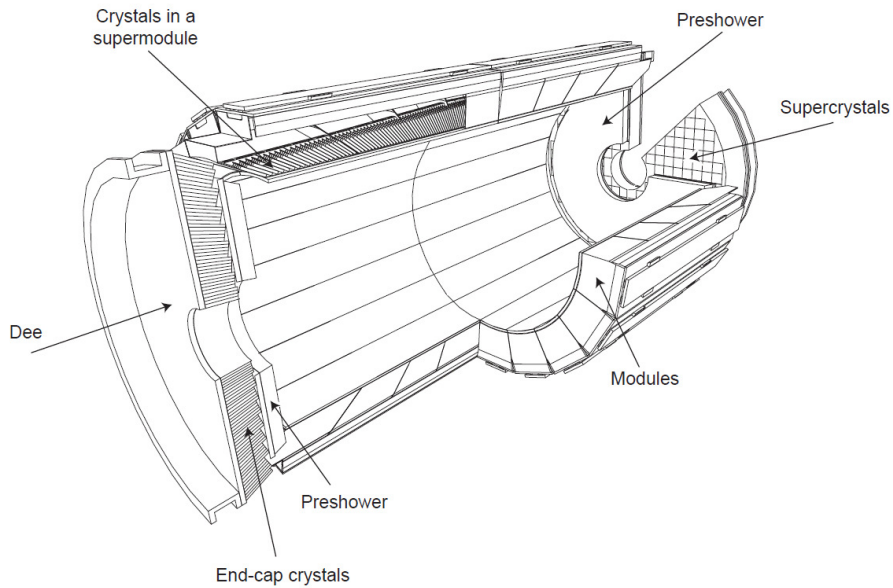


Figure 3.7. Layout of the CMS electromagnetic calorimeter [45].

Surrounding the tracker, the ECAL (shown in Figure 3.7) is composed of transparent lead-tungstate scintillator crystals. Lead-tungstate was chosen because of its radiation hardness, fast light output (80 % in 25 ns), short radiation length (the mean distance an incoming electron traverses before losing all but $1/e$ of its energy) of 0.89 cm, and a small Molière radius (the radius of a cylinder containing 90 % of the shower energy) of 2.2 cm. A short radiation length and small Molière radius is desirable in order to fully absorb incoming particles with a compact design. The electromagnetic shower induced in the crystals causes the energized scintillator atoms to emit light in proportion to the size of the shower. This scintillation light is collected by avalanche photodiodes in the barrel and photomultiplier devices called vacuum phototriodes in the endcaps. There are 61 000 crystals in the ECAL barrel region, which covers $|\eta| < 1.479$, and 7 324 crystals in each of the ECAL endcaps, covering $1.479 < |\eta| < 3$. The crystals vary between 22–23 cm in length.

A finer-grained detector, the ECAL preshower (ES), is placed in front of each endcap. It has a total thickness of about 20 cm and uses a sampling design consisting of two layers of lead radiators to induce showers, interleaved with two layers of silicon strip sensors to measure the deposited energy. The intended function of the ES was to help distinguish $\pi^0 \rightarrow \gamma\gamma$ decays from single high-energy prompt photons. Unfortunately, the large number of neutral pions produced through hadronic interactions with the tracker material in practice degrades the ES capabilities and thus the energy deposits in the ES are simply added to corresponding deposits in the ECAL.

Over time, radiation induces lattice damage in the lead tungstate crystals and reduces their transparency. Consequently, a calibration system is necessary. This system uses the LHC abort gaps to measure the baseline electronic noise (*pedestal*) levels and fire laser or LED pulses into the crystals to measure their transparency at regular intervals during data collection (on the order of once per hour).

3.2.5.2 Hadronic calorimeter

The HCAL is a sampling calorimeter composed of interleaved layers of brass absorber and tiles of plastic scintillator material. Brass is an ideal material for this purpose because it is non-magnetic and has a relatively short interaction length (16.4 cm), which is the average distance a particle will travel before interacting with the absorber. The HCAL surrounds the ECAL, and extends radially between 1.77–2.95 m from the beamline, providing a minimum of six interaction lengths for hadrons passing through it.

Incident hadrons cause showers of hadrons and leptons in the brass, which causes light to be produced in the scintillating layers in proportion to the amount of energy deposited. The light from the tiles is collected by wavelength-shifting fibers that are grouped into readout towers before being transmitted to hybrid photodiodes for amplification and readout.

The HCAL barrel (HB) region covers $|\eta| < 1.3$, and there are two HCAL end caps (EC), with each covering $1.3 < |\eta| < 3$. The EC and HB sections do not fully absorb all hadronic showers; any components passing beyond the HCAL are referred to as *hadronic punchthrough*. To measure shower energy deposited outside the HB, there is an extra layer of scintillator tiles located outside the solenoid, concentric to the HB, called the HCAL outer, or tail catcher. This outer layer uses the solenoid coil itself as an absorber. At either end of CMS, there are additional sections called the HCAL forward (HF) that covers $4.5 < |\eta| < 5.2$ and absorbs particles emerging from the IP at shallow angles. The HF absorbs the majority of the energy from the collision and so is made of steel absorber plates interleaved with layers of quartz fibers; interactions in the absorber produce showers of particles that pass through the quartz fibers, generating flashes of Cerenkov light.

3.2.6 Muon subsystem

Muons lose less energy due to bremsstrahlung than electrons because of their higher mass and do not interact via the strong force. Consequently the CMS muon subsystem [53] is the outermost detector since muons are the only charged particle that makes it that far. Since

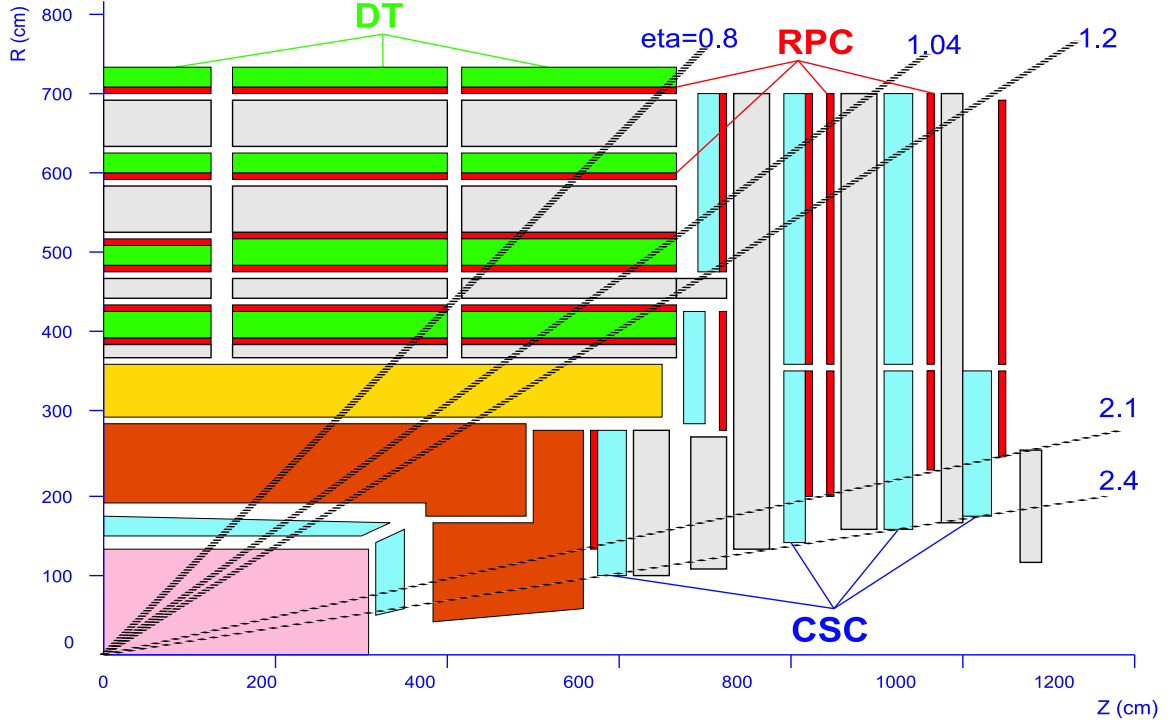


Figure 3.8. An $R-Z$ cross section of one quadrant of the CMS muon system [43].

the mass of the particle is known, it is only necessary to measure its charge and momentum.

The muon tracker (MT) subsystem consists of a cylindrical barrel region and two endcaps. The barrel is divided along the beam axis into five separate wheels with four concentric layers, alternating with layers of the return yoke. Throughout Run 1, the endcaps had three disks (*stations*) each; during LS1 a fourth disk was installed [54].

Because it is so far from the IP, the MT must cover a much greater area than the inner trackers to be hermetic. Consequently it must be less expensive per unit area to be practical; however, it does not require as high a resolution as the inner detectors. Gas ionization chambers of three types were selected.

Drift tubes (DTs) are used in the barrel region because it has lower muon rates. There are four layers of stations covering $|\eta| < 1.2$. The DTs consist of $13 \text{ mm} \times 42 \text{ mm} \times 2.4 \text{ m}$ cells filled with a mixture of carbon dioxide and argon gas. An anode wire held at a positive potential

of 3.6 kV runs the length of the tube, while electrode strips on the 42 mm walls are held at 1.8 kV and on the 13 mm walls are held at -1.2 kV. A muon passing through the tube ionizes the gas, producing free electrons and positively charged ions that drift to the wire or strips, creating a signal that is amplified and measured. Groups of four cells are stacked, staggered by half a cell to eliminate dead spots, to form a superlayer (SL). The DTs are arranged in four concentric cylinders around the beam axis, called stations, which alternate with layers of the steel plates that form the magnet return yoke. The three inner stations have three SLs each; two have wires running parallel to the beamline in order to measure $r-\phi$ and one has wires perpendicular to the beamline in order to measure η . The outer station has two SLs with wires that are parallel to the beam and only measure $r-\phi$.

Cathode strip chambers (CSCs) with higher resolution than the drift tubes are used in the endcaps, where the magnetic field is nonuniform and the muon rates are higher. Each endcap has four disks of CSCs covering $1.2 < |\eta| < 2.4$. As with the barrel, the disks alternate with steel plates of the return yoke. Each CSC is trapezoidal in shape with seven layers of cathode panels that have strips at constant ϕ intervals pointing radially out from the beam. These are interleaved with six panels of anode wires, oriented roughly perpendicularly to the cathode strips, to measure the radial position of the hit. Between the panels is a mixture of carbon dioxide, argon, and carbon tetrafluoride. CSCs use a mechanism similar to the DTs. An electrostatic potential is maintained between the anodes and cathodes; incoming muons ionize gas molecules, producing ions that drift towards the wire or strips, creating a small current pulse that is amplified and recorded.

Resistive plate chambers (RPCs) are gaseous parallel-plate detectors. The mechanism of operation is similar to the CSCs, except that there is a very narrow gap of 2 mm between the charged plates. This gap means they have a lower spatial resolution, but a very fast response time of only one nanosecond. Because the CSCs and DTs alone are not fast enough to unambiguously associate muons with a specific bunch crossing, RPCs are used throughout the system, covering the entire range of pseudorapidity.

The spatial resolution achieved per chamber is 80–120 μm in the DTs, 40–150 μm in the CSCs, and 8–12 mm in the RPCs. The efficiency for reconstructing hits and track segments in the muon system is in the range 95–98 % [55].

3.2.7 Trigger and data acquisition

Each crossing between two bunches of protons at the IP in which protons collide is called an *event*. During Run 1, the LHC normally provided pp collisions with a bunch spacing of 50 ns. In Run 2, the bunch spacing was reduced to 25 ns. This corresponds to a peak (in other words, excluding abort gaps) crossing frequency of 20 MHz and 40 MHz for Run 1 and Run 2, respectively. The raw information read out by the detector for each event is $\mathcal{O}(500 \text{ kB})$, meaning that CMS produces tens of terabytes of data per second, far exceeding the pace at which the data can be recorded. This rate necessitates a *trigger* system that quickly selects the most interesting events for storage. Discarded data cannot be recovered, so a properly functioning trigger system is critical to the success of the experiment. Periods during which events are not recorded because systems are busy processing existing data are referred to as *dead time* and should be minimized. The event rate is reduced in two stages: first, a faster, hardware-based Level 1 (L1) trigger reduces the event rate to 80–100 kHz. Next, a slower, software-based High Level Trigger (HLT) further reduces the event rate to approximately 1 kHz.

The time between the collision and when the L1 delivers a final decision is referred to as the *latency*. During this time, all of the data from the detector are buffered in a pipeline before being either discarded or forwarded on to the HLT. The L1 latency is 4 μsec , which coincides with the length of the LHC abort gap, making it possible to operate with zero dead time. This goal is accomplished with fast electronics (field programmable gate arrays) that construct local *trigger primitives* from the raw output of the calorimeters and muon systems (track reconstruction is too computationally expensive to complete on the L1 timescale). The calorimeter output channels are grouped into trigger towers. Trigger primitive generators calculate E_T sums for each tower and send them to the Regional Calorimeter Trigger, which

combines these values with quality information about the shower pattern to identify candidate electrons and photons, along with E_T sums for groups of 4×4 towers. These quantities are then passed to the Global Calorimeter Trigger (GCT), which calculates the sum E_T, p_T^{miss} and the scalar sum of transverse energy H_T . Track segments and hit patterns from the DT and CSC local triggers are sent in parallel to their respective track finders, which sort the trigger primitives by p_T and quality. Information is shared between the two track finders so that overlapping track segments can be linked. The RPCs assemble muon candidates from regional hit patterns and, because they are much faster than the CSCs and DTs, provide unambiguous bunch crossing identification. The best candidates from all three systems are passed on to the Global Muon Trigger (GMT), which combines the information for improved momentum resolution and efficiency.

The highest-quality objects (i.e., electrons/photons, muons, jets, or τ leptons) from the GCT and GMT are forwarded to the Global Trigger (GT). A *path* is a set of algorithms and requirements on objects in an event. L1 paths have simple criteria; for example, the event must contain at least one muon exceeding a given threshold, or the event must contain combinations of several objects. A trigger menu with a maximum of 128 different paths could be specified during Run 1. During LS1, the L1 trigger was upgraded [56], which added several L1 improvements, including adding fast pileup subtraction and expanded the menu size to a maximum of 256 paths.

If an event satisfies the criteria of any path in the L1 menu, the GT sends an L1 Accept signal to the Timing, Trigger, and Control (TTC) system which controls readout of the subdetector frontend buffers to the front-end drivers (FEDs). FED fragments are subsequently merged by the Event Builder and passed to the HLT. The HLT is a version of the sophisticated event reconstruction software used in offline analysis and achieves similar quality reconstruction. The HLT had a per event time budget of 175 sec during 2012. This longer timescale is made possible by a computer cluster with 13 000 cores which can process multiple events simultaneously. The HLT menu is composed of more than 400 paths.

HLT paths are more complex than at L1, consisting of a sequence of reconstruction and filtering steps. Products are filtered as algorithms are run. In this way, selections relying on calorimeter and muon system data, which are run first, reduce the number of events over which the computationally expensive tracker reconstruction is performed.

More details on the trigger system can be found in [57].

CHAPTER 4

OBJECT RECONSTRUCTION AND SELECTION

The CMS subdetectors produce raw electrical signals, from which *physics objects* are inferred as software representations of particle observables. This process is called *reconstruction* and is described in Section 4.1. In Section 4.2, the selection criteria for physics objects used in this work are discussed. Except where noted, features are common to both the 8 TeV measurement (presented in Chapter 5) and the 13 TeV measurement (presented in Chapter 6).

4.1 Particle flow

CMS uses a reconstruction approach called *Particle Flow* (PF), which takes advantage of information provided by each of the CMS subdetectors, as well as the correlations between these observations to construct a global picture of each event. The following description of the CMS PF algorithms is based on reference [50].

The reconstructed origin of a set of physics objects is called a *vertex*. An example of a vertex is the origin of the physics objects emerging from a proton-proton collision. As shown in Figure 3.3, each bunch crossing results in multiple proton-proton collisions. The majority of these are pileup. The vertex for which $\sum p_T^2$ is highest is most likely to be associated with the production of heavy particles, and is chosen to be the *primary vertex* (PV). For the 8 TeV analysis, the sum is taken over all charged particles that were used in the vertex reconstruction. For the 13 TeV analysis, the sum is taken over all charged leptons and jets formed from charged tracks used in the vertex reconstruction, in addition to the p_T^{miss} , which is computed as the magnitude vector sum of p_T from all PF candidates.

An iterative approach is used [46] to reconstruct particle trajectories in the inner tracker. Each iteration begins with seeds generated from a few tracker hits. Next, trajectories are built by adding hits from successive track layers. Finally, track-fitting using a Kalman filter [58] is performed to calculate the track origin and direction, and the p_T . In each iteration, tracks are subject to quality requirements on the seeds, χ^2 fit, vertex compatibility, p_T , $|\eta|$, and number of hits. After each iteration, hits that have been assigned to tracks are removed from the available hit pool. The first iteration applies stringent criteria, which produces a low rate of misreconstructed tracks with moderate tracking efficiency. Over 10 total iterations, the criteria are progressively relaxed as the algorithms become more complex and time-consuming. After the size of the hit pool has been sufficiently reduced, later iterations can perform more computationally intensive searches for displaced tracks with secondary vertices outside the interaction region (mainly due to photon conversions or nuclear interactions in the tracker material, b jets, Δ decays, etc.)

Most electrons radiate bremsstrahlung photons due to the coulomb field of atomic nuclei as they pass through the tracker. These photons are emitted in a characteristic arc due to the azimuthal bending of the electron track in the magnetic field. To account for these interactions, tracks likely to be electrons are selected after the iterative procedure described above based on their number of hits and χ^2 fit. The track-fitting is then repeated with a Gaussian Sum Filter (GSF) [59], which approximates the electron energy loss as a sum of Gaussian distributions instead of a single Gaussian, resulting in an improved momentum resolution.

The determination of particle energy measured in the calorimeters is accomplished with a *clustering* algorithm, which takes as input a map of individual energy deposits and outputs groupings of energy deposits that can be associated with candidate physics objects. First, *cluster seeds* are determined by identifying calorimeter cells with energies exceeding a predefined threshold. Next, clusters are built up by adding adjacent cells with energy deposits exceeding typical electronic noise levels.

After the PF elements have been assembled, a *linking* algorithm evaluates the degree

of the compatibility between different subdetector elements that are nearest neighbors in the (ϕ, η) plane to determine which are likely to be associated with the same particle. Tracks are extrapolated from the last hit measured in the tracker to the PS and subsequently to the ECAL and HCAL. If the extrapolation falls within the cluster boundaries, a link is established. ECAL clusters that are consistent with the extrapolation of tangents from points in which inner tracks intersect with tracker layers are linked as possible bremsstrahlung photons. Photons have a substantial probability to pair produce in the tracker material, so a similar process is used to identify possible photon conversions. If the (ϕ, η) cluster position in the ECAL is within the cluster envelope in the HCAL, a link is established. *Global muons* are formed if tracks in the inner tracker and tracks in the muon system can be linked by interpolating towards each other with an acceptable χ^2 . *Tracker muons* are formed if the extrapolation out to the muon track from the inner track is sufficiently consistent. The highest-quality links are used to form sets of PF elements that are linked, or have common elements that are linked. These sets are called *blocks*.

Following linking, PF elements are assigned particle identifications. To reduce combinatorics, elements are removed from blocks after assignment. Muons are assigned first because of the high purity with which they can be reconstructed, and they are required to satisfy any of three sets of criteria. Particle *isolation* quantifies the degree to which a particle is separated from others. Isolated muons are identified if the sum of the track p_T and transverse energy of calorimeter deposits in a cone with $R=0.3$ does not exceed 10 % of the muon p_T . Nonisolated muons are identified if they pass a more stringent selection [55] optimized for muons within jets. Finally, muons that fail the tight selection because of problems with the inner track reconstruction, but which have a sufficiently high-quality fit in the muon system, are assigned as PF muons and removed from their blocks.

Because of the tendency for electrons to emit bremsstrahlung photons, and for photons to pair produce electrons in the tracker material, electrons and isolated photons are reconstructed together, following muon reconstruction. Electron candidates are seeded with

tracks in the inner tracker linked to sufficiently large energy deposits in the ECAL. Photon candidates are seeded from similar energy deposits that have no linked corresponding inner tracks. Electron and photon candidates must meet requirements on the distribution of energy deposits in the calorimeters. Photons additionally must be isolated from other tracks and energy deposits. A multivariate discriminator based on track variables and energy distribution patterns is used to select electrons.

After muons, electrons, and isolated photons have been classified and removed from their blocks, hadrons and nonisolated photons are identified. The majority of jet energy is carried by charged hadrons, then photons, and finally neutral hadrons. For this reason, within the tracker acceptance, ECAL clusters that are not linked to inner tracks are classified as photons, and HCAL clusters that are not linked to inner tracks are classified as neutral hadrons. If the calorimeter clusters are linked to tracks, each track is assumed to give rise to a charged hadron. If the calibrated calorimeter energy is greater than the total track momenta by more than what is expected from the energy resolution to measure charged hadrons, the surplus is assumed to be carried by photons and neutral hadrons: up to the ECAL energy, the surplus is assigned as a photon, and any energy above that is assigned as a neutral hadron. Outside of the tracker acceptance, where charged and neutral hadrons cannot be distinguished, ECAL clusters with no associated HCAL cluster are classified as photons, while ECAL clusters that are linked to HCAL clusters are classified as arising from the same hadronic shower. The response produced in the ECAL to photons and charged hadrons with equivalent energy is different and must be corrected for after particle identification has been made.

The anti- k_t algorithm [60] is used to identify which PF candidate hadrons belong to the same jet. The quantity

$$d_{ij} = \frac{\Delta_{ij}^2}{R^2 \max(k_{ti}, k_{tj})} \quad (4.1)$$

is calculated for all possible candidate pairs i, j , where R is a characteristic radius parameter,

$\Delta_{ij}^2 = (y_i - y_j)^2 + (\phi_i - \phi_j)^2$, and k_{ti} , y_i , and ϕ_i are respectively the transverse momentum, rapidity, and ϕ -coordinate of candidate i . The pair that minimizes d_{ij} is replaced by a clustered candidate with the sum of their four-momenta. If all possible $d_{ij} > 1/k_{ti}$ for any candidate, it is promoted to a jet and removed from consideration. The process is repeated until no candidates remain. If there are many soft particles and only one hard particle within a radius $2R$, the minimum d_{ij} will be dominated by the momentum of the "seed" hard particle, which will accumulate soft particles, resulting in a jet that is a circle around the hard particle in the $\phi - \eta$ plane of radius R . If there are two hard particles that are more than R apart but less than $2R$, there will be two jets, with more of the soft particles naturally assigned to the higher- k_t particle. If there are two hard particles less than R apart, they will be merged into a single jet.

Anti- k_t jets with $R = 0.5$ are used in the 8 TeV analysis. Because the average jet energy increases with increasing center-of-mass energy, resulting in narrower jets, the cone size was reduced to $R = 0.4$ in the 13 TeV analysis.

Because top quarks almost always decay to b quarks, identifying jets from b quarks (called *b-tagging*) is critical for selecting $t\bar{t}W$ and $t\bar{t}Z$ events. Because the b quark is the lighter of the third generation quark doublet, it decays via a generation-changing process to c or u quarks, which are off-diagonal CKM matrix elements and therefore suppressed. This results in a substantially longer b quark lifetime than expected for their mass. On the other hand, it is short enough that decays still occur inside the detector. The long lifetime of b hadrons in jets originating from the hadronization of b quarks allows them to travel $\mathcal{O}(1 \text{ mm})$ before decaying at a displaced *secondary vertex*. In the 8 TeV analysis, a Combined Secondary Vertex (CSV) b-tagging algorithm was used [61]. The CSV algorithm uses variables related to the distance from a track's point of origin to the PV (the impact parameter, or IP) along with secondary vertex information to produce a continuous score between zero and one that is assigned to each jet. Higher scores indicate that the jet is more likely to have originated from a b quark than a light-flavor quark or gluon. In the 13 TeV analysis, improvements were made to allow for the inclusion of additional variables, improving the

tagging efficiency by several percentage points, and is referred to as CSVv2 [62].

4.2 Object selection

Our analyses, like most on CMS, rely on the PF and jet clustering procedure described above. Additional requirements, optimized for selecting $t\bar{t}W$ and $t\bar{t}Z$ events, are discussed below.

4.2.1 Leptons

For the $t\bar{t}W$ and $t\bar{t}Z$ measurements, differentiating *prompt* from *nonprompt* leptons is critical. Prompt leptons are muons or electrons arising either directly from a W or Z boson parent, or from a tau lepton which arise directly from a W or Z boson parent. Prompt leptons that have misidentified charge are called *charge flip* leptons. Nonprompt leptons arise mainly from b hadron decays, misidentified jets, or photon conversions.

Prompt leptons are generally more isolated from other objects in the event than nonprompt leptons, which are often produced in hadronic decays. To assess the degree of isolation, the scalar sum of PF particles within a cone of $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} = 0.4$ (or $\Delta R = 0.3$ for electrons in the 13 TeV analysis) around the lepton direction is calculated. To account for the charged component of pileup, the sum p_T of charged particles not originating from the PV is subtracted. To account for the neutral component, half of the charged PU contribution is subtracted in the 8 TeV analysis. In the 13 TeV analysis, a $\rho(\eta)$ correction is used as described in Section 4.2.2. The ratio of this corrected sum to the lepton p_T is called the relative isolation. Prompt leptons are also characterized by low minimum displacement of the track from the vertex in the transverse (d_{xy}) and longitudinal (d_z) planes, and by the ratio of the three-dimensional impact parameter to its uncertainty (SIP_{3D}). Electrons are additionally required to pass cuts on the score from an ElectronID MVA classifier which combines discriminating power from a set of shower-shape, track-cluster consistency, and track quality variables. This classifier was re-optimized for the 13 TeV analysis.

Analysis-specific criteria, along with a description of the different criteria categories used for event selection and data-driven background estimations, are summarized below for the 8 TeV and 13 TeV analyses.

4.2.1.1 8 TeV

The 8 TeV analysis imposes requirements related to the properties of the nearest jet. A nearby jet with a high CSV value indicates that the lepton is likely to be nonprompt, originating from b-hadron decay. Furthermore, the ratio of lepton p_T to that of the nearest jet tends to be lower for such leptons.

Four sets of lepton selection criteria are defined: preselected, loose, tight, and good charge. Preselected leptons, with criteria designed to select both prompt (selected with ~ 100 % efficiency) and nonprompt leptons, are used in the data-driven background estimations described in Section 5.2. Tight leptons, used to select events in the SS dilepton and three lepton channels, must pass more stringent criteria designed to accept prompt and reject nonprompt leptons. The efficiency for accepting prompt leptons passing the tight criteria is 68–98 % for muons and 49–93 % for electrons, while 80 % of nonprompt leptons are rejected. Loose leptons, selected with a less stringent set of criteria, are used to select events in the OS dilepton and four lepton channels, where there are fewer nonprompt leptons. These cuts accept 93–99 % of prompt muons and 89–96 % of prompt electrons while rejecting ~ 50 % of nonprompt leptons. In addition to the preselected, loose, or tight criteria, additional charge ID requirements are imposed to reject charge flip leptons. This charge ID cut has 99 % efficiency for right charge muons and rejects ~ 100 % of charge flip muons; for electrons, the efficiency ranges from 85–100 %, while more than 97 % of charge flip electrons are rejected. These four sets of selection criteria are summarized in Table 4.1.

TABLE 4.1

LEPTON SELECTION CRITERIA (8 TeV)

Lepton selection criteria	Preselected		Loose		Tight		Charge ID	
Lepton flavor	e	μ	e	μ	e	μ	e	μ
p_T (GeV)	> 10	> 10	> 10	> 10	> 10	> 10		
$ \eta $	< 2.5	< 2.4	< 2.5	< 2.4	< 2.5	< 2.4		
Relative isolation	< 0.4	< 0.4	< 0.4	< 0.4	< 0.4	< 0.4		
Charged relative isolation			< 0.15	< 0.20	< 0.05	< 0.15		
Ratio of lepton p_T to jet p_T					> 0.6	> 0.6		
CSV of nearest jet	< 0.679	< 0.679	< 0.679	< 0.679	< 0.679	< 0.679		
d_{xy} (mm)	< 5	< 5	< 5	< 5	< 5	< 5		
d_z (mm)	< 10	< 10	< 10	< 10	< 10	< 10		
$ \text{IP} $ (mm)					< 0.15			
$\text{SIP}_{3\text{D}}$	< 10	< 10	< 10	< 4	< 10	< 4		
Inner tracker hits								> 5
Missing inner tracker hits	< 2		< 2		< 2		0	
Tracker charge – ECAL charge							0	
Electron conversion veto								Pass
ElectronID MVA ^a	$> \begin{cases} 0.5 \\ 0.12 \\ 0.6 \end{cases}$		$> \begin{cases} 0.5 \\ 0.12 \\ 0.6 \end{cases}$		$> \begin{cases} 0.5 \\ 0.12 \\ 0.6 \end{cases}$			

^a Corresponds (top to bottom) to $|\eta| < 0.8$, $0.8 < |\eta| < 1.479$, and $1.479 < |\eta|$.

4.2.1.2 13 TeV

The 13 TeV analysis defines two sets of lepton selection criteria. Loose leptons are used in the data-driven background estimations. Tight leptons are used to select signal events. Additional quality criteria are imposed, corresponding to the medium working point recommended by the muon Physics Object Group (POG). These selection criteria are summarized in Table 4.2.

TABLE 4.2

LEPTON SELECTION CRITERIA (13 TeV)

Lepton selection criteria	Loose		Tight	
Lepton flavor	e	μ	e	μ
p_T (GeV)	> 10	> 10	> 10	> 10
$ \eta $	< 2.5	< 2.4	< 2.5	< 2.4
d_{xy} (mm)	< 5	< 5	< 5	< 5
d_z (mm)	< 10	< 10	< 10	< 10
SIP _{3D}	4	4	4	4
Relative isolation	< 1	$< \begin{cases} 1^a \\ 0.7^b \end{cases}$	0.1	$< \begin{cases} < 0.15^a \\ < 0.25^{b,c} \end{cases}$
ElectronID MVA ^d	$> \begin{cases} -0.56 \\ -0.72 \\ -0.1 \end{cases}$		$> \begin{cases} 0.837 \\ 0.715 \\ 0.357 \end{cases}$	
Muon working point ^e	medium		medium	

^a Corresponds to the SS dilepton analysis.

^b Corresponds to the three-lepton analysis.

^c Corresponds to the four-lepton analysis.

^d Corresponds, top to bottom, to cut application regions: $|\eta| < 0.8$, $0.8 < |\eta| < 1.479$, and $1.479 < |\eta|$.

^e This corresponds to a set of quality requirements recommended by the CMS Muon Physics Object Group.

4.2.2 Jets and missing energy

Jet energy corrections (JECs), described in detail in reference [63], are required to calibrate the reconstructed jet energy to match the true parton energy. To mitigate the effects of pileup, charged hadrons that do not originate from the PV are removed before jet clustering as described in reference [64]. To account for neutral hadrons from pileup, the average energy density as a function of the pseudorapidity $\rho(\eta)$ is calculated for each event and, for each jet, multiplied by the effective jet area and subtracted. Jets coming from pileup vertices are removed using a multivariate discriminator [65] based on tracking information and jet shape, with sensitivity driven by the ratio of the total p_T from charged candidates that do not originate from the PV to the total p_T of all charged candidates in the jet and the average spread in p_T among jet PF candidates. Additional corrections, parameterized in η and jet p_T , are applied to account for residual differences between data and simulation, and the nonlinear detector response to hadrons. To reject fake jets from misreconstruction and instrumental noise, jets must have at least two PF constituents and the fraction of their energy from the ECAL or HCAL deposits must not exceed 99 %.

Jets with $|\eta| < 2.4$ and $p_T > 25$ GeV ($p_T > 30$ GeV) are selected for the 8 TeV (13 TeV) analysis. To prevent double counting, jets must be separated from the selected leptons by $\Delta R > 0.5$ for the 8 TeV analysis and $\Delta R > 0.4$ for the 13 TeV analysis.

Loose and medium working points for the b-tagging output discriminant are used, defined to operate with a ≈ 10 % and ≈ 1 % chance, respectively to incorrectly tag light quark or gluon jets (i.e., to *mistag*). This approach corresponds to an efficiency of ≈ 85 % to correctly tag b jets for the loose working point and ≈ 70 % for the medium working point, depending on the jet p_T and η . To satisfy the loose or medium working point, $CSV > 0.244$ and $CSV > 0.679$ were required respectively for the 8 TeV analysis. For the 13 TeV analysis, this requirement corresponded to $CSV_2 > 0.5426$ and $CSV_2 > 0.8484$.

The p_T^{miss} is defined as the magnitude of the sum of the negative transverse momenta of all of the PF particles. Because pileup can contribute to p_T^{miss} , the 8 TeV analysis makes use

of the variable H_T^{miss} , computed as the sum of the negative transverse momenta of selected jets and leptons. The H_T^{miss} is more robust than p_T^{miss} at the expense of reduced resolution.

CHAPTER 5

OBSERVATION OF TOP QUARK PAIRS PRODUCED IN ASSOCIATION WITH A VECTOR BOSON IN PROTON–PROTON COLLISIONS AT $\sqrt{s} = 8$ TeV

This chapter summarizes $t\bar{t}W$ and $t\bar{t}Z$ cross-section measurements that were completed using 8 TeV pp collision data corresponding to an integrated luminosity of 5519.5 fb^{-1} collected during Run 1 of the LHC. The $t\bar{t}Z$ analysis was the first to exceed the 5σ significance threshold, which is considered sufficient to constitute the discovery of a hypothesized process. The reinterpretation of this measurement within the context of effective field theory will be described in Chapter 7.

More details on this analysis can be found in reference [66].

5.1 Event selection

We searched for $t\bar{t}W$ and $t\bar{t}Z$ events by optimizing selection criteria for each of several possible decay channels. At least one of several triggers, all based on lepton energies, had to be satisfied for data collection: either a dilepton trigger ($ee, e\mu, \mu\mu$) with p_T thresholds of 17

TABLE 5.1

TARGETED DECAY MODES AND BRANCHING FRACTIONS

Channel	SS $t\bar{t}W$	3 ℓ $t\bar{t}W$	OS $t\bar{t}Z$	3 ℓ $t\bar{t}Z$	4 ℓ $t\bar{t}Z$
Decay	$(bjj)(b\ell\nu)(\ell)$	$(b\ell)(b\ell)(\ell)$	$(bjj)(bjj)(\ell\ell)$	$(b\ell)(bjj)(\ell\ell)$	$(b\ell)(b\ell)(\ell\ell)$
Br. fraction	5.0 %	1.8 %	3.9 %	2.8 %	0.5 %

and 8 GeV or a trielectron trigger with thresholds of 15, 8, and 5 GeV. To ensure that events are recorded with the triggers operating near their maximum efficiency, selected leptons must have $p_T > 10$ GeV, with at least one $p_T > 20$ GeV. To reduce selection of Υ and J/ψ events the invariant mass of any pair of leptons generated by the event must exceed 12 GeV.

The decay of a W or Z boson yields two fermions, a particle and an antiparticle, that are either both leptons or both hadrons (i.e., quarks). Decays yielding top quarks are excluded by conservation of energy. The hadronic decay of either W or Z produces a quark and an antiquark, which are detected as jets, and Z decay produces a particle and its own antiparticle. Consequently, leptonic Z decay produces a pair of charged leptons of opposite sign (OS) or a pair of neutrinos; the latter case is not detected. Leptonic W decay yields a single charged lepton and a neutrino of the same generation. The leptonic decay of either W or Z can produce tau leptons, which usually decay between the collision point and the tracker. About 35 % of the time, taus decay leptonically to either $(e, \bar{\nu}_e, \nu_\tau)$ or $(\mu, \bar{\nu}_\mu, \nu_\tau)$. In this dissertation, ℓ refers to an electron, a muon, or a tau that decays leptonically.

Possible final states are further characterized by the decay mode of the $t\bar{t}$ pair. The decay of each top quark produces a quark (usually a b quark) and a W boson. If both W bosons decay into leptons, the decay is referred to as leptonic. If one W decays leptonically and one hadronically, only one lepton is produced and the decay is referred to as semileptonic. If both W bosons decay hadronically, the decay is referred to as hadronic.

If exactly two charged leptons are detected, they will either be OS or same sign (SS). Recall that the lepton pair from a Z decay will be OS. If the lepton pair results from the leptonic decay of both W bosons from a $t\bar{t}$ pair, it will also be OS because the electric charges of the top quarks have opposite signs. SS leptons can arise when both the associated W or Z boson and a W boson from the $t\bar{t}$ pair produce leptons. The final states examined in this analysis are described below and summarized in Table 5.1. The corresponding selection criteria are summarized in Table 5.2.

Hadronic $t\bar{t}$ decay For the $t\bar{t}Z$ process, the hadronic decay of the $t\bar{t}$ pair and leptonic decay

of the Z results in two OS leptons in the final state. This series of events is designated as the OS $t\bar{t}Z$ channel. The main background for this channel comes from Z production with extra radiated partons, and $t\bar{t}$ decaying leptonically to produce an OS lepton pair.

Loose selection criteria are sufficient for the leptons because an OS lepton pair is required with $m_{\ell\ell}$ within 10 GeV of the Z boson rest mass. At least five jets are required, with at least one medium b-tagged jet. The channel is categorized into events with exactly five jets and events with six or more jets, with the latter having a higher signal-to-background ratio. True $t\bar{t}Z$ events must have OS same-flavor (OSSF) lepton pairs, so the channel is further split into $e\mu$ and $ee/\mu\mu$ categories. The $e\mu$ category is used to calibrate the $t\bar{t}$ background. Note that the $e\mu$ category includes $\sim 6\%$ of true OS $t\bar{t}Z$ events in which the Z boson decays to a pair of τ leptons, with one decaying to a muon and the other to an electron.

Semileptonic $t\bar{t}$ decay For the $t\bar{t}W$ process, the leptonic decay of the associated W and semileptonic decay of the $t\bar{t}$ pair produces two leptons. The lepton charges are independent, and in half of such cases, they will be SS. The most significant background is from leptonic $t\bar{t}$ events in which one lepton has misreconstructed charge or semileptonic $t\bar{t}$ with an extra nonprompt lepton. To suppress this background, tight SS leptons that pass the charge identification are required. Misreconstruction is more likely for electrons, so if the SS leptons are electrons, they must have $|M_{ee} - M_Z| > 10 \text{ GeV}$. (This requirement rejects events with Z boson decays with one misidentified electron charge.) Events must have at least three jets, with ≥ 1 medium or ≥ 2 loose b-tagged jets. The channel is categorized by lepton flavor (ee , $e\mu$, or $\mu\mu$) and whether the event has three jets or ≥ 4 jets.

For the $t\bar{t}Z$ process, the semileptonic decay of the $t\bar{t}$ pair produces three leptons in the final state, which is designated as the 3ℓ $t\bar{t}Z$ channel. The dominant backgrounds are leptonically decaying $t\bar{t}$ production with an extra nonprompt lepton and single Z and WZ production with extra partons including heavy flavor (HF). Events must have an OSSF lepton pair with an invariant mass within 10 GeV of the Z boson mass. Lepton charges must add up to ± 1 , and the SS leptons must pass the tight identification and charge identification. Events must have at least one medium or two loose b-tagged jets, and they are categorized based on whether they have exactly three jets or ≥ 4 jets.

Leptonic $t\bar{t}$ decay For the $t\bar{t}W$ process, leptonic decay of the $t\bar{t}$ pair also produces three leptons in the final state. Dominant backgrounds are leptonic $t\bar{t}$ with an extra nonprompt lepton and single Z events with extra partons including HF and an extra nonprompt lepton. Lepton charges must add up to ± 1 , and the SS leptons must pass the tight identification and charge identification. Events must have at least one medium or two loose b-tagged jets and are categorized according to whether the event has exactly one jet or ≥ 2 jets. To reduce background from $t\bar{t}Z$ and Z production, events with OSSF lepton pairs with an invariant mass within 10 GeV of the Z boson mass, which are already included in the 3ℓ $t\bar{t}Z$ channel, are rejected.

For the $t\bar{t}Z$ process, the leptonic decay of the $t\bar{t}$ pair produces four leptons in the final state. The most significant background comes from ZZ with extra radiated partons

TABLE 5.2

EVENT SELECTION CRITERIA

Channel	OS $t\bar{t}Z$	SS $t\bar{t}W$	3 ℓ $t\bar{t}W$	3 ℓ $t\bar{t}Z$	4 ℓ $t\bar{t}Z$
Lepton flavor ^a	$ee/\mu\mu, e\mu$	$ee, e\mu,$	Any	Any	Any
Lepton ID	2 loose	2 tight	SS tight	SS tight	4 loose
Lepton charge ID	-	2 pass	SS pass	SS pass	4 pass
$Z \rightarrow \ell\ell$ candidates ^b	1	0	0	≥ 1	1, 2
Number of jets ^a	5, ≥ 6	3, ≥ 4	1, ≥ 2	3, ≥ 4	≥ 1
Number of b tags	≥ 1 medium	≥ 2 loose or ≥ 1 medium	≥ 2 loose or ≥ 1 medium	≥ 2 loose or ≥ 1 medium	≥ 1 loose
Other	-	$Z \rightarrow ee$ veto	-	$H_T^{\text{miss}} > 30$ GeV	-
Subchannels ^c	4	6	2	2	2

^a Entries separated by commas indicate events are categorized according to each listed criteria.

^b A candidate is defined as an OSSF lepton pair where $|M_{\ell\ell} - M_Z| < 10$ GeV.

^c Each subchannel corresponds to a unique combination of categories.

(ZZ+jets). Four loose leptons that pass the charge identification and have charges adding up to zero are required. At least one pair of the four leptons must have an invariant mass within 10 GeV of the Z boson mass; events are categorized according to whether they have one or two such pairs, to help separate $t\bar{t}Z$ and ZZ. Further suppression of the ZZ+jets background (with no neutrinos in the final state) is achieved by requiring that H_T^{miss} exceeds 30 GeV. One loose b-tagged jet is required.

5.2 Event modeling

The event selection criteria detailed in the preceding section are optimized to reduce the backgrounds as much as possible. Unfortunately, a substantial amount of contamination persists after the selection. To estimate how many of the passing events are actual $t\bar{t}W$ or

$t\bar{t}Z$ events, both the quantity and kinematic features of signal and background processes must be accurately modeled.

5.2.1 Prompt backgrounds and signal processes

Background events, which include prompt leptons from W or Z decays, that satisfy the lepton selection criteria are considered prompt backgrounds. Prompt backgrounds and signal processes were modeled using Monte Carlo (MC) simulation, which was normalized using the inclusive cross section. Minor prompt backgrounds, which produce fewer expected events than signal, include $W^\pm W^\pm$ and rare processes including triboson production (WWW and WWZ), associated production of a Z boson with a single top quark (tbZ), and $t\bar{t}$ with an on or off-shell photon ($t\bar{t}\gamma/t\bar{t}\gamma^*$) or two W bosons ($t\bar{t}WW$). The dominant prompt backgrounds are Z boson and $t\bar{t}$ production (in OS $t\bar{t}Z$), WZ (in the SS $t\bar{t}W$ and 3ℓ channels), and ZZ events (in the 3ℓ and 4ℓ channels). These processes were generated using the MADGRAPH 5.1.3 [67] tree-level matrix element generator followed by PYTHIA 6.4 [68] for the parton shower and hadronization. The associated production of a top quark pair with a Higgs boson is expected to produce a small fraction of events in signal regions and was simulated with PYTHIA, assuming a Higgs boson mass of 125 GeV. In all samples with top quarks, a top quark mass of 172.5 GeV was assumed. GEANT4 software [69] was used to model the CMS detector response. MC simulation was required to pass the same trigger as data, and was reconstructed via the same algorithms. The CTEQ6L1 PDF set [70] it was used for all samples.

5.2.1.1 Simulation corrections

Simulation does not always yield a perfectly accurate description of the data. To enhance the agreement between the data and simulation, corrections were derived as follows. The binned distribution of the quantity to be corrected for in the data is divided by the corresponding distribution in the simulation. Each bin of the resulting ratio corresponded to a correction scale factor (SF) for that bin. An overview of the specific corrections that

were applied is presented below.

Pileup distribution The distribution of pileup collisions in events depends on the distribution of luminosities delivered by the LHC. The latter must be estimated when the simulation is produced, which may be before or during data collection. An event-level correction is applied to account for differences between the estimation and the true distribution. A histogram of the number of pileup vertices in data events is divided by a histogram of the number of pileup vertices in simulated events to obtain a ratio histogram of correction SFs as a function of the number of pileup vertices. For example, if there are twice as many events that have 20 pileup vertices in the data as in the simulation, then simulated events with 20 pileup vertices will be weighted by an SF of two.

Jet CSV A correction was made to account for differences in the distribution of CSV values for the bottom and light-flavor (uds or gluon) jets between the data and the simulation. Charm jets received no SF. The SFs were parameterized by jet parton flavor, CSV value, p_T , and η . Events were weighted according to the product of corrections evaluated on each jet. The procedure is described in detail in [71].

Lepton trigger, identification, and isolation efficiency Event-level corrections were applied to account for differences in trigger efficiency between the data and the simulation. These corrections were parameterized by lepton flavor and η . The event weight is the product of SFs evaluated on the selected leptons. Event-level weights to correct the lepton identification and isolation efficiency were also determined from the product of corrections calculated for each of the selected leptons. The corrections were calculated from $Z \rightarrow \ell\ell$ events and parameterized by lepton flavor, η , and p_T .

Top quark p_T The simulated p_T distribution in $t\bar{t}$ events tended to be higher than the distribution observed in the data. For this reason, a correction, parameterized as a function of the top quark p_T , was applied. More details can be found in [72].

Prompt backgrounds with extra HF jets The single Z, WZ, and ZZ processes are significant backgrounds that can pass the final selection criteria owing to extra radiated partons. The simulation of these processes was produced with fewer partons than are required in the final selection, leading to large uncertainties in their contribution to signal channels. Therefore, SFs were derived to ensure good agreement between the data and the simulation. Most of the extra radiated partons are gluons and light-flavor quarks, so the SF was determined in a region with fewer b-tags than the signal region and applied to the signal region. Events were selected that contained an OSSF lepton pair consistent with a Z boson decay but without a requirement on the number of jets, and with no medium b-tagged jets. This process yielded a very pure sample of about 5000 $Z \rightarrow \ell\ell$ data events, which was used to derive a correction SF parameterized by the number of jets. The SF was derived for four-jet events with 0 medium b-tags and applied to events with ≥ 1 medium b-tags. The SFs ranged from 1.35 to 1.7; we assigned an uncertainty of 30 % to the SFs based on the agreement between the data and the simulation. Because the simulation in OS dilepton events with ≥ 4 jets (excluding the signal region) was found to underestimate the data, additional uncertainties were

assessed on the p_T^{miss} distribution in Z+jets events, as well as the η distribution of jets in Z boson and $t\bar{t}$ simulations. WZ and ZZ events were only simulated with up to two extra partons. To determine how well the simulation and data agreed, we derived SFs for diboson events. We selected events using an identical selection as the $3\ell t\bar{t}Z$ channel, except that no requirement was made on the number of jets, and the b-tagging cut was inverted; in other words, we required zero medium and no more than one loose b-tagged jets. This selection yielded 80 data events, from which we obtained SFs of 1.4 for 3-jet and 1.6 for ≥ 4 -jet events. These values are assigned 40 % and 60 % uncertainties respectively, because of the small number of data events used in the derivation.

A technique for correcting for differences between the data and the simulation in $t\bar{t} + \text{HF jets}$ has been developed for the $t\bar{t}H$ analysis, as described in [23]. A similar approach was used here; $t\bar{t}$, WZ, and ZZ events were classified according to whether they had one or two extra b jets or one or two extra c jets; such events were assigned an extra rate uncertainty of 30 %. A similar approach was used for the single Z boson simulation, but the uncertainty was reduced to 30 %, using a sideband region with the criteria for $t\bar{t}Z$ selection other than having zero medium b-tagged jets. This approach was validated in events with exactly four jets and low p_T^{miss} .

5.2.2 Nonprompt backgrounds

Backgrounds with at least one nonprompt lepton were estimated from the data. Some nonprompt leptons pass the tight criteria. The rate at which any preselected nonprompt lepton in an event passes the tight criteria is called the misidentification rate, while the rate per lepton is designated f . Nonprompt backgrounds in the signal regions were estimated by selecting events that met the signal selection requirements, except that one or more leptons failed the tight lepton criteria, and then weighting those events according to f .

We measured the misidentification rate separately for electrons and muons in five p_T bins. We used two selection criteria: SS events with ≥ 2 jets (dominated by $t\bar{t}$ decays with a nonprompt lepton and W+jets) and 3ℓ events with ≤ 2 jets, a lepton pair consistent with Z boson decay, and low p_T^{miss} (dominated by Z boson production with an additional nonprompt lepton). In both cases, there was usually one prompt and one nonprompt lepton, but we did

not know which was which. (If we had known, we could have trivially calculated f as the ratio of tight, nonprompt leptons to preselected leptons.) We used the tag-and-probe method, in which the prompt lepton is tagged with the tight lepton selection, and the fraction of preselected probe leptons passing the tight selection measures f . In other words, the *tag* lepton has to be tight in the numerator and denominator, and the *probe* lepton has to be tight in the numerator but only preselected in the denominator. We use the subscripts T and P to indicate the number of leptons passing the tight and preselected criteria, respectively, with i indicating the bin in lepton flavor and p_T . The numerator contains a term for the yield where the tagged lepton is actually prompt and the probe passes the tight criteria, N_{TT}^i . The denominator contains a term for the yield where the tag is prompt and the probe passes the preselected criteria, N_{PT}^i . To account for contamination from tag leptons that were actually nonprompt, we subtract $N_{TF}^i \frac{f^i}{1-f^i}$ in the numerator and $N_{PF}^i \frac{f^i}{1-f^i}$ in the denominator. Then f^i is given by

$$f^i = \frac{N_{TT}^i - N_{TF}^i \frac{f^i}{1-f^i}}{N_{PT}^i - N_{PF}^i \frac{f^i}{1-f^i}}$$

Because the preceding calculations assume that each event has at least one nonprompt lepton, we estimated the number of events with two prompt SS leptons from simulation and explicitly subtracted it. Note that f^i appears on both sides of the equation, it cannot be explicitly solved for. We instead performed a χ^2 simultaneous fit over all bins i to find the set of f^i that minimized the residuals between the predicted yields and the data in the SS and 3ℓ derivation regions. The results are summarized in Table 5.3.

The transformation between the number of leptons that pass the loose criteria but fail the tight criteria in the sideband region and the number of events that have one or more nonprompt leptons that pass the tight criteria can be expressed in terms of f as a system of linear equations. This system can be simplified under the approximation that prompt leptons failing the tight selection are negligible compared with the quantity of nonprompt leptons that pass the tight selection. Using this simplification, events in the sideband region

TABLE 5.3

MISIDENTIFICATION RATES FOR MUONS AND ELECTRONS

μ				e		
10–15 GeV	15–20 GeV	20–30 GeV	>30 GeV	10–20 GeV	20–40 GeV	>40 GeV
0.17	0.20	0.21	0.36	0.15	0.15	0.15

are weighted according to

$$w = (-1)^{j+1} \prod_{j=1}^N \frac{f_j}{1-f_j}, \quad (5.1)$$

where there are $N = 1, 2$, or 3 leptons that pass the loose criteria but fail the tight criteria and f_j corresponds to f evaluated according to the flavor and p_T of the j th lepton. The negative weights account for events with two nonprompt leptons contaminating the sideband region.

There were too few events in the 4ℓ channel for the data-driven method for modeling the contribution from nonprompt leptons. Instead, yields from $t\bar{t}$ (and two nonprompt leptons), Z boson production (and two nonprompt leptons), and WZ (and one nonprompt lepton) were estimated from the MC simulation. To correct the normalization of these samples, an SF of 2 was derived for the number of nonprompt leptons passing the loose criteria, using $t\bar{t}$ and Z boson MC with exactly three loose leptons, one or two jets, with ≥ 1 medium b tag. After applying this SF, we found good agreement in three-lepton events with at least one medium b-tagged jet. Because of the relatively small number of selected events used in the derivation, this SF was applied in 4ℓ events with 100 % rate uncertainty.

5.2.3 Charge-misidentified backgrounds

In the SS $t\bar{t}W$ channel, a major source of background comes from OS events in which the charge of one lepton is misidentified. The main contribution of this type comes from $t\bar{t}$ and

single Z production events. A data-driven approach was used to estimate the rate of charge misidentification. It was more straightforward to calculate the charge misidentification rate than f . In the latter case, we knew the number of leptons passing the tight and preselected criteria, but we wished to calculate the rate at which nonprompt leptons passed the tight criteria; we needed to determine the mapping between nonprompt and prompt leptons, and tight and preselected leptons. In the former case, we could select events such that *all* lepton pairs should be OS, and no such mapping was necessary: we simply calculated the ratio of SS to OS events that passed the charge identification requirement.

The probability that a charge-misidentified lepton nevertheless passed the charge identification requirement was measured by selecting dilepton events with $|M_{ee} - M_Z| < 10 \text{ GeV}$, ≤ 3 jets, and no b-tags, and then calculating the ratio of SS to OS events passing this selection. For muons, this rate was negligible. For electrons the rate was derived as a function of η ; it ranged from 0.003 % in the central region of the detector to 0.1 % in the endcaps. The charge misidentification rate for a SS pair was equal to the charge misidentification rate for a single electron multiplied by the number of electrons in the pair (i.e., two.) We assigned a 30 % uncertainty per electron as the final charge misidentification rate based on the agreement between predicted and observed SS ee events with an invariant mass close to the Z mass and two or three jets.

5.3 Full event reconstruction

In most channels, there are still more background events than signal events, so additional separation of signal and background is needed. This separation is accomplished by *full reconstruction* of the event, in which the fundamental particles are reconstructed from their decay products to leverage additional kinematic information to distinguish signal and background.

Consider the OS $t\bar{t}Z$ channel as an example in which the top quark pair decays hadronically. A major background comes from $t\bar{t}$ pairs decaying leptonically, producing two OS leptons. Both processes produce real $t\bar{t}$ pairs, but the decay modes are different. We exploited these differences to improve discrimination between signal and background.

Signal events are heavier systems than major backgrounds. Particles decay isotropically, so heavy particles produce decay products with much higher average transverse momentum than the uninteresting particles that are sprayed down the beamline following a collision. In addition, radiated partons from the collision tend to have lower p_T than the decay products of the top and W bosons, which get a boost from their parent particle.

In the decay $t \rightarrow bW \rightarrow b(q\bar{q})$, three quarks are expected, with an invariant mass equal to the top quark mass and two of the quarks forming a W mass. The jet CSV value provides information about the flavor of a jet's parent quark, while the electric charge of the parent quark can be inferred from the jet's constituent hadrons. In a top decay, if the W decays leptonically, then the sum of the transverse component of the invariant mass (M_T) of the b quark, lepton, and missing transverse momenta (from the neutrino) will be less than the top mass. This information can be used to match reconstructed particles with their parent particles.

Based on the $t\bar{t}$ simulation, in which the true parentage of leptons and jets was known, we derived a linear discriminant that determined the most likely pairing between objects and their parents. A full list of input variables is given in Table 5.4. For each channel, we first produced a histogram of the distribution of each input variable for correctly matched objects (e.g., the invariant mass of a lepton and a b-tagged jet from a top quark decay). We then made the same distribution, but without any matching requirement, for all possible objects or combinations of objects in the event. Finally, we took the ratio of the two histograms and normalized it such that correctly matched objects had a mean value of one. A comparison of the W dijet mass in semileptonic $t\bar{t}$ decays for any pair of jets and for matched jets, along with the corresponding ratio histogram, is shown in Figure 5.1.

The ratio histograms from the simulation, in which the correct matching was known, were used to reconstruct the $t\bar{t}$ system in the data, where the correct matching was unknown, as follows. First, we removed the decay products from the associated boson. In the $t\bar{t}W$ channels, we assumed the lepton with the worst matching to a $t\bar{t}$ decay was from the W, and in the $t\bar{t}Z$ channels, the leptons whose invariant mass most closely matched the Z boson

mass were removed. Next, for each possible pairing between parent and daughter leptons and jets in the $t\bar{t}$ system, we calculated the product of the corresponding values for all bins of the ratio histogram. The permutation with the highest product discriminant value was considered to be the most probable reconstruction of the $t\bar{t}$ system. To get a more convenient range of values, we took the log of this discriminant so that correctly matched permutations centered around 0. More details on this procedure can be found in [66].

This approach produced the correct assignment 75 % of the time for semileptonic $t\bar{t}$ decays in events with exactly four jets, all from the $t\bar{t}$ system. For events with more than four jets (only four of which are from the $t\bar{t}$ system), the correct assignment was obtained 40 % of the time owing to the additional permutations. We also attempted partial reconstructions with one or two jets missing because some signal jets failed to be reconstructed or may not have passed the selection cuts.

Recall that in OS $t\bar{t}Z$ events, the $t\bar{t}$ system decays hadronically, while the main background events come from leptonic $t\bar{t}$ decays, which produce OS leptons. Similarly, in SS $t\bar{t}W$ and 3ℓ $t\bar{t}Z$ events, the $t\bar{t}$ system decays semileptonically, and in 3ℓ $t\bar{t}W$ events, the $t\bar{t}$ system decays leptonically. Meanwhile in SS and 3ℓ $t\bar{t}$ events, an additional nonprompt lepton, usually from a b-hadron decay, is present. Because the parent particles for the $t\bar{t}$ system differ between background and signal, match scores computed for the $t\bar{t}$ system in $t\bar{t}W$, $t\bar{t}Z$, and $t\bar{t}$ decays provide discrimination between signal and background.

5.4 Signal extraction

At this point, we had a set of events, each associated with a list of measured input variables. We wished to assign each event to either signal or background based on these measured parameters. To accomplish this end, we used a method called a Boosted Decision Tree (BDT) [73].

A decision tree divides a set of events into two groups based on the cutoff values for one or more parameters. The full sample, comprising equal parts signal and background events,

TABLE 5.4

INPUT VARIABLES TO THE FULL EVENT RECONSTRUCTION

Reconstructed event	Decay products of the $t\bar{t}$ system				
	$bq\bar{q}\bar{b}q\bar{q}$	$b\ell\nu\bar{b}q\bar{q}$	$b\ell\nu\bar{b}\ell\nu$	$\ell_b q\bar{q}\bar{b}\ell\nu$	$\ell_b \ell\nu\bar{b}\ell\nu$
OS dilepton $t\bar{t}Z$	X				
SS dilepton $t\bar{t}W$		X			
3ℓ $t\bar{t}Z$		X			
3ℓ $t\bar{t}W$			X		
OS dilepton $t\bar{t}$			X		
SS dilepton $t\bar{t}$				X	
3ℓ $t\bar{t}$					X
Input variables ^a					
b jet CSV	X	X	X	X	X
Higher jet CSV from $W \rightarrow q\bar{q}$	X	X		X	
Lower jet CSV from $W \rightarrow q\bar{q}$	X	X		X	
Charge of b jet from t	X		X		
Charge of b jet from $\bar{t}$	X		X		
Charge of b jet from $t \rightarrow b\ell\nu$		X			
Charge of b jet from $t \rightarrow bq\bar{q}$		X			
Charge of b jet not decaying to a lepton				X	X
Sum of charges of jets from $W \rightarrow q\bar{q}$	X	X		X	
Mass of lepton and b jet from t			X		
Mass of lepton and b jet from $\bar{t}$			X		
Mass of lepton and b jet from $t \rightarrow b\ell\nu$		X		X	X
Mass of leptons from $t \rightarrow \ell_b \ell\nu$					X
M_T of $\vec{p}_T^{\text{miss}}$ and $\vec{p}_T$ of lepton and b jet from t		X		X	
Mass of two jets from $W \rightarrow q\bar{q}$	X	X		X	
Mass of b jet and quark jets from $t \rightarrow bq\bar{q}$	X	X			
Mass of lepton from b and jets from $t \rightarrow \ell_b q\bar{q}$				X	
Ratio of M_T to mass for jets from t or W	X	X	X		

^a In $t\bar{t}$ events, ℓ_b denotes a lepton from b-hadron decay.

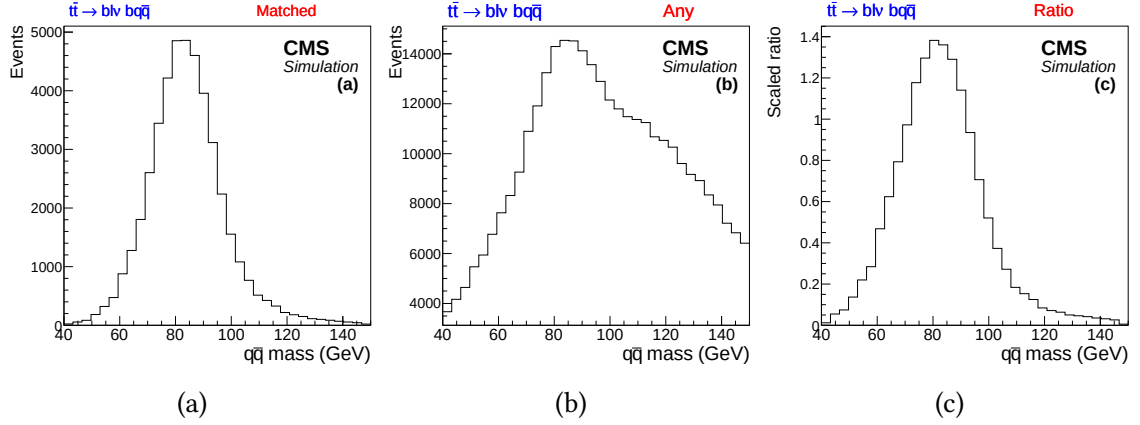


Figure 5.1. The dijet mass in semileptonic $t\bar{t}$ decays. In panel (a), only pairs of jets arising from the decay of the same W are included. In panel (b), no particular parentage is required; all dijet combinations are included. The ratio of the histograms in panels (a) and (b) is presented in (c).

makes up the first *node* of the tree. The sample is then divided using the variable and cutoff value that provides the best discrimination between signal and background events to create two *branches*. These branches become two new nodes, and the process is repeated until one of the stopping criteria is satisfied; usually, a specified maximum number of final branches (or *leaves*) is obtained, a specified minimum number of events per branch is reached, or a leaf is entirely signal or background. Finally, the tree (*trained* with simulated data, for which signal and background events are labeled) can be used to assign a score to real (unlabeled) data events. Events are fed through the tree, with each event following a path according to whether it passes or fails each established cut. The event is then assigned a score based on the final leaf it lands on, usually according to the purity (defined below) of the terminal leaf, such that larger scores correspond to more signal-like events.

The quality of separation between signal and background events, designated *purity*, is defined as $P = \frac{\sum_s W_s}{\sum_s W_s + \sum_b W_b}$, where each event has weight W_i and $\sum_s$ refers to the sum over signal events and $\sum_b$ refers to the sum over background events. Let the Gini index G be defined as $G = \sum_{i=1}^n W_i P(1 - P)$. Note that $G = 0$ if a sample is either all signal or all

background. To separate signal and background maximally, an optimal cutoff value for a given input variable can be found by minimizing $G_{\text{pass}} + G_{\text{fail}}$, where G_{pass} (G_{fail}) refers to the Gini index of the events that pass (fail) the cut. The input variables can be ranked by finding the maximum value of $G_{\text{parent}} - G_{\text{pass}} - G_{\text{fail}}$ for each one.

To improve performance, a *boosting* procedure can be performed. After the initial tree is built, the weight of misclassified events is increased and an additional tree is constructed; this process is repeated many times to obtain a *forest* of trees. Events are fed through each tree, and the final discriminant score is evaluated by summing the scores from all trees.

A total of 10 BDTs were generated, one for each jet category and channel, using the “gradient boosting” implementation in the Toolkit for Multivariate Analysis package [74]. Simulated $t\bar{t}W$ signal events were used with simulated $t\bar{t}$ background events for training of the SS $t\bar{t}W$ and 3ℓ $t\bar{t}W$ BDTs. Simulated $t\bar{t}Z$ events were used with simulated $t\bar{t}$ and WZ events for the 3ℓ $t\bar{t}Z$ BDTs. For the OS $t\bar{t}Z$ channel, we initially trained a BDT with $t\bar{t}Z$ events against a $t\bar{t}$ background, then used that output as an input variable to a final BDT, which was trained with $t\bar{t}Z$ signal against $t\bar{t}$ and Z boson simulation. The 4ℓ channel had too few events to train a BDT, so the number of medium b -tagged jets was used as a discriminant instead. The match scores and other input variables for all BDTs are listed in Tables 5.5 to 5.9.

5.5 Statistical procedure

The procedure used to determine cross sections for $t\bar{t}W$ and $t\bar{t}Z$ is similar to that used for the LHC Higgs boson analysis [14, 75]. The statistical analysis was based on previously described methods [76, 77, 78].

The quantum mechanical nature of events produced at the LHC is such that the outcome is characterized by *random variables*; that is, the result of a single event cannot be predicted deterministically even if the parameters, H , of the Lagrangian describing the system are completely known. With complete knowledge of H , however, accurate predictions can be made about the statistical distribution of the outcomes of a large number of such experiments. The

TABLE 5.5

INPUT VARIABLE QUALITY RANKING¹: SS $t\bar{t}W$ VS. $t\bar{t}$ BDT

	Rank (3 jet)	Rank (≥ 4 jets)
M_T of $\vec{p}_T^{\text{miss}}$ and $\vec{p}_T$ of leptons and jets	1	1
p_T^{miss}	4	2
Second-highest lepton p_T	6	3
Match score for $t\bar{t} \rightarrow \ell_b q \bar{q} \bar{\ell} \nu$	2	4
Highest lepton p_T	5	5
Second-highest CSV value of a jet	8	6
$t\bar{t}$ matched top quark M_T from $b\ell\nu$	7	7
Match score for $t\bar{t}W \rightarrow b\ell\nu\bar{b}q$	9	8
Match score for $t\bar{t}W \rightarrow b\ell\nu\bar{b}q\bar{q}$	-	9
$t\bar{t}$ matched top quark mass from $\ell_b q\bar{q}$	3	-

¹ Rank refers to variable's position in the list of variables, ordered by the quality of the signal to background separation.

description of the statistical distribution of a random variable is called a probability density function; an integral of a probability density function for some variable over some interval results in the probability of the variable having a value that falls within that interval.

In experimental studies, however, we are often faced with the opposite situation. We can conduct an experiment hundreds or thousands of times and observe the statistical distribution of the outcome. However we do not, *a priori*, know the parameters H of the hypothesized mechanism. How can we decide what value for H represents the best possible description of the physical system?

To answer this question, we can utilize two similar but distinct functions. Consider a hypothesis characterized by one or more parameters H . Given H , the probability of obtaining

TABLE 5.6

INPUT VARIABLE QUALITY RANKING¹: 3ℓ $t\bar{t}W$ VS. $t\bar{t}$ BDT

	Rank (1 jet)	Rank ≥ 2 jets
Second-highest CSV value of a jet	-	1
M_T of $\vec{p}_T^{\text{miss}}$ and $\vec{p}_T$ of leptons and jets	1	2
Match score for $t\bar{t}W \rightarrow \ell\nu b\bar{\ell}\nu\bar{b}\ell\nu$	-	3
Second-highest SS lepton p_T	4	4
$t\bar{t}$ matched top quark mass from ℓ_W and ℓ_b	-	5
Highest SS lepton p_T	3	6
Match score for $t\bar{t}W \rightarrow \ell\nu b\bar{\ell}\nu\ell\nu$	2	-
p_T^{miss}	5	-
Jet p_T	6	-

¹ Rank refers to variable's position in the list of variables, ordered by the quality of the signal to background separation.

an outcome x is known as the probability density function $p(x|H)$. Often, however, H is unknown; we instead observe x and try to determine the most plausible H (note that we are now viewing x as a function of H). A natural choice is to choose H such that the probability of observing x is maximized; that is, we maximize the likelihood function $L(H)$, defined as $L(H) = p(x|H)$.

A likelihood function is constructed as follows. Denote the expected number of signal events predicted by the SM as s , and the expected number of background events as b . We introduce the signal strength modifier $\mu = \frac{\sigma}{\sigma_{\text{SM}}}$, such that $\mu = 0$ corresponds to the background-only hypothesis, and $\mu = 1$ corresponds to the SM hypothesis. The total number of expected events is $\mu s + b$. The likelihood function is the conditional probability of obtaining the observed number of data events n_{obs} given that we expect $\mu s + b$ events.

TABLE 5.7

INPUT VARIABLE QUALITY RANKING¹: 3ℓ $t\bar{t}Z$ VS. $t\bar{t}$ AND WZ BDT

	Rank (3 jet)	Rank (≥ 4 jets)
Match score for $t\bar{t}Z \rightarrow \ell\ell b\ell\nu\bar{b}q$	1	1
Match score for $t\bar{t}Z \rightarrow \ell\ell b\ell\nu\bar{b}q\bar{q}$	-	2
Match score for $t\bar{t}Z \rightarrow \ell\ell \ell\nu\bar{b}q\bar{q}$	8	3
Match score for $t\bar{t}Z \rightarrow \ell\ell b\ell\nu q\bar{q}$	9	4
Number of medium b-tagged jets	3	5
Mass of lepton pair matched to Z boson	7	6
M_T of $\vec{p}_T^{\text{miss}}$ and $\vec{p}_T$ of leptons and jets	4	7
Match score for $t\bar{t}Z \rightarrow \ell\ell b\ell\nu\bar{b}$	2	-
Match score for $t\bar{t}Z \rightarrow \ell\ell \ell\nu\bar{b}q$	5	-
Match score for $t\bar{t}Z \rightarrow \ell\ell b\ell\nu q$	6	-

¹ Rank refers to variable's position in the list of variables, ordered by the quality of the signal to background separation.

TABLE 5.8

INPUT VARIABLE QUALITY RANKING^A: OS $t\bar{t}Z$ VS. $t\bar{t}$ BDT^B

	Rank (5 jet)	Rank (≥ 6 jets)
ΔR between leptons	1	1
p_T of dilepton system	2	2
Dilepton invariant mass	3	3
H_T^{miss}	4	4
Match score for $t\bar{t} \rightarrow b\ell\nu\bar{b}\ell\nu$	5	5
Number of jets with $p_T > 40$ GeV	9	6
Match score for $t\bar{t}Z \rightarrow \ell\ell b\bar{q}\bar{q}\bar{b}q\bar{q}$	-	7
Match score for $t\bar{t}Z \rightarrow \ell\ell b\bar{q}\bar{q}\bar{b}q$	8	8
Match score for $t\bar{t}Z \rightarrow \ell\ell b\bar{q}\bar{q}\bar{b}q$	7	9
Ratio of M_T to mass of jets	6	10
CSV of jet matched to b from $t\bar{t}$	11	11
CSV of jet matched to $\bar{b}$ from $t\bar{t}$	10	12

^A Rank refers to variable's position in the list of variables, ordered by the quality of the signal to background separation.

^B This BDT was used as input to the final discriminant.

TABLE 5.9

INPUT VARIABLE QUALITY RANKING^A: OS $t\bar{t}Z$ VS. Z AND $t\bar{t}$ BDT^B

	Rank (5 jet)	Rank (≥ 6 jets)
OS $t\bar{t}Z$ vs. $t\bar{t}$ BDT	1	1
Match score for $t\bar{t}Z \rightarrow \ell\ell b\bar{q}b\bar{q}$	3	2
Match score for $t\bar{t}Z \rightarrow \ell\ell b\bar{q}\bar{b}q$	4	3
Match score for $t\bar{t}Z \rightarrow \ell\ell b\bar{q}\bar{b}q\bar{q}$	-	4
Minimum χ^2 for $t\bar{t}Z \rightarrow \ell\ell b\bar{q}\bar{b}q\bar{q}$	-	5
Number of jets with $p_T > 40$ GeV	6	6
Fifth-highest jet p_T	5	7
Ratio of M_T to mass of jets and leptons	2	8
Second-highest jet CSV	7	9
Highest jet CSV	8	10

^A Rank refers to variable's position in the list of variables, ordered by the quality of the signal to background separation.

^B This BDT was used as the final discriminant.

The number of observed events n_i for each bin i of the final discriminant in each channel follows a Poisson distribution with mean $\mu s_i + b_i$. The n_i are statistically independent, so the likelihood function¹ is the product of Poisson probabilities over M bins:

$$L(\mu) = \prod_{i=1}^M \frac{(\mu s_i + b_i)^{n_i}}{n_i!} e^{-(\mu s_i + b_i)}.$$

Various uncertainties are associated with our model. We introduce a *nuisance parameter* to parameterize each uncertainty and denote the set as θ ; the expected signal and background yields become $s(\theta)$ and $b(\theta)$. When channels have the same uncertainty, it is represented with the same nuisance parameter. This correlation allows for bins in *control regions* with many data events but few expected signal events to constrain large uncertainties. The conditional probability to measure a nuisance parameter to be $\tilde{\theta}_i$, given that the true value is θ_i , is encoded in the probability density function $\rho(\tilde{\theta}_i|\theta_i)$. For uncertainties that must be positive (like cross sections and luminosities), a log-normal distribution for ρ is typically used. For statistical uncertainties on the number of events in a control region, we usually use the gamma distribution. Denoting the probability density function for the full set of nuisance parameters as $\rho(\tilde{\theta}|\theta)$, we have the following:

$$L(\mu, \theta) = \mathcal{P}(\text{data}|\mu, \theta) \rho(\tilde{\theta}|\theta) = \prod_{i=1}^M \frac{(\mu s_i + b_i)^{n_i}}{n_i!} e^{-(\mu s_i + b_i)} \rho(\tilde{\theta}|\theta). \quad (5.2)$$

We denote $\hat{\theta}$ and $\hat{\mu}$ to be the values of θ and μ that globally maximize L , referred to also as the best fit (or post-fit) values. But $\hat{\mu}$ by itself cannot reveal the whole story. We have found the

¹Technically, this is the extended likelihood function [78]. The standard maximum likelihood method maximizes $L = \prod_{i=1}^M p(x_i; a_1..a_n)$ where p is the probability density (normalized to 1), M is the number of events, x is the measured quantity, and a_i are the parameters to be determined. The fit only determines the shape and indicates nothing about the number of events. In the extended maximum likelihood, the predicted N is a function of the parameters, and the fit determines shape and size: p is replaced by P with $\int P(x_i|a_1..a_n) dx_i = N(a_1..a_n)$. For example, say the lifetime of a light bulb is modeled by an exponential distribution with lifetime λ . The lifetime x of M light bulbs can be tested and the result used to calculate the maximum likelihood estimate of λ . In this case, M is not dependent on λ , and the probability $p(x|\lambda) = \lambda e^{-\lambda x}$ must be normalized to 1 (if a measurement is made, *some* lifetime will be observed). For our cross section measurements, on the other hand, the number of observed events itself is relevant to the cross section being measured, so it should appear in the likelihood function.

best μ *relative* to the other possible ones our model permits, but it may just be the best choice from among very bad choices. We wish to define some scalar function of the data whose value encodes the level of compatibility between the data and a hypothesized value of μ . Furthermore, we wish to remove the dependence on the nuisance parameters. We can accomplish this approximately by making use of the *profile likelihood ratio*, which is only a function of μ :

$$\lambda(\mu) = \frac{L(\mu, \hat{\theta}(\mu))}{L(\hat{\mu}, \hat{\theta})} \quad (5.3)$$

where $\hat{\theta}(\mu)$ is the value of θ that maximizes L for a given μ (and thus a function of μ itself). The profile likelihood can take on values $0 \leq \lambda \leq 1$, with higher λ implying better agreement with the data. Noting that the log of a function will be maximized at the same point as the function itself, we can equivalently define the test statistic² $t_\mu = -2 \ln \lambda(\mu)$, which, according to Wilks' theorem, has the advantage of approaching a χ^2 distribution for large data samples independently of θ [77]. As t_μ gets larger, the incompatibility between the data and μ increases. In this analysis, we wish to determine if we have discovered a signal. We consider μ to be physically bounded by $\mu \geq 0$, and we want to reject the background-only hypothesis that $\mu = 0$, so we define

$$q_0 = \begin{cases} -2 \ln \lambda(0) = -2 \ln \frac{L(\hat{\theta}(0))}{L(\hat{\mu}, \hat{\theta})} & \hat{\mu} \geq 0, \\ 0 & \hat{\mu} < 0, \end{cases}$$

where $\hat{\theta}(0)$ correspond to the nuisance parameter values that maximize L under the background-only hypothesis. As the yield gets larger than the background, the incompatibility between the data and the background-only hypothesis increases. We now calculate the p-value

$$p_0 = P(q_0 \geq q_0^{\text{obs}}) = \int_{q_0^{\text{obs}}}^{\infty} f(q_0|0) dq_0, \quad (5.4)$$

²Sometimes the quantity $\ln \lambda = \ln L(\mu, \hat{\theta}(\mu)|\text{data}) - \ln L(\hat{\mu}, \hat{\theta}|\text{data})$ is referred to as $\Delta \ln L$.

which reveals what the probability that q_0 would be at least as large as the one we observe, q_0^{obs} , under the background-only hypothesis; in other words, it is the probability that the background events look as signal-like as those in the data. We define the significance Z such that the probability is p_0 to find a Gaussian distributed variable Z standard deviations σ above the mean:

$$p_0 = \int_Z^\infty \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx.$$

In high energy physics, the convention is to consider $Z = 5$, corresponding to $p = 2.87 \times 10^{-7}$, as the threshold for discovery. In other words, a discovery corresponds to an observation that, to be consistent with the background-only hypothesis, would require a statistical fluctuation that is expected to occur fewer than three times out of every 10 million otherwise identical experiments.

To evaluate Equation (5.4), we need to know $f(q_0|0)$. We could generate MC pseudo-data by sampling from $\mathcal{P}(\text{data}|\mu, \theta)$ and $\rho(\tilde{\theta}|\theta)$ around $\mu=0$; however, we instead use the approximation from [77], which draws on the fact that $f(q_0|0)$ approximates a χ^2 distribution for large datasets.

We would also like to determine confidence intervals; in other words, ranges of possible values of μ that have a given probability of containing the true value of μ . Quantifying the absence of a signal will be required, so we consider a closely related test statistic:

$$q_\mu = \begin{cases} -2 \ln \lambda(\mu) & \hat{\mu} \leq \mu, \\ 0 & \hat{\mu} > \mu. \end{cases} \quad (5.5)$$

We require that $q_\mu = 0$ when $\hat{\mu} > \mu$ so that upward fluctuations of the data such that $\hat{\mu} > \mu$ are not considered evidence against the signal hypothesis. Denoting the cumulative distribution function of q_μ as $F(q_\mu|\mu)$, the p-value is

$$p_\mu = P(q_\mu \geq q_\mu^{\text{obs}}) = 1 - F(q_\mu|\mu) = \int_{q_\mu^{\text{obs}}}^\infty f(q_\mu|\mu) dq_\mu.$$

If $p_\mu \leq \alpha$ for a specified probability $1 - \alpha$, then the corresponding μ is said to be excluded with a confidence level (CL) of $1 - \alpha$, and the set of points that is not excluded create the $1 - \alpha$ CL interval. We can find the endpoints of the interval by setting $p_\mu = \alpha$ and solving for μ . Recall that $2 \ln \lambda = 2\Delta \ln L$ approaches the χ^2 distribution with n degrees of freedom for large samples, where n is the number of parameters being estimated:

$$1 - \alpha = 1 - p_\mu = F(q_\mu | \mu) \approx (\chi^2; n). \quad (5.6)$$

We can then find the interval endpoints by evaluating the quantile function (the inverse of the cumulative distribution function) $F^{-1}(1 - \alpha)$. For estimations with one degree of freedom, a 68.27 % CL corresponds to $-2\Delta \ln L = 1.00$ and a 95 % CL corresponds to $-2\Delta \ln L = 3.84$.

5.6 Systematic uncertainties

The use of nuisance parameters for encoding uncertainties was introduced in Section 5.5. In the current section, the various sources of systematic uncertainties are discussed in more detail. Unlike statistical uncertainties, systematic uncertainties do not get smaller as more data are accrued. Systematic uncertainties arise from uncertainties in the measurement process and theoretical models, for example, because of imperfectly known detector performance, discriminant efficiencies, and SM parameters. We must examine the associated uncertainties and how they affect the final determination of cross sections of the various channels. For $t\bar{t}W$, the largest uncertainties come from the CSV shape, theoretical uncertainties in the signal modeling, and the rate of nonprompt backgrounds. For $t\bar{t}Z$, the largest uncertainties are associated with the CSV shape, signal modeling, and the rates of nonprompt backgrounds and prompt Z, WZ, and ZZ with extra jets. *Rate uncertainties* affect the rate of a process, which alters each bin of the final discriminant by the same value. *Shape uncertainties* affect the shape of variables and can alter bins separately, changing the shape of the discriminant.

Integrated luminosity and pileup There was a 2.6 % uncertainty on the integrated lu-

minosity [79], which was correlated across the entire analysis. The total inelastic proton–proton cross section was varied up and down by 5 %, which affected the number of pileup vertices, and was propagated to the output distributions [80].

Jet energy scale To account for uncertainty on the energy assigned to jets (the jet energy scale, or JES [81]), we computed the MC with the JES shifted up and down by one standard deviation and used the resulting rates and distributions to define the uncertainty.

b tagging efficiency We estimated the number of non-b jets in the data region used to derive the b jet CSV shape. We also estimated the number of bottom and charm jets in the data region used to derive the event weights for light-flavor jets. The uncertainties associated with those estimates were assessed by shifting the expected yields of contaminated jets up and down by one standard deviation. An additional source of uncertainty was associated with the finite number of data events used to calculate the individual weights in each bin of p_T , η , CSV, and flavor. We approximated these by using an *envelope* of maximum linear and quadratic shifts. The linear shift for a given bin was given by $f_l = \sigma(1 - 2x)$, where x was the central CSV value of the bin and σ represents the statistical uncertainty on that bin. The quadratic shift corresponded to $f_q = \sigma(1 - 6x + 6x^2)$. The SFs for each bin were then shifted within the envelope defined by $\pm f_l$ and $\pm f_q$. Uncertainties were not separately derived for c jets, rather they were assigned an uncertainty twice that of the b jets.

Selection efficiency for prompt leptons There was uncertainty associated with the SFs used to correct for differences in reconstruction and selection efficiency for prompt leptons between the data and MC. Because the calibration region had so many events, these uncertainties were small. There was uncertainty associated with the efficiency to reconstruct and select prompt leptons of 1.5 % per lepton. The dilepton trigger efficiencies had an associated 3 % uncertainty in the rate, while three-lepton events had an associated 1 % rate uncertainty. Four-lepton events had no trigger efficiency uncertainty.

Selection efficiency for nonprompt leptons A 40 % uncertainty was assigned to the rate of nonprompt electrons and a 60 % uncertainty to the rate of nonprompt muons passing the tight selection, based on the agreement between expected and observed yields in control regions. Most nonprompt leptons have low p_T , so uncertainties of 50 % were applied to the rates of 20–30 GeV muons and 20–40 GeV electrons, and an additional 100 % to muons with $p_T > 30$ GeV and electrons with $p_T > 40$ GeV. Because the SS and 3ℓ channels may have different sources of nonprompt leptons, the uncertainties were applied separately for electrons and muons, and were uncorrelated between the SS and 3ℓ channels. The final discriminant contained bins with mostly nonprompt backgrounds; such bins were useful in order to constrain these uncertainties. The final fit constrained them to 10–15 %.

Rate of charge misidentified electrons In the SS channel, a 30 % rate uncertainty was assessed on the rate of charge misidentified electrons based on the level of agreement in the control region of SS ee events consistent with a Z boson decay.

Top quark p_T reweighting To assess the uncertainty due to the top quark p_T reweighting of $t\bar{t}$ MC, we obtained a lower bound by removing the p_T weighting, and an upper bound by assessing twice the weight.

Prompt backgrounds with extra HF jets Rate uncertainties were applied to account for the SFs applied to correct the yields in Z events with five or more jets and in WZ and ZZ events with three or more jets; these uncertainties were relatively large because of the relatively small number of events in the light-flavor control region. Events with a Z boson plus five jets or six or more jets received uncorrelated 30 % rate uncertainties. WZ and ZZ events received a 40 % uncertainty if they had three jets, and a 60 % uncertainty if they had four or more jets. The $Z+b\bar{b}$, $Z+b$, and $Z+b\bar{b}$ uncertainties were constrained to 30 % each, owing to good agreement in dileptonic Z boson events with four jets. To account for the deviations between data and MC seen in Z boson events with four or more jets (excluding the $t\bar{t}Z$ signal region) in the distribution of H_T^{miss} , an additional shape uncertainty was added by weighting events by $w = 1 \pm 0.005(H_T^{\text{miss}} - 30)$. Mismodeling was also observed in the ratio between the M_T and the invariant mass of the reconstructed jets and leptons in an event. This variable measured the centrality of the Z boson decay products. To account for this deviation, an additional shape uncertainty was added by weighting events by $w = 1.0 \pm 0.7(\frac{M_T}{m_{\text{mass}}} - 0.65)$. A control region of OS $e\mu$ events with five or more jets showed similar mismodeling in the M_T -to-mass ratio, so the same shape uncertainty was applied to $t\bar{t}$ events.

PDF uncertainties A rate uncertainty associated with the choice of PDF of 7.2 % for $t\bar{t}W$ and 8.2 % for $t\bar{t}Z$ was assessed [82, 83]. In addition, shape uncertainties derived by using different PDF sets and PYTHIA tunes were applied to $t\bar{t}W$, $t\bar{t}Z$, and $t\bar{t}H$ using envelope shape uncertainties of 10–11 %.

Renormalization and factorization scales There was a rate uncertainty of 10 % for $t\bar{t}W$ and 11 % for $t\bar{t}Z$, associated with the choice of renormalization and factorization scales [84].

Rare processes Expected yields are small for rare processes, including WWW, WWZ, $t\bar{t}Z$, $t\bar{t}\gamma/t\bar{t}\gamma^*$, and $t\bar{t}WW$. These have not yet been measured, and some have not been calculated at NLO. Consequently, they received conservative 50 % rate uncertainties.

5.7 Results

The expected signal and background yields, obtained by performing the fit described in Section 5.5, are presented in Tables 5.10 to 5.12.

We performed a one-dimensional fit for the $t\bar{t}W$ cross sections, with the $t\bar{t}Z$ cross section set to the SM prediction with an uncertainty equal to the uncertainty in the theory prediction, and vice versa. The $t\bar{t}Z$ cross section was measured to be 242^{+65}_{-55}fb with a significance of

6.4 standard deviations from the background-only hypothesis, which agrees well with the SM prediction. The $t\bar{t}W$ cross section was higher than expected, measured to be $382^{+117}_{-102}\text{fb}$ with a significance of 4.8 standard deviations from the background-only hypothesis. The discrepancy was driven by an excess of SS dimuon data events. A similar excess was seen in the CMS $t\bar{t}H$ search [23], also mostly due to the same dimuon events. No evidence of mismodeling was found. The best fit values were compatible with the SM expectation at the 13 % CL for $t\bar{t}W$ and at the 60 % level for $t\bar{t}Z$. These results are summarized in Sections 5.7 and 5.7. The post-fit plots, which have the nuisances and signal strengths set to their best fit values, are presented in Figures 5.2 and 5.3.

We also used all channels to perform a fit to both of the $t\bar{t}W$ and $t\bar{t}Z$ cross sections simultaneously. The two-dimensional likelihood scan in $(\sigma_{t\bar{t}Z}, \sigma_{t\bar{t}W})$ plane is presented in Figure 5.4. The best fit values from the simultaneous fit are close to those for the one-dimensional fits and compatible with the SM at the 15 % CL.

TABLE 5.10

POST-FIT YIELDS FOR THE OS $t\bar{t}Z$ CHANNEL

	$ee / \mu\mu$		$e\mu$	
	5 jets	≥ 6 jets	5 jets	≥ 6 jets
Z + light flavor jets	265 ± 57	93 ± 20	< 0.1	< 0.1
Z + $b\bar{b}$ jets	341 ± 74	106 ± 23	< 0.1	< 0.1
Z + b jet	236 ± 59	68 ± 18	< 0.1	< 0.1
Z + $b\bar{b}$ jets	378 ± 72	136 ± 25	< 0.1	< 0.1
$t\bar{t}$ + light flavor jets	188 ± 19	58.4 ± 7.3	180 ± 16	57.8 ± 6.4
$t\bar{t}$ + heavy flavor jets	57 ± 16	30.6 ± 8.3	52 ± 15	27.3 ± 7.3
tbZ/ $t\bar{t}$ WW	4.2 ± 1.8	1.8 ± 0.7	< 0.1	< 0.1
$t\bar{t}H$	1.4 ± 0.1	1.0 ± 0.2	1.0 ± 0.1	0.6 ± 0.1
Background total	1 470 ± 135	494 ± 45	233 ± 21	85.8 ± 9.7
$t\bar{t}Z$	24.0 ± 5.5	28.2 ± 6.8	1.3 ± 0.3	0.8 ± 0.2
$t\bar{t}W$	1.1 ± 0.2	0.5 ± 0.1	1.2 ± 0.2	0.8 ± 0.2
Expected total	1 495 ± 135	523 ± 45	236 ± 21	87.4 ± 9.7
Data	1 493	526	251	78

TABLE 5.11

POST-FIT YIELDS FOR THE SS $t\bar{t}W$ CHANNEL

	ee		e/μ		$\mu\mu$	
	3 jets	≥ 4 jets	3 jets	≥ 4 jets	3 jets	≥ 4 jets
Non-prompt	16.0 ± 3.7	12.9 ± 3.1	57.0 ± 5.4	40.5 ± 4.2	29.0 ± 4.7	26.0 ± 4.4
Charge-misidentified	3.3 ± 1.6	1.7 ± 0.8	2.9 ± 0.7	1.6 ± 0.4	-	-
WZ	1.6 ± 0.5	0.9 ± 0.3	4.5 ± 1.4	2.2 ± 0.8	3.1 ± 1.0	1.3 ± 0.5
ZZ	0.2 ± 0.1	0.1 ± 0.1	0.3 ± 0.1	0.2 ± 0.1	0.2 ± 0.1	0.1 ± 0.1
Multiboson ^a	0.8 ± 0.3	0.5 ± 0.2	1.5 ± 0.5	1.2 ± 0.4	1.2 ± 0.5	1.1 ± 0.4
$t\bar{t}Z/t\bar{t} + X^b$	1.4 ± 0.4	2.5 ± 1.3	4.1 ± 1.4	5.8 ± 2.2	0.9 ± 0.3	1.2 ± 0.4
$t\bar{t}H$	0.3 ± 0.1	1.4 ± 0.2	1.1 ± 0.1	4.0 ± 0.5	0.7 ± 0.1	3.0 ± 0.5
Background total	23.7 ± 4.1	20.1 ± 3.5	71.4 ± 5.8	55.4 ± 4.9	35.1 ± 4.8	32.8 ± 4.5
$t\bar{t}W$	5.5 ± 1.4	8.1 ± 1.9	13.9 ± 3.7	25.2 ± 5.5	10.4 ± 2.8	17.7 ± 4.0
$t\bar{t}Z$	0.4 ± 0.1	1.3 ± 0.3	1.1 ± 0.2	3.0 ± 0.6	0.7 ± 0.1	2.1 ± 0.4
Expected total	29.6 ± 4.4	29.4 ± 4.0	86.4 ± 6.9	83.6 ± 7.3	46.2 ± 5.6	52.6 ± 6.0
Data	31	32	89	69	47	61

^a Includes WWW, WWZ, and $W^\pm W^\pm$.^b Includes $t\bar{t}\gamma$, $t\bar{t}\gamma^*$, and $t\bar{t}WW$.

TABLE 5.12

POST-FIT YIELDS FOR THE 3ℓ $t\bar{t}W$, 3ℓ $t\bar{t}Z$, AND 4ℓ $t\bar{t}Z$ CHANNELS

	3ℓ $t\bar{t}W$		3ℓ $t\bar{t}Z$		4ℓ $t\bar{t}Z$	
	1 jet	≥ 2 jets	3 jets	≥ 4 jets	≥ 1 jet + Z^a	≥ 1 jet + Z -veto ^b
Non-prompt	44.6 ± 5.3	54.8 ± 6.4	8.2 ± 2.8	5.4 ± 2.1	-	-
Non-prompt WZ/Z	-	-	-	-	< 0.1	< 0.1
Non-prompt $t\bar{t}$	-	-	-	-	< 0.1	0.2 ± 0.2
WZ	3.2 ± 0.8	8.0 ± 1.7	11.7 ± 2.9	5.4 ± 1.6	-	-
ZZ	1.0 ± 0.2	1.5 ± 0.3	1.6 ± 0.4	0.9 ± 0.3	3.3 ± 0.5	1.8 ± 0.3
Multiboson ^c	0.1 ± 0.1	0.4 ± 0.2	0.5 ± 0.2	0.5 ± 0.2	< 0.1	0.3 ± 0.1
$t\bar{b}Z/t\bar{t} + X^d$	0.4 ± 0.1	3.8 ± 1.1	1.6 ± 0.6	0.7 ± 0.3	< 0.1	< 0.1
$t\bar{t}H$	0.2 ± 0.1	4.7 ± 0.4	0.3 ± 0.1	0.4 ± 0.1	< 0.1	0.2 ± 0.1
Background total	49.5 ± 5.4	73.1 ± 6.7	23.9 ± 4.1	13.3 ± 2.7	3.3 ± 0.5	2.4 ± 0.4
$t\bar{t}W$	2.5 ± 0.8	18.8 ± 4.7	0.5 ± 0.1	0.2 ± 0.1	-	-
$t\bar{t}Z$	0.3 ± 0.1	7.5 ± 1.2	8.8 ± 1.9	16.9 ± 3.6	0.4 ± 0.1	4.3 ± 1.0
Expected total	52.3 ± 5.4	99.4 ± 8.3	33.2 ± 4.5	30.4 ± 4.5	3.7 ± 0.5	6.7 ± 1.1
Data	51	97	32	30	3	6

^a Requires exactly two lepton pairs consistent with a Z boson decay.^b Requires exactly one lepton pair consistent with a Z boson decay.^c Includes WWW, WWZ.^d Includes $t\bar{t}\gamma$, $t\bar{t}\gamma^*$, and $t\bar{t}WW$.

TABLE 5.13

MEASURED¹ CROSS SECTIONS AND SIGNIFICANCES FOR $t\bar{t}W$

	Cross section (fb)		Signal strength (μ)		Significance (σ)	
	Expected	Observed	Expected	Observed	Expected	Observed
SS	203^{+88}_{-73}	414^{+135}_{-112}	$1.00^{+0.45}_{-0.36}$	$2.04^{+0.74}_{-0.61}$	3.4	4.9
3ℓ	203^{+215}_{-194}	210^{+225}_{-203}	$1.00^{+1.09}_{-0.96}$	$1.03^{+1.07}_{-0.99}$	1.0	1.0
SS + 3ℓ	203^{+84}_{-71}	382^{+117}_{-102}	$1.00^{+0.43}_{-0.35}$	$1.88^{+0.66}_{-0.56}$	3.5	4.8

¹ Includes 68 % CL ranges.

TABLE 5.14

MEASURED¹ CROSS SECTIONS AND SIGNIFICANCES FOR $t\bar{t}Z$

	Cross section (fb)		Signal strength (μ)		Significance (σ)	
	Expected	Observed	Expected	Observed	Expected	Observed
OS	206^{+142}_{-118}	257^{+158}_{-129}	$1.00^{+0.72}_{-0.57}$	$1.25^{+0.76}_{-0.62}$	1.8	2.1
3ℓ	206^{+79}_{-63}	257^{+85}_{-67}	$1.00^{+0.42}_{-0.32}$	$1.25^{+0.45}_{-0.36}$	4.6	5.1
4ℓ	206^{+153}_{-109}	228^{+150}_{-107}	$1.00^{+0.77}_{-0.53}$	$1.11^{+0.76}_{-0.52}$	2.7	3.4
OS + 3ℓ + 4ℓ	206^{+62}_{-52}	242^{+65}_{-55}	$1.00^{+0.34}_{-0.27}$	$1.18^{+0.35}_{-0.29}$	5.7	6.4

¹ Includes 68 % CL ranges.

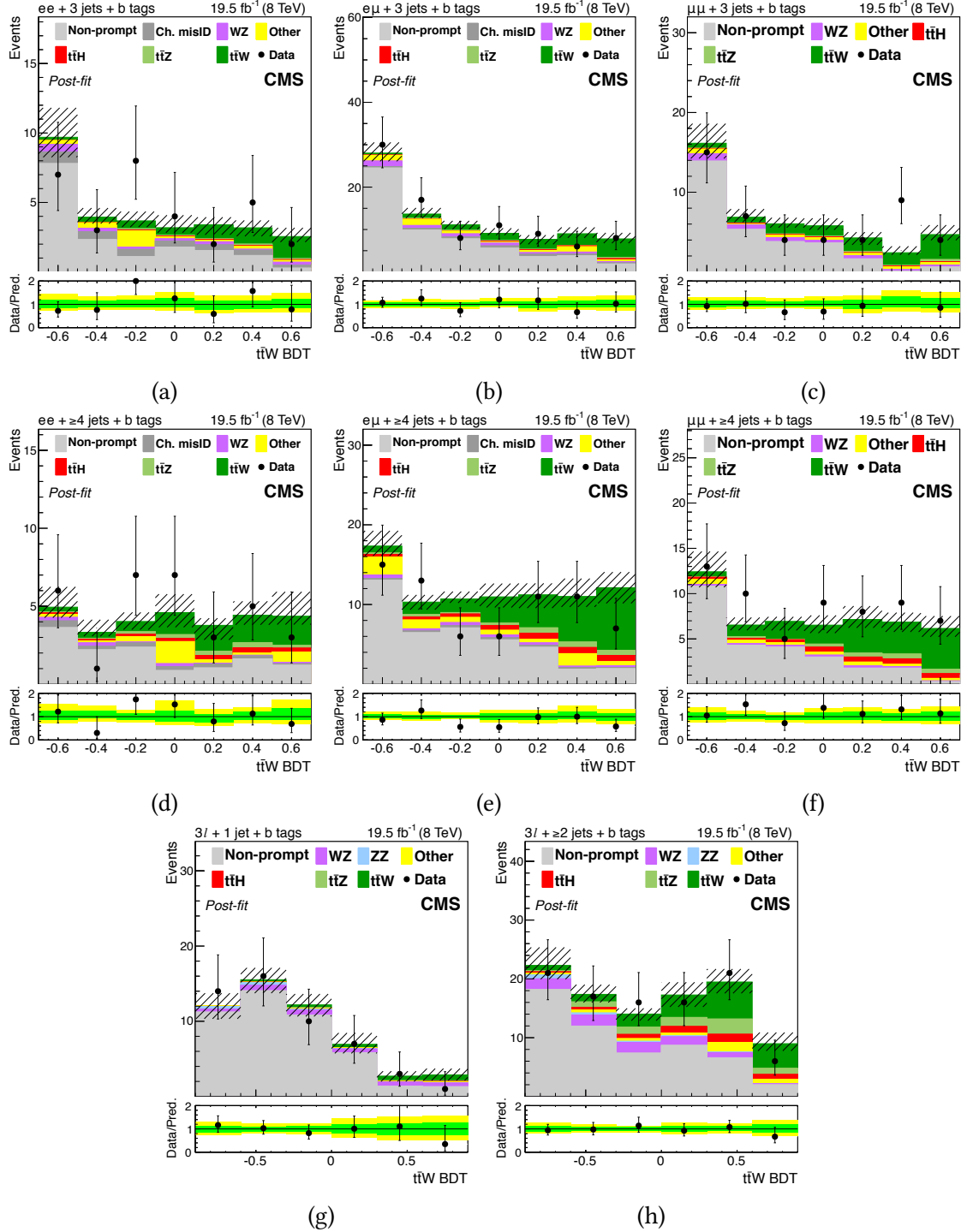


Figure 5.2. The post-fit final discriminant for events in the SS $t\bar{t}W$ channel with 3 jet (top row) and ≥ 4 jets (center row), for lepton flavors ee (a, d), $e\mu$ (b, e), and $\mu\mu$ (c, f). The post-fit final discriminant for events in the $3l$ $t\bar{t}W$ channel with 1 jet (g) and ≥ 2 jets (h). The hashed regions in the stack histogram indicate 68 % CL uncertainty on the signal plus background. The green and yellow shaded regions on the data-to-prediction ratio plot indicate 68 % and 95 % CL bands, respectively.

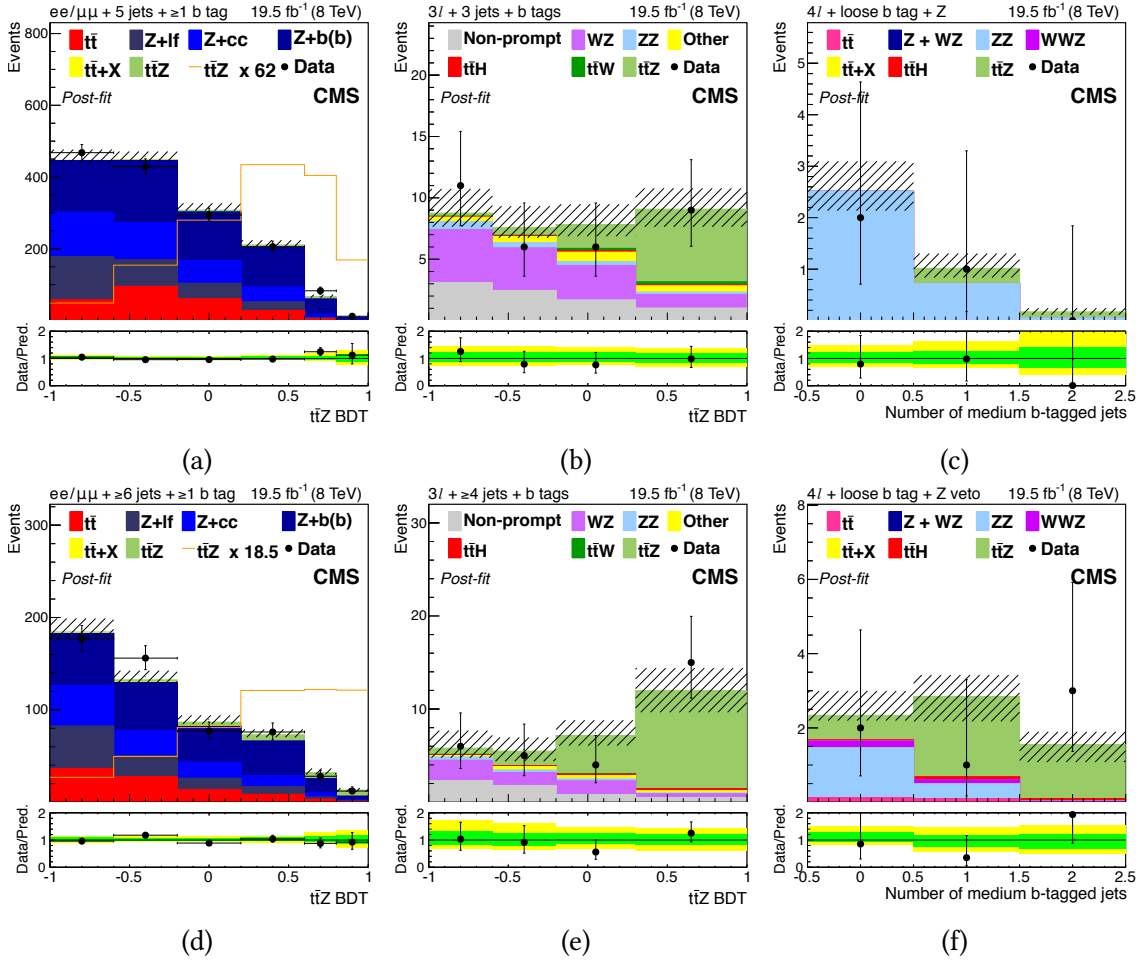


Figure 5.3. The post-fit final discriminant for events in the OS $t\bar{t}Z$ channel with two OS leptons and 5 jets (a) or ≥ 6 jets (d), three leptons and 3 jets (b) or ≥ 4 jets (e), or four leptons and two lepton pairs (c) or exactly one lepton pair (f) consistent with a $Z \rightarrow \ell\ell$ decay. The hashed regions in the stack histogram indicate 68 % CL uncertainty on the signal plus background. The green and yellow shaded regions on the data-to-prediction ratio plot indicate 68 % and 95 % CL bands, respectively.

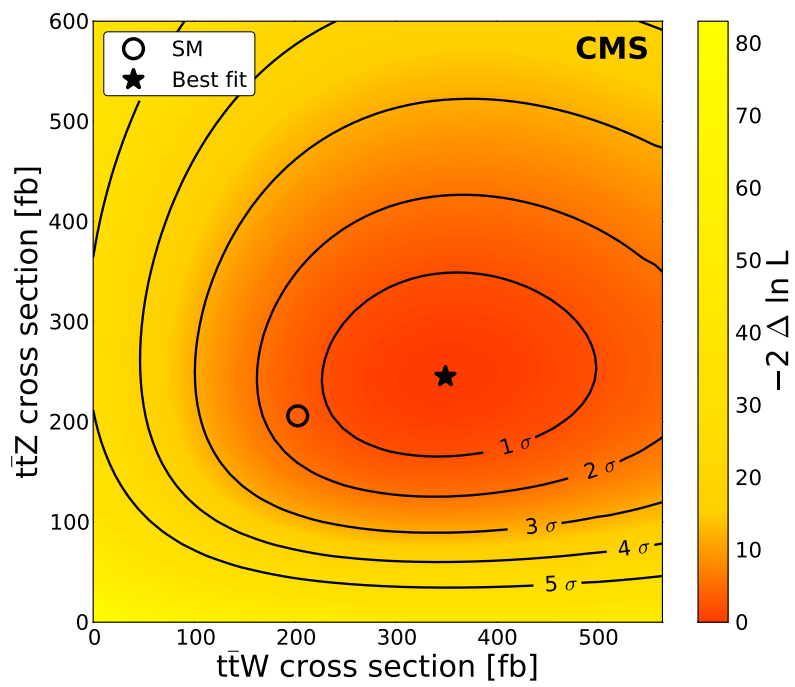


Figure 5.4. Profile likelihood as a function of $\sigma_{t\bar{t}W}$ and $\sigma_{t\bar{t}Z}$. Lines denote the 1 to 5 standard deviation (σ) CL contours.

CHAPTER 6

CROSS SECTION MEASUREMENT FOR TOP QUARK PAIR PRODUCTION IN ASSO- CIATION WITH A W OR Z BOSON IN PROTON-PROTON COLLISIONS AT $\sqrt{s}=13$ TeV

This chapter presents measurements of the $t\bar{t}W$ and $t\bar{t}Z$ cross sections performed using 35.9 fb^{-1} of 13 TeV pp collision data from the LHC. As in the analysis described in Chapter 5, $t\bar{t}W$ and $t\bar{t}Z$ events were targeted where the associated boson produces electrons or muons. The theoretical $t\bar{t}W$ and $t\bar{t}Z$ cross sections at 13 TeV are 3–4 times higher than at 8 TeV. This, combined with the higher integrated luminosity permitted a more precise measurement to be made, while simplifying some aspects: the OS $t\bar{t}Z$ and 3ℓ $t\bar{t}W$ channels present in the earlier analysis were not targeted, a full event reconstruction is not carried out, and a BDT was used only in the SS $t\bar{t}W$ channel.

The reinterpretation of this measurement within the context of effective field theory is described in Chapter 7.

6.1 Samples

The data-driven estimation of backgrounds due to nonprompt leptons relies on control regions of data collected, with single lepton triggers requiring the presence of at least one electron or muon with $p_T > 27 \text{ GeV}$ or $p_T > 24 \text{ GeV}$. To simplify the treatment of this background, data collected with the same triggers were used for signal regions. The trigger efficiencies exceeded 95 % for the SS $t\bar{t}W$ channel and 98 % for the 3ℓ and 4ℓ $t\bar{t}Z$ channels.

Expected signal events and some backgrounds were modeled with MC simulation. The simulated datasets and their cross sections are summarized in Table 6.1. The samples produced at LO in QCD used the NNPDF3.0LO [85] PDF, while those at NLO in QCD used

the NNPDF3.0NLO [85] PDF. Parton showering, hadronization, and the underlying event were simulated using PYTHIA v8.212 [86, 87] with the CUETP8M1 tune [88, 89]. GEANT4 software [69] was used to model the CMS detector response. The pileup distribution was simulated to match the one observed in the data.

6.2 Event selection

We searched for $t\bar{t}W$ and $t\bar{t}Z$ events by optimizing the selection process for three exclusive final states: $SS\ t\bar{t}W$, $3\ell\ t\bar{t}Z$, and $4\ell\ t\bar{t}Z$. These final states were described in Section 5.1 and have the same backgrounds and similar selection criteria. The selection criteria for each final state are summarized in Table 6.2. Each channel was divided into categories according to the number of jets and b-tagged jets, and for the $SS\ t\bar{t}W$ channel, additionally according to the score from a BDT and the lepton charges. The $t\bar{t}W$ BDT and categorization scheme for each channel are described in more detail below.

BDTs were described in Section 5.4. In the $SS\ t\bar{t}W$ channel, a BDT with gradient boosting [73] is used to help distinguish signal $t\bar{t}W$ events from backgrounds. The BDT is trained with events from MC $t\bar{t}W$ and $t\bar{t}$ samples, divided into equal samples for training and testing. The following variables were used as inputs: the number of jets, N_{jets} ; the number of b jets, N_b ; the scalar sum of p_T of the jets, H_T ; p_T^{miss} ; the highest- p_T (leading) and the lowest- p_T (trailing) lepton p_T ; the invariant mass calculated using p_T^{miss} and p_T of each lepton, M_T ; the highest (leading) and second highest (subleading) jet p_T ; and the separation ΔR between the trailing lepton and the nearest selected jet.

Events were divided into categories according to the BDT output score: $D < 0$ was used as a control region to constrain uncertainties, and the signal region is divided into $0 < D < 0.6$ and $D > 0.6$. These definitions were determined to provide the optimal expected sensitivity. Signal region events were additionally categorized according to whether they had two, three, or more than three jets; those with three or more jets were further divided according to whether they had one or more than one b-tagged jet. Because we are studying pp collisions,

TABLE 6.1

SIMULATED DATASETS AND CROSS SECTIONS

Process	QCD order	Generator ^a
$W\gamma^*$	LO	MG5_AMC@NLO
WW	LO	MG5_AMC@NLO
tWZ	LO	MG5_AMC@NLO
t $\bar{t}$ VV (V = W, Z, or H)	LO	MG5_AMC@NLO
t $\bar{t}$ Z	NLO	MG5_AMC@NLO
t $\bar{t}$ W	NLO	MG5_AMC@NLO
tZq	NLO	MG5_AMC@NLO
tHq	NLO	MG5_AMC@NLO
tHW	NLO	MG5_AMC@NLO
WZ	NLO	MG5_AMC@NLO
WWZ	NLO	MG5_AMC@NLO
WZZ	NLO	MG5_AMC@NLO
ZZZ	NLO	MG5_AMC@NLO
t $\bar{t}\gamma^*$	NLO	MG5_AMC@NLO
$Z\gamma^*$	NLO	MG5_AMC@NLO
t $\bar{t}$ H [90]	NLO	POWHEG
$q\bar{q} \rightarrow ZZ$ [91, 92]	NLO ^b	POWHEG
$gg \rightarrow ZZ$	LO ^c	MCFM

^a Software versions corresponded to MG5_AMC@NLO [93] v2.2.2 or v2.3.3, POWHEG v2 [94], and MCFM v7.0 [95].

^b Scaled to NNLO ($K = 1.1$) [96].

^c Scaled to NLO ($K = 1.7$) [97].

TABLE 6.2

SUMMARY OF SELECTION REQUIREMENTS FOR EACH CHANNEL

	SS $t\bar{t}W$	3ℓ $t\bar{t}Z$	4ℓ $t\bar{t}Z$
lepton p_T	$> \begin{cases} 27/40 \text{ GeV}(e) \\ 25 \text{ GeV}(\mu) \end{cases}$	$> 10/20/40 \text{ GeV}$	$> 10/10/10/40 \text{ GeV}$
jets ^a	2, 3, > 3	2, 3, > 3	≥ 2
b-tagged jets ^a	1, > 1	0, 1, > 1	0, ≥ 1
Z window	$ M_{ee} - M_Z > 15 \text{ GeV}$	OSSF pair with $ M_{\ell\ell} - M_Z < 10 \text{ GeV}$	OSSF pair with $ M_{\ell\ell} - M_Z < 20 \text{ GeV}$ and veto on additional OSSF pair
loose lepton veto			
charge ^a	$\ell^+ \ell^+, \ell^- \ell^-$		
p_T^{miss}	$> 30 \text{ GeV}$		
BDT ^a	$D > 0, D < 0$		

^a Entries separated by commas indicate events are categorized according to each listed criteria.

there is a charge asymmetry between $t\bar{t}W^+$ and $t\bar{t}W^-$ production, while the majority of the backgrounds produce charge-symmetric dileptons. We took advantage of this circumstance by further splitting the $D > 0$ events according to the lepton charge ($\ell^+ \ell^+$ or $\ell^- \ell^-$).

In the 3ℓ $t\bar{t}Z$ channel, events were divided into nine exclusive categories: events with two, three, or more than three jets, with each jet multiplicity being further split according to whether it had zero, one, or more than one b tag. The two jet category provided a background-dominated control region, which constrained uncertainties. Although signal $t\bar{t}Z$ events have at least four jets, we found that the expected sensitivity was improved by

including the three-jet categories, which have larger background contamination but allow for the recovery of events with merged jets or jets that fall outside the acceptance limits.

Events in the 4ℓ $t\bar{t}Z$ channel were categorized according to whether they contained zero or at least one b-tagged jet.

6.3 Event modeling

The background sources were described in Section 5.2. The modeling for each background is described below.

6.3.1 Nonprompt backgrounds

We used a data-based *tight-to-loose* method to model background from nonprompt leptons (introduced in Section 4.2.1), which are an important background source in the SS $t\bar{t}W$ and 3ℓ channels.

We measured the rate¹ f_{TL} at which loose leptons (see Table 4.2) also pass the tight criteria in a multijet control region enriched with nonprompt leptons (the *measurement region*). We selected events with a single lepton and one or more jets, where the lepton and jets were separated by $\Delta R > 1$. Contamination from prompt leptons (mainly from W +jets) was suppressed by requiring $p_{\text{T}}^{\text{miss}} < 20$ GeV and $M_{\text{T}} < 20$ GeV; any remaining contamination (with a maximum contribution on the order of a few percent for high- p_{T} leptons) was estimated from the simulation and subtracted.

Once f_{TL} was obtained, it was applied to a second control region (the *application region*). In the application region, events must pass the full selection criteria except that one lepton fails the tight selection but passes the loose requirements (as defined in Section 4.2.1.2). A weight for events in this region is calculated from f_{TL} to extrapolate to the signal region.

¹Note this rate is similar to f measured using a modified tag-and-probe technique and used in the 8 TeV analysis as described in Section 5.2.2. The different terminology is used to emphasize the different definitions involved.

Variations in the value of f_{TL} are driven by characteristics of the parent parton. To mitigate the effects of such differences, the values of f_{TL} are parameterized by lepton $p_{\text{T}}^{\text{corr}}$ and η , measured separately for electrons and muons. The corrected $p_{\text{T}}^{\text{corr}}$, preferred because it is highly correlated with the parent parton p_{T} , is defined as the sum of the lepton p_{T} and the energy in the isolation cone, which exceeds the isolation threshold value. The measured value varied between 2–16 %.

Similarly to Equation (5.1), events in the application region are weighted according to

$$w = (-1)^{i+1} \prod_{i=1}^N \frac{f_{\text{TL},i}}{1 - f_{\text{TL},i}},$$

where there are $N = 1, 2$, or 3 leptons that pass the loose criteria but fail the tight criteria and $f_{\text{TL},i}$ corresponds to f_{TL} evaluated according to the flavor, $p_{\text{T}}^{\text{corr}}$, and η of the i th lepton. The negative weights account for events with two nonprompt leptons contaminating the application region.

6.3.2 Charge-misidentified backgrounds

For the SS $t\bar{t}W$ final state, there is a nonnegligible background from isolated OS leptons (usually from $t\bar{t}$ or DY production) where the charge of one of the leptons has been misidentified. The charge mismeasurement rate is negligible for muons. A partially data-based approach is used to estimate the contribution to channels with one or more electrons. The charge mismeasurement rate, parameterized in p_{T} and η , is calculated from simulated DY and $t\bar{t}$ events. The rate is calculated as the number of SS events with a lepton that have different reconstructed and generator-level charges divided by the number of OS events. The contribution to the signal regions is then estimated by weighting OS ee or $e\mu$ events that pass the full kinematic selection.

6.3.3 WZ background

WZ is an important background for the $3\ell\ \bar{t}tZ$ analysis, especially in the bins with no b-tagged jets. This background was modeled using simulated events. To check the quality of our modeling, we examined a WZ-enriched control region in the data, in which we required events to have three leptons with $p_T > 40, 20, 10$ GeV. Two leptons had to form an OSSF pair with $|M(\ell\ell) - M_Z| < 10$ GeV. There also had to be fewer than two jets and no b-tagged jets. A leptonically decaying W produces a neutrino, so we required $p_T^{\text{miss}} > 30$ GeV and the transverse mass calculated with p_T^{miss} and the lepton that was not included in the $M(\ell\ell)$ calculation had to exceed 50 GeV. The purity of WZ events in this region was expected to be 85 %. Contamination from nonprompt leptons was modeled using the same data-driven method as in the signal region, and other backgrounds were modeled using simulation. The data and simulation were found to be in good agreement, with a ratio of 0.94 ± 0.07 (stat). Based on this good agreement, we did not apply any correction factor but propagated the statistical uncertainty on the ratio to the final estimation.

SFs, described in more detail in Refs. [98, 62], are applied to account for differences in b-tagging efficiencies and misidentification rates between the data and the simulation. To check the quality of our corrected model, we examined events with two OSSF lepton pairs that had $|M(\ell\ell) - M_Z| < 20$ GeV, both in the data and in the simulated production of a Z boson and extra partons. Based on the good agreement we observed, we assigned a 10 % systematic uncertainty on the WZ background estimation. An additional uncertainty of 20 % was applied for events with more than three jets in the $3\ell\ \bar{t}tZ$ analysis.

6.3.4 Rare backgrounds and $t(\bar{t})X$

Backgrounds that have at least one top quark in the final state are grouped under the label $t(\bar{t})X$, which includes $t\bar{t}H$, tWZ , tqZ , tHq , tHW , $t\bar{t}VV$, and $t\bar{t}t\bar{t}$. All others are grouped as rare backgrounds, which includes WW , ZZ , $W\gamma^*$, $Z\gamma^*$, WWW , WWZ , WZZ , and ZZZ . Rare and $t(\bar{t})X$ backgrounds were modeled using simulated events scaled by their NLO cross

sections and normalized to the integrated luminosity from data.

6.4 Systematic uncertainties

Sources of systematic uncertainties were similar to those in Section 5.6 and are described in further detail below.

Integrated luminosity and pileup There was a 2.5 % uncertainty based on the integrated luminosity [99]. This was correlated across the entire analysis. The uncertainty based on the total inelastic proton–proton cross section was 5 % [100], which affected the number of pileup vertices, and the propagated uncertainty based on the expected yields was 1–2 %.

Selection efficiency for prompt leptons The efficiency to reconstruct and select leptons, parameterized in p_T and η , was measured using a tag-and-probe [101, 102] method. It was found to exceed 65 % for electrons and 96 % for muons. Measurements were made separately in the data and simulation and were found to agree within 1–4 % per lepton. The associated systematic uncertainties were between 2–7 %. The trigger efficiencies were measured in a control region in the data as well as the simulation, separately for each channel and parametrized as a function of lepton p_T and η . The efficiency exceeded 95 % for $t\bar{t}W$ and 98 % for the 3ℓ and 4ℓ $t\bar{t}Z$ final states. The results from the data and simulation agreed to within 1 % in all channels except the SS dimuon channel. In the SS dimuon channel, SFs were applied to correct the discrepancy, which was as large as 3 %. The systematic uncertainty due to this scaling varied between 2–4 %.

Jet energy scale and resolution To account for uncertainty on the JES, we compute the MC with the JES shifted up and down by one standard deviation (corresponding to 2–5 % [103, 104], depending on p_T and η), and used the resulting rates and distributions to define the uncertainty. Uncertainties associated with the JER were estimated with the same approach and found to vary between 1–6 %.

b tagging efficiency Differences in the efficiency to tag b jets between the data and the simulation were corrected for with SFs derived from the data. More details about this correction can be found in reference [98, 62]. To assess the uncertainty on predicted yields, the SFs were shifted up and down by one standard deviation according to the gen-matched parton flavor. The propagated uncertainties on the predicted yields varied between 2–5 %.

Rate of nonprompt leptons To evaluate the systematic uncertainty on estimating the nonprompt background, the f_{TL} was calculated from simulated multijet events and applied in simulated $t\bar{t}$ and Z +jets events. The magnitude and distribution of the nonprompt contribution was found to be well-reproduced. Performance of the method is additionally checked using nonprompt-enriched control regions in data. The region with $D < 0$ (also used in the final fit to constrain uncertainties) is used for the SS

$t\bar{t}W$ channel. Systematic differences between performance in the $D < 0$ and $D > 0$ regions were studied in simulation and found to be negligible. For the $3\ell t\bar{t}Z$ channel, a control region for validation is defined by vetoing events with OSSF pairs (targeting leptonically-decaying $t\bar{t}$ events with a nonprompt lepton), and separately by requiring an OSSF pair, and no b-tagged jets present (targeting DY production). Based on the good agreement between predicted and observed yields that was observed in these control regions, a systematic uncertainty of 30 % is assigned to the estimated contribution from nonprompt backgrounds.

Rate of charge-misidentified electrons To evaluate the systematic uncertainty on the estimation of the background due to charge-misidentified electrons, the misidentification rate was measured independently in the data and DY simulation events which has two electrons with an invariant mass around the Z boson mass window $76 \text{ GeV} < M(\ell\ell) < 106 \text{ GeV}$ as the ratio of SS to OS events. The mismeasurement rate increased with electron p_T and ranged from 4×10^{-5} in the barrel region to 4×10^{-3} in the endcaps. Based on the good agreement between the methods, a 20 % uncertainty was assigned to the estimation of this contribution.

Renormalization and factorization scales The uncertainty on the signal acceptance due to the renormalization and factorization scales was assessed by varying each parameter independently up and down by a factor of two using alternative weights that were saved in the simulation. The corresponding low and high yields in each signal region were assigned as the uncertainty, which did not exceed 2 %.

PDF uncertainties PDF replicas are produced by varying α_s and rederiving the PDFs with MC pseudodata generated about the experimental input data. Uncertainties associated with the PDF were derived by applying event weights corresponding to different replicas of the NNPDF30 PDF set [85] to estimate the uncertainty on final acceptance for $t\bar{t}W$ and $t\bar{t}Z$, which were typically less than 1 %.

Theory (background yields) There was an ≈ 11 % [105] uncertainty on the $t\bar{t}H$, $t\bar{t}Z$, and $t\bar{t}W$ cross sections associated with the QCD scale, choice of PDF, and electroweak corrections. For the WW and ZZ backgrounds, this uncertainty was 10 %. A conservative 50 % uncertainty was assigned for rare backgrounds.

We estimated the impact of each source of systematic uncertainty by fixing the associated nuisance parameter (introduced in Section 5.5) to -1σ and $+1\sigma$ around its best fit value, with all other parameters profiled as normal, and evaluating the shift in the best fit signal strength $\hat{\mu}$. The results tabulated for all uncertainty sources are summarized in Table 6.3. The largest effects on both the $t\bar{t}W$ and $t\bar{t}Z$ cross section measurements are due to uncertainties on the integrated luminosity, lepton identification, trigger selection efficiencies, nonprompt leptons, and $t(\bar{t})X$ backgrounds.

TABLE 6.3

SYSTEMATIC UNCERTAINTY IMPACTS

Source	Uncertainty from each source (%)	Impact on the measured $t\bar{t}W$ cross section (%)	Impact on the measured $t\bar{t}Z$ cross section (%)
Integrated luminosity	2.5	4	3
Jet energy scale and resolution	2–5	3	3
Trigger	2–4	4–5	5
B tagging	1–5	2–5	4–5
PU modeling	1	1	1
Lepton ID efficiency	2–7	3	6–7
Choice in μ_R and μ_F	1	< 1	1
PDF	1	< 1	1
Non-prompt background	30	4	< 2
WZ cross section	10–20	< 1	2
ZZ cross section	20	-	1
Charge misidentification	20	3	-
Rare SM background	50	2	2
$t(\bar{t})X$ background	10–15	4	3
Stat. unc. in non-prompt background	5–50	4	2
Stat. unc. in rare SM backgrounds	20–100	1	< 1
Total systematic uncertainty	-	14	12

6.5 Results

The statistical procedure used to perform the final fit is described in Section 5.5. Post-fit yields, with signal strengths and nuisance parameters set to their best fit values, for all processes are shown in Figures 6.1 and 6.2 and tabulated in Tables 6.4 to 6.7. Agreement between the observed and expected yields is good, with the exception of 3ℓ $t\bar{t}Z$ categories requiring two or three jets and more than one b-tagged jet, where a small excess is observed. Comprehensive studies of the excess were performed, and no evidence of mismodeling was found. Consequently, we ascribe this excess to a statistical fluctuation in the data.

One-dimensional fits were performed to find the best fit signal strength parameter for $t\bar{t}W$ using all SS $t\bar{t}W$ categories, while the $t\bar{t}W^+$ and $t\bar{t}W^-$ measurements used the $\ell^+\ell^+$ and $\ell^-\ell^-$ categories, respectively. The one-dimensional fit for the best fit signal strength parameter for $t\bar{t}Z$ was performed using the 3ℓ and 4ℓ categories. When the $t\bar{t}W$ cross section was being measured, the $t\bar{t}Z$ cross section was set to the expectation from theory, and vice versa. The $t\bar{t}W$ process was observed with a significance of 5.3σ , which is slightly higher than the expected significance of 4.5σ . The expected (observed) significances for the $t\bar{t}W^+$ and $t\bar{t}W^-$ processes were found to be 4.2σ (5.5σ) and 2.4σ (2.3σ), respectively. In the 4ℓ $t\bar{t}Z$ channel, the expected (observed) signal significance is 4.7σ (4.5σ). In the 3ℓ $t\bar{t}Z$ channel, the expected and observed significances were found to be much larger than 5 standard deviations. These results are summarized in Table 6.8. The signal strength parameters and their associated uncertainties were measured to be

$$\mu_{t\bar{t}W} = 1.23^{+0.19}_{-0.18}(\text{stat})^{+0.20}_{-0.18}(\text{syst})^{+0.13}_{-0.12}(\text{theory})$$

$$\mu_{t\bar{t}Z} = 1.17^{+0.11}_{-0.10}(\text{stat})^{+0.14}_{-0.12}(\text{syst})^{+0.11}_{-0.12}(\text{theory}).$$

These were multiplied by the SM predictions [105] to obtain the cross sections:

$$\sigma(\text{pp} \rightarrow \text{t}\bar{\text{t}}\text{W}) = 0.77^{+0.12}_{-0.11}(\text{stat})^{+0.13}_{-0.12}(\text{syst}) \text{ pb},$$

$$\sigma(\text{pp} \rightarrow \text{t}\bar{\text{t}}\text{Z}) = 0.99^{+0.09}_{-0.08}(\text{stat})^{+0.12}_{-0.10}(\text{syst}) \text{ pb}.$$

The measured cross sections for the $\text{t}\bar{\text{t}}\text{W}^+$ and $\text{t}\bar{\text{t}}\text{W}^-$ processes were

$$\sigma(\text{pp} \rightarrow \text{t}\bar{\text{t}}\text{W}^+) = 0.58 \pm 0.09 (\text{stat})^{+0.09}_{-0.08} (\text{syst}) \text{ pb},$$

$$\sigma(\text{pp} \rightarrow \text{t}\bar{\text{t}}\text{W}^-) = 0.19 \pm 0.07 (\text{stat}) \pm 0.06 (\text{syst}) \text{ pb}.$$

We also used all channels to perform a fit to both of the $\text{t}\bar{\text{t}}\text{W}$ and $\text{t}\bar{\text{t}}\text{Z}$ cross sections simultaneously using the SS $\text{t}\bar{\text{t}}\text{W}$, 3 ℓ $\text{t}\bar{\text{t}}\text{Z}$, and 4 ℓ $\text{t}\bar{\text{t}}\text{Z}$ channels together. The result of this fit and the corresponding 68 % and 95 % CL contours, along with the one-dimensional fits in the $(\sigma_{\text{t}\bar{\text{t}}\text{Z}}, \sigma_{\text{t}\bar{\text{t}}\text{W}})$ plane, are presented in Figure 6.3.

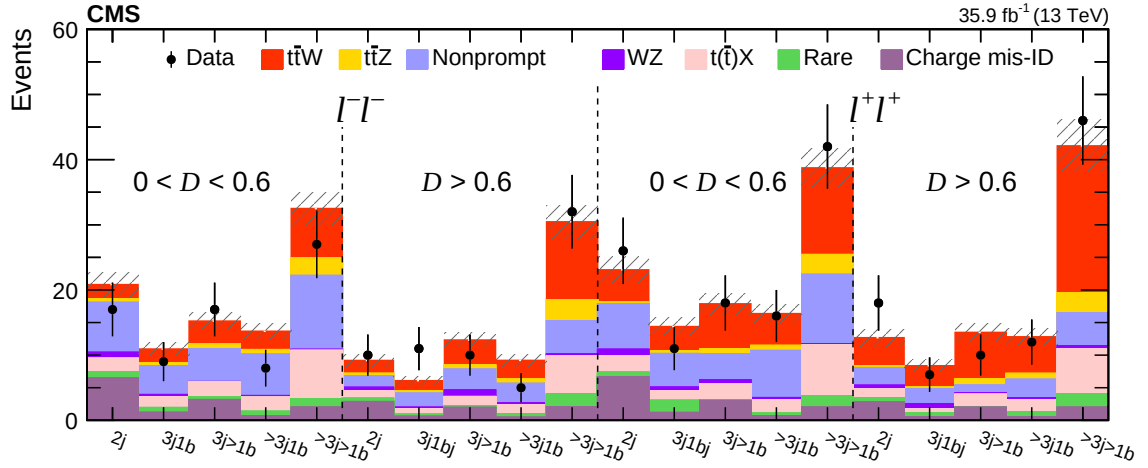


Figure 6.1. Predicted and observed post-fit yields for each category in the SS $t\bar{t}W$ channel. The total post-fit uncertainty is shown as a hatched band.

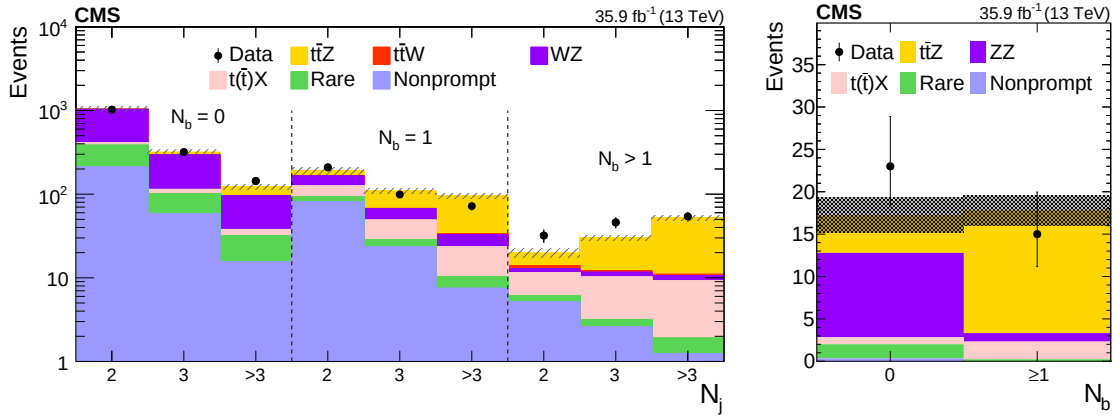


Figure 6.2. Predicted and observed post-fit yields for each category in the $3\ell t\bar{t}Z$ channel (left) and the $4\ell t\bar{t}Z$ channel (right). The total post-fit uncertainty is shown as a hatched band.

TABLE 6.4

POST-FIT YIELDS FOR THE SS $t\bar{t}W$ CHANNEL ($D < 0$)

Process	$N_{\text{jets}} = 2$	$N_{\text{jets}} = 3$	$N_{\text{jets}} > 3$
Nonprompt	136.5 ± 13.9	110.3 ± 11.3	57.3 ± 6.1
Total background	192.1 ± 15.6	137.7 ± 11.7	74.0 ± 6.4
$t\bar{t}W$	13.1 ± 0.3	17.6 ± 0.3	13.8 ± 0.3
$t\bar{t}Z$	1.6 ± 0.4	3.1 ± 0.7	4.4 ± 1.0
Total	206.8 ± 15.7	158.4 ± 11.8	92.3 ± 6.5
Observed	229	144	92

TABLE 6.5

POST-FIT YIELDS FOR THE SS $t\bar{t}W$ CHANNEL ($D > 0$)

		N_{jets}	N_b	Background	$t\bar{t}W$	$t\bar{t}Z$	Total	Data
$\ell^-\ell^-$	$0 < D < 0.6$	2	>0	18.1 ± 1.8	2.2 ± 0.4	0.5 ± 0.1	20.8 ± 1.9	17
		3	1	8.3 ± 0.9	2.1 ± 0.4	0.5 ± 0.1	10.9 ± 0.9	9
		3	>1	10.9 ± 1.1	3.5 ± 0.6	0.8 ± 0.1	15.2 ± 1.3	17
		>3	1	10.1 ± 1.1	2.8 ± 0.5	0.7 ± 0.2	13.7 ± 1.3	8
		>3	>1	22.2 ± 2.0	7.6 ± 1.2	2.7 ± 0.4	32.5 ± 2.4	27
$\ell^-\ell^-$	$D > 0.6$	2	>0	6.8 ± 0.9	2.0 ± 0.3	0.4 ± 0.1	9.2 ± 0.9	10
		3	1	4.1 ± 0.6	1.6 ± 0.3	0.3 ± 0.1	6.1 ± 0.6	11
		3	>1	7.8 ± 0.9	3.8 ± 0.6	0.7 ± 0.1	12.3 ± 1.1	10
		>3	1	5.6 ± 0.7	2.9 ± 0.5	0.7 ± 0.2	9.2 ± 0.9	5
		>3	>1	15.3 ± 1.5	12.0 ± 1.9	3.2 ± 0.5	30.5 ± 2.5	32
$\ell^+\ell^+$	$0 < D < 0.6$	2	>0	17.9 ± 1.8	4.9 ± 0.8	0.3 ± 0.1	23.1 ± 2.0	26
		3	1	10.2 ± 1.3	3.7 ± 0.6	0.4 ± 0.1	14.4 ± 1.4	11
		3	>1	10.2 ± 1.2	6.9 ± 1.1	0.8 ± 0.2	17.9 ± 1.6	18
		>3	1	10.7 ± 1.2	4.9 ± 0.8	0.8 ± 0.2	16.4 ± 1.4	16
		>3	>1	22.4 ± 2.0	13.3 ± 2.2	3.0 ± 0.5	38.7 ± 3.0	42
$\ell^+\ell^+$	$D > 0.6$	2	>0	8.0 ± 1.1	4.3 ± 0.7	0.4 ± 0.1	12.7 ± 1.3	18
		3	1	4.8 ± 0.7	3.2 ± 0.5	0.3 ± 0.1	8.4 ± 0.9	7
		3	>1	5.4 ± 0.7	7.1 ± 1.2	1.0 ± 0.2	13.5 ± 1.4	10
		>3	1	6.3 ± 0.8	5.6 ± 0.9	0.9 ± 0.2	12.8 ± 1.2	12
		>3	>1	16.5 ± 1.5	22.5 ± 3.7	3.1 ± 0.5	42.1 ± 4.0	46

TABLE 6.6

POST-FIT YIELDS FOR THE 3ℓ $t\bar{t}Z$ CHANNEL

N_b	N_{jets}	Background	$t\bar{t}W$	$t\bar{t}Z$	Total	Data
0	2	1032.8 ± 77.1	0.9 ± 0.1	18.2 ± 3.2	1051.9 ± 77.2	1022
	3	293.5 ± 21.4	0.4 ± 0.1	22.3 ± 3.9	316.3 ± 21.8	318
	>3	95.4 ± 7.4	0.3 ± 0.1	26.1 ± 4.6	121.8 ± 8.7	144
1	2	164.6 ± 17.8	1.9 ± 0.3	24.3 ± 4.3	190.7 ± 18.3	209
	3	66.6 ± 6.7	0.9 ± 0.2	41.2 ± 7.2	108.7 ± 9.8	99
	>3	32.8 ± 3.3	0.8 ± 0.1	61.3 ± 10.8	94.9 ± 11.3	72
>1	2	12.9 ± 2.4	1.0 ± 0.2	5.9 ± 1.0	19.8 ± 2.6	32
	3	11.6 ± 1.7	0.6 ± 0.1	17.9 ± 3.2	30.1 ± 3.6	46
	>3	10.6 ± 1.6	0.4 ± 0.1	41.0 ± 7.2	52.0 ± 7.4	54

TABLE 6.7

POST-FIT YIELDS FOR THE 4ℓ $t\bar{t}Z$ CHANNEL

Process	$N_b = 0$	$N_b > 0$
Backgrounds	12.8 ± 2.0	3.3 ± 0.3
$t\bar{t}Z$	4.5 ± 0.6	14.5 ± 1.8
Total	17.2 ± 2.0	17.8 ± 1.8
Data	23	15

TABLE 6.8

EXPECTED AND OBSERVED SIGNIFICANCES¹

Channel	Expected	Observed
SS dilepton $\ell^-\ell^-$ ($t\bar{t}W^-$)	2.4	2.3
SS dilepton $\ell^+\ell^+$ ($t\bar{t}W^+$)	4.2	5.5
SS dilepton $\ell^\pm\ell^\pm$ ($t\bar{t}W^\pm$)	4.5	5.3
Three-lepton ($t\bar{t}Z$)	> 5.0	> 5.0
Four-lepton ($t\bar{t}Z$)	4.7	4.5
Three- and four-lepton combined ($t\bar{t}Z$)	> 5.0	> 5.0

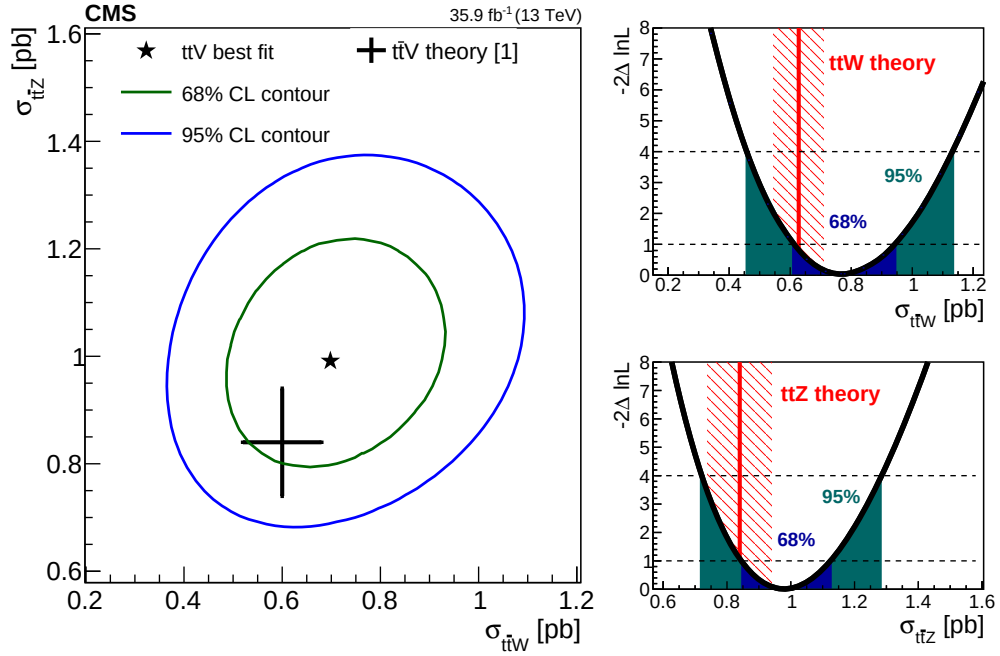
¹ All numbers are in standard deviations.

Figure 6.3. Left: Result of the simultaneous fit for the $t\bar{t}W$ and $t\bar{t}Z$ cross sections, shown as a star, along with the corresponding 68 and 95 % CL contours. Right: likelihood scan for the one-dimensional fit for $t\bar{t}W$ (top) and $t\bar{t}Z$ (bottom), along with the 68 and 95 % CL intervals and the theory prediction [105].

CHAPTER 7

EFFECTIVE FIELD THEORY INTERPRETATION FOR MEASUREMENTS OF TOP QUARK PAIR-PRODUCTION IN ASSOCIATION WITH A W OR Z BOSON

The utility of EFT as a tool for searching for NP in a model-independent way was introduced in Section 2.2.3. Some dimension-six operators would affect the $t\bar{t}W$ and $t\bar{t}Z$ cross sections; therefore, the cross section measurements of the $t\bar{t}W$ and $t\bar{t}Z$ processes at 8 TeV (described in Chapter 5) and 13 TeV (described in Chapter 6) can be used to constrain possible contributions from NP. Section 7.1 explains the general strategy used for all analyses. Section 7.2 describes the 8 TeV EFT analysis, and Section 7.3 describes the 13 TeV EFT analysis. In carrying out the 8 TeV analysis and from the feedback we received from it, our understanding of how to do this type of analysis improved; subsequently, a number of details evolved from the first to the second version. Finally, Section 7.4 presents an outlook on possibilities for future work.

7.1 General strategy

To investigate the effects of NP on $t\bar{t}W$ and $t\bar{t}Z$ production, calculate the expected $t\bar{t}W$ and $t\bar{t}Z$ cross section as a function of the Wilson coefficients is necessary. The matrix element can be written as the sum of the SM and NP components:

$$\mathcal{M} = \mathcal{M}_{\text{SM}} + \sum c_i \mathcal{M}_i.$$

As a first step, we vary each coefficient independently. For clarity, we denote the set of Wilson coefficients as c_i , and a particular Wilson coefficient that varies independently (i.e.,

with all other coefficients set to 0) as c_1 :

$$\mathcal{M} = \mathcal{M}_{\text{SM}} + c_1 \mathcal{M}_1.$$

The cross section is proportional to the matrix element modulus squared, which can be parameterized as a second-order polynomial [106]:

$$\begin{aligned} \sigma(c_1) &= \sigma_{\text{SM+NP}}(c_1) \propto |\mathcal{M}|^2 \\ &\propto |\mathcal{M}_{\text{SM}} + c_1 \mathcal{M}_1|^2 \\ &\propto s_0 + s_1 c_1 + s_2 c_1^2. \end{aligned} \tag{7.1}$$

The above relation allows for efficient modeling of NP effects on cross sections. Rather than calculating the cross section at each possible Wilson coefficient value that is studied, we can instead calculate it for three values of c_1 and solve the system of linear equations to determine the coupling structure constants s_0 , s_1 , and s_2 . In the SM case $c_1 = 0$, implying that s_0 corresponds to the SM cross section. We evaluate the cross sections expected from both SM and NP effects at leading order in QCD using the Higgs Effective Lagrangian `FEYNRULES` [107] implementation from reference [32] with `MADGRAPH` [108] at a set of values for each Wilson coefficient and fit the points with a quadratic function to determine $\sigma_{\text{SM+NP}}(c_1)$. No constraints are imposed on the number of allowed QCD or electroweak vertices. The scaling due to NP effects can be understood in terms of the signal scaling function

$$\mu(c_1) = \frac{\sigma_{\text{SM+NP}}(c_1)}{\sigma_{\text{SM}}} = \frac{\sigma_{\text{SM+NP}}(c_1)}{\sigma_{\text{SM+NP}}(0)}. \tag{7.2}$$

Note that it is closely related to the signal strength modifier described in Chapters 5 and 6, but in this context it is called the signal scaling function to emphasize that it is a function of the Wilson coefficients. The signal scaling functions for $t\bar{t}W$, $t\bar{t}Z$, and $t\bar{t}H$, evaluated at several possible Wilson coefficient values, are shown along with the corresponding quadratic fits

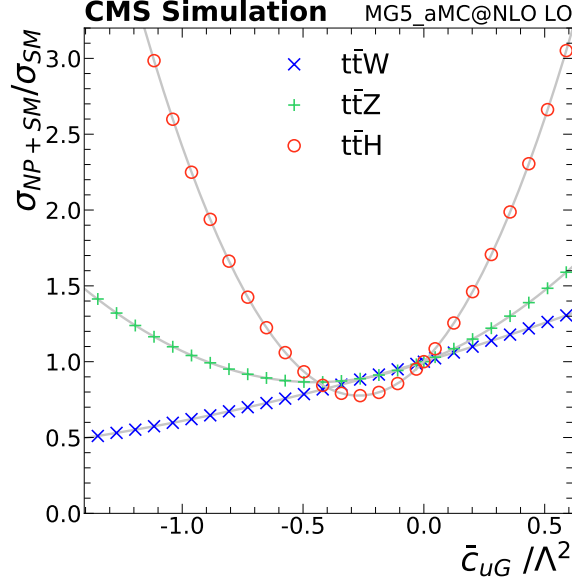


Figure 7.1. Expected cross section scaling at 13 TeV as a function of $\bar{c}_{uG}/\Lambda^2$ for $t\bar{t}W$ (blue crosses), $t\bar{t}Z$ (green plus signs) and $t\bar{t}H$ (red circles). The quadratic fit is shown as a gray line.

in Figure 7.1. Good agreement is observed. Note that although $\sigma(c_1)$ is always quadratic, the minimum is not constrained to appear at the SM value ($c_1 = 0$), and in the case of destructive interference with the SM, it is possible to have $\sigma_{SM+NP}(c_1) < \sigma_{SM}$.

The statistical procedure used to set limits is similar to that described in Section 5.5. The signal scaling, which encodes the scaling due to NP effects as a function of the Wilson coefficients, $\mu_{t\bar{t}Z}(c_1)$, $\mu_{t\bar{t}W}(c_1)$, and $\mu_{t\bar{t}H}(c_1)$, is defined as described above. From this, we construct a profile likelihood that is similar to Equation (5.2), except that now the signal strength modifier is now replaced by the signal scaling function, which is a function of c_1 and is constrained by Equation (7.2). The parameter of interest that we fit for is c_1 , and not μ as in Equation (5.2):

$$L(c_1, \theta) = \mathcal{P}(\text{data}|c_1, \theta) \rho(\tilde{\theta}|\theta) = \prod_{i=1}^M \frac{(\sum_k \mu_k(c_1) s_i + b_i)^{n_i}}{n_i!} e^{-(\sum_k \mu_k(c_1) s_i + b_i)} \rho(\tilde{\theta}|\theta), \quad (7.3)$$

where the sum is over k processes for which we are considering the NP effects. Similarly,

the likelihood ratio in Equation (5.3) becomes

$$\lambda(c_1) = \frac{L(c_1, \hat{\theta}(c_1))}{L(\hat{c}_1, \hat{\theta})},$$

from which the test statistic is defined as in Equation (5.5). To find the best fit c_1 , we maximize the likelihood using the Higgs Combine tool [109].

7.2 8 TeV analysis

The strategy outlined above was used for the 8 TeV analysis. Cross sections were computed, assuming flavor-independent couplings, for the production of $t\bar{t}$, a Higgs boson, $t\bar{t}Z$, and $t\bar{t}W$, sampling 20 points for each c_1 , in the range $-1 < c_1 < 1$. From this survey, we select five operators as being of particular interest because they have a small effect on inclusive Higgs boson and $t\bar{t}$ production and a large effect on $t\bar{t}Z$, $t\bar{t}W$, or both: $\bar{c}_{uB}$, $\bar{c}_{HQ}$, $\bar{c}'_{HQ}$, $\bar{c}_{Hu}$ and $\bar{c}_{3W}$. The lagrangian terms these operators correspond to are presented in Table 7.2. The signal scalings as a function of these selected Wilson coefficients for $t\bar{t}$, a Higgs boson, $t\bar{t}Z$, and $t\bar{t}W$ are summarized in Table 7.12.

Each coupling was scanned with all other couplings set to their SM value. Profile likelihood scans are shown in Figure 7.3. Best fit values, along with 1σ and 2σ CL ranges, are summarized in Table 7.3. Operators for which the minima in $\mu_{t\bar{t}Z}(c_1)$ and $\mu_{t\bar{t}W}(c_1)$ coincide have symmetric likelihood distributions because the cross section scaling as a function of the Wilson coefficient is quadratic; thus, there are two coefficient values corresponding to the same scaling.

TABLE 7.1

SIGNAL SCALING FOR SELECTED PROCESSES (8 TEV)

Coefficient	Process	$\mu(0.17)$	$\mu(0.34)$	$\mu(0.51)$	$\mu(0.68)$	$\mu(0.85)$
$\bar{c}_{3W}$	H		1.0	1.0	1.0	1.0
	$t\bar{t} + 1 \text{ parton}$		1.0	1.0	1.0	1.0
	$t\bar{t}W$		2.1	3.4	5.3	7.6
	$t\bar{t}Z$		1.0	1.0	1.1	1.1
$\bar{c}_{HQ}$	H		1.0	1.0	1.0	1.0
	$t\bar{t} + 1 \text{ parton}$		1.0	1.0	1.0	1.0
	$t\bar{t}W$		1.0	1.0	1.0	1.0
	$t\bar{t}Z$		1.0	1.4	2.0	2.9
$\bar{c}_{Hu}$	H		1.0	1.0	1.0	1.0
	$t\bar{t} + 1 \text{ parton}$		1.0	1.0	1.0	1.0
	$t\bar{t}W$		1.0	1.0	1.0	1.0
	$t\bar{t}Z$		1.7	2.4	3.2	4.2
$\bar{c}'_{HQ}$	H		1.0	1.0	1.0	1.0
	$t\bar{t} + 1 \text{ parton}$		1.0	1.0	1.0	1.0
	$t\bar{t}W$		3.7	6.7	11	15
	$t\bar{t}Z$		2.3	3.4	4.8	6.4
$\bar{c}_{uB}$	H		1.0	1.0	1.0	1.0
	$t\bar{t} + 1 \text{ parton}$		1.1	1.3	1.6	1.9
	$t\bar{t}W$		1.1	1.2	1.4	1.6
	$t\bar{t}Z$		6.0	12	21	32

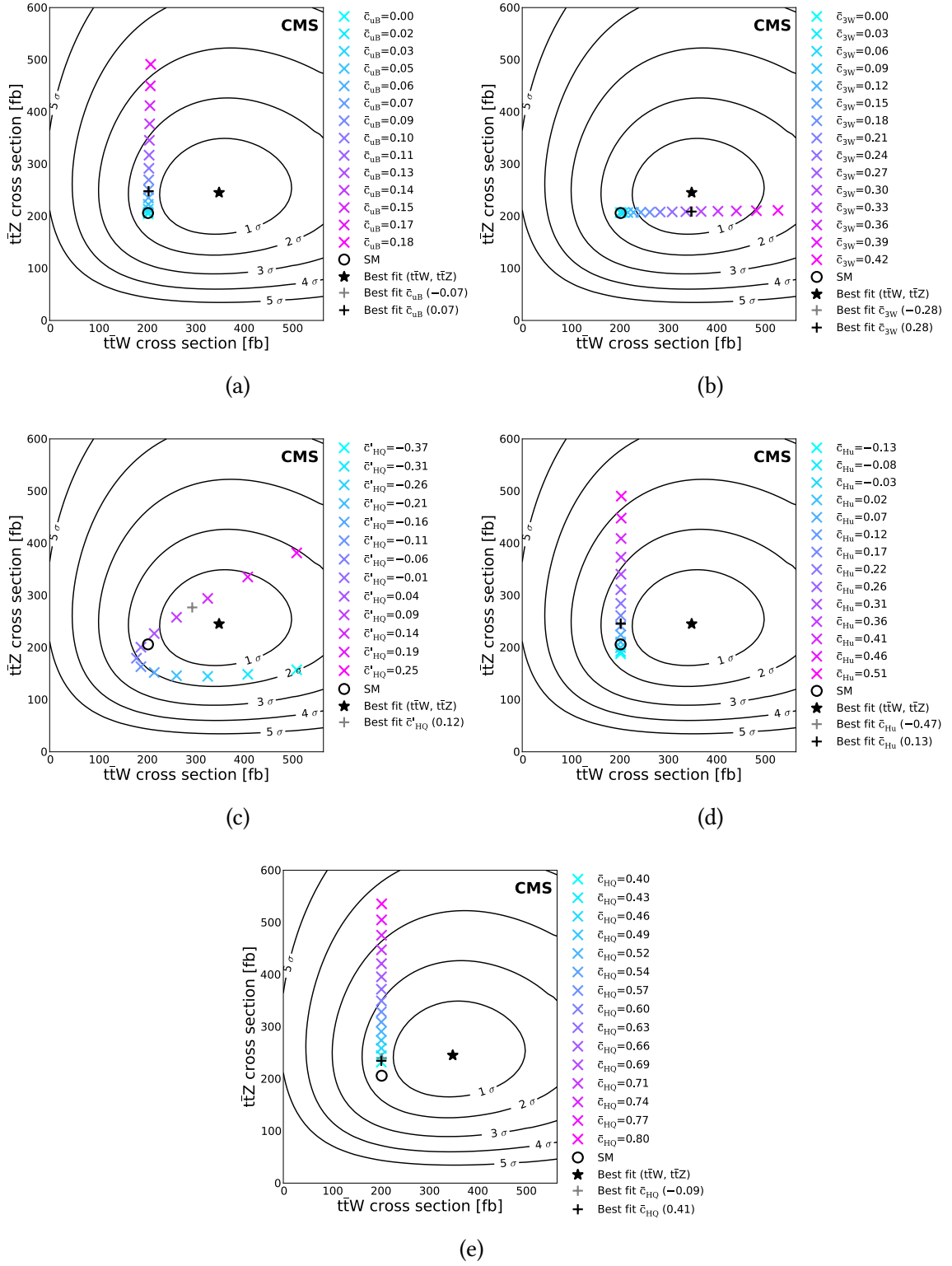


Figure 7.2. Sampled coefficient values (crosses) and best fit value (plus signs) for $\bar{c}_{ub}$ (a), $\bar{c}_{3W}$ (b), $\bar{c}'_{HQ}$ (c), $\bar{c}_{Hu}$ (d), and $\bar{c}_{HQ}$ (e), plotted in the $\sigma_{t\bar{t}W}$, $\sigma_{t\bar{t}Z}$ plane, for the 8 TeV analysis. The simultaneous best fit to $\mu_{t\bar{t}W}$ and $\mu_{t\bar{t}Z}$ is shown as a star.

TABLE 7.2

LAGRANGIAN TERMS FOR SELECTED OPERATORS

Lagrangian term	Notation
$\frac{\tilde{c}_H}{2v^2} \partial^\mu [\Phi^\dagger \Phi] \partial_\mu [\Phi^\dagger \Phi]$ $\frac{\tilde{c}_{HQ}}{v^2} [\bar{Q}_L \gamma^\mu Q_L] [\Phi^\dagger \tilde{D}_\mu \Phi]$ $\frac{4i\tilde{c}'_{HQ}}{v^2} [\bar{Q}_L \gamma^\mu T_{2k} Q_L] [\Phi^\dagger T_2^k \tilde{D}_\mu \Phi]$ $\frac{i\tilde{c}_{Hu}}{v^2} [\bar{u}_R \gamma^\mu u_R] [\Phi^\dagger \tilde{D}_\mu \Phi]$ $\frac{2g' \tilde{c}_{uB}}{m_W^2} y_u \Phi^\dagger \cdot \bar{Q}_L \gamma^{\mu\nu} u_R B_{\mu\nu}$ $\frac{4g\tilde{c}_{uW}}{m_W^2} y_u \Phi^\dagger \cdot (\bar{Q}_L T_{2k}) \gamma^{\mu\nu} u_R W_{\mu\nu}^k$ $\frac{g^3 \tilde{c}_{3W}}{m_W^2} \epsilon_{ijk} W_{\mu\nu}^i W_{\rho}^{\nu j} W^{\rho\mu k}$ $\frac{g_s^3 \tilde{c}_{3G}}{m_W^2} f_{abc} G_{\mu\nu}^a G_{\rho}^{\nu b} \tilde{G}^{\rho\mu c}$ $\frac{g_s^3 \tilde{c}_{3G}}{m_W^2} f_{abc} G_{\mu\nu}^a G_{\rho}^{\nu b} G^{\rho\mu c}$ $\frac{\tilde{c}_{2G}}{m_W^2} D^\mu G_{\mu\nu}^a D_\rho G_a^{\rho\nu}$	We follow the notation described in reference [32]: B_μ, W_μ^k, and G_μ^a (g', g, and g_s) refer to the $U(1)_Y$, $SU(2)_L$, and $SU(3)_c$ gauge vector fields (coupling constants), respectively. The corresponding $B_{\mu\nu}$, $W_{\mu\nu}^k$, and $G_{\mu\nu}^a$ field tensors are defined in Equation (2.8) of reference [32]. The left-handed $Q_L = \begin{pmatrix} u_L \\ d_L \end{pmatrix}$ and right-handed up-type u_R quark fields refer to the three generations in the context of the 8 TeV analysis and to the third generation in the context of the 13 TeV analysis. The Higgs doublet is referred to as Φ. The Hermitian derivative operators $\tilde{D}_\mu$ are defined as $\Phi^\dagger \tilde{D}_\mu \Phi = \Phi^\dagger D^\mu \Phi - D_\mu \Phi^\dagger \Phi$. The Higgs quartic coupling and vacuum expectation value is given by λ and v, respectively. The 3×3 Yukawa coupling matrix in flavor space is given by y_u. The generators of $SU(2)$ are given by $T_{2K} = \frac{\sigma_K}{2}$, where σ_k are the Pauli matrices.

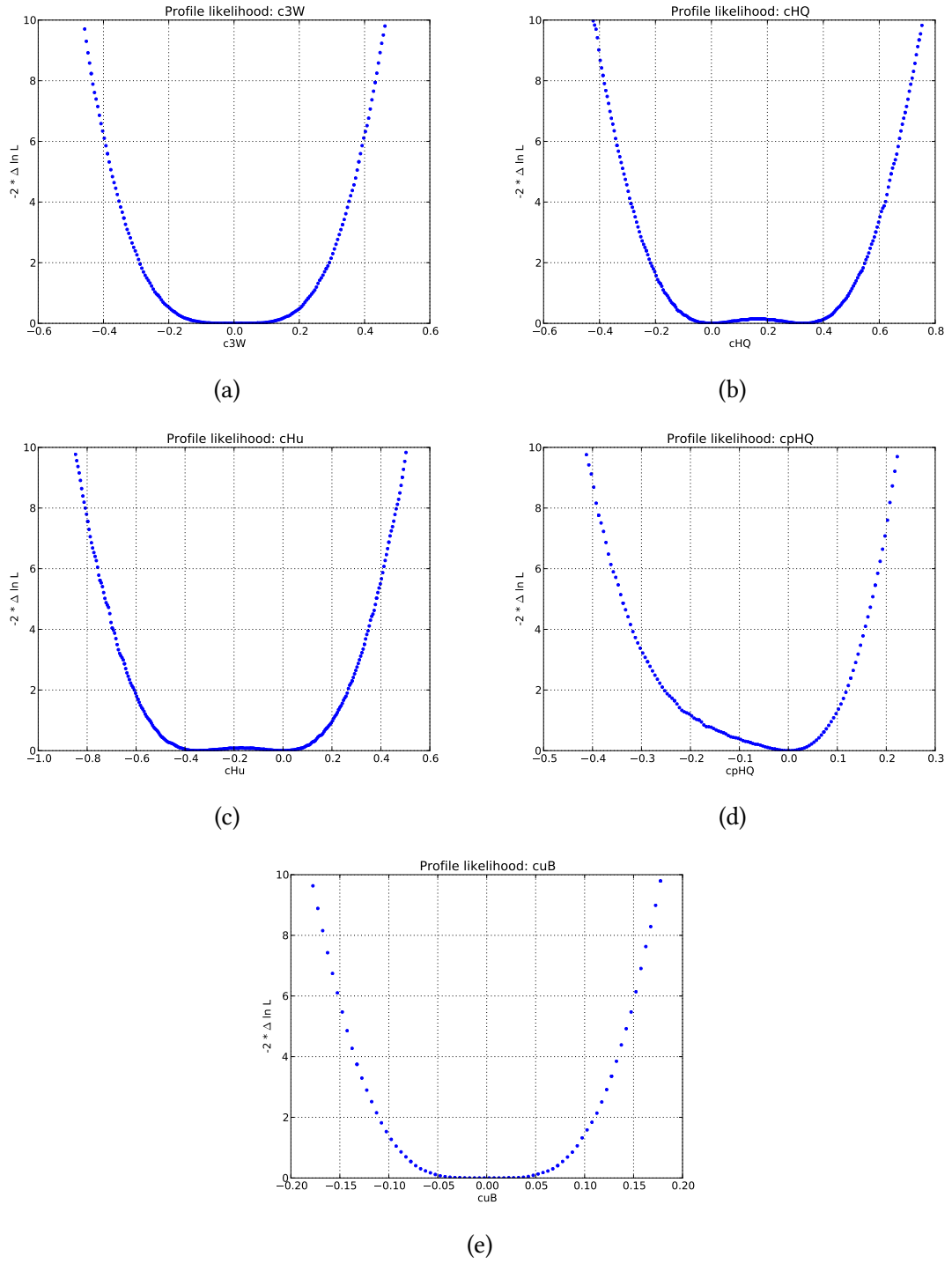


Figure 7.3. Profile likelihood scans for $\bar{c}_{3W}$ (a), $\bar{c}_{HQ}$ (b), $\bar{c}'_{HQ}$ (c), $\bar{c}_{uB}$ (d), and $\bar{c}_{Hu}$ (e) for the 8 TeV analysis.

TABLE 7.3
CONSTRAINTS ON SELECTED WILSON COEFFICIENTS (8 TeV)

Coefficient	Best fit point(s)	68 % CL	95 % CL
$\bar{c}_{uB}$	-0.07 and 0.07	$[-0.11, 0.11]$	$[-0.14, 0.14]$
$\bar{c}_{3W}$	-0.28 and 0.28	$[-0.36, -0.18]$ and $[0.18, 0.36]$	$[-0.43, 0.43]$
$\bar{c}_{Hu}$	-0.47 and 0.13	$[-0.60, -0.23]$ and $[-0.11, 0.26]$	$[-0.71, 0.37]$
$\bar{c}_{HQ}$	-0.09 and 0.41	$[-0.22, 0.08]$ and $[0.24, 0.54]$	$[-0.31, 0.63]$
$\bar{c}'_{HQ}$	0.12	$[-0.07, 0.18]$	$[-0.33, -0.24]$ and $[-0.02, 0.23]$

7.3 13 TeV analysis

Building on the work described in Section 7.2, the 13 TeV data described in Chapter 6 was re-interpreted within the framework of EFT. A strategy similar to that used for the 8 TeV analysis was employed, but several improvements were incorporated.

7.3.1 Independent constraints

The FEYNRULES implementation from reference [32] used in the 8 TeV analysis assumes flavor-independent fermion couplings. Because the Z and W boson coupling to light quarks is highly constrained by other measurements, we removed the couplings to the first two generations with help from Martin [110]. For this reason, results from the 8 TeV analysis and 13 TeV analysis cannot be directly compared. This modified implementation was then used to determine $\mu(c_1)$, as described in Section 7.1, again considering the NP effects of one operator at a time. The strategy was more careful than what was used in the 8 TeV analysis. If the MADGRAPH calculation were infinitely precise, only three points would be required to completely specify the system of equations and calculate the three structure constants exactly. We would then be free to choose any three values when setting up the system of equations. Due to the finite precision of the MADGRAPH calculation, however, care must be taken in choosing the sample points: they should adequately cover the parameter space where the fit will be evaluated. This approach prevents noise in the sampled region from having a disproportionate effect in the regions to which the fit is extrapolated. To choose which points to sample, a *coarse scan* was first run for each coefficient using MADGRAPH to evaluate the cross section at ten points sampled using a pseudo-random number generator in the interval $-4\pi < c_1 < 4\pi$. These points were then divided by σ_{SM} and fit with a quadratic function to determine $\mu_{\text{coarse}}(c_1)$. Next, 30 values of c_1 were sampled using a pseudo-random number generator in the range corresponding to $\mu_{\text{coarse}}(c_1) < 10$, and a quadratic fit to these points was used to determine $\mu(c_1)$. This strategy was found to improve the quality of the fit. Next, we used $\mu(c_1)$ to construct a profile likelihood test statistic in terms of c_1 . The likelihood was then maximized using

the Higgs Combine tool to find the best fit c_1 for the 39 Wilson coefficients included in the FEYNRULES model. Each coupling was scanned with all other couplings set to their SM values.

In the 8 TeV analysis, only the scaling of $t\bar{t}W$ and $t\bar{t}Z$ due to NP effects was included. In the 13 TeV analysis, the NP effects on $t\bar{t}H$ were considered as well because $t\bar{t}H$ is a sizable and irreducible background to $t\bar{t}W$, and many operators produce similar cross section scaling of both processes.

The strategy used to select which operators to study was also more sophisticated in the 13 TeV analysis. First, we eliminated any operators that did not affect $t\bar{t}W$, $t\bar{t}Z$, or $t\bar{t}H$. To determine which operators to discard, we fit $\sigma(c_1) = s_0 + s_1 c_1 + s_2 c_1^2$ and eliminated operators where s_1 and s_2 were both less than 10^{-5} . This requirement excluded the operators proportional to the following Wilson coefficients: $\bar{c}_1$, $\bar{c}_{1B}$, $\bar{c}_d$, $\bar{c}_{2B}$, $\bar{c}_{lW}$, $\bar{c}_{dB}$, and $\bar{c}_{dW}$.

We also eliminated operators with strong effects on already very precisely measured processes such as $t\bar{t}$ production, because significant deviations would have been apparent from previous measurements. We defined the extreme signal scaling $\mu_e(c_1)$ to be $\mu(c_1)$ evaluated at the c_1 which that $|\mu(c_1) - 1|$, the deviation from the SM, within the range of c_1 corresponding to 2σ sensitivity. Operators producing $|\mu_e - 1| > 0.7$ for $t\bar{t}$, inclusive Higgs, WW, and WZ were excluded. The $\mu_e(c_1)$ values for all operators that satisfied the requirement for s_1 and s_2 but failed to meet the requirement on $\mu_e(c_1)$ are presented in Table 7.4.

Sixteen operators satisfied all the requirements on sensitivity and μ_e : $\bar{c}_{uW}$, $\bar{c}_H$, $\tilde{c}_{3G}$, $\bar{c}_{3G}$, $\bar{c}_{uG}$, $\bar{c}_{Hu}$, $\bar{c}_{2G}$, $\bar{c}_{uB}$, $\bar{c}_{HB}$, $\tilde{c}_{HW}$, $\bar{c}_{Hud}$, $\bar{c}_{HQ}$, $\bar{c}_B$, $\tilde{c}_A$, $\bar{c}'_{HQ}$, and $\bar{c}_u$. For each of these operators, we estimated the effect of NP on yields by scaling the yields for each process by $\mu_e(c_1)$. Because the nonprompt background is dominated by $t\bar{t}$, we used $\mu_{e, t\bar{t}}$ to estimate the impact of NP effects on that background. For rare backgrounds, we made the simplifying assumption that the baseline yields were due to equal contributions from WZZ, ZZZ, WWW, and WWZ. Table 7.5 summarizes the μ_e due to operators that had significant effects on background yields, mainly due to the tqZ , tHq , VH (WH or ZH), and triboson processes. The corresponding difference in yields due to these operators is tabulated in Table 7.7. These operators were excluded

TABLE 7.4

EXTREME SIGNAL SCALING¹

	$\mu_{e,t\bar{t}Z}$	$\mu_{e,t\bar{t}W}$	$\mu_{e,t\bar{t}H}$	$\mu_{e,t\bar{t}}$	$\mu_{e,H}$	$\mu_{e,DY}$	$\mu_{e,ZZ}$	$\mu_{e,WZ}$	$\mu_{e,WW}$
$\tilde{c}_{HB}$	1.3	1.1	1.0	1.0	1.0	1.0	1.1	5.0	4.6
$\bar{c}_G$	1.2	0.6	1.8	0.6	1610.7	1.1	5.2	1.0	2.5
$\tilde{c}_G$	1.3	1.0	1.8	1.0	739.4	1.0	2.9	1.0	1.7
$\bar{c}_{Hd}$	1.3	1.0	1.4	1.0	1.0	1.3	1.1	1.0	2.8
$\bar{c}_{HW}$	1.1	1.5	1.1	1.0	1.0	1.0	1.0	1.7	1.1
$\tilde{c}_{3W}$	1.0	1.5	1.0	1.0	1.0	1.0	1.0	7.9	2.6
$\bar{c}_{3W}$	1.0	1.5	1.0	1.0	1.0	1.0	1.0	7.9	2.6
$\bar{c}_{WW}$	1.0	1.5	1.0	1.0	1.0	1.0	1.0	2.0	1.2

¹ Only the μ_e due to operators which failed the requirement that $|\mu_e - 1| < 0.7$ for $t\bar{t}$, inclusive Higgs, WW, or WZ are tabulated.

from further study. Table 7.6 and Table 7.8 summarize the $\mu_e(c_1)$ and yields, respectively, due to operators for which the effect on background yields was considered acceptably small. The operators that were excluded by each requirement are summarized in table Table 7.9. Eight operators satisfied all requirements: $\bar{c}_H$, $\tilde{c}_{3G}$, $\bar{c}_{3G}$, $\bar{c}_{uG}$, $\bar{c}_{uB}$, $\bar{c}_{Hu}$, $\bar{c}_{uW}$, and $\bar{c}_{2G}$.

TABLE 7.5
THE μ_e DUE TO ELIMINATED OPERATORS¹ (13 TeV)

	$\mu_{e,\bar{t}\bar{t}Z}$	$\mu_{e,\bar{t}\bar{t}W}$	$\mu_{e,\bar{t}\bar{t}H}$	$\mu_{e,WWWW}$	$\mu_{e,WZZ}$	$\mu_{e,ZZZ}$	$\mu_{e,WWZ}$	$\mu_{e,VH}$	$\mu_{e,tZq}$	$\mu_{e,tHq}$	$\mu_{e,tHW}$	$\mu_{e,t\bar{t}\bar{t}}$	$\mu_{e,tWZ}$	$\mu_{e,tG}$
$\tilde{c}_{HW}$	1.2	1.4	1.0	12.0	22.1	27.9	8.1	108.4	1.1	35.4	10.1	1.0	1.2	1.1
$\bar{c}_u$	1.0	1.0	2.9	1.0	1.0	1.0	1.0	1.0	1.0	22.5	15.8	3.9	1.0	1.0
$\tilde{c}_A$	1.3	1.0	1.1	1.0	1.2	57.4	15.3	63.8	1.0	1.0	1.0	1.0	1.0	1.0
$\bar{c}_{Hud}$	1.3	1.0	1.5	1.0	1.0	1.0	1.3	1.0	7.8	76.6	69.9	1.2	11.3	2.5
$\bar{c}'_{HQ}$	0.4	1.0	2.0	1.0	1.0	3.4	1.9	4.5	10.1	101.4	81.2	2.6	13.1	0.0
$\bar{c}_B$	1.3	1.0	1.5	1.0	1.0	21.8	14.0	14.1	1.0	1.0	1.0	1.4	1.0	1.0

¹ Only the μ_e due to operators which pass the requirement that $|\mu_e - 1| < 0.7$ for $\bar{t}\bar{t}$, inclusive Higgs, WW, or WZ, but have an unacceptably large effect on background yields are tabulated.

TABLE 7.6
THE μ_e DUE TO SELECTED OPERATORS¹ (13 TeV)

	$\mu_{e,t\bar{t}Z}$	$\mu_{e,t\bar{t}W}$	$\mu_{e,t\bar{t}H}$	$\mu_{e,WWW}$	$\mu_{e,WZZ}$	$\mu_{e,ZZZ}$	$\mu_{e,WWZ}$	$\mu_{e,VH}$	$\mu_{e,tZq}$	$\mu_{e,tHq}$	$\mu_{e,tHW}$	$\mu_{e,t\bar{t}\bar{t}}$	$\mu_{e,tWZ}$	$\mu_{e,tG}$
$\bar{c}_{uW}$	1.3	1.1	1.1	1.0	1.0	1.0	1.0	1.0	1.4	5.7	3.6	1.1	1.9	1.2
$\bar{c}_{uB}$	1.3	1.0	1.1	1.0	1.0	1.0	1.0	1.0	1.1	1.0	1.0	1.2	1.2	1.1
$\bar{c}_H$	1.0	1.0	2.9	1.2	1.3	1.4	1.2	3.1	1.0	3.1	2.9	4.0	1.0	1.0
$\tilde{c}_{3G}$	1.3	1.0	1.7	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	2.8	1.0	1.0
$\bar{c}_{3G}$	1.3	1.0	2.2	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	3.7	1.0	1.0
$\bar{c}_{Hu}$	1.3	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.1	1.0	1.0	1.9	1.3	1.0
$\bar{c}_{2G}$	1.3	1.0	2.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	43.1	1.0	1.0

¹ Only the μ_e due to operators which pass the requirement that $|\mu_e - 1| < 0.7$ for $t\bar{t}$, inclusive Higgs, WW, or WZ, and have an acceptably small effect on background yields are tabulated.

TABLE 7.7

YIELD¹ Δ DUE TO NP FOR ELIMINATED OPERATORS² (13 TeV)

	$T_{\text{SM}}^3 = 443$			$T_{\text{SM}} = 218$					$T_{\text{SM}} = 175$					$T_{\text{SM}} = 1176$		
	$\Delta_{\text{t}\bar{\text{t}}\text{Z}}$	$\Delta_{\text{t}\bar{\text{t}}\text{W}}$	$\Delta_{\text{t}\bar{\text{t}}\text{H}}$	Δ_{WWWW}	Δ_{WZZ}	Δ_{ZZZ}	Δ_{WWZ}	Δ_{VH}	Δ_{tZq}	Δ_{tHq}	Δ_{tHW}	$\Delta_{\text{t}\bar{\text{t}}\text{t}\bar{\text{t}}}$	Δ_{tWZ}	Δ_{tG}	Δ_{WZ}	Δ_{ZZ}
$\bar{c}_{\text{HB}}$	135	0.3	36.5	59.5	8.6	272.8	1079.5	3355.3	47.8	0.0	0.0	1.0	0.3	0.0	113.3	1.3
$\tilde{c}_{\text{HW}}$	75.1	103.5	2.1	2052.7	950.4	416.7	1017.2	36709.2	37.0	804.6	47.3	0.0	0.4	0.0	260.2	8.6
$\bar{c}_{\text{Hud}}$	136	0.0	35.6	0.0	0.0	0.0	42.7	0.0	1818.3	1676.5	339.6	0.5	19.5	0.0	0.0	0.0
$\bar{c}_{\text{HQ}}$	-175	2.7	3.3	0.0	0.0	41.2	224.7	1280.5	257.7	0.0	0.0	3.4	4.2	0.0	0.0	90.1
$\bar{c}_{\text{B}}$	136	0.0	35.1	0.0	0.6	304.2	1750.1	4177.8	0.0	0.0	0.0	1.0	0.0	0.0	0.0	0.8
$\tilde{c}_{\text{A}}$	146	0.0	8.2	0.0	8.8	835.2	1958.5	20517.7	0.0	0.0	0.0	0.1	0.0	0.0	0.0	9.3
$\bar{c}'_{\text{HQ}}$	-150	2.5	47.5	0.0	0.0	24.4	81.3	772.0	1687.7	1503.6	269.0	2.5	15.8	0.0	0.0	17.8
$\bar{c}_{\text{u}}$	1.00	0.1	81.0	0.0	0.0	0.0	0.0	0.0	0.0	292.0	44.7	4.1	0.0	0.0	0.0	0.0

¹ Yields differences are calculated as $\Delta = \sum_i N_{\text{SM}}^i (\mu_e - 1)$, where N_{SM}^i refers to the expected SM yield in a particular channel, and the sum is over i channels.² Only scaling effects due to operators which pass the requirement that $|\mu_e - 1| < 0.7$ for $\text{t}\bar{\text{t}}$, inclusive Higgs, WW, and WZ, but have an unacceptably large effect on background yields are tabulated.³ T_{SM} is the sum of expected SM yields over channels and the given process group: $T_{\text{SM}} = \sum_{i,p} N_{\text{SM}}^{i,p}$.

TABLE 7.8

YIELD¹ Δ DUE TO NP FOR SELECTED OPERATORS² (13 TeV)

	$T_{\text{SM}}^3 = 443$			$T_{\text{SM}} = 218$					$T_{\text{SM}} = 175$					$T_{\text{SM}} = 1176$		
	$\Delta_{\text{t}\bar{\text{t}}\text{Z}}$	$\Delta_{\text{t}\bar{\text{t}}\text{W}}$	$\Delta_{\text{t}\bar{\text{t}}\text{H}}$	Δ_{WWW}	Δ_{WZZ}	Δ_{ZZZ}	Δ_{WWZ}	Δ_{VH}	Δ_{tZq}	Δ_{tHq}	Δ_{tHW}	$\Delta_{\text{t}\bar{\text{t}}\text{t}\bar{\text{t}}}$	Δ_{tWZ}	Δ_{tG}	Δ_{WZ}	Δ_{ZZ}
$\bar{c}_{\text{uW}}$	153.9	28.5	5.4	0.0	0.0	0.0	5.5	0.0	103.0	108.9	12.9	0.2	1.6	0.0	0.0	0.0
$\bar{c}_{\text{H}}$	2.1	0.4	81.0	20.1	9.2	3.7	16.5	428.1	0.0	29.3	6.2	4.2	0.0	0.0	0.0	1.2
$\bar{c}_{3\text{G}}$	128.7	0.0	46.7	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	4.0	0.0	0.0	0.0	0.0
$\bar{c}_{3\text{G}}$	138.8	0.0	66.6	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	5.1	0.0	0.0	0.0	0.0
$\bar{c}_{\text{uG}}$	109.2	30.1	53.6	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.8	0.8	0.1	0.0	0.0	0.0
$\bar{c}_{\text{Hu}}$	148.3	4.1	1.1	0.0	0.0	0.0	0.7	0.0	15.9	0.0	0.0	1.6	0.5	0.0	0.0	0.0
$\bar{c}_{2\text{G}}$	139.9	0.0	58.3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	82.3	0.0	0.0	0.0	0.0
$\bar{c}_{\text{uB}}$	147.1	1.2	6.3	0.0	0.0	0.0	0.2	0.0	20.2	0.0	0.0	0.4	0.5	0.0	0.0	0.0

¹ Yields differences are calculated as $\Delta = \sum_i N_{\text{SM}}^i (\mu_e - 1)$, where N_{SM}^i refers to the expected SM yield in a particular channel, and the sum is over i channels.

² Only scaling effects due to operators which pass the requirement that $|\mu_e - 1| < 0.7$ for $\text{t}\bar{\text{t}}$, inclusive Higgs, WW, and WZ, and have an acceptably small effect on background yields are tabulated.

³ T_{SM} is the sum of expected SM yields over channels and the given process group: $T_{\text{SM}} = \sum_{i,p} N_{\text{SM}}^{i,p}$.

TABLE 7.9

OPERATOR SELECTION REQUIREMENTS (13 TeV)

Requirement	Wilson coefficient of eliminated operator
No effect on $t\bar{t}H$, $t\bar{t}Z$, $t\bar{t}W$	$\bar{c}_l, \bar{c}_{lB}, \bar{c}_d,$ $\bar{c}_{2B}, \bar{c}_{lW}, \bar{c}_{dB}, \bar{c}_{dW}, \bar{c}'_{HL}, \bar{c}_{He}, \bar{c}_{dG}, \bar{c}_6, \bar{c}_{HL}$
$ \mu_{WZ}-1 > 0.7$	$\tilde{c}_{HB}, \bar{c}_T, \tilde{c}_{3W}, \bar{c}_{3W}, \bar{c}_{WW}, \bar{c}_{HW}$
$ \mu_{ZZ}-1 > 0.7$	$\bar{c}_G, \tilde{c}_G, \bar{c}_T$
$ \mu_{WW}-1 > 0.7$	$\tilde{c}_{HB}, \bar{c}_G, \tilde{c}_G, \bar{c}_T, \bar{c}_{Hd}, \tilde{c}_{3W}, \bar{c}_{3W}$
$ \mu_{tt}-1 > 0.7$	$\bar{c}_{uG}, \bar{c}_G$
$ \mu_H-1 > 0.7$	$\bar{c}_G, \tilde{c}_G$
Large effect on background yields	$\bar{c}_{HB}, \tilde{c}_{HW}, \bar{c}_{Hud}, \bar{c}_{HQ}, \bar{c}_B, \tilde{c}_A, \tilde{c}'_{HQ}, \bar{c}_u$

To obtain $\sigma(c_1)$ from the post-fit values of the Wilson coefficients and nuisance parameters, we used the following transformation:

$$\sigma(c_1) = \sigma_{\text{NLO}} \cdot \mu(c_1) \cdot \kappa_j^{\theta_j},$$

where θ_j corresponds to the post-fit nuisance parameters due to the scale and PDF uncertainties. These nuisances were parameterized by using a log-normal probability density function characterized by $\kappa_j > 1$, which encodes the spread in the distribution, with $(\kappa_j - 1)$ corresponding roughly to the relative uncertainty [111]. To determine the 1σ and 2σ contours in the $\sigma_{t\bar{t}Z}/\sigma_{t\bar{t}W}$ plane, we sampled the parameters randomly from the post-fit covariance matrix, then drew contours containing 68.27 % and 95.45 % of the samples. The result is presented in the right panels of Figures 7.4 to 7.11. The likelihood scans are presented in the center panels of Figures 7.4 to 7.11. We removed any assumptions about the energy scale

of the NP made in reference [32] and report the ratio c_i/Λ^2 . In cases in which $\sigma_{\text{SM+NP}}(c_i)$ had the same minimum for all three processes, the profile likelihood was symmetric around this point, and we present the results for $|c_i - c_{i,\text{min}}|$ to make this symmetry explicit.

The scaling due to NP effects, evaluated at the best fit value of the Wilson coefficient for $t\bar{t}Z$, $t\bar{t}W$, and $t\bar{t}H$, is tabulated in Table 7.12. In Figures 7.12 to 7.15, the ratio between the yields resulting from a fit with two free parameters for $\mu_{t\bar{t}Z}$ and $\mu_{t\bar{t}W}$ and the yields resulting from the fit with one free parameter for c_1 , which was subject to the constraints of $\mu_{t\bar{t}Z}(c_1)$, $\mu_{t\bar{t}W}(c_1)$, and $\mu_{t\bar{t}H}(c_1)$, is shown for each coefficient. The small variations between categories can be explained by slight differences in the best fit nuisance parameters. For $\bar{c}_{uW}$ and $\bar{c}_{uG}$, which are proportional to operators that affect all three processes, the observed excess was accommodated reasonably well in both the $t\bar{t}Z$ and $t\bar{t}W$ channels. It is evident (see Figure 7.12) that the sensitivity was driven by the $t\bar{t}Z$ channel because in both cases, the ratio of the yields was approximately one for $t\bar{t}Z$ and 0.91–0.95 for $t\bar{t}W$. This indicates that the fit was more sensitive to the $t\bar{t}Z$ channel and thus matched it more closely than the $t\bar{t}W$ channel. For the operators that are proportional to $\bar{c}_{3G}$, $\bar{c}_{3G}$, $\bar{c}_{uB}$, and $\bar{c}_{2G}$, which affect the $t\bar{t}Z$ and $t\bar{t}H$ processes but not the $t\bar{t}W$ process, the one-dimensional (1D) fit again yielded a coefficient value that produced the same $t\bar{t}Z$ scaling as in the two-dimensional (2D) fit. This result is shown in the corresponding right panels of Figures 7.6 to 7.9, where the 1D best fit in the $\sigma_{t\bar{t}Z}$, $\sigma_{t\bar{t}W}$ plane agrees closely with 2D best fit. The $t\bar{t}H$ yield was higher due to the scaling of the $t\bar{t}H$ process, and the $t\bar{t}W$ yield was lower because the $t\bar{t}W$ process is not affected by variations in this coefficient. Similarly, $\bar{c}_H$ only affects $t\bar{t}H$; consequently, the ratios for both the $t\bar{t}W$ and $t\bar{t}Z$ channels are below one. The $t\bar{t}H$ ratio is also below one, although we expected the fit to yield a higher scaling for $t\bar{t}H$ as a substitute for $t\bar{t}W$. This effect was studied carefully; it can be explained by a difference in the discriminant shapes of the $t\bar{t}W$ and $t\bar{t}H$ channels, in combination with the higher sensitivity to the $t\bar{t}Z$ channel. The fit adjusted nuisance parameters to improve the agreement for the $t\bar{t}Z$ channel rather than scale the $t\bar{t}H$ channel. The degree to which the nuisances affected the fit can be seen in Figure 7.10; in the absence of nuisance parameters, the 1D best fit

cross would coincide with the theory point because variations in $\bar{c}_H$ do not affect $t\bar{t}W$ or $t\bar{t}Z$ production. For $\bar{c}_{Hu}$, which only affects $t\bar{t}Z$ production, the best fit point matched the 2D case in the $t\bar{t}Z$ channel and was unable to accommodate the excess observed in the $t\bar{t}W$ channel.

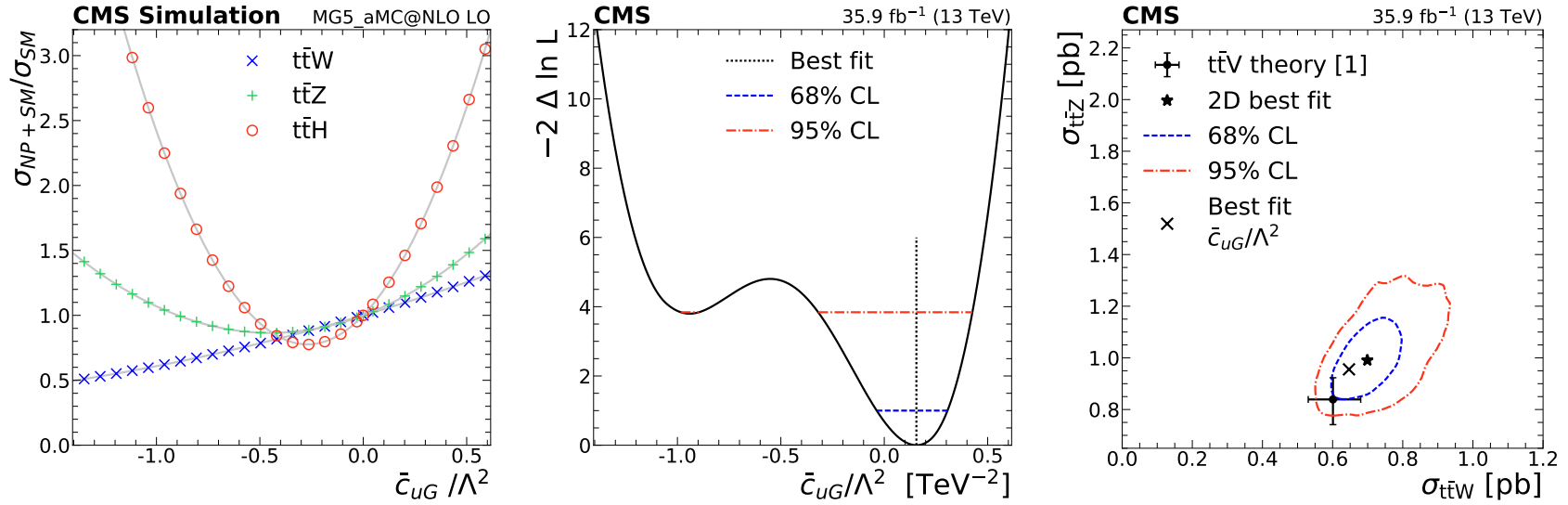


Figure 7.4. Left: signal scaling as a function of the c_1 for $t\bar{t}W$ (crosses), $t\bar{t}Z$ (pluses), and $t\bar{t}H$ (circles) for $\bar{c}_{uG}$ (13 TeV analysis). Center: the test statistic $q(c_i)$ scan as a function of c_1 , profiling all other nuisance parameters. The best fit value is indicated by a dotted line. Dashed and dash-dotted lines indicate 68 % and 95 % CL intervals, respectively. Right: The best fit c_1 value (shown as a cross), along with the corresponding 68 % (dashed) and 95 % (dash-dotted) contours in the $\sigma_{t\bar{t}Z}$, $\sigma_{t\bar{t}W}$ plane. The two-dimensional best fit to the $t\bar{t}W$ and $t\bar{t}Z$ cross sections is given by the star. The theory predictions [105] are shown as a dot with bars representing their respective uncertainties.

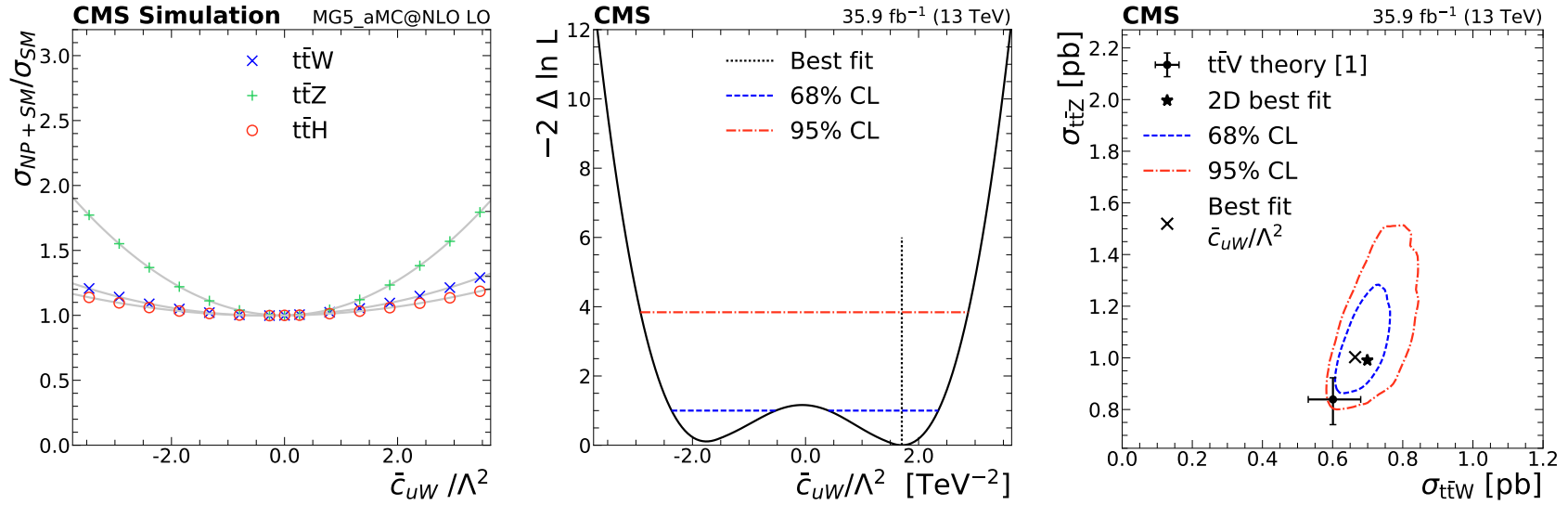


Figure 7.5. Left: signal scaling as a function of the c_1 for $t\bar{t}W$ (crosses), $t\bar{t}Z$ (pluses), and $t\bar{t}H$ (circles) for $\tilde{c}_{uW}$ (13 TeV analysis). Center: the test statistic $q(c_i)$ scan as a function of c_1 , profiling all other nuisance parameters. The best fit value is indicated by a dotted line. Dashed and dash-dotted lines indicate 68 % and 95 % CL intervals, respectively. Right: The best fit c_1 value (shown as a cross), along with the corresponding 68 % (dashed) and 95 % (dash-dotted) contours in the $\sigma_{t\bar{t}Z}$, $\sigma_{t\bar{t}W}$ plane. The two-dimensional best fit to the $t\bar{t}W$ and $t\bar{t}Z$ cross sections is given by the star. The theory predictions [105] are shown as a dot with bars representing their respective uncertainties.

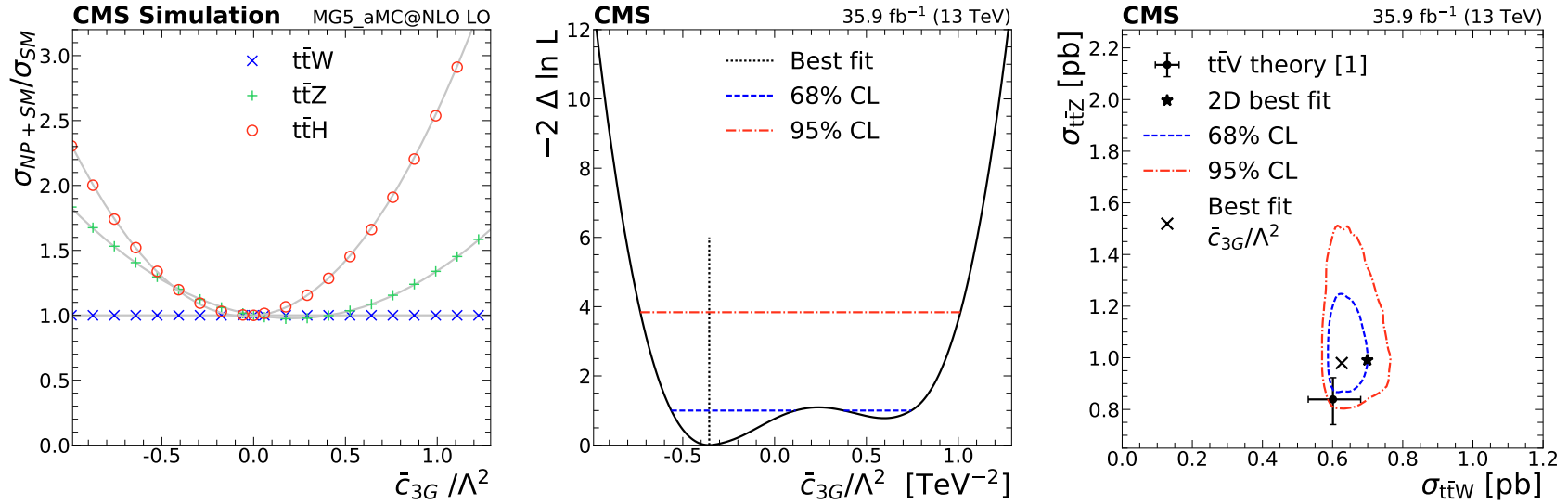


Figure 7.6. Left: signal scaling as a function of the c_1 for $t\bar{t}W$ (crosses), $t\bar{t}Z$ (pluses), and $t\bar{t}H$ (circles) for $\bar{c}_{3G}$ (13 TeV analysis). Center: the test statistic $q(c_i)$ scan as a function of c_1 , profiling all other nuisance parameters. The best fit value is indicated by a dotted line. Dashed and dash-dotted lines indicate 68 % and 95 % CL intervals, respectively. Right: The best fit c_1 value (shown as a cross), along with the corresponding 68 % (dashed) and 95 % (dash-dotted) contours in the $\sigma_{t\bar{t}Z}$, $\sigma_{t\bar{t}W}$ plane. The two-dimensional best fit to the $t\bar{t}W$ and $t\bar{t}Z$ cross sections is given by the star. The theory predictions [105] are shown as a dot with bars representing their respective uncertainties.

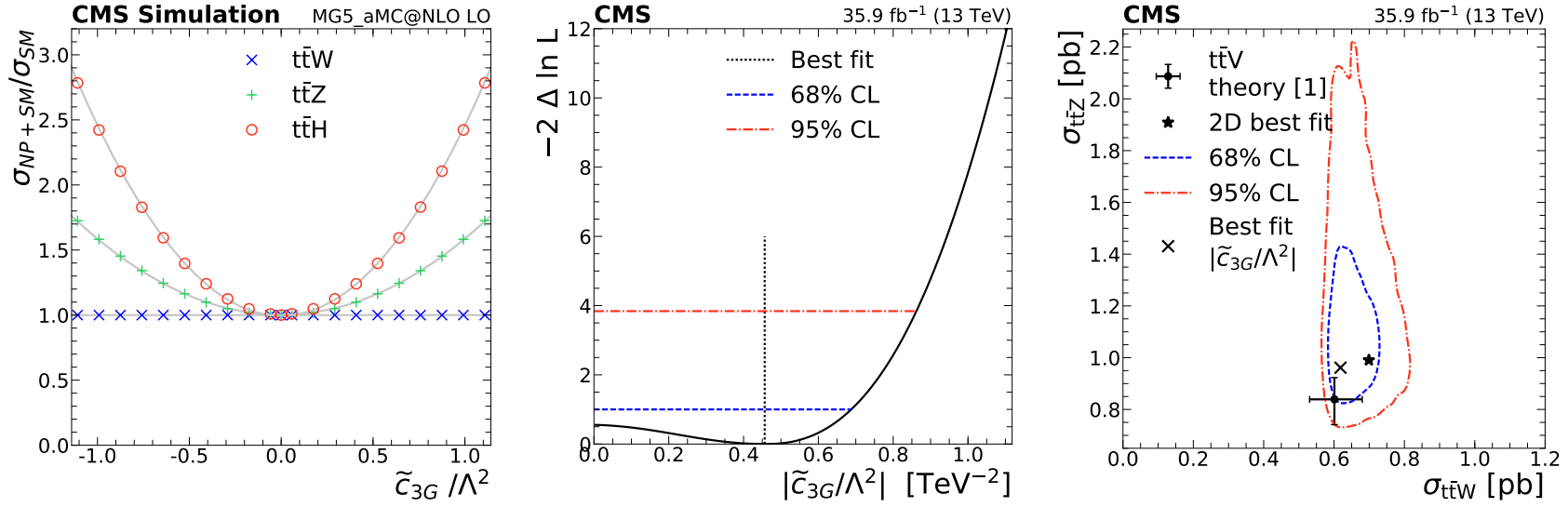


Figure 7.7. Left: signal scaling as a function of the c_1 for $t\bar{t}W$ (crosses), $t\bar{t}Z$ (pluses), and $t\bar{t}H$ (circles) for $\tilde{c}_{3G}$ (13 TeV) analysis). Center: the test statistic $q(c_i)$ scan as a function of c_1 , profiling all other nuisance parameters. The best fit value is indicated by a dotted line. Dashed and dash-dotted lines indicate 68 % and 95 % CL intervals, respectively. Right: The best fit c_1 value (shown as a cross), along with the corresponding 68 % (dashed) and 95 % (dash-dotted) contours in the $\sigma_{t\bar{t}Z}$, $\sigma_{t\bar{t}W}$ plane. The two-dimensional best fit to the $t\bar{t}W$ and $t\bar{t}Z$ cross sections is given by the star. The theory predictions [105] are shown as a dot with bars representing their respective uncertainties.

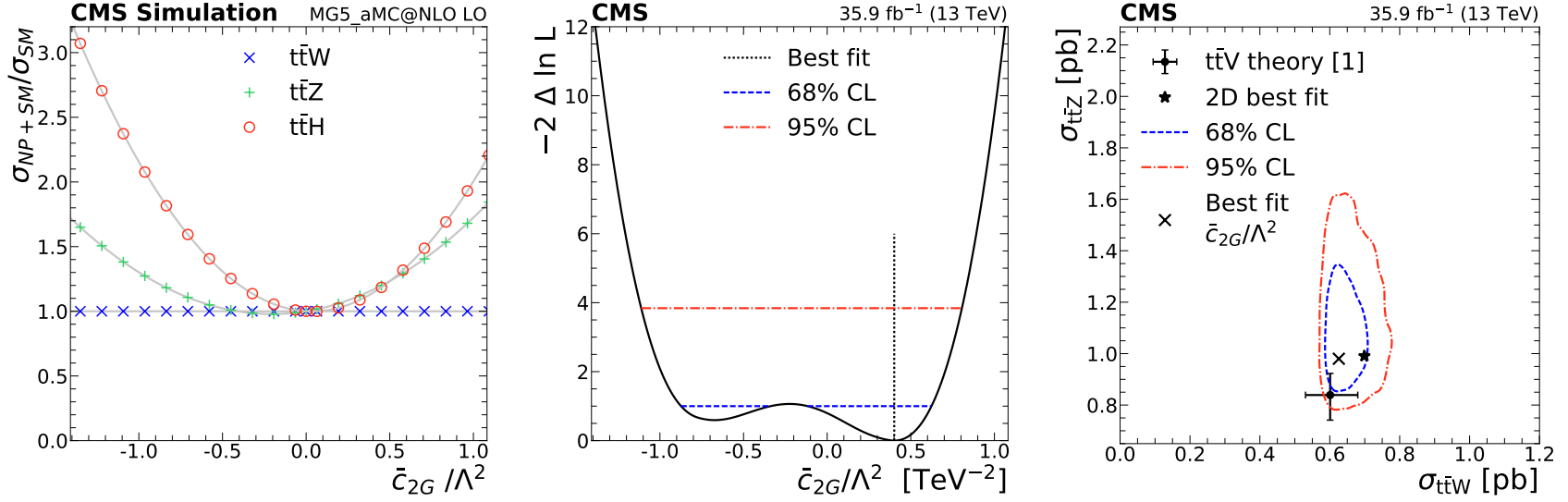


Figure 7.8. Left: signal scaling as a function of the c_1 for $t\bar{t}W$ (crosses), $t\bar{t}Z$ (pluses), and $t\bar{t}H$ (circles) for $\bar{c}_{2G}$ (13 TeV analysis). Center: the test statistic $q(c_i)$ scan as a function of c_1 , profiling all other nuisance parameters. The best fit value is indicated by a dotted line. Dashed and dash-dotted lines indicate 68 % and 95 % CL intervals, respectively. Right: The best fit c_1 value (shown as a cross), along with the corresponding 68 % (dashed) and 95 % (dash-dotted) contours in the $\sigma_{t\bar{t}Z}$, $\sigma_{t\bar{t}W}$ plane. The two-dimensional best fit to the $t\bar{t}W$ and $t\bar{t}Z$ cross sections is given by the star. The theory predictions [105] are shown as a dot with bars representing their respective uncertainties.

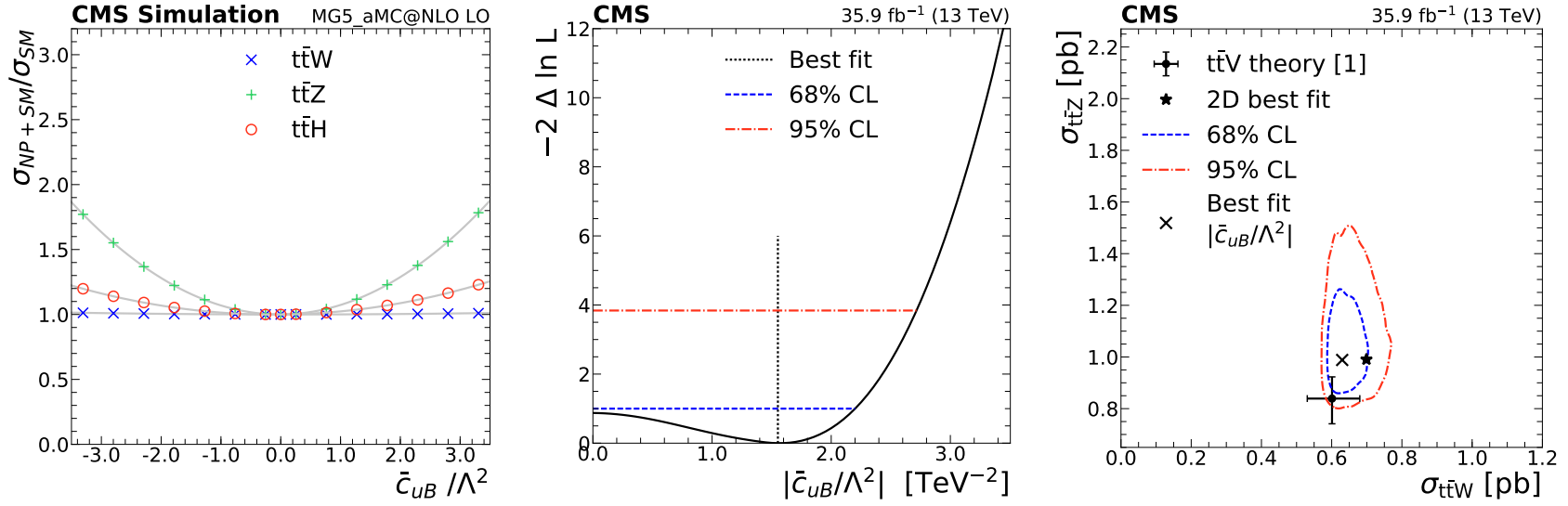


Figure 7.9. Left: signal scaling as a function of the c_1 for $t\bar{t}W$ (crosses), $t\bar{t}Z$ (pluses), and $t\bar{t}H$ (circles) for $\bar{c}_{uB}$ (13 TeV analysis). Center: the test statistic $q(c_i)$ scan as a function of c_1 , profiling all other nuisance parameters. The best fit value is indicated by a dotted line. Dashed and dash-dotted lines indicate 68 % and 95 % CL intervals, respectively. Right: The best fit c_1 value (shown as a cross), along with the corresponding 68 % (dashed) and 95 % (dash-dotted) contours in the $\sigma_{t\bar{t}Z}$, $\sigma_{t\bar{t}W}$ plane. The two-dimensional best fit to the $t\bar{t}W$ and $t\bar{t}Z$ cross sections is given by the star. The theory predictions [105] are shown as a dot with bars representing their respective uncertainties.

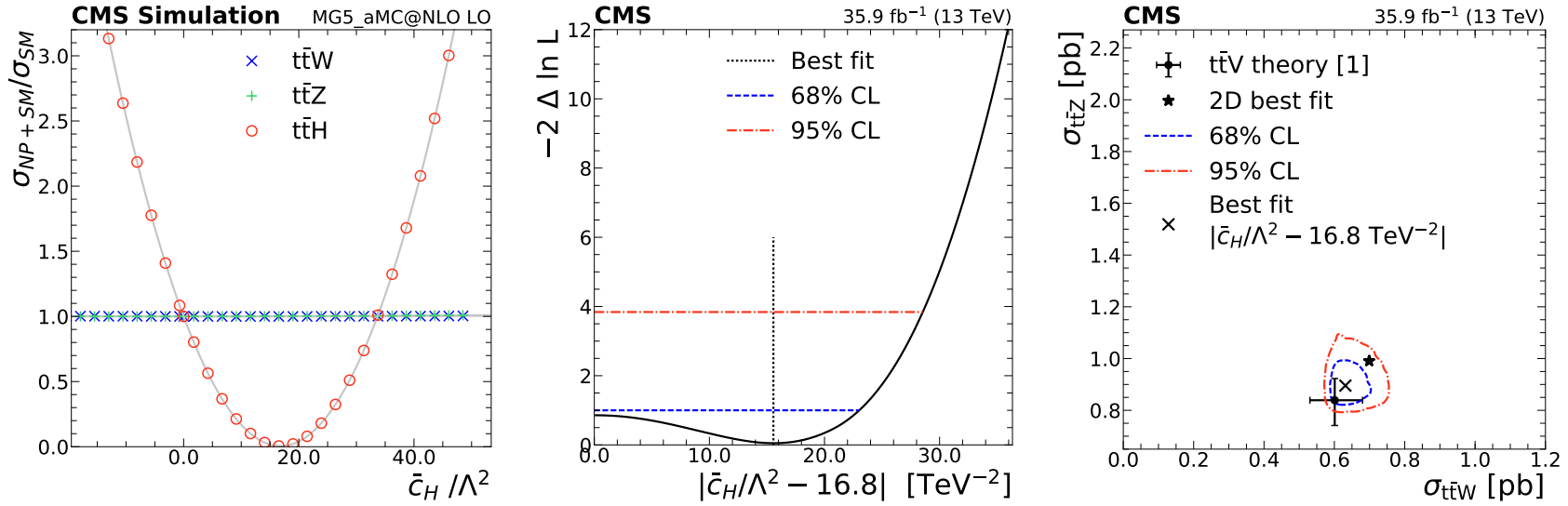


Figure 7.10. Left: signal scaling as a function of the c_1 for $t\bar{t}W$ (crosses), $t\bar{t}Z$ (pluses), and $t\bar{t}H$ (circles) for $\tilde{c}_H$ (13 TeV analysis). Center: the test statistic $q(c_i)$ scan as a function of c_1 , profiling all other nuisance parameters. The best fit value is indicated by a dotted line. Dashed and dash-dotted lines indicate 68 % and 95 % CL intervals, respectively. Right: The best fit c_1 value (shown as a cross), along with the corresponding 68 % (dashed) and 95 % (dash-dotted) contours in the $\sigma_{t\bar{t}Z}$, $\sigma_{t\bar{t}W}$ plane. The two-dimensional best fit to the $t\bar{t}W$ and $t\bar{t}Z$ cross sections is given by the star. The theory predictions [105] are shown as a dot with bars representing their respective uncertainties.

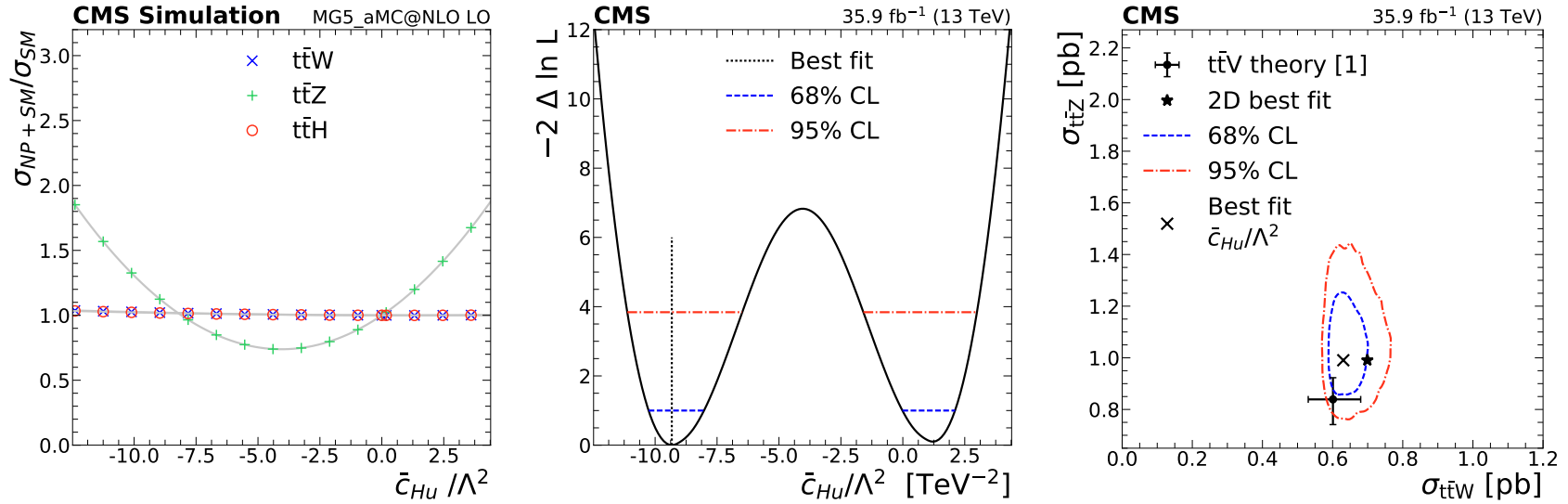


Figure 7.11. Left: signal scaling as a function of the c_1 for $t\bar{t}W$ (crosses), $t\bar{t}Z$ (pluses), and $t\bar{t}H$ (circles) for $\bar{c}_{Hu}$ (13 TeV analysis). Center: the test statistic $q(c_i)$ scan as a function of c_1 , profiling all other nuisance parameters. The best fit value is indicated by a dotted line. Dashed and dash-dotted lines indicate 68 % and 95 % CL intervals, respectively. Right: The best fit c_1 value (shown as a cross), along with the corresponding 68 % (dashed) and 95 % (dash-dotted) contours in the $\sigma_{t\bar{t}Z}$, $\sigma_{t\bar{t}W}$ plane. The two-dimensional best fit to the $t\bar{t}W$ and $t\bar{t}Z$ cross sections is given by the star. The theory predictions [105] are shown as a dot with bars representing their respective uncertainties.

TABLE 7.10

EXPECTED CL INTERVALS (13 TeV)

Wilson coefficient	68 % CL TeV^{-2}	95 % CL TeV^{-2}
$\bar{c}_{uW}/\Lambda^2$	$[-1.6, 1.5]$	$[-2.2, 2.2]$
$ \bar{c}_H/\Lambda^2 - 16.8 \text{ TeV}^{-2} $	$[3.7, 23.4]$	$[0, 28.7]$
$\tilde{c}_{3G}/\Lambda^2$	$[-0.5, 0.5]$	$[-0.7, 0.7]$
$\bar{c}_{3G}/\Lambda^2$	$[-0.3, 0.7]$	$[-0.5, 0.9]$
$\bar{c}_{uG}/\Lambda^2$	$[-0.9, -0.8] \cup [-0.3, 0.2]$	$[-1.1, 0.3]$
$ \bar{c}_{uB}/\Lambda^2 $	$[0.0, 1.5]$	$[0.0, 2.1]$
$\bar{c}_{Hu}/\Lambda^2$	$[-9.2, -6.5] \cup [-1.6, 1.1]$	$[-10.1, 2.0]$
$\bar{c}_{2G}/\Lambda^2$	$[-0.7, 0.4]$	$[-0.9, 0.6]$

7.3.2 Preliminary simultaneous constraints

The results of the previous sections, obtained by considering the effects of each dimension-six operator with the others fixed to zero, are an important first step. They rely on the assumption that NP processes only enhance certain Wilson coefficients; however, this case is special. Additional work is needed to remove this assumption and generalize the results by performing a simultaneous fit to all Wilson coefficients, which would provide a more meaningful description of the underlying physics. Fortunately, adapting Equation (7.1), used in the previous sections to parameterize the scaling due one operator at a time, for two operators is straightforward:

$$\begin{aligned}
\sigma(c_1, c_2) &\propto |\mathcal{M}_{\text{SM}} + c_1 \mathcal{M}_1 + c_2 \mathcal{M}_2|^2 \\
&\propto \underbrace{s_0}_{\text{constant}} + \underbrace{s_1 c_1 + s_2 c_2}_{\text{linear}} + \underbrace{s_3 c_1^2 + s_4 c_2^2}_{\text{quadratic}} + \underbrace{s_5 c_1 c_2}_{\text{mixed}}.
\end{aligned} \tag{7.4}$$

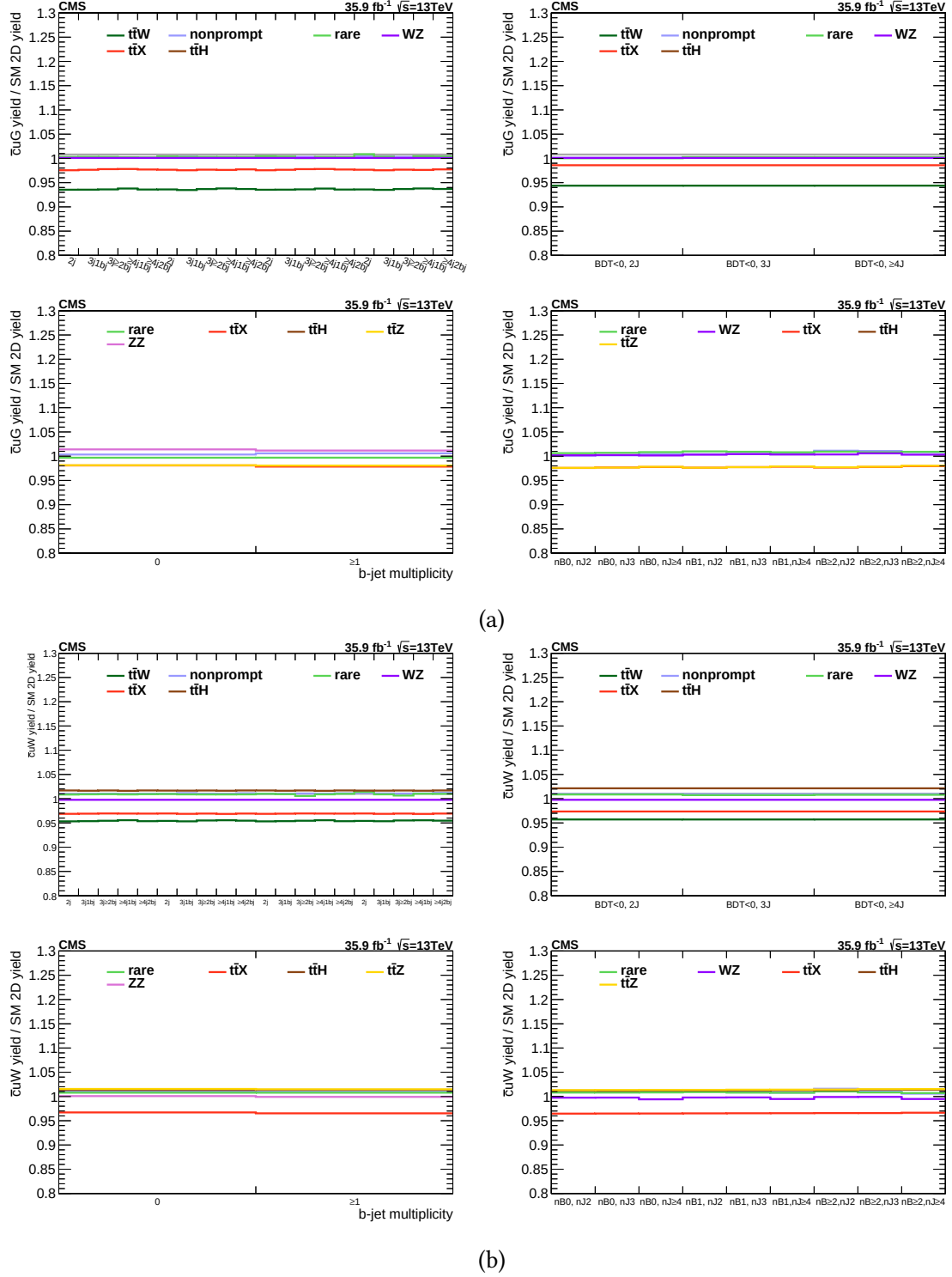


Figure 7.12. Ratio of yields resulting from the fit with two free parameters for μ_{ttW} and μ_{ttZ} to those from the fit with one free parameter for c_1 . In (a), panels refer to (clockwise from top left) the SS ttW signal region with $D > 0$ and control region with $D < 0$, the 3ℓ ttZ channel, and the 4ℓ ttZ channel for $\bar{c}_{\text{uG}}$, and similarly in (b) for

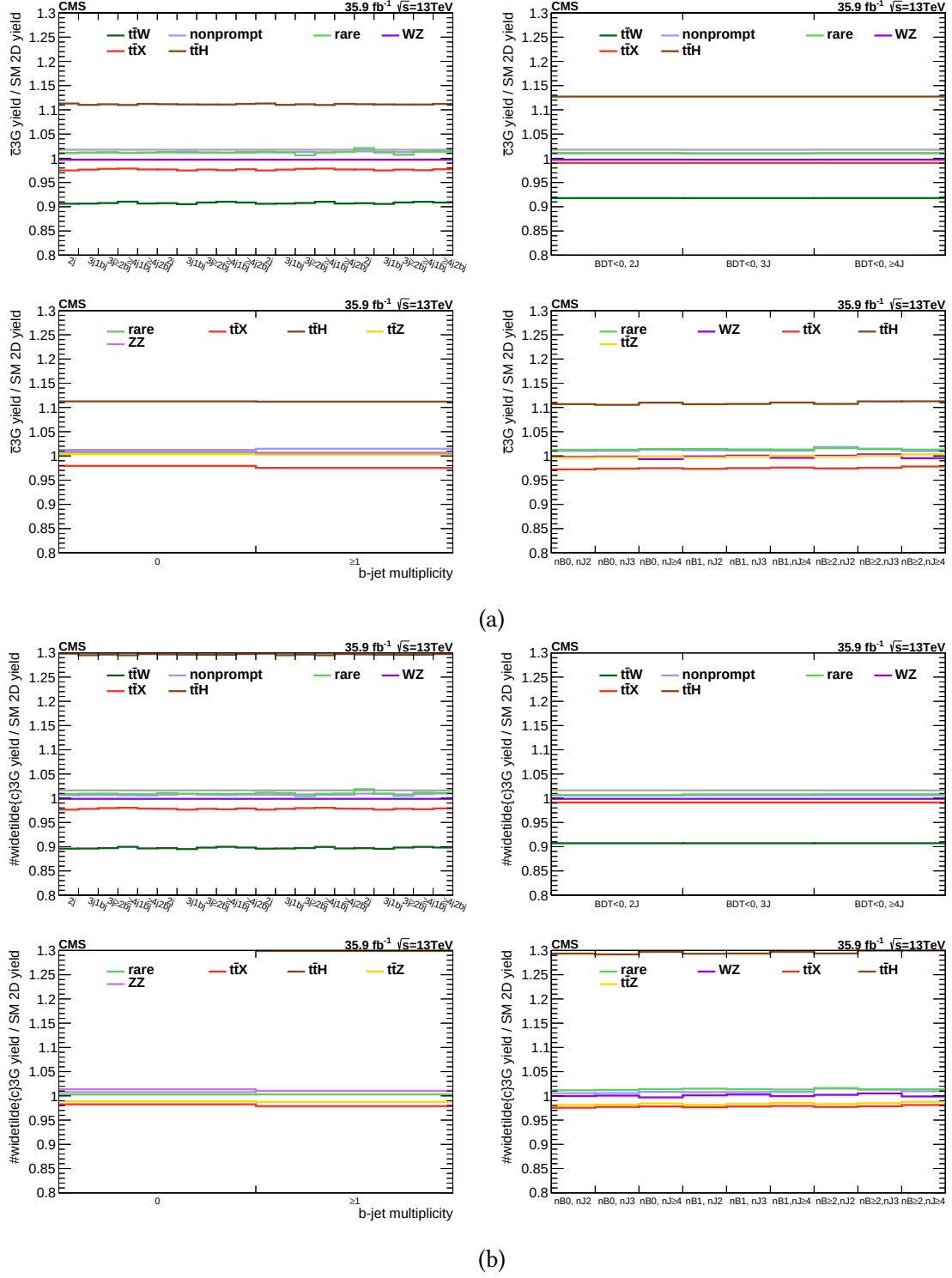


Figure 7.13. Ratio of yields resulting from the fit with two free parameters for μ_{ttW} and μ_{ttZ} to those from the fit with one free parameter for c_1 . In (a), panels refer to (clockwise from top left) the SS ttW signal region with $D > 0$ and control region with $D < 0$, the 3ℓ ttZ channel, and the 4ℓ ttZ channel for $\tilde{c}_{3G}$, and similarly in (b) for $\tilde{c}_{3G}$.

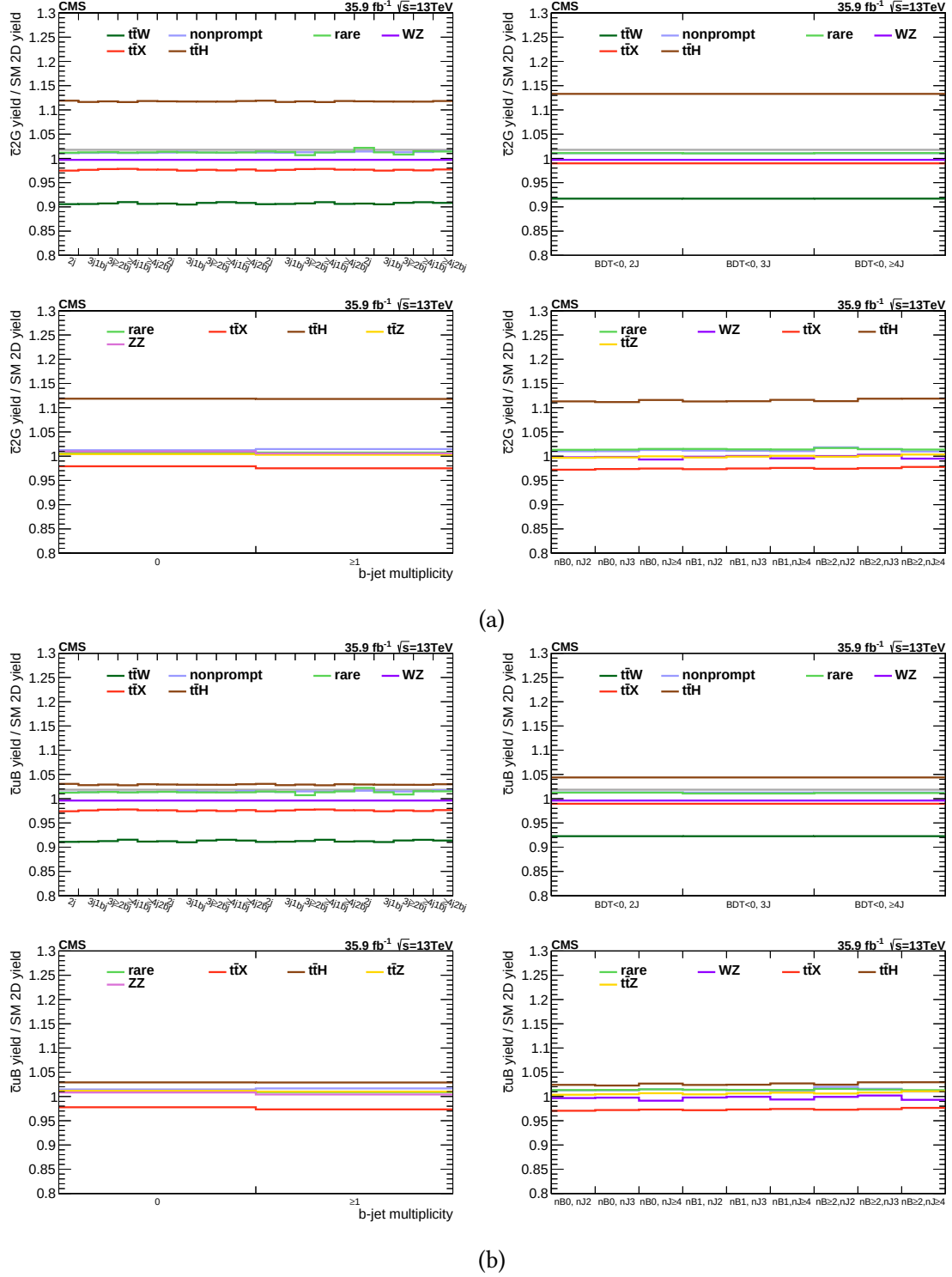


Figure 7.14. Ratio of yields resulting from the fit with two free parameters for μ_{ttW} and μ_{ttZ} to those from the fit with one free parameter for c_1 . In (a), panels refer to (clockwise from top left) the SS ttW signal region with $D > 0$ and control region with $D < 0$, the 3ℓ ttZ channel, and the 4ℓ ttZ channel for $\bar{c}_{2G}$, and similarly in (b) for

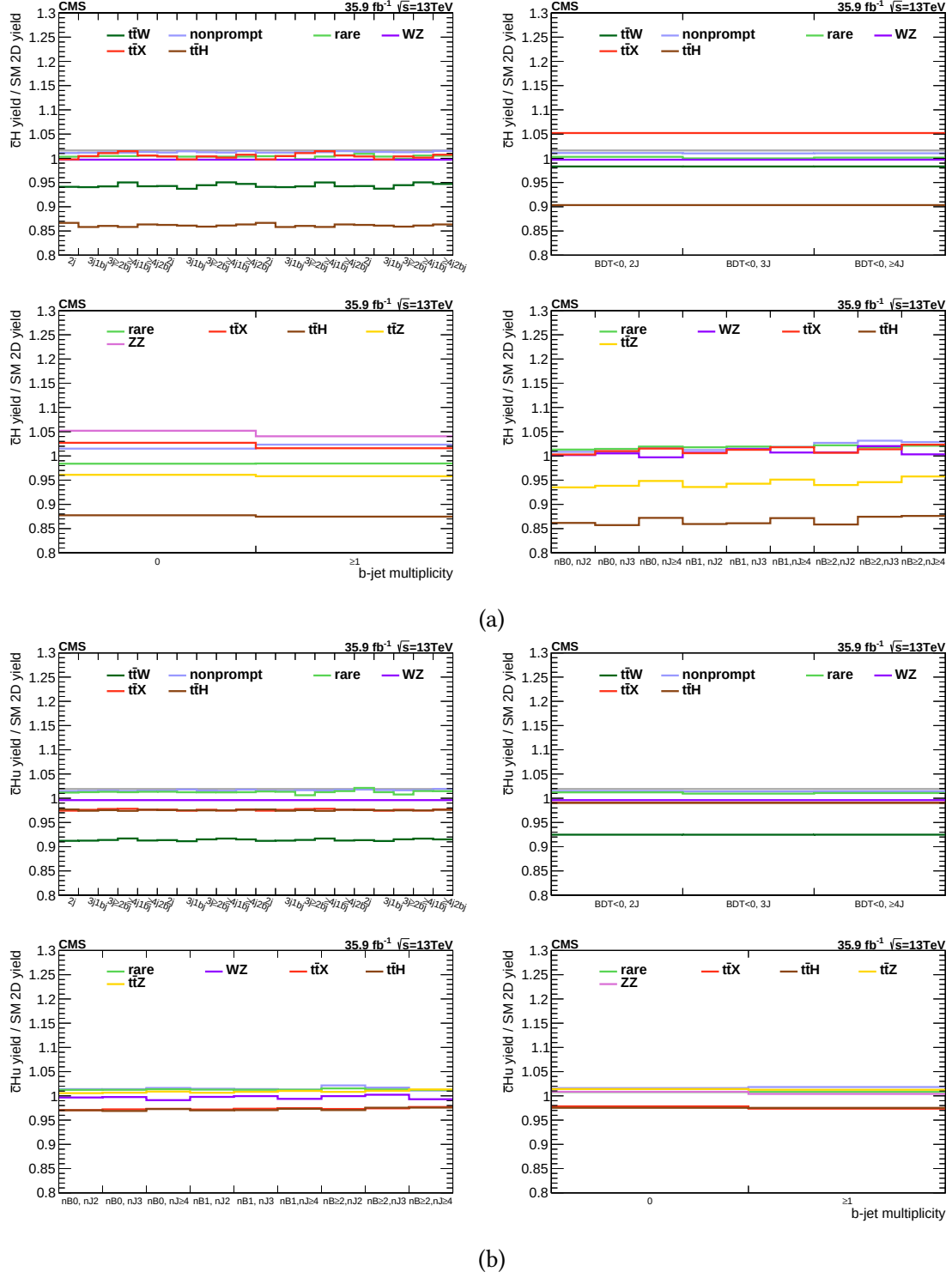


Figure 7.15. Ratio of yields resulting from the fit with two free parameters for μ_{ttW} and μ_{ttZ} to those from the fit with one free parameter for c_1 . In (a), panels refer to (clockwise from top left) the SS ttW signal region with $D > 0$ and control region with $D < 0$, the 3ℓ ttZ channel, and the 4ℓ ttZ channel for $\bar{c}_H$, and similarly in (b) for

TABLE 7.11

OBSERVED BEST FIT VALUES¹ AND CL INTERVALS (13 TeV)

Wilson coefficient	Best fit [TeV ⁻²]	68 % CL [TeV ⁻²]	95 % CL [TeV ⁻²]
$\bar{c}_{uW}/\Lambda^2$	1.7	$[-2.4, -0.5] \cup [0.4, 2.4]$	$[-2.9, 2.9]$
$ \bar{c}_H/\Lambda^2 - 16.8 \text{ TeV}^{-2} $	15.6	$[0, 23.0]$	$[0, 28.5]$
$ \tilde{c}_{3G}/\Lambda^2 $	0.5	$[0, 0.7]$	$[0, 0.9]$
$\bar{c}_{3G}/\Lambda^2$	-0.4	$[-0.6, 0.1] \cup [0.4, 0.7]$	$[-0.7, 1.0]$
$\bar{c}_{uG}/\Lambda^2$	0.2	$[0, 0.3]$	$[-1.0, -0.9] \cup [-0.3, 0.4]$
$ \bar{c}_{uB}/\Lambda^2 $	1.6	$[0, 2.2]$	$[0, 2.7]$
$\bar{c}_{Hu}/\Lambda^2$	-9.3	$[-10.3, -8.0] \cup [0, 2.1]$	$[-11.1, -6.5] \cup [-1.6, 3.0]$
$\bar{c}_{2G}/\Lambda^2$	0.4	$[-0.9, -0.3] \cup [-0.1, 0.6]$	$[-1.1, 0.8]$

¹ In some cases the profile likelihood shows another local minimum that cannot be excluded; the number reported here is the global minimum.

And for three operators:

$$\begin{aligned}
\sigma(c_1, c_2, c_3) &\propto |\mathcal{M}_{\text{SM}} + c_1 \mathcal{M}_1 + c_2 \mathcal{M}_2 + c_3 \mathcal{M}_3|^2 \\
&\propto \underbrace{s_0}_{\text{constant}} + \underbrace{s_1 c_1 + s_2 c_2 + s_3 c_3}_{\text{linear}} + \underbrace{s_4 c_1^2 + s_5 c_2^2 + s_6 c_3^2}_{\text{quadratic}} + \underbrace{s_7 c_1 c_2 + s_8 c_1 c_3 + s_9 c_2 c_3}_{\text{mixed}}.
\end{aligned}$$

where the structure constants have been re-numbered in order for clarity. There are only constant, linear, and quadratic terms. For two operators, we can solve for the six structure constants by evaluating the cross section at six points (sets of c_1 and c_2 values); for three operators, we need 10 sets of c_1 , c_2 , and c_3 values. In general, if there are d operators under consideration, there will be N_s structure constants, where

$$N_s = 1 + 2d + \frac{d}{2}(d-1). \quad (7.5)$$

TABLE 7.12

SCALING DUE TO NP EFFECTS AT THE BEST-FIT c_1 (13 TeV)

Wilson coefficient	$\mu_{\text{t}\bar{\text{t}}\text{Z}}(c_{1,\text{bf}})$	$\mu_{\text{t}\bar{\text{t}}\text{W}}(c_{1,\text{bf}})$	$\mu_{\text{t}\bar{\text{t}}\text{H}}(c_{1,\text{bf}})$
$\bar{c}_{\text{uW}}$	1.20	1.08	1.05
$\bar{c}_{\text{uB}}$	1.18	1.00	1.06
$\bar{c}_{\text{H}}$	1.00	1.00	0.85
$\tilde{c}_{3\text{G}}$	1.14	1.00	1.33
$\bar{c}_{3\text{G}}$	1.17	1.00	1.14
$\bar{c}_{\text{Hu}}$	1.18	1.00	1.00
$\bar{c}_{2\text{G}}$	1.17	1.00	1.14
$\bar{c}_{\text{uG}}$	1.11	1.08	1.33

We need to evaluate the cross section at N_s sets of Wilson coefficient values to completely specify the system of equations. There is no a priori limit on the number of operators for which their effects can be parameterized with this procedure; more operators simply correspond to larger N_s . Restricting ourselves to $d = 2$ as an example, and dividing by σ_{SM} to obtain the signal scaling function, the system of equations corresponding to Equation (7.4) can be written as $AX = B$:

$$\begin{bmatrix} 1 & c_{11} & c_{12} & c_{11}^2 & c_{12}^2 & c_{11}c_{12} \\ 1 & c_{21} & c_{22} & c_{21}^2 & c_{22}^2 & c_{21}c_{22} \\ 1 & c_{31} & c_{32} & c_{31}^2 & c_{32}^2 & c_{31}c_{32} \\ 1 & c_{41} & c_{42} & c_{41}^2 & c_{42}^2 & c_{41}c_{42} \\ 1 & c_{51} & c_{52} & c_{51}^2 & c_{52}^2 & c_{51}c_{52} \\ 1 & c_{61} & c_{62} & c_{61}^2 & c_{62}^2 & c_{61}c_{62} \end{bmatrix} \begin{bmatrix} s_0 \\ s_1 \\ s_2 \\ s_3 \\ s_4 \\ s_5 \end{bmatrix} = \begin{bmatrix} \mu(c_{11}, c_{12}) \\ \mu(c_{21}, c_{22}) \\ \mu(c_{31}, c_{32}) \\ \mu(c_{41}, c_{42}) \\ \mu(c_{51}, c_{52}) \\ \mu(c_{61}, c_{62}) \end{bmatrix},$$

where we sample six c_1, c_2 points and evaluate $\mu(c_1, c_2)$; c_{11} and c_{12} correspond to the first point, c_{21} and c_{22} correspond to the second point, and so forth. In practice, because the MADGRAPH calculation has finite precision, solving the system of equations exactly is usually not possible. Therefore, we use an overdetermined least-squares fit to $N_d > N_s$ points, where we solve $AX = B$ to determine X by minimizing $\|B - AX\|^2$. We choose randomly distributed points in the range of coefficient values corresponding to $\mu(c_1) < 10$ for the 1D case. An example implementation in Python is presented in Figure 7.16.

Determining simultaneous constraints on operators using the approach described in this work critically depends on the determination of a reliable parameterization of the NP effects of the operators under consideration. To study the performance of higher dimensional simultaneous parameterizations, we started with the eight operators selected in the 13 TeV analysis. To determine the structure constants of the parameterization, we could scan each pair of coefficients separately, with all other coefficients set to zero (2D fit), or we could scan all eight of the coefficients simultaneously. If the MADGRAPH calculation was infinitely precise, the system of equations would be well-determined and the two approaches would be equivalent. We find, however, that under realistic conditions, scanning coefficients simultaneously leads to improved performance of the fit. The relative difference between the signal scaling calculated from MADGRAPH and that predicted from the fit for both approaches is given as a percent error in Figure 7.17. The 8D test yielded better results than the 2D test, which is likely a statistical effect. All 45 structure constants (see Equation (7.5)) have uncertainties, which are assumed to be randomly distributed, so in the simultaneous test, all the uncertainties are combined. The standard deviation of the sum of set of random variables decreases as the number of variables increases, producing a narrower distribution.

The dependence of the fit performance on the number of points included in the fit is shown in Figure 7.18 for $d = 2$ and $d = 8$, which correspond to $N_s = 6$ and $N_s = 45$, respectively. For the $d = 2$ case, all $\binom{8}{2} = 28$ combinations of Wilson coefficients are included. The fit performance improves sharply at N_s and continues to improve as more points are added.

```

import numpy as np
import itertools

# sampled coefficient values (two coefficients, six points)
points = np.array([
    [0.78, 0.41],
    [0.95, 0.46],
    [0.98, 0.21],
    [0.83, 0.85],
    [0.76, 0.03],
    [0.46, 0.29]
])

# NP scaling per point (from MadGraph calculation)
B = np.array([4.45, 5.55, 4.53, 7.02, 2.96, 2.74])

rows, dim = points.shape
pairs = list(itertools.combinations(range(0, dim), 2))
left = [x0 for x0, x1 in pairs]
right = [x1 for x0, x1 in pairs]

constant = np.array([[1.0]] * rows)
linear = points
quad = points * points
mixed = points[:, left] * points[:, right]

A = np.hstack([constant, linear, quad, mixed])

# let the fit constants we wish to calculate be represented by X;
# this solves  $AX = B$  by computing the  $X$  which minimizes  $\|B - AX\|^2$ 
X, _, _, _ = np.linalg.lstsq(A, B)

```

Figure 7.16. Example Python [112] code for parameterizing the NP effects on cross sections using NumPy [113]. Here, each column in `points` refers to a Wilson coefficient and each row corresponds to a set of Wilson coefficient values corresponding to the NP signal scaling in the corresponding row of `B`. The example includes six points in the parameter space of two Wilson coefficients, but is a general solution that would work for arbitrary-size input array `points`.

For $d = 2$, no improvement from the additional points is seen after about 15 points. For $d = 8$, the fit performance continues to improve slowly until around 500 points. After 45 points, which is required to completely determine the system of equations for the 8D fit, the performance of the 8D fit exceeds that of the 2D fit. This behavior is in agreement with the better performance of the 8D fit observed in Figure 7.17.

To visualize how pairs of operators interfere with each other, we plot the signal scaling in the plane formed from their Wilson coefficients. An example is shown for $\bar{c}_H$ and $\bar{c}_{2G}$ in Figures 7.19(a) and 7.19(b) and for $\bar{c}_H$ and $\bar{c}_{uG}$ in Figure 7.19(c). The NP effects on the $t\bar{t}Z$, $t\bar{t}H$, and $t\bar{t}W$ processes are shown in separate panels. Purple hues represent scaling that produces an expected cross section that is smaller than the SM, and green hues represent scaling that produces an expected cross section that is larger than the SM. Points that are approximately equivalent to the SM ($|\frac{\sigma_{NP+SM}}{\sigma_{SM}} - 1| \leq 0.1$) are plotted in gray. An entirely gray plane indicates that none of the included Wilson coefficient values are associated with significant scaling of the process under consideration, as is the case for $t\bar{t}W$ in the $\bar{c}_{uG}$, $\bar{c}_{2G}$ plane. Note that the origin, where both coefficients are zero, corresponds to the SM case, and will always be gray. In Figure 7.19(a) the top panels, each colored circle corresponds to one MADGRAPH calculation. In Figures 7.19(b) and 7.19(c), the color is interpolated using the simultaneous parameterization for eight operators. For all plots in Figure 7.19, all other Wilson coefficients are fixed to zero. Good agreement is seen between the MADGRAPH calculation and the parameterized fit for all pairs of coefficients.

Depending on the structure constants of the parameterization, the surface describing the NP scaling will be either an elliptic or hyperbolic paraboloid. Contours of equal scaling in the plane formed by the two coefficients will be either an ellipse or a hyperbola. In cases in which a large degree of interference exists between the two operators, the structure constant corresponding to their mixed term will be large, resulting in a tilted ellipse or parabola. An example of each can be seen in Figure 7.19(c).

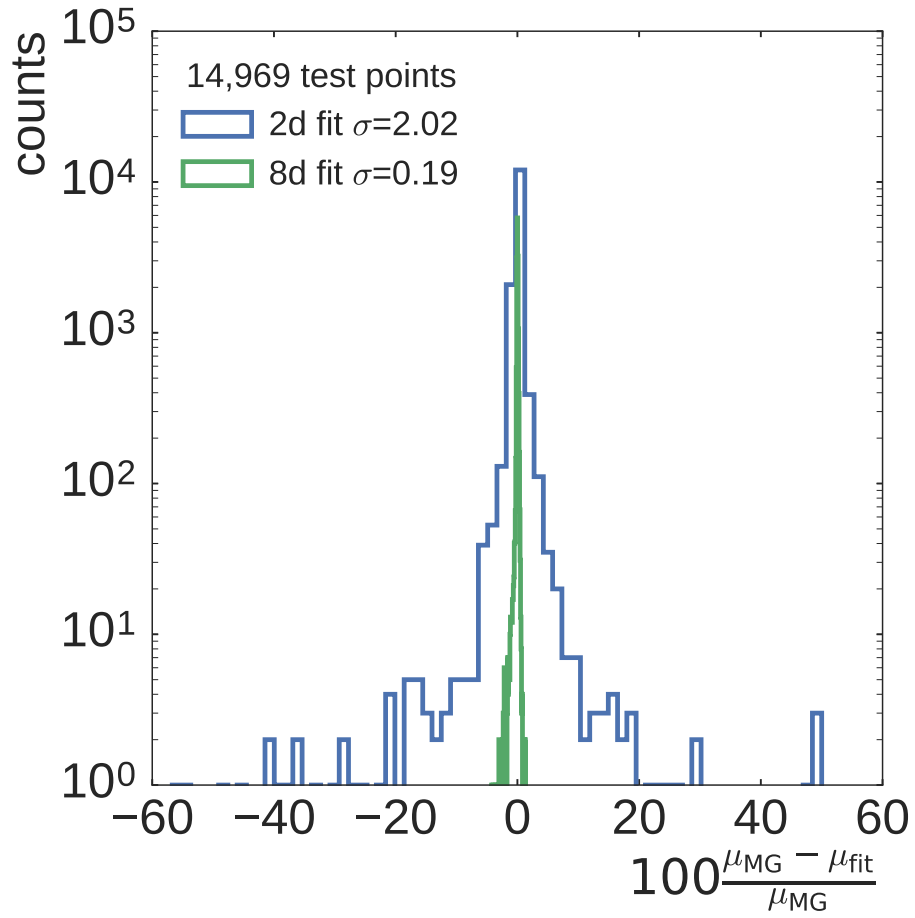


Figure 7.17. The percent error calculated as the relative difference between the signal scaling obtained from MADGRAPH (μ_{MG}) and that predicted from the fit (μ_{fit}). For the 2D fit, fit constants are determined with points from a 2D scan (performing $\binom{8}{2} = 28$ separate fits, each on a pair of two coefficients), or determining all fit constants simultaneously from points sampled in the 8D phase space of coefficient values (8D fit).

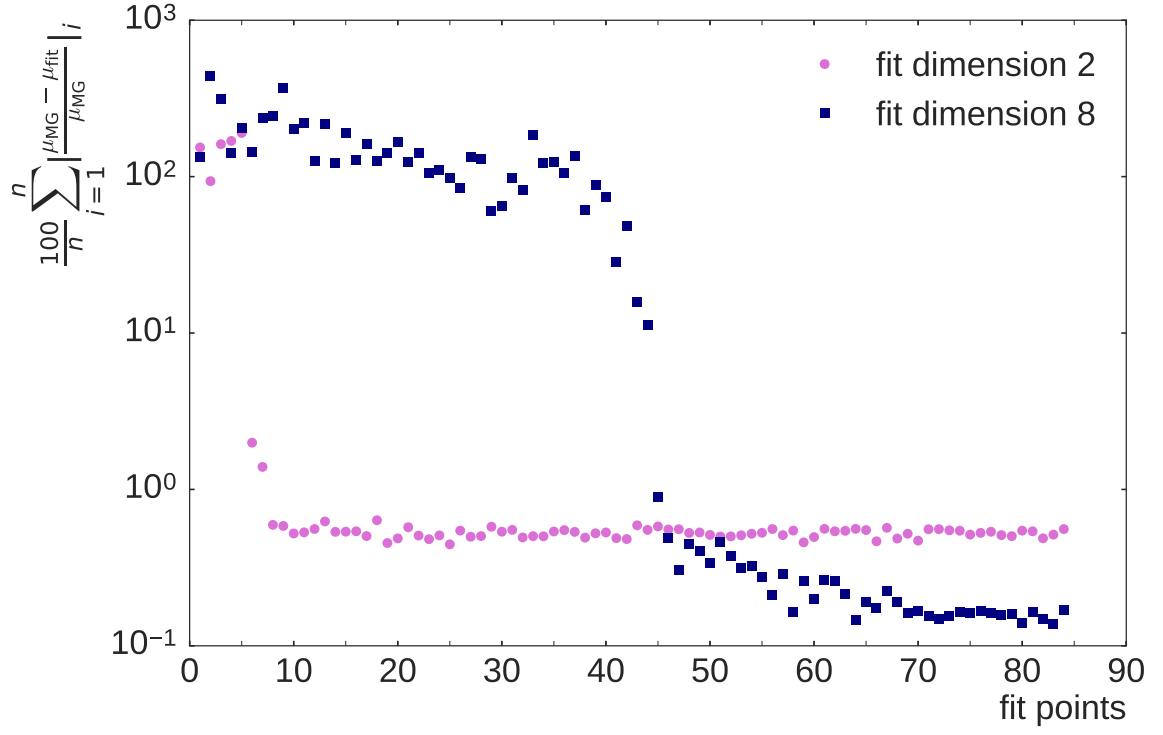


Figure 7.18. The average absolute percent error in the signal scaling parameterizations, as a function of the number of points in Wilson coefficient parameter space included in the fit. Parameterizations of the effects of two and eight operators are shown as purple circles and navy squares, respectively. The 14 969 test points are divided approximately equally among parameterizations for the $t\bar{t}H$, $t\bar{t}W$, and $t\bar{t}Z$ processes.

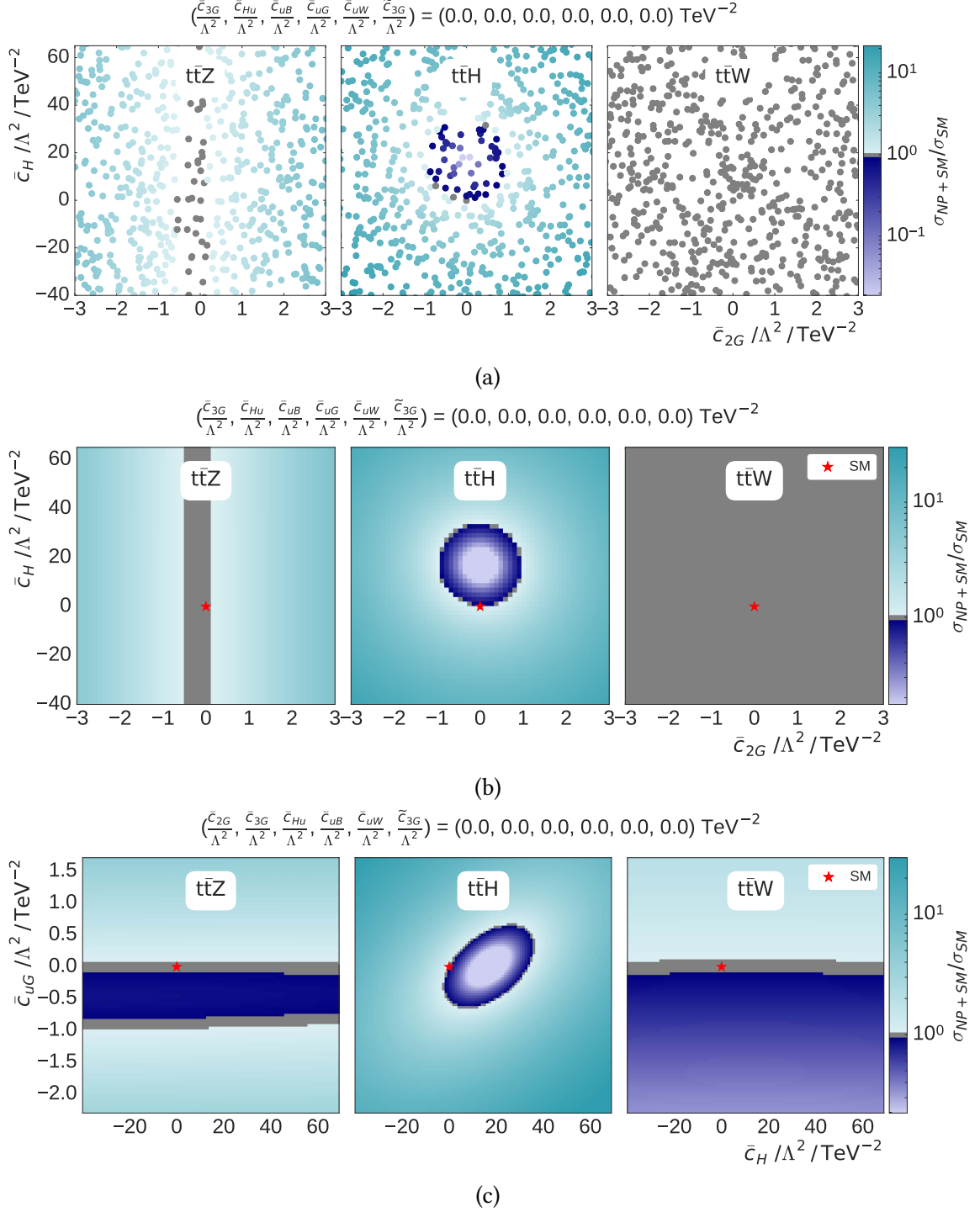


Figure 7.19. Cross section scaling shown in the $\tilde{c}_H$, $\tilde{c}_{2G}$ plane (a) and (b) and $\tilde{c}_H$, $\tilde{c}_{uG}$ plane (c) for $t\bar{t}Z$ (left), $t\bar{t}H$ (center), and $t\bar{t}W$ (right). The color represents the scaling ($\sigma_{NP+SM}/\sigma_{SM}$) due to NP effects. The star represents the SM point in which all $c_i = 0$. In (a), each colored circle corresponds to one MADGRAPH calculation.

In (b) and (c), the color is interpolated using the parameterized 8D fit.

TABLE 7.13
STRUCTURE CONSTANTS FOR $t\bar{t}Z$

	$\frac{\tilde{c}_{2G}}{\Lambda^2}$	$\frac{\tilde{c}_{3G}}{\Lambda^2}$	$\frac{\tilde{c}_H}{\Lambda^2}$	$\frac{\tilde{c}_{Hu}}{\Lambda^2}$	$\frac{\tilde{c}_{uB}}{\Lambda^2}$	$\frac{\tilde{c}_{uG}}{\Lambda^2}$	$\frac{\tilde{c}_{uW}}{\Lambda^2}$	$\frac{\tilde{c}_{3G}}{\Lambda^2}$
1.0	$2.11(1) \times 10^{-1}$	$-2.44(2) \times 10^{-1}$	$5(7) \times 10^{-5}$	$1.321(3) \times 10^{-1}$	$2.1(7) \times 10^{-3}$	$6.03(3) \times 10^{-1}$	$3.2(6) \times 10^{-3}$	$-1(30) \times 10^{-4}$
$\frac{\tilde{c}_{2G}}{\Lambda^2}$	$5.070(5) \times 10^{-1}$	$-1.110(1)$	$2(30) \times 10^{-6}$	$2.358(6) \times 10^{-2}$	$-5.1(1) \times 10^{-3}$	$3.46(1) \times 10^{-1}$	$4.7(1) \times 10^{-3}$	$1(70) \times 10^{-5}$
$\frac{\tilde{c}_{3G}}{\Lambda^2}$	-	$5.981(8) \times 10^{-1}$	$-2(50) \times 10^{-6}$	$-2.727(7) \times 10^{-2}$	$5.8(3) \times 10^{-3}$	$-3.88(1) \times 10^{-1}$	$-5.3(3) \times 10^{-3}$	$-7(800) \times 10^{-6}$
$\frac{\tilde{c}_H}{\Lambda^2}$	-	-	$3(1) \times 10^{-6}$	$8(5) \times 10^{-6}$	$-5(700) \times 10^{-8}$	$-4.6(6) \times 10^{-4}$	$2(6) \times 10^{-6}$	$3(30) \times 10^{-6}$
$\frac{\tilde{c}_{Hu}}{\Lambda^2}$	-	-	-	$1.618(1) \times 10^{-2}$	$-2(2) \times 10^{-5}$	$4.89(2) \times 10^{-2}$	$-6(1) \times 10^{-5}$	$-2(10) \times 10^{-5}$
$\frac{\tilde{c}_{uB}}{\Lambda^2}$	-	-	-	-	$7.759(7) \times 10^{-2}$	$-1.87(2) \times 10^{-2}$	$-1.313\,5(5) \times 10^{-1}$	$2(20) \times 10^{-5}$
$\frac{\tilde{c}_{uG}}{\Lambda^2}$	-	-	-	-	-	$6.73(2) \times 10^{-1}$	$1.72(5) \times 10^{-2}$	$-2(20) \times 10^{-4}$
$\frac{\tilde{c}_{uW}}{\Lambda^2}$	-	-	-	-	-	-	$6.527(6) \times 10^{-2}$	$-9(300) \times 10^{-6}$
$\frac{\tilde{c}_{3G}}{\Lambda^2}$	-	-	-	-	-	-	-	$5.98(1) \times 10^{-1}$

TABLE 7.14

STRUCTURE CONSTANTS FOR $t\bar{t}W$

	$\frac{\tilde{c}_{2G}}{\Lambda^2}$	$\frac{\tilde{c}_{3G}}{\Lambda^2}$	$\frac{\tilde{c}_H}{\Lambda^2}$	$\frac{\tilde{c}_{Hu}}{\Lambda^2}$	$\frac{\tilde{c}_{uB}}{\Lambda^2}$	$\frac{\tilde{c}_{uG}}{\Lambda^2}$	$\frac{\tilde{c}_{uW}}{\Lambda^2}$	$\frac{\tilde{c}_{3G}}{\Lambda^2}$
1.0	$-1(40) \times 10^{-5}$	$2(70) \times 10^{-5}$	$-6(4) \times 10^{-5}$	$-6.5(8) \times 10^{-4}$	$-3(2) \times 10^{-4}$	$4.68(1) \times 10^{-1}$	$1.22(3) \times 10^{-2}$	$7(700) \times 10^{-6}$
$\frac{\tilde{c}_{2G}}{\Lambda^2}$	$8(200) \times 10^{-6}$	$-2(40) \times 10^{-5}$	$1(100) \times 10^{-7}$	$-3(200) \times 10^{-7}$	$-2(40) \times 10^{-6}$	$6(400) \times 10^{-6}$	$-3(50) \times 10^{-6}$	$1(20) \times 10^{-5}$
$\frac{\tilde{c}_{3G}}{\Lambda^2}$	-	$1(30) \times 10^{-5}$	$-9(2000) \times 10^{-8}$	$4(200) \times 10^{-7}$	$2(80) \times 10^{-6}$	$-2(400) \times 10^{-6}$	$3(100) \times 10^{-6}$	$-2(20) \times 10^{-5}$
$\frac{\tilde{c}_H}{\Lambda^2}$	-	-	$2.6(8) \times 10^{-6}$	$-2(100) \times 10^{-8}$	$-2(200) \times 10^{-8}$	$-1(20) \times 10^{-6}$	$3.2(3) \times 10^{-5}$	$8(1000) \times 10^{-8}$
$\frac{\tilde{c}_{Hu}}{\Lambda^2}$	-	-	-	$1.90(4) \times 10^{-4}$	$1.05(5) \times 10^{-4}$	$-3(40) \times 10^{-6}$	$-2.220(5) \times 10^{-3}$	$2(300) \times 10^{-7}$
$\frac{\tilde{c}_{uB}}{\Lambda^2}$	-	-	-	-	$1.12(2) \times 10^{-3}$	$2(70) \times 10^{-6}$	$-1.0(2) \times 10^{-3}$	$-8(500) \times 10^{-7}$
$\frac{\tilde{c}_{uG}}{\Lambda^2}$	-	-	-	-	-	$7.87(6) \times 10^{-2}$	$1(2000) \times 10^{-7}$	$1(50) \times 10^{-5}$
$\frac{\tilde{c}_{uW}}{\Lambda^2}$	-	-	-	-	-	-	$2.047(3) \times 10^{-2}$	$2(100) \times 10^{-6}$
$\frac{\tilde{c}_{3G}}{\Lambda^2}$	-	-	-	-	-	-	-	$1(30) \times 10^{-5}$

TABLE 7.15

STRUCTURE CONSTANTS FOR $t\bar{t}H$

	$\frac{\tilde{c}_{2G}}{\Lambda^2}$	$\frac{\tilde{c}_{3G}}{\Lambda^2}$	$\frac{\tilde{c}_H}{\Lambda^2}$	$\frac{\tilde{c}_{Hu}}{\Lambda^2}$	$\frac{\tilde{c}_{uB}}{\Lambda^2}$	$\frac{\tilde{c}_{uG}}{\Lambda^2}$	$\frac{\tilde{c}_{uW}}{\Lambda^2}$	$\frac{\tilde{c}_{3G}}{\Lambda^2}$
1.0	$-7.7(4) \times 10^{-2}$	$8.9(5) \times 10^{-2}$	$-1.189(3) \times 10^{-1}$	$-4(4) \times 10^{-4}$	$4(1) \times 10^{-3}$	1.660(7)	$6.4(9) \times 10^{-3}$	$-3(6) \times 10^{-3}$
$\frac{\tilde{c}_{2G}}{\Lambda^2}$	1.10(2)	-2.410(5)	$4.7(2) \times 10^{-3}$	$-6(20) \times 10^{-5}$	$-3(6) \times 10^{-4}$	1.50(5)	$4(6) \times 10^{-4}$	$2(4) \times 10^{-3}$
$\frac{\tilde{c}_{3G}}{\Lambda^2}$	-	1.370(3)	$-5.4(2) \times 10^{-3}$	$1(30) \times 10^{-5}$	$4(8) \times 10^{-4}$	-1.750(6)	$-3(7) \times 10^{-4}$	$-2(4) \times 10^{-3}$
$\frac{\tilde{c}_H}{\Lambda^2}$	-	-	$3.536(7) \times 10^{-3}$	$3(1) \times 10^{-5}$	$-2.6(4) \times 10^{-4}$	$-9.94(3) \times 10^{-2}$	$-4.2(3) \times 10^{-4}$	$-2(20) \times 10^{-5}$
$\frac{\tilde{c}_{Hu}}{\Lambda^2}$	-	-	-	$1.4(2) \times 10^{-4}$	$7.4(6) \times 10^{-4}$	$-4(4) \times 10^{-4}$	$-2.10(6) \times 10^{-3}$	$2(3) \times 10^{-4}$
$\frac{\tilde{c}_{uB}}{\Lambda^2}$	-	-	-	-	$2.08(1) \times 10^{-2}$	$-2(1) \times 10^{-3}$	$1.8(2) \times 10^{-3}$	$3(8) \times 10^{-4}$
$\frac{\tilde{c}_{uG}}{\Lambda^2}$	-	-	-	-	-	3.10(6)	$1.2(9) \times 10^{-3}$	$3(6) \times 10^{-3}$
$\frac{\tilde{c}_{uW}}{\Lambda^2}$	-	-	-	-	-	-	$1.333(9) \times 10^{-2}$	$-2(6) \times 10^{-4}$
$\frac{\tilde{c}_{3G}}{\Lambda^2}$	-	-	-	-	-	-	-	1.370(3)

The structure constants for the parameterized fit to the eight selected Wilson coefficients are given for $t\bar{t}Z$, $t\bar{t}W$, and $t\bar{t}H$, respectively, in Table 7.13, Table 7.14, and Table 7.15. Each cell in the table corresponds to a term in the parameterization. The cell in the first row and first column corresponds to the term that is linear in $\bar{c}_{2G}$; the cell in the second row and column corresponds to the $\bar{c}_{2G}$, $\bar{c}_{3G}$ mixed term; the cell in third row and second column corresponds to the term that is quadratic in $\bar{c}_{3G}$; and so forth. Now that we have parameterized the cross section scaling due to the NP effects of all eight selected operators, we can replace

$$\mu(c_1) \rightarrow \mu(\bar{c}_{uW}, \bar{c}_H, \bar{c}_{3G}, \tilde{c}_{3G}, \bar{c}_{uG}, \bar{c}_{uB}, \bar{c}_{Hu}, \bar{c}_{2G})$$

in Equation (7.3) and proceed as before to maximize the likelihood. Recall from Equation (5.6) that when fitting a single parameter, the $1 - \alpha$ CL interval is a range of parameter values with endpoints that can be calculated from the quantile function of a χ^2 distribution with 1 degree of freedom as $-2\Delta \ln L = F^{-1}(1 - \alpha)$. When the two parameters are being fit, the $1 - \alpha$ CL region is defined by contours of constant $-2\Delta \ln L = F^{-1}(1 - \alpha)$ with $n = 2$. We perform fits considering the $\binom{8}{2} = 28$ pairs of coefficients, with the other six coefficients fixed to zero. These likelihood scans are included in Appendix B. When $n > 2$, the $1 - \alpha$ CL is defined by an n -dimensional surface of constant $-2\Delta \ln L = F^{-1}(1 - \alpha)$. There is a probability of $1 - \alpha$ that the true value of all n parameters lies inside that surface. Such a surface is not readily visualized. Instead, in Table 7.16, the “bounding box” of the 8D CL surface is presented for a fit to all eight coefficients. The range for each coefficient corresponds to the dimensions along that axis of a box that would contain the CL surface.

7.4 Future directions

The approach taken in the current work to constrain dimension-six operators by parameterizing the cross section scaling due to NP effects as a function of the Wilson coefficients is an important first step, but it has several serious shortcomings. In considering one operator

TABLE 7.16

CL SURFACE BOUNDARIES FOR 8D SIMULTANEOUS FIT

	0.68 % CL / TeV^{-2}	0.95 % CL / TeV^{-2}
$\bar{c}_{2G}/\Lambda^2$	[-10, 10]	[-12, 12]
$\bar{c}_{3G}/\Lambda^2$	[-9.2, 8.6]	[-10, 9.9]
$\bar{c}_H/\Lambda^2$	[-25, 52]	[-31, 57]
$\bar{c}_{Hu}/\Lambda^2$	[-13, 6.2]	[-14, 7.3]
$\bar{c}_{uB}/\Lambda^2$	[-7.8, 7.3]	[-8.5, 7.9]
$\bar{c}_{uG}/\Lambda^2$	[-1.4, 1.0]	[-1.6, 1.1]
$\bar{c}_{uW}/\Lambda^2$	[-7.5, 6.8]	[-8.2, 7.4]
$\tilde{c}_{3G}/\Lambda^2$	[-1.3, 1.3]	[-1.4, 1.4]

a time, the Wilson coefficient has two values that correspond to the same scaling of the cross section. As the number of operators considered grows, the surfaces of equal scaling grow; this is a fundamental degeneracy. To keep the complexity of the analysis manageable, operators that had large effects on backgrounds to $t(\bar{t})X$ and $t\bar{t}Z$ – mostly on $t(\bar{t})X$, VH , and triboson processes– were eliminated from consideration. In general, the effect on these processes varies from operator to operator, so a global fit that considers the NP effects on all background processes simultaneously will aid in resolving these degeneracies. Additionally, this work relies on the assumption that NP effects do not significantly alter acceptances and efficiencies. With infinite computing power, we could produce MC simulation data at full detector level, covering the entire phase space of possible Wilson coefficient values. In this way, we could properly account for acceptance and efficiency effects due to NP. NP effects on kinematic distributions would further resolve degeneracies. In practice, combinatorics renders this approach futile for large-dimensional studies and makes it very difficult to use

advanced analysis techniques, such as MVAs. The matrix element reweighting technique for obtaining simulated data in a large parameter phase space is presented in the next section, and a novel (to our knowledge) prescription for extending that approach with the parameterization techniques presented in this work is described in Section 7.4.2.

7.4.1 Matrix element reweighting

To address the difficulty in obtaining simulated data in large parameter phase spaces, a technique called matrix element reweighting can be employed. This technique allows for fully simulated MC samples to be “re-used” for arbitrary models by taking advantage of the fact that the parton-level generation, parton-shower and hadronization, detector simulation, and reconstruction stages of the simulation are all independent. It is therefore possible to transform a sample from representing a reference model hypothesis to an alternative one by adjusting event-level weights. The weights are proportional to the modulus of the matrix element squared evaluated at the corresponding phase space point, and at leading order, they can be transformed according to:

$$w_{\text{new}} = w_{\text{orig}} \frac{|\mathcal{M}_{\text{new}}|^2}{|\mathcal{M}_{\text{orig}}|^2} \quad (7.6)$$

where $\mathcal{M}_{\text{new}}$ and w_{new} correspond to the new matrix element and weight, and $\mathcal{M}_{\text{orig}}$ and w_{orig} correspond to the original one. In the limit of infinite MC samples, reweighting reproduces traditional simulation exactly. Under realistic conditions, successful application of this technique requires that the reference model be chosen with care. It is critical that the phase space of the new model be a subset of the reference model. If there is a large contribution to the new model in a region of phase space where the reference model has a small number of events, the statistical uncertainty of the new model will be large. This is demonstrated in Figure 7.20, which shows, for three different reference models, a comparison between the signal scaling as a function of $\bar{c}_{\text{uB}}$ evaluated using independent samples and using one

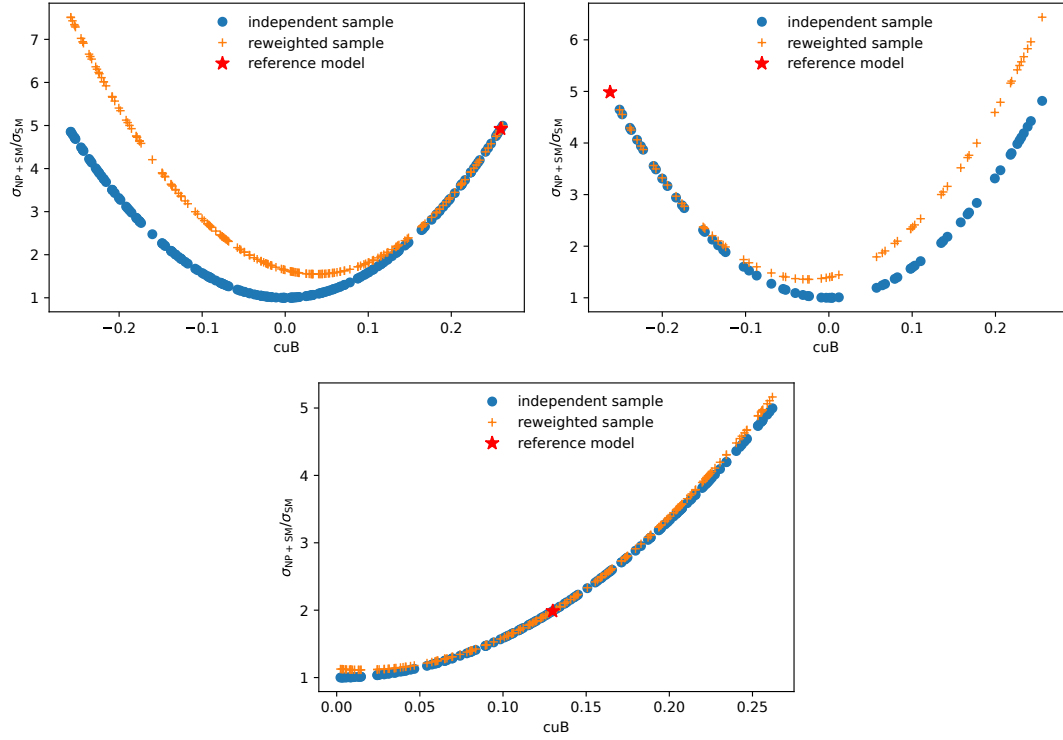


Figure 7.20. Signal scaling as a function of $\bar{c}_{uB}$, for independent samples (circles), and one sample reweighted to each value of $\bar{c}_{uB}$ (crosses). The reference model is indicated by the red star.

sample, reweighted to each $\bar{c}_{uB}$ value.

A more detailed description of the matrix element reweighting method can be found in references [114, 115].

7.4.2 Parameterized matrix element reweighting

Matrix element reweighting is an extremely powerful tool, but by itself it does not alleviate the “combinatoric explosion” of high-dimensionality parameter phase spaces. Though it saves a huge amount of computing power, computations in large dimensional spaces still quickly become unfeasible. Assuming the use of a single-precision floating point format, requiring 32 bits of memory per weight, even storing the weights for one event evaluated at just 10 values per parameter in a 15 dimensional parameter space would require $10^{15} \times 32$ bits =

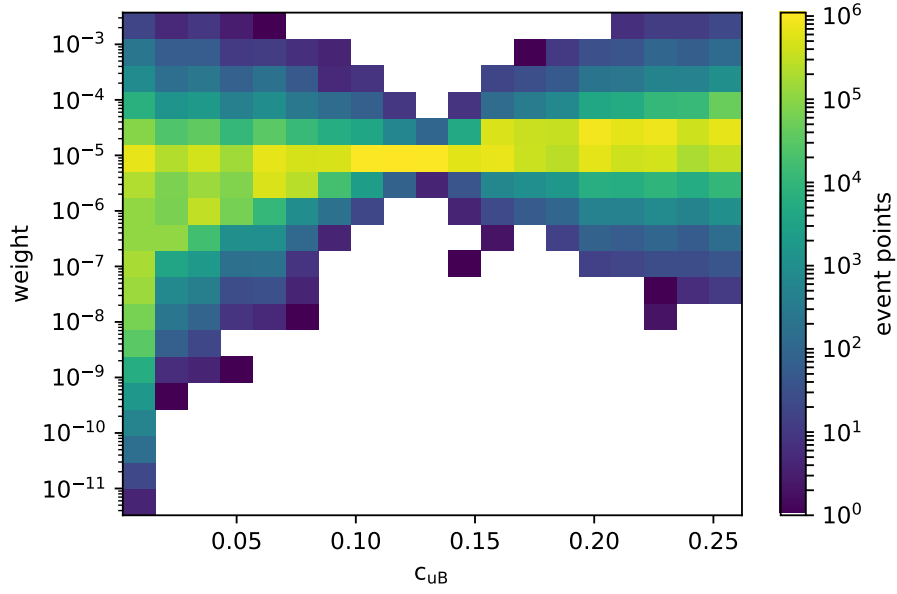
4 000 TB. Note that Equation (7.6) is proportional to the matrix element squared: the structure of the transformation between new and old weights can be parameterized analogously to what is described in the previous sections, which for one non-zero Wilson coefficient is

$$\begin{aligned} w_{\text{new}}(c_1) &= w_{\text{orig}} \frac{|\mathcal{M}_{\text{SM}} + c_1 \mathcal{M}|^2}{|\mathcal{M}_{\text{orig}}|^2} \\ &= w_{\text{orig}} \frac{s_0 + s_1 c_1 + s_2 c_1^2}{|\mathcal{M}_{\text{orig}}|^2}. \end{aligned}$$

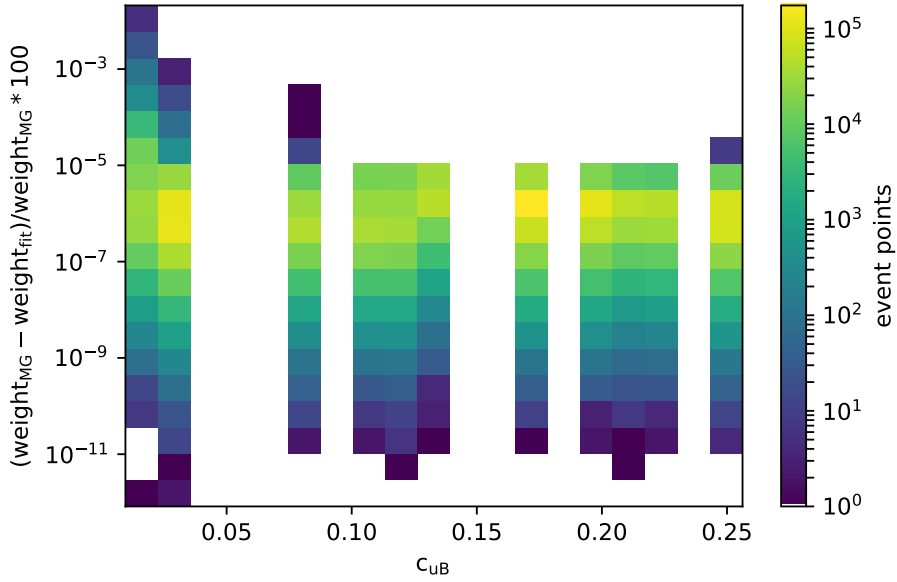
This result can be generalized to parameterize the effects of an arbitrary number of dimension-six operators. Previously, we evaluated the cross section at a number of points in Wilson coefficient phase space, then performed a least-squares fit to obtain a parameterization of the cross section per process. In this case, we evaluate the weight at a number of points in Wilson coefficient phase space and parameterize the weights per event. The required number of parameterizations is larger, but each one is computationally cheap. For the 8D cross section parameterization case, the improvement to the fit from additional points in 8D phase space asymptotes around 500 points. Performing 100 000 fits in 1D (one for each event) to 134 parameter points took approximately 34 seconds running in a single-threaded process, while performing 50 000 fits in 8D to 200 parameter points took approximately 100 seconds. The top panel of Figure 7.21 presents an example showing the distribution of event weights as a function of $\bar{c}_{\text{uB}}$ for a sample of 100 000 events that was produced with a reference model of $\bar{c}_{\text{uB}} = 0.1$ and then subsequently reweighted for 134 points randomly sampled between $-0.41 < \bar{c}_{\text{uB}} < 0.41$. The largest weights are associated with the most extreme values of $\bar{c}_{\text{uB}}$. In the bottom panel of Figure 7.21, the percent error between the weight obtained from MADGRAPH and the weight calculated from the parameterized fits is shown. The plot only includes a set of test points for each event that were not included in the fit. The distribution of errors is reasonably consistent throughout the range of $\bar{c}_{\text{uB}}$. In Figure 7.22 the performance of 50 000 fits to 227 points in the 8D phase space of parameters selected in the previous section is presented. The coefficients were sampled randomly in the range corresponding

to $\mu_{\text{tZ}}(c_1) < 10$ for the one-dimensional scaling due to each coefficient c_1 . The top panel of Figure 7.22 shows the relative error as a function of the weight for a set of test points that were not included in the fit. A trend towards larger relative errors for smaller weights is apparent, which suggests some constant source of absolute error. The center panel shows the weight as a function of the cross section scaling associated with the point. The trend towards larger weights for larger cross section scaling means that the error associated with these points is higher. In the bottom panel, the distribution of the weights is shown for the parameterized fit overlaid with the calculation from MADGRAPH, for a selection of Wilson coefficient points that were not included in the fit. Good agreement is observed.

These results demonstrate that parameterizing matrix element weights in high-dimensional spaces is feasible. Future work must determine an optimal procedure for choosing the reference model, evaluate how uncertainties in the parameterization of the weights should be properly handled, and understand the conditions under which weighted samples are sufficiently close approximations of independent ones. The large range of weights, corresponding to very different statistical uncertainties, will be a challenge. If those problems can be addressed satisfactorily, then the parameterization of the weights for each event can be determined and saved when an event is simulated. Detector-level analyses can proceed in the usual way, without restrictions regarding the use of advanced techniques such as MVAs. As a final step, the discriminant histograms can be reweighted to reflect a desired point in high-dimensional parameter space. The cost of this operation will be low enough that it could be performed for each step of the likelihood minimization.

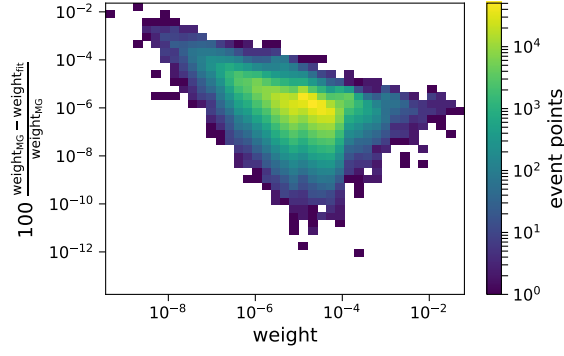


(a)

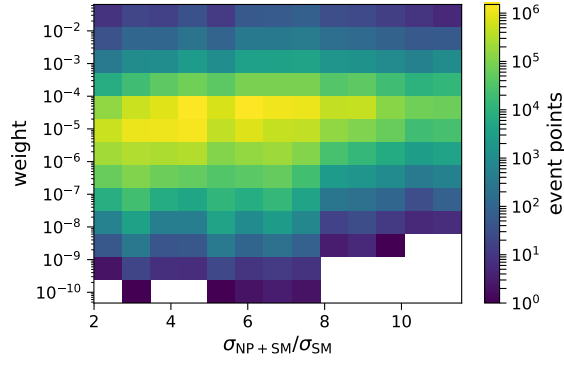


(b)

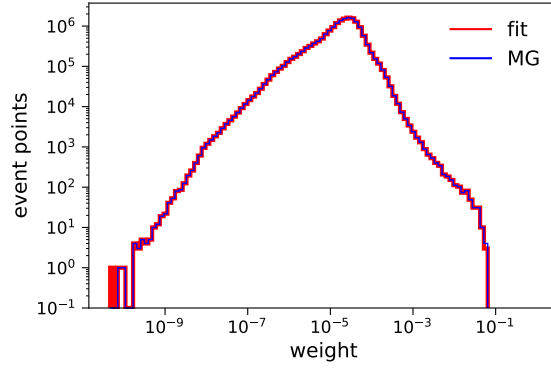
Figure 7.21. Top: the distribution of weights as a function of $\bar{c}_{uB}$, for a reference model with $\bar{c}_{uB} = 0.13$. There are 100 000 events in the sample, reweighted to 172 different values of $\bar{c}_{uB}$. Bottom: the distribution of percent error between the weight reported from MADGRAPH and the result of the parameterization, for a selection of test points which were not included in the parameterization.



(a)



(b)



(c)

Figure 7.22. Top: The percent error between the weight reported from MADGRAPH and the result of the parameterization, as a function of the event weight. Center: The weight as a function of the signal scaling. Bottom: the distribution of weights, calculated from MADGRAPH (blue) and from the fit (red), for test points which were not included in the fit. All panels are derived from a 100 000 event sample which was reweighted at 168 randomly sampled points in 8D phase space. The reference model was $(\bar{c}_H, \bar{c}_{Hu}, \bar{c}_{uB}, \bar{c}_{uW}, \bar{c}_{uG}, \bar{c}_{3G}, \bar{c}_{2G}, \bar{c}_{3G}) = (8.7 \times 10^{-1}, 8.9 \times 10^{-1}, 1.2 \times 10^{-4}, -5.9 \times 10^{-2}, 3.2 \times 10^{-3}, -4.7 \times 10^{-3}, -1.0 \times 10^{-2}, 2.2 \times 10^{-3})$.

CHAPTER 8

CONCLUSIONS

In the face of limited collider energy, and as researchers looking for striking evidence of NP in traditional “bump hunts” at the LHC continue to come up empty-handed, it is necessary to find new ways of searching for NP. To that end, the main goal of this dissertation was to interpret measurements of $t\bar{t}W$ and $t\bar{t}Z$ within the model-independent framework of effective field theory to constrain the Wilson coefficients of dimension-six operators that would affect $t\bar{t}W$ and $t\bar{t}Z$. Effective field theory interpretations are challenging because the large phase space of possible Wilson coefficient values make traditional modes of analysis untenable. In this dissertation, this difficulty was managed by making the simplifying assumption that NP has a negligible effect on the kinematics of studied processes and only scales their overall production. We determined a parameterization of the cross section scaling due to NP effects which could be evaluated as a substitute for producing dedicated samples at each studied point in the phase space of possible Wilson coefficient values. The first iteration of this analysis used LHC Run 1 proton-proton data at $\sqrt{s} = 8$ TeV to set constraints on the Wilson coefficients of five dimension-six operators. Building on lessons learned during the 8 TeV analysis, we used LHC Run 2 proton-proton data at $\sqrt{s} = 13$ TeV to set constraints on eight dimension-six operators. The second iteration of the analysis incorporated several improvements. We used a more sophisticated method of selecting which operators to study, made improvements that increased the quality of the parameterization, and removed couplings to the first two generations from the NP model. Both the 8 TeV and the main part of the 13 TeV analyses studied the special case in which only one Wilson coefficient is enhanced at a time, while the others are fixed at zero. As a

first step towards global constraints, preliminary results of simultaneous fits for pairs of coefficients and to the Wilson coefficients of all eight selected operators were also presented.

For a given process, many combinations of Wilson coefficient values may correspond to identical scaling of the cross section. These surfaces of equal scaling are a fundamental degeneracy that cannot be resolved using the approach taken in this dissertation. Nevertheless, our work has hinted at promising directions for future investigation. We carefully studied the effects of various operators on $t\bar{t}W$ and $t\bar{t}Z$ background processes and we identified several operators that had a large effect on expected background yields. We found particularly large effects for triboson processes, ZH , WH , and the associated production of a single top quark and a Z or Higgs boson. Future work should add dedicated signal regions for these processes. Because the NP effects vary from process to process, additional sensitivity would be gained from performing a simultaneous fit with these additional signal regions.

Building on our work developing reliable parameterizations of the scaling effects of NP on cross sections, we have presented a technique for parameterizing event weights that can be used to model the effects of various combinations of Wilson coefficient values with a single simulation sample. We have demonstrated good agreement between the weights derived from the parameterization and the weights calculated in the traditional manner. Future work could use this strategy to perform a wide variety of detector-level analyses, taking advantage of sophisticated techniques for exploiting kinematic discriminants which would further aid in resolving degeneracies. As data accumulate and analysis techniques are refined, stronger limits on NP coupling parameters or the discovery of NP effects will be possible.

APPENDIX A

REPRODUCIBILITY

Reproducibility is critical to the scientific method as it is taught, but the reality is far from this ideal. Could someone read this dissertation in fifty years and reproduce the analysis with only the description it contains? Could a junior student in my group extend this work to test something new? These are difficult and important questions. In this appendix, I describe the modest progress I have made on the least ambitious aspect of reproducibility: Can I reliably reproduce my own work?

Analytical research progresses like the growth of a tree, with a multitude of related investigations frequently branching off in new directions. A researcher focuses on one area, makes improvements, then moves on to another portion of the work. Changes in one aspect of the analysis often affect others, but disentangling the cause of changes is not always straightforward. Sometimes when the researcher returns to a portion of the analysis after some time, it is difficult to precisely reproduce the most recent output of that phase of the work.

At the start of my graduate studies, my code was a mess¹ of scripts, many inherited from previous graduate students and in various stages of decay, which I used to produce files with names like²

```
test_new_June3_v13_justOne_v2_lepCut_final_really_final.root.
```

As a student, I had the good fortune of participating in the High Energy Physics (HEP) computing group, which was a collaboration at Notre Dame between members of the experimental HEP group, the Center for Research Computing, and Professor Douglas Thain's

¹It is still a mess, but it is a slightly more orderly mess.

²This is an actual example I found deep in the detritus of old working directories.

Cooperative Computing Laboratory in the Computer Science department. This fruitful collaboration yielded important progress in the study of reproducibility³ [123, 124, 125]. To address the problem of being able to reliably reproduce my own work, I made use of a software tool called Makeflow [126] developed at The Cooperative Computing Laboratory.

The code for the EFT analysis described in Chapter 7 is structured such that there is a single entry point executable to which the location of a configuration file is passed as an argument. The configuration file specifies the location of all inputs, a list of the desired output plots and tables, and a path to an output directory. Calling this executable produces an output directory at the requested location, where it

1. copies all of the inputs⁴ (including the configuration file);
2. writes a Makeflow script (which specifies all commands to be run to produce the outputs);
3. writes a git patch, allowing one to exactly reproduce the version of the code Makeflow invokes to produce the outputs (even if the researcher has forgotten to commit changes); and
4. writes instructions for using these pieces to produce (or reproduce) the outputs, which are saved in the output directory.

The advantage of this approach is that it allows for versioning not just of the code (which is already possible with version control software such as git), but also of the inputs and outputs. This makes comparing the results at various stages in analysis development relatively straightforward. If any intermediate results in the analysis chain (specified in the Makeflow script) in a particular output directory are changed, the Makeflow script can be rerun and Makeflow will intelligently only update any dependent results. A new git patch

³Another notable product of this collaboration was a software tool for harnessing non-dedicated computing resources for HEP workflows called Lobster [116, 117, 118, 119, 120, 121], which is built on top of Work Queue [122] and Parrot, both developed at the Cooperative Computing Laboratory. Lobster was critical to the analysis described in Chapter 7, which would not have been feasible with conventional means of accessing computing resources.

⁴In this case, only the result of the SM analysis, which can be stored in small text files, is versioned. Applying this approach to the SM analysis, which processes much larger input files, would require a more sophisticated method of ensuring preservation of the correct input file versions.

will be copied such that the output directory always contains the necessary information to reproduce the analysis code. This gives users fine-grained control of how often to save versions of their output.

I have associated Digital Object Identifiers (DOIs) with the analysis code and input file versions used in this dissertation [127, 128] and uploaded them to the Zenodo repository. Zenodo claims it will maintain access to the code for at least 20 years. This is not sufficient to easily reproduce the results of the EFT analysis: even restricting our consideration to this simpler case, setting up the proper software dependencies and execution environment will still be required. Nevertheless it may be a step in the right direction, providing an example which could be extended in future analyses as new tools to aid in making scientific work more reproducible are developed.

APPENDIX B

PRELIMINARY CROSS SECTION

SCALING AND LIKELIHOOD SCANS FOR PAIRS OF SELECTED COEFFICIENTS

This appendix presents the cross section scaling due to all $\binom{8}{2} = 28$ combinations of Wilson coefficients proportional to the eight selected operators, for $t\bar{t}W$, $t\bar{t}Z$, and $t\bar{t}H$. Additionally, the profile likelihood scans for these combinations is presented.

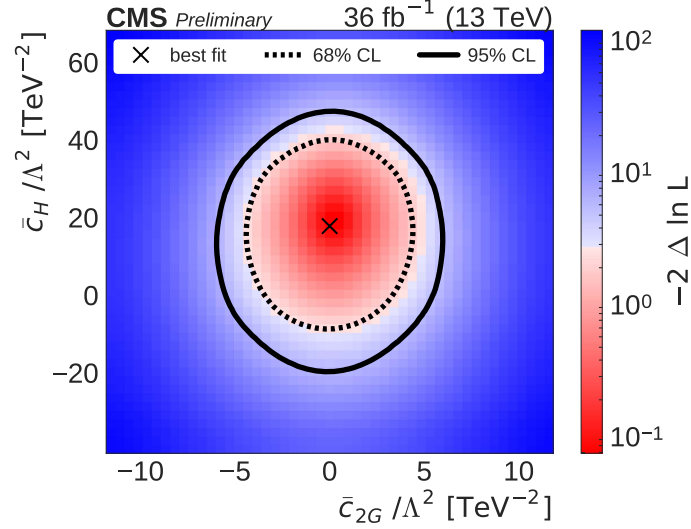
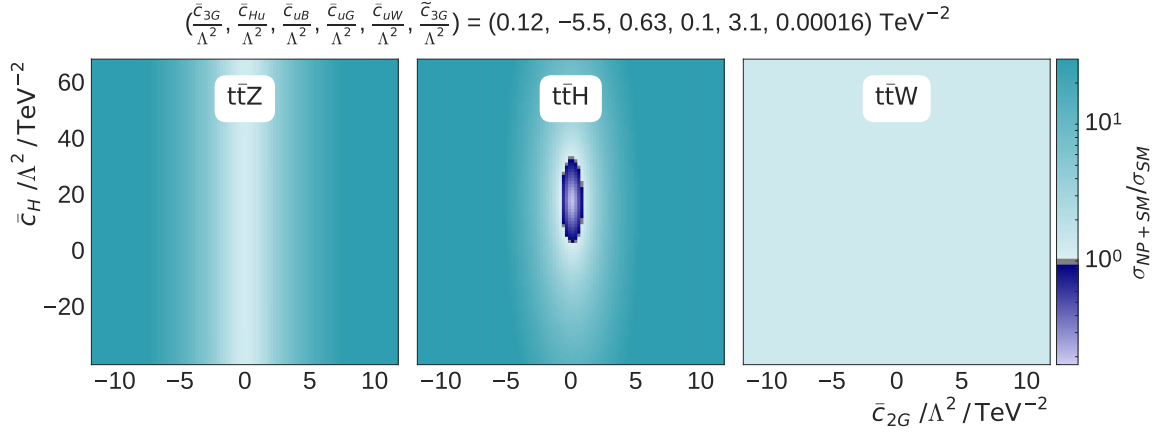
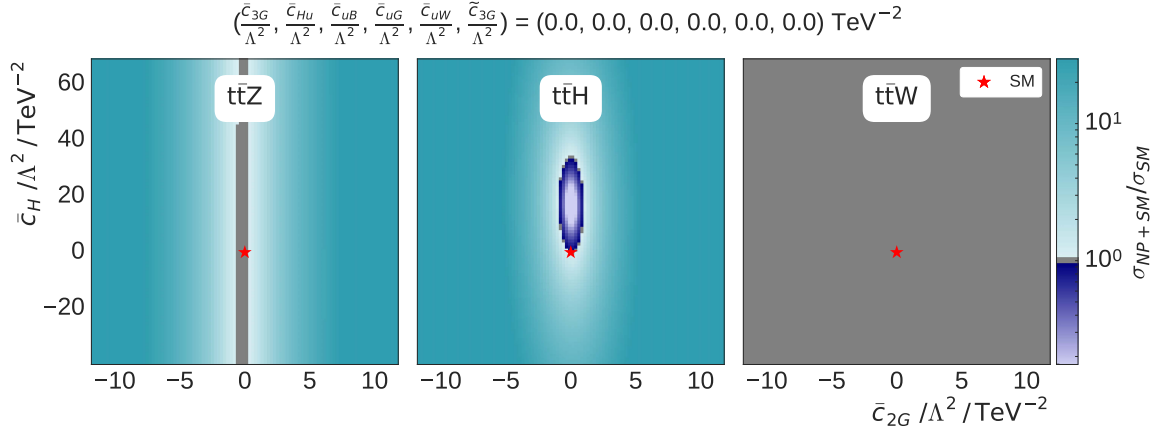


Figure B.1. Signal scaling shown in the $\bar{c}_H$, $\bar{c}_{2G}$ plane with all other coefficients fixed to zero (a) to their best-fit values (b) for $t\bar{t}Z$ (left), $t\bar{t}H$ (center), and $t\bar{t}W$ (right). The color represents the scaling ($\sigma_{\text{NP} + \text{SM}}/\sigma_{\text{SM}}$) due to NP effects. The star represents the SM point in which all $c_i = 0$. The negative log likelihood is shown in (c). The best fit is represented by a cross. The 68 % and 95 % CL contours are shown with dashed and solid lines, respectively.

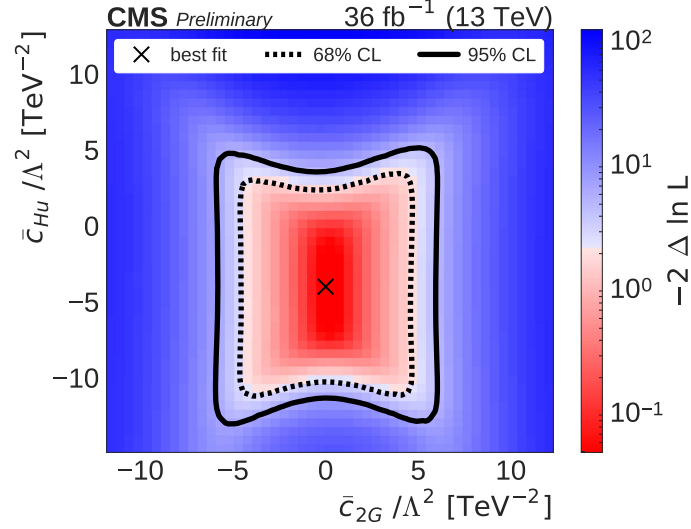
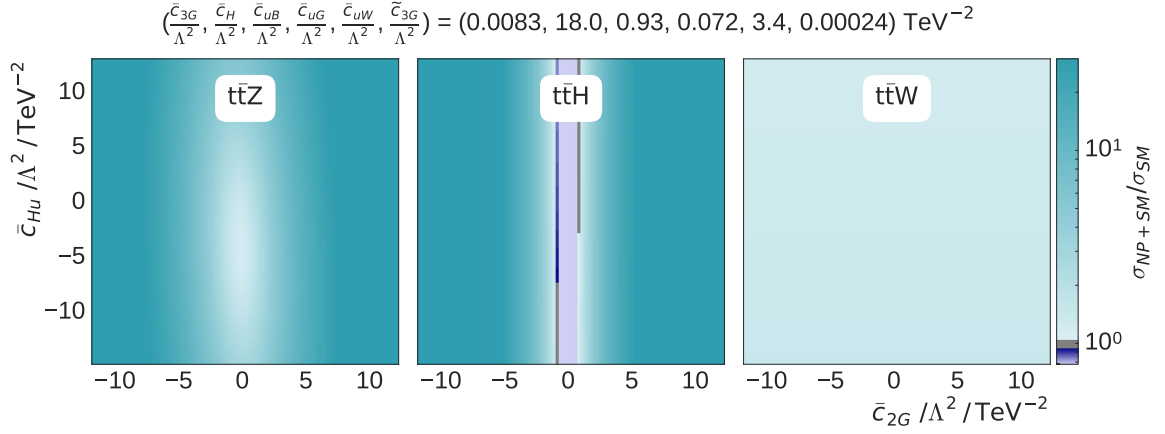
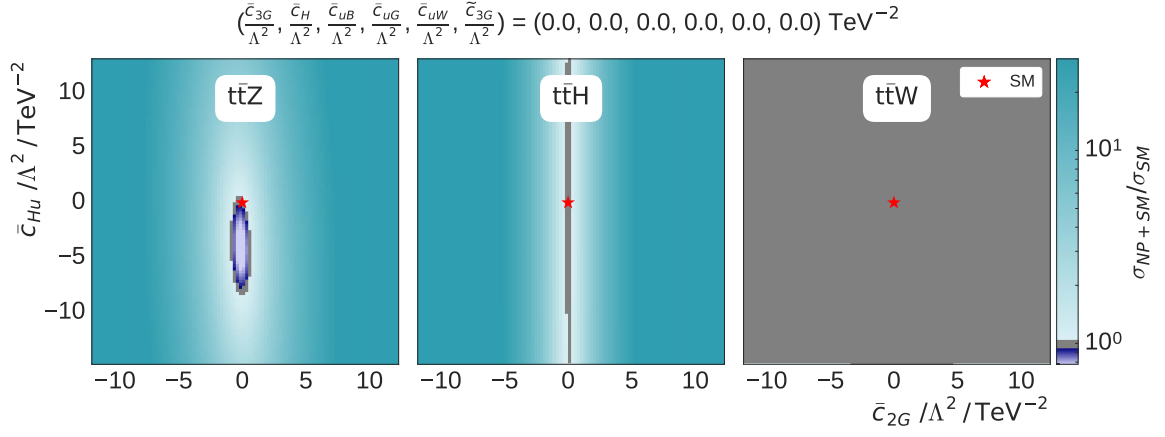


Figure B.2. Signal scaling shown in the $\bar{c}_{Hu}$, $\bar{c}_{2G}$ plane with all other coefficients fixed to zero (a) to their best-fit values (b) for $t\bar{t}Z$ (left), $t\bar{t}H$ (center), and $t\bar{t}W$ (right). The color represents the scaling ($\sigma_{NP+SM}/\sigma_{SM}$) due to NP effects. The star represents the SM point in which all $c_i=0$. The negative log likelihood is shown in (c). The best fit is represented by a cross. The 68 % and 95 % CL contours are shown with dashed and solid lines, respectively.

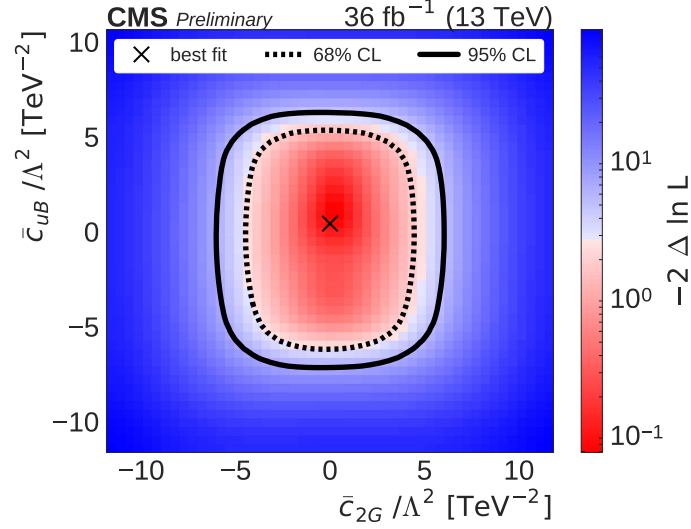
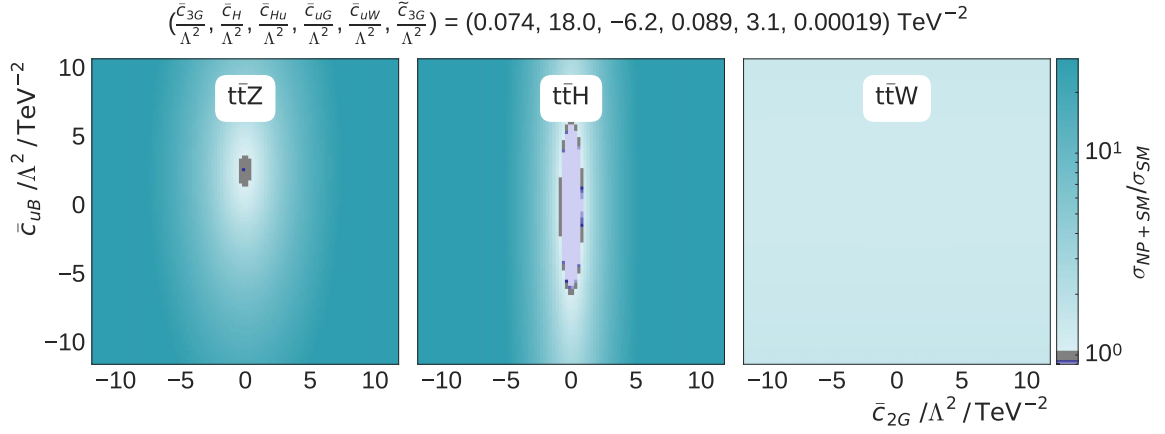
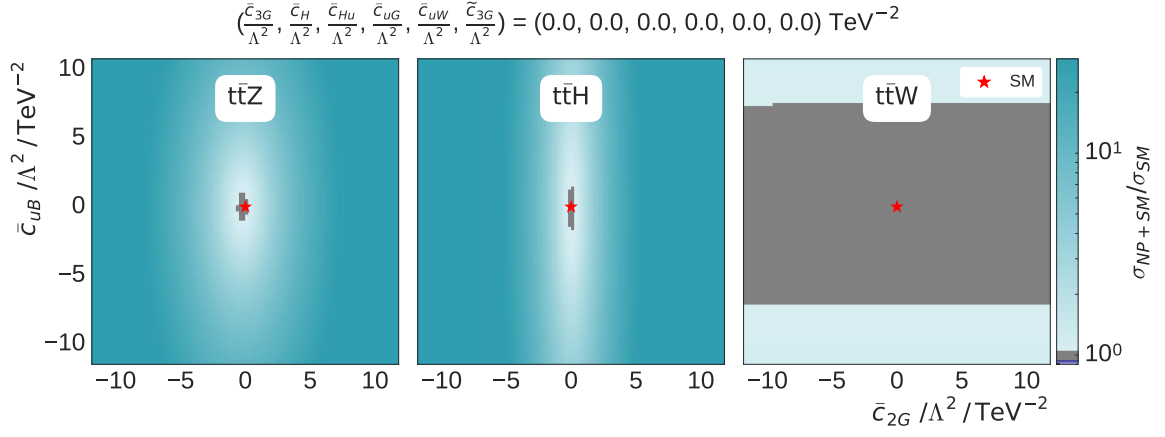


Figure B.3. Signal scaling shown in the $\tilde{c}_{2G}$, $\tilde{c}_{uB}$ plane with all other coefficients fixed to zero (a) or their best-fit values (b) for $t\bar{t}Z$ (left), $t\bar{t}H$ (center), and $t\bar{t}W$ (right). The color represents the scaling ($\sigma_{\text{NP}+\text{SM}}/\sigma_{\text{SM}}$) due to NP effects. The star represents the SM point in which all $c_i=0$. The negative log likelihood is shown in (c). The best fit is represented by a cross. The 68 % and 95 % CL contours are shown with dashed and solid lines, respectively.

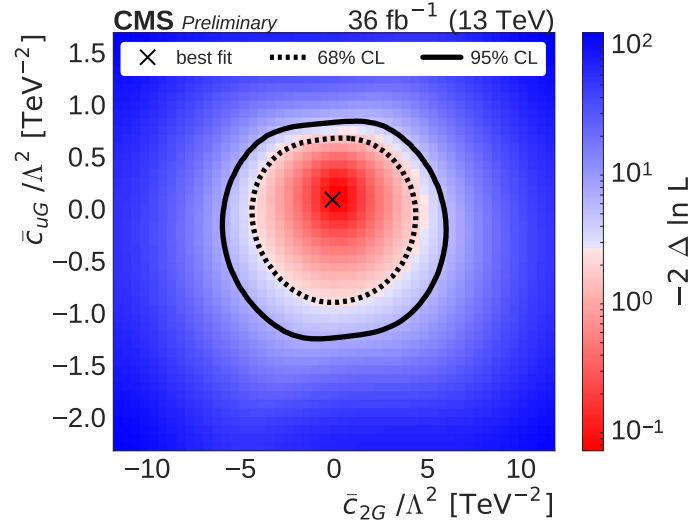
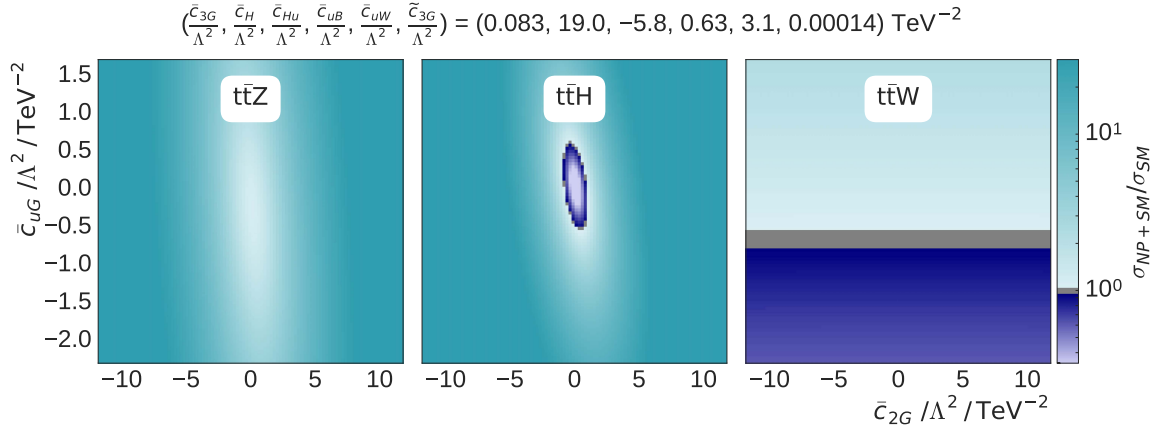
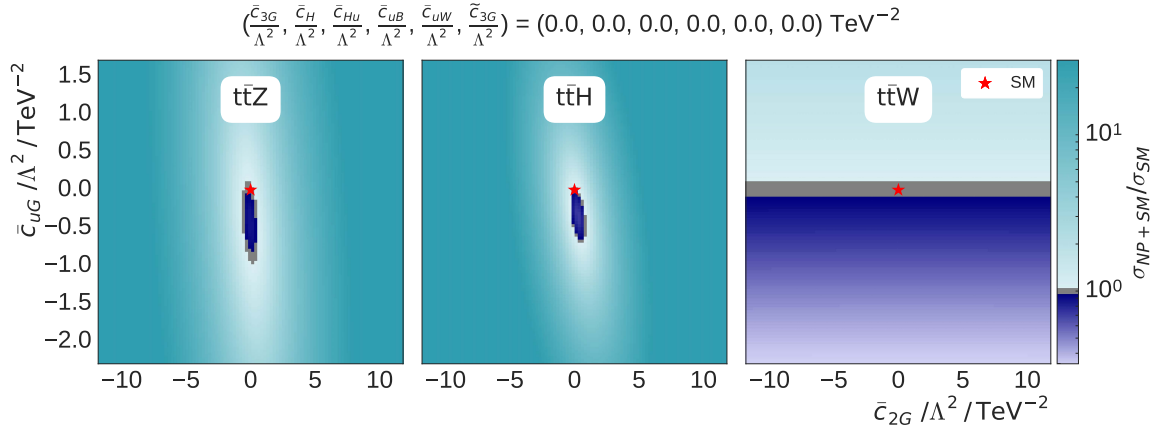


Figure B.4. Signal scaling shown in the $\bar{c}_{uG}$, $\bar{c}_{2G}$ plane with all other coefficients fixed to zero (a) or their best-fit values (b) for $t\bar{t}Z$ (left), $t\bar{t}H$ (center), and $t\bar{t}W$ (right). The color represents the scaling ($\sigma_{NP+SM}/\sigma_{SM}$) due to NP effects. The star represents the SM point in which all $c_i=0$. The negative log likelihood is shown in (c). The best fit is represented by a cross. The 68 % and 95 % CL contours are shown with dashed and solid lines, respectively.

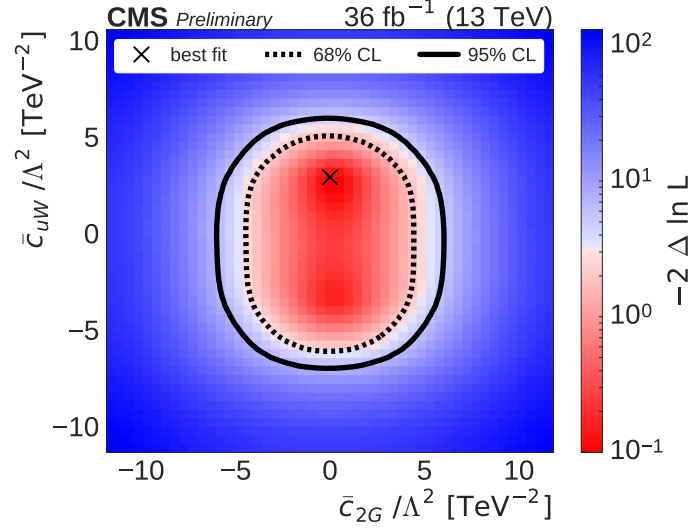
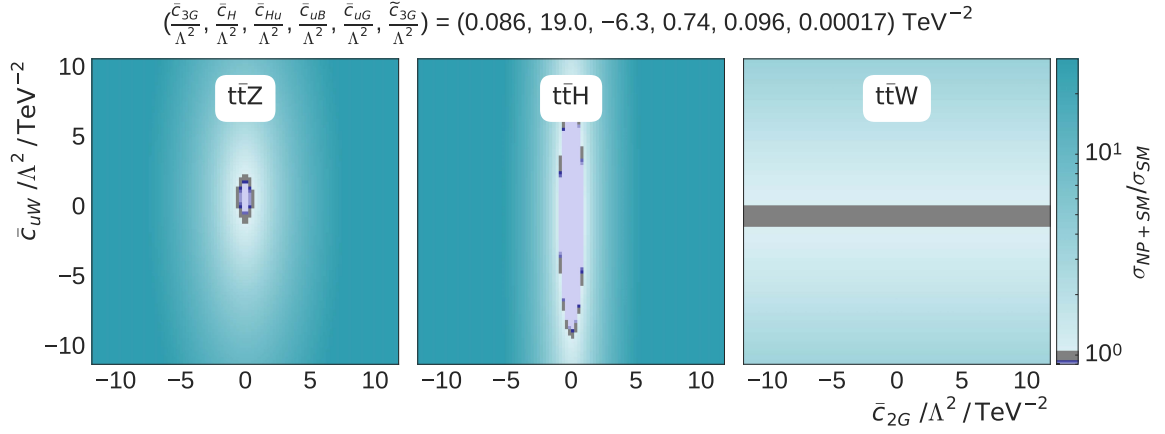
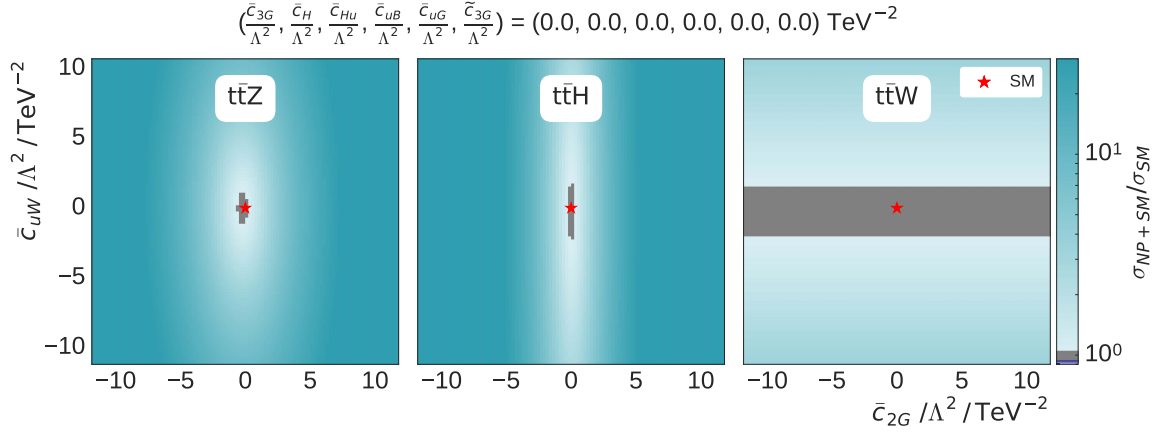


Figure B.5. Signal scaling shown in the $\bar{c}_{uW}$, $\bar{c}_{2G}$ plane with all other coefficients fixed to zero (a) or their best-fit values (b) for $t\bar{t}Z$ (left), $t\bar{t}H$ (center), and $t\bar{t}W$ (right). The color represents the scaling ($\sigma_{\text{NP} + \text{SM}}/\sigma_{\text{SM}}$) due to NP effects. The star represents the SM point in which all $c_i = 0$. The negative log likelihood is shown in (c). The best fit is represented by a cross. The 68 % and 95 % CL contours are shown with dashed and solid lines, respectively.

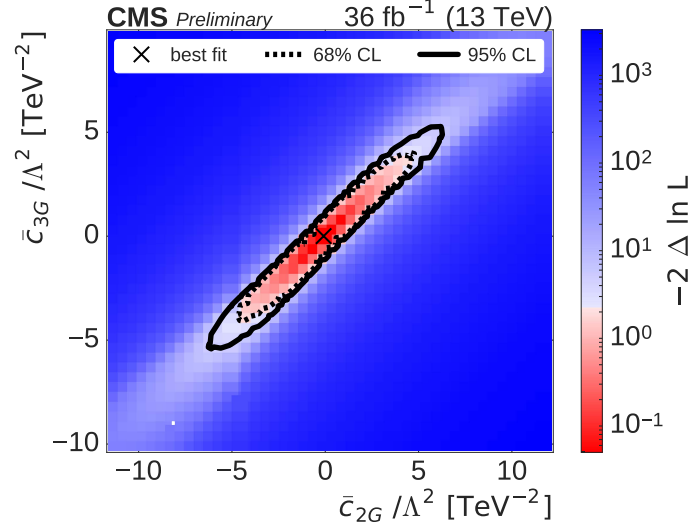
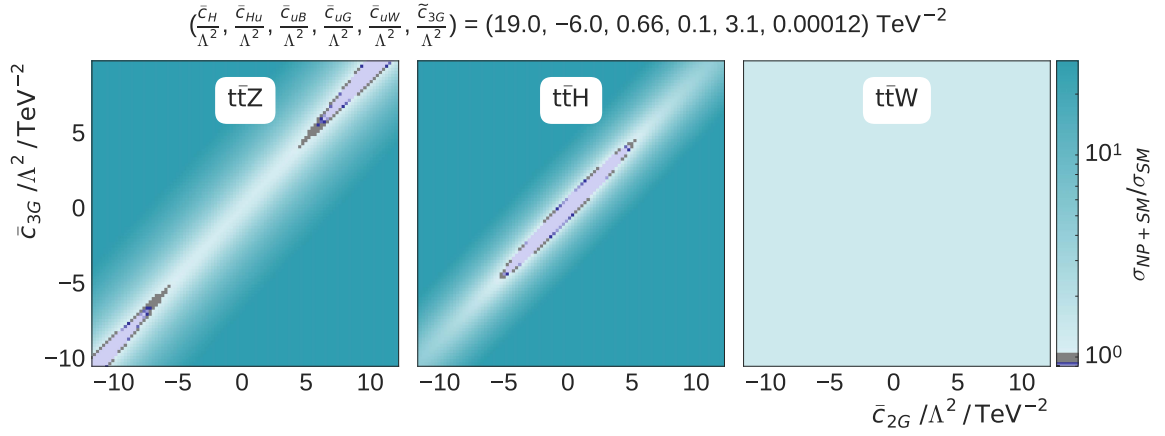
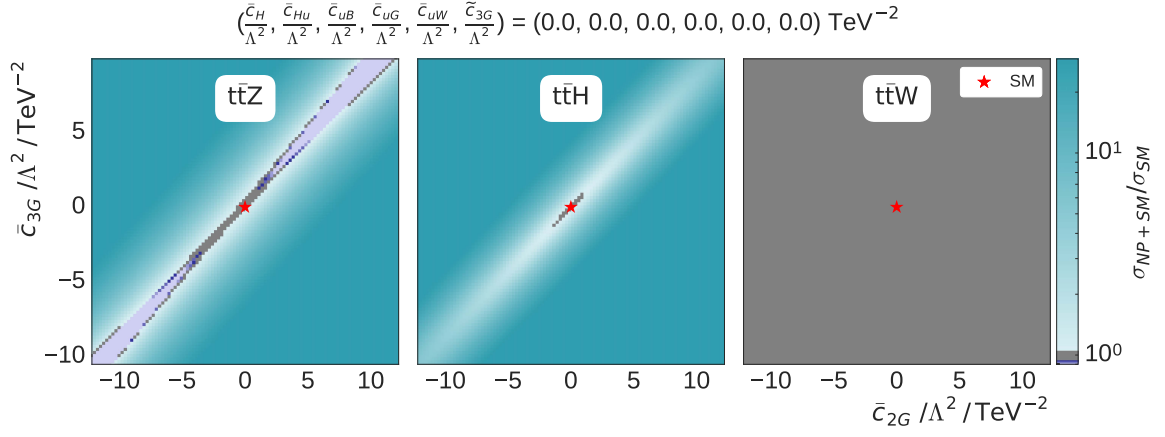


Figure B.6. Signal scaling shown in the $\bar{c}_{3G}$, $\bar{c}_{2G}$ plane with all other coefficients fixed to zero (a) or their best-fit values (b) for $t\bar{t}Z$ (left), $t\bar{t}H$ (center), and $t\bar{t}W$ (right). The color represents the scaling ($\sigma_{\text{NP}+\text{SM}}/\sigma_{\text{SM}}$) due to NP effects. The star represents the SM point in which all $c_i=0$. The negative log likelihood is shown in (c). The best fit is represented by a cross. The 68 % and 95 % CL contours are shown with dashed and solid lines, respectively.

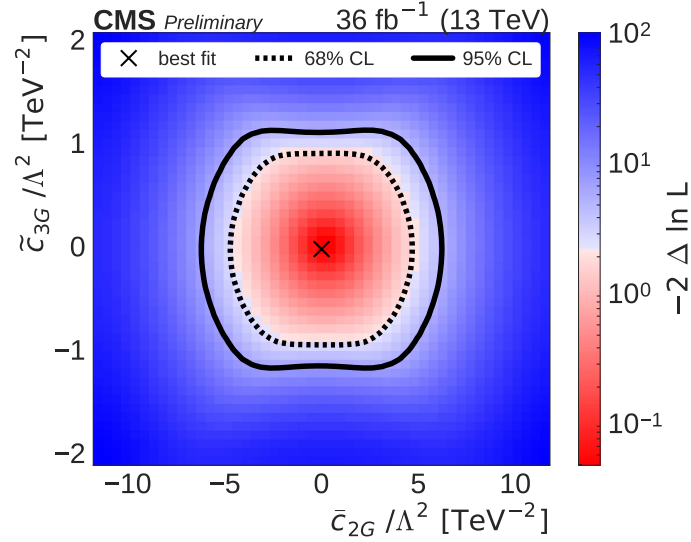
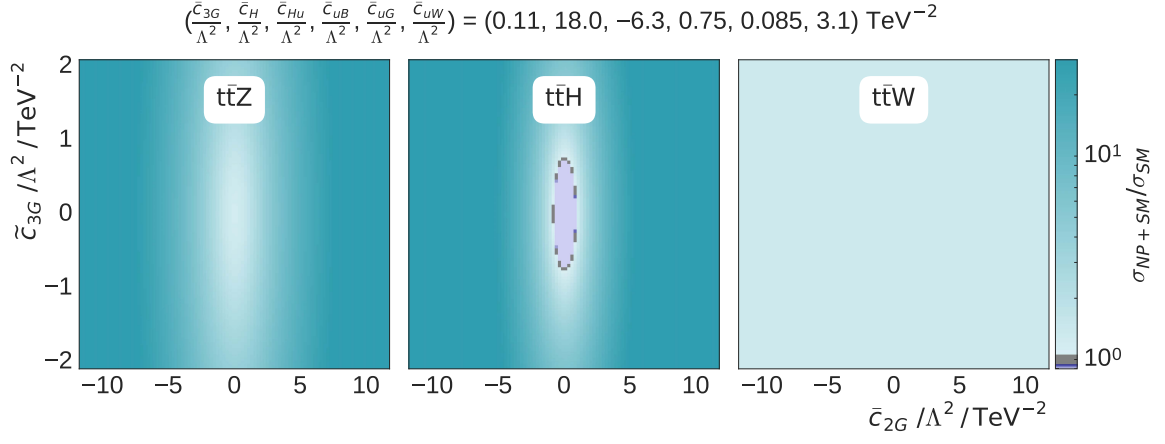
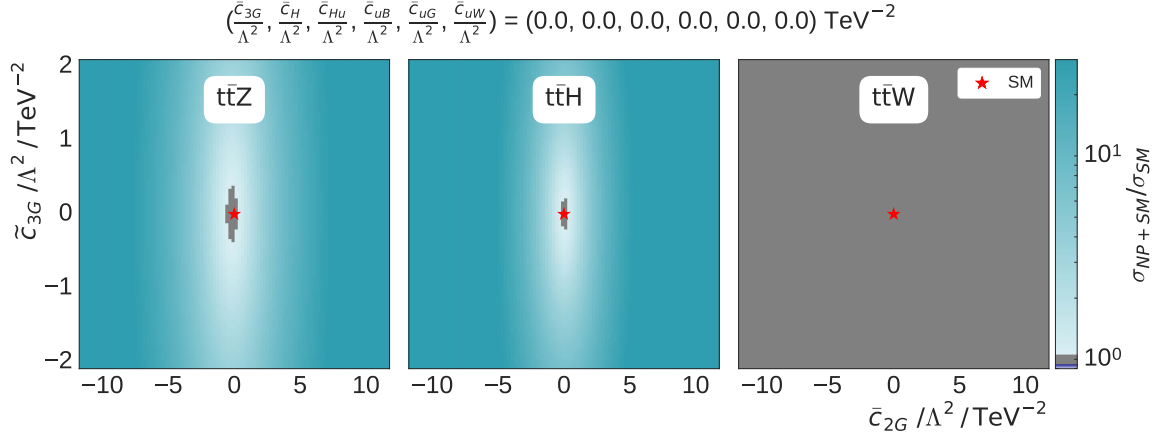


Figure B.7. Signal scaling shown in the $\tilde{c}_{3G}$, $\tilde{c}_{2G}$ plane with all other coefficients fixed to zero (a) or their best-fit values (b) for $t\bar{t}Z$ (left), $t\bar{t}H$ (center), and $t\bar{t}W$ (right). The color represents the scaling ($\sigma_{NP+SM}/\sigma_{SM}$) due to NP effects. The star represents the SM point in which all $c_i=0$. The negative log likelihood is shown in (c). The best fit is represented by a cross. The 68 % and 95 % CL contours are shown with dashed and solid lines, respectively.

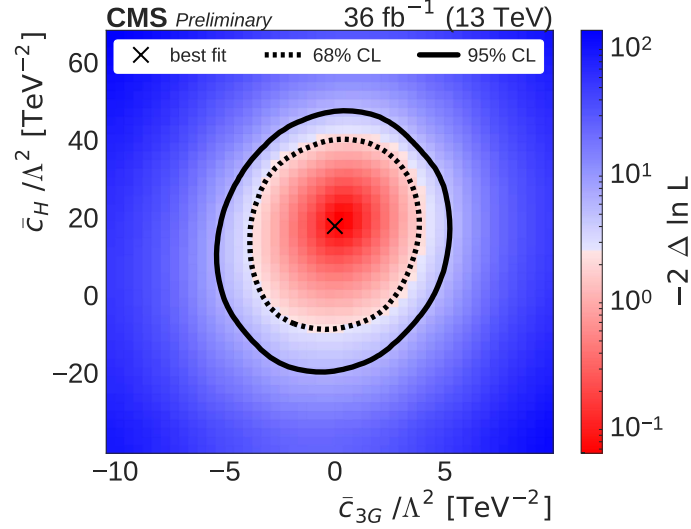
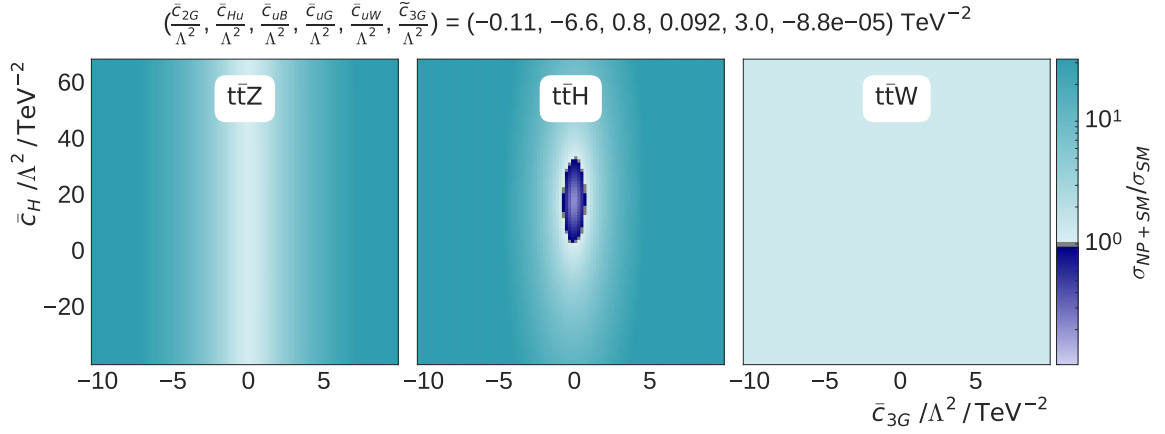
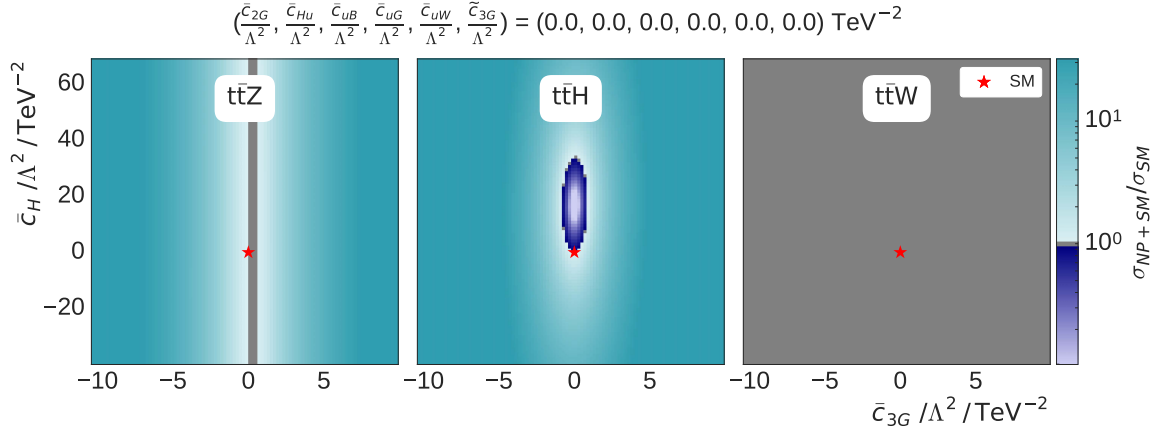


Figure B.8. Signal scaling shown in the $\bar{c}_H$, $\bar{c}_{3G}$ plane with all other coefficients fixed to zero (a) or their best-fit values (b) for $t\bar{t}Z$ (left), $t\bar{t}H$ (center), and $t\bar{t}W$ (right). The color represents the scaling ($\sigma_{\text{NP}+\text{SM}}/\sigma_{\text{SM}}$) due to NP effects. The star represents the SM point in which all $c_i=0$. The negative log likelihood is shown in (c). The best fit is represented by a cross. The 68 % and 95 % CL contours are shown with dashed and solid lines, respectively.

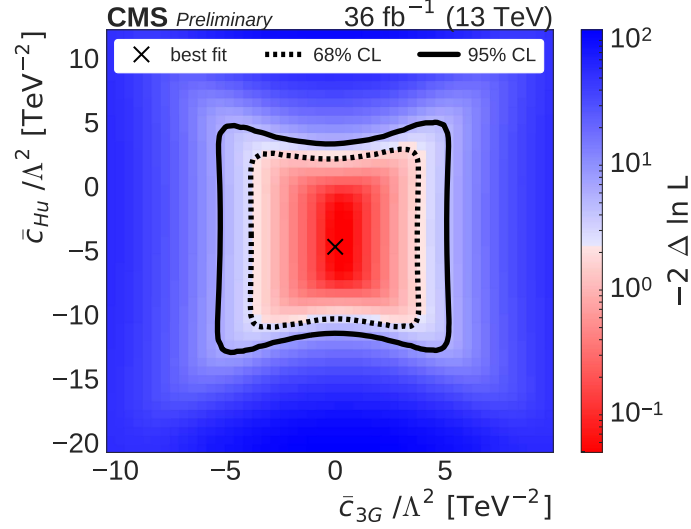
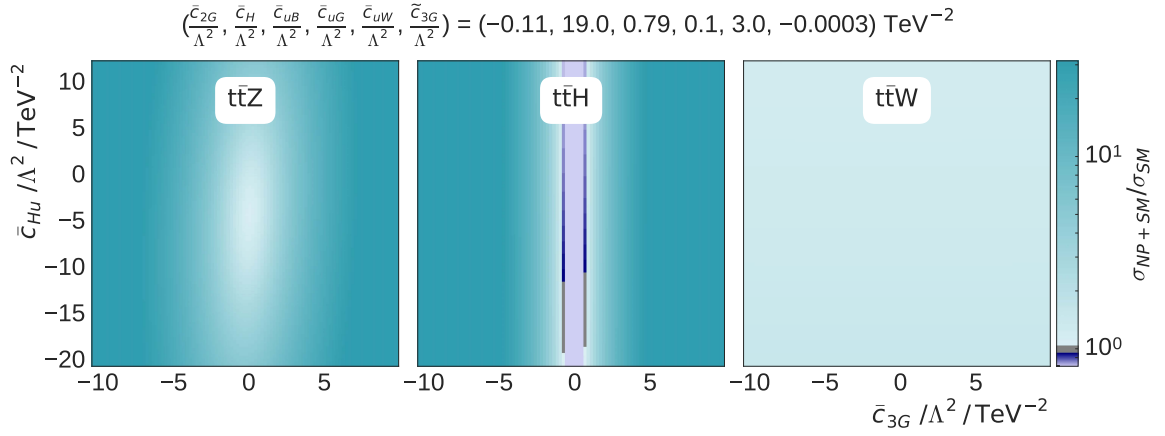
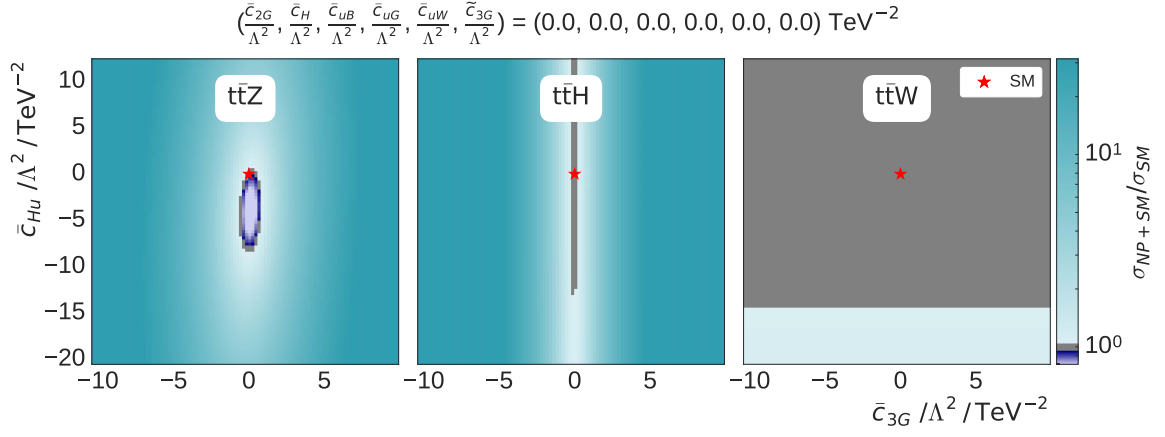


Figure B.9. Signal scaling shown in the $\bar{c}_{Hu}$, $\bar{c}_{3G}$ plane with all other coefficients fixed to zero (a) or their best-fit values (b) for $t\bar{t}Z$ (left), $t\bar{t}H$ (center), and $t\bar{t}W$ (right). The color represents the scaling ($\sigma_{NP+SM}/\sigma_{SM}$) due to NP effects. The star represents the SM point in which all $c_i=0$. The negative log likelihood is shown in (c). The best fit is represented by a cross. The 68 % and 95 % CL contours are shown with dashed and solid lines, respectively.

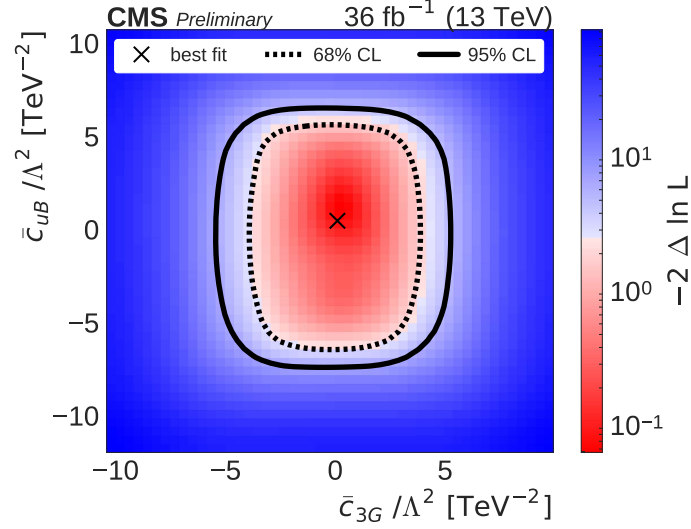
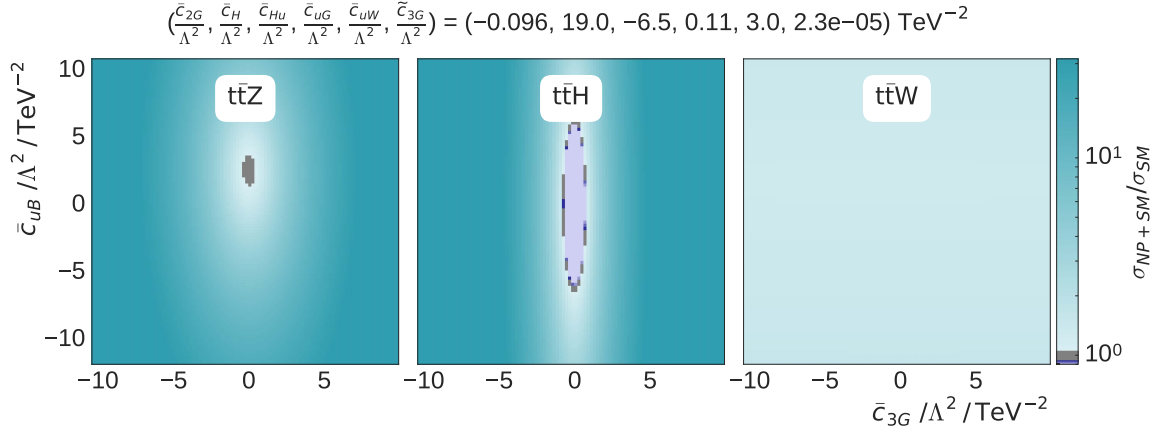
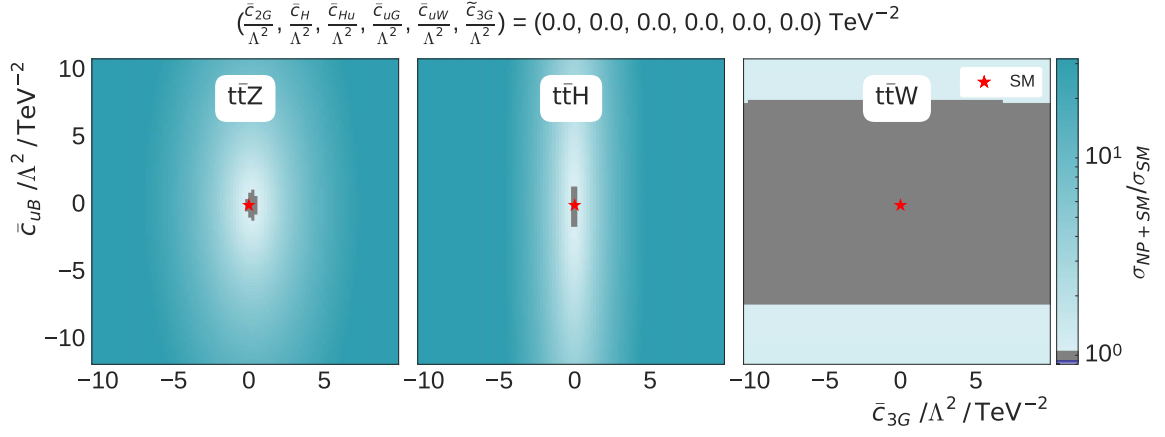


Figure B.10. Signal scaling shown in the $\bar{c}_{3G}$, $\bar{c}_{uB}$ plane with all other coefficients fixed to zero (a) or their best-fit values (b) for $t\bar{t}Z$ (left), $t\bar{t}H$ (center), and $t\bar{t}W$ (right). The color represents the scaling ($\sigma_{\text{NP}+\text{SM}}/\sigma_{\text{SM}}$) due to NP effects. The star represents the SM point in which all $c_i=0$. The negative log likelihood is shown in (c). The best fit is represented by a cross. The 68 % and 95 % CL contours are shown with dashed and solid lines, respectively.

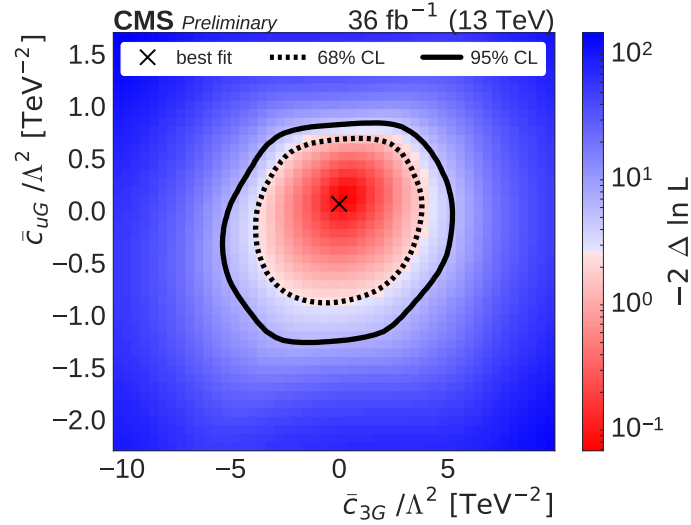
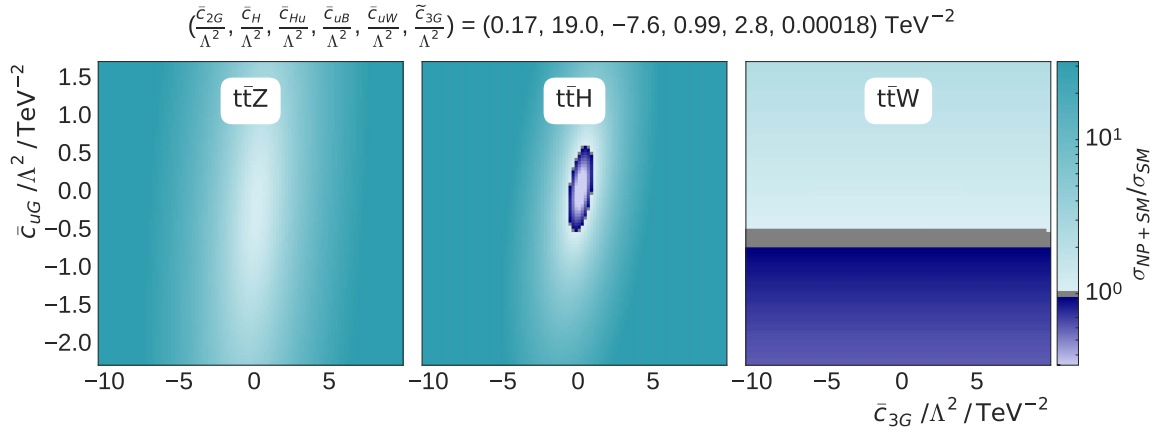
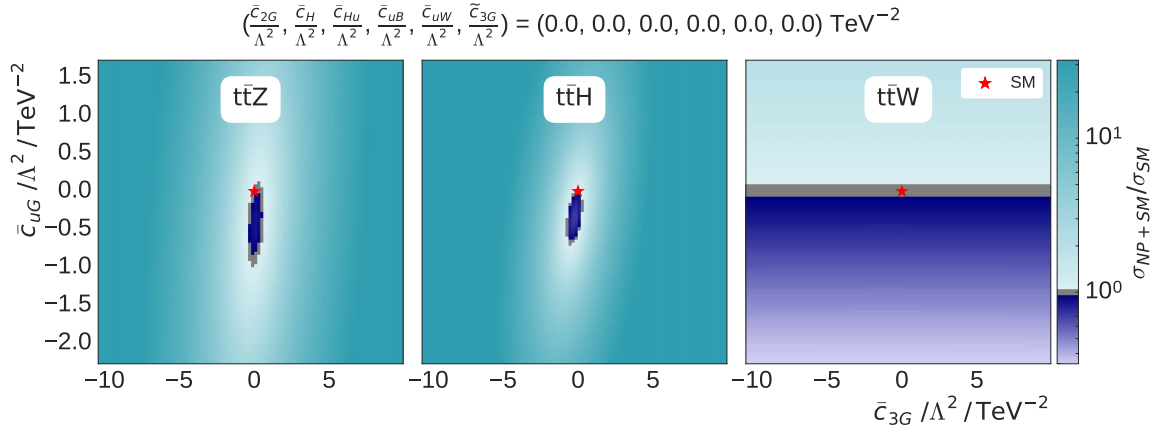


Figure B.11. Signal scaling shown in the $\bar{c}_{3G}$, $\bar{c}_{uG}$ plane with all other coefficients fixed to zero (a) or their best-fit values (b) for $t\bar{t}Z$ (left), $t\bar{t}H$ (center), and $t\bar{t}W$ (right). The color represents the scaling ($\sigma_{NP+SM}/\sigma_{SM}$) due to NP effects. The star represents the SM point in which all $c_i=0$. The negative log likelihood is shown in (c). The best fit is represented by a cross. The 68 % and 95 % CL contours are shown with dashed and solid lines, respectively.

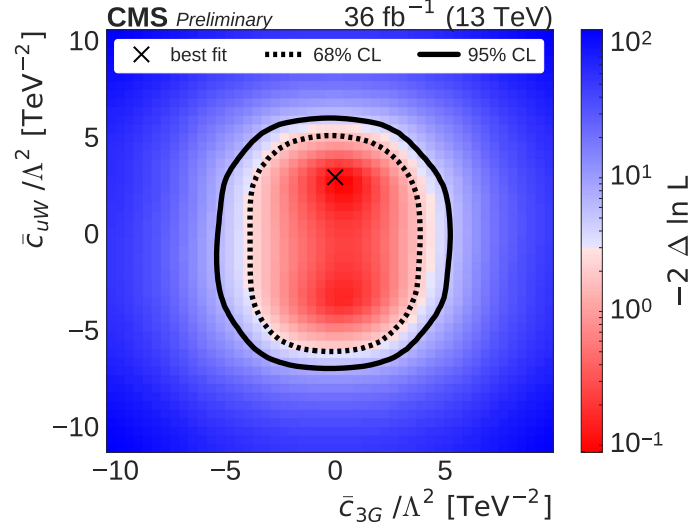
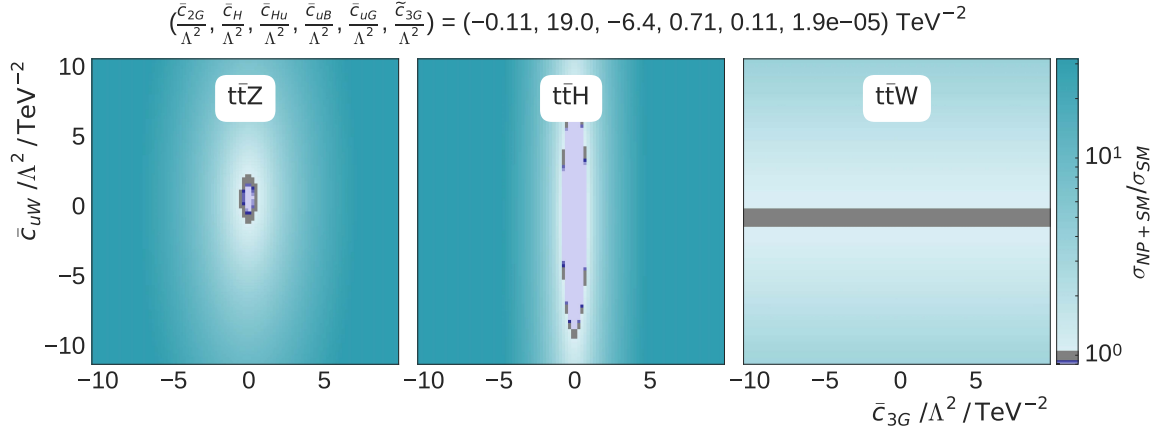
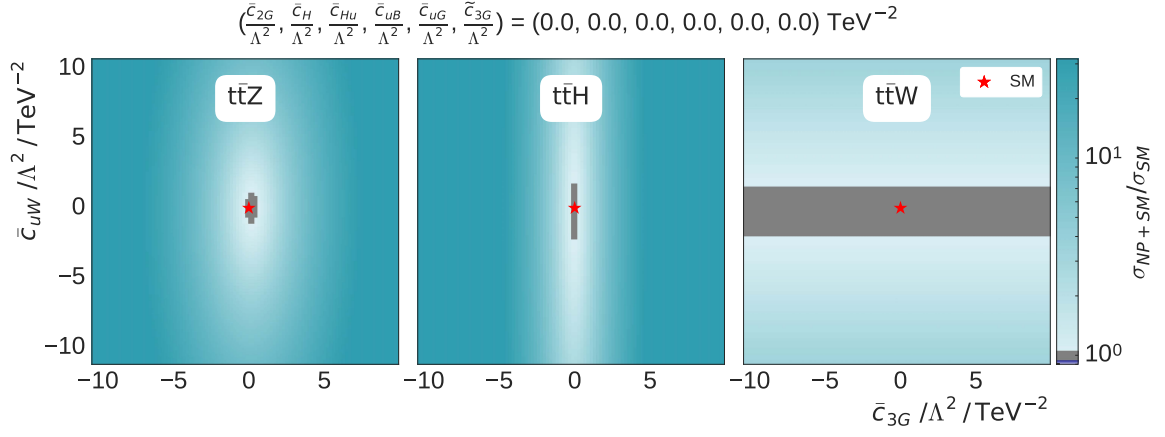


Figure B.12. Signal scaling shown in the $\bar{c}_{3G}$, $\bar{c}_{uW}$ plane with all other coefficients fixed to zero (a) or their best-fit values (b) for $t\bar{t}Z$ (left), $t\bar{t}H$ (center), and $t\bar{t}W$ (right). The color represents the scaling ($\sigma_{\text{NP}+\text{SM}}/\sigma_{\text{SM}}$) due to NP effects. The star represents the SM point in which all $c_i=0$. The negative log likelihood is shown in (c). The best fit is represented by a cross. The 68 % and 95 % CL contours are shown with dashed and solid lines, respectively.

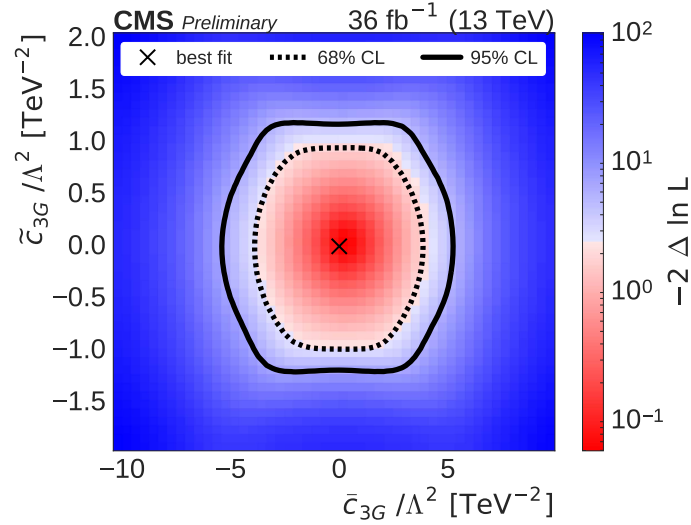
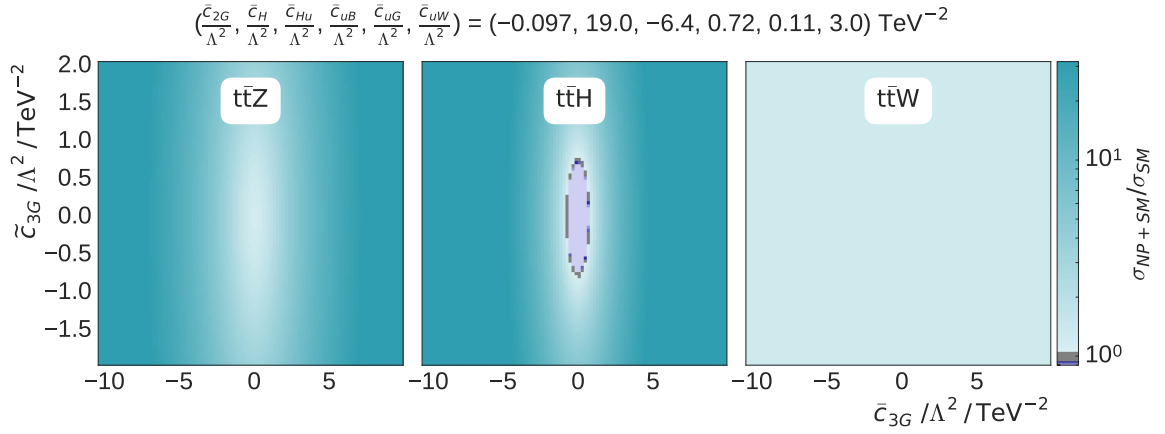
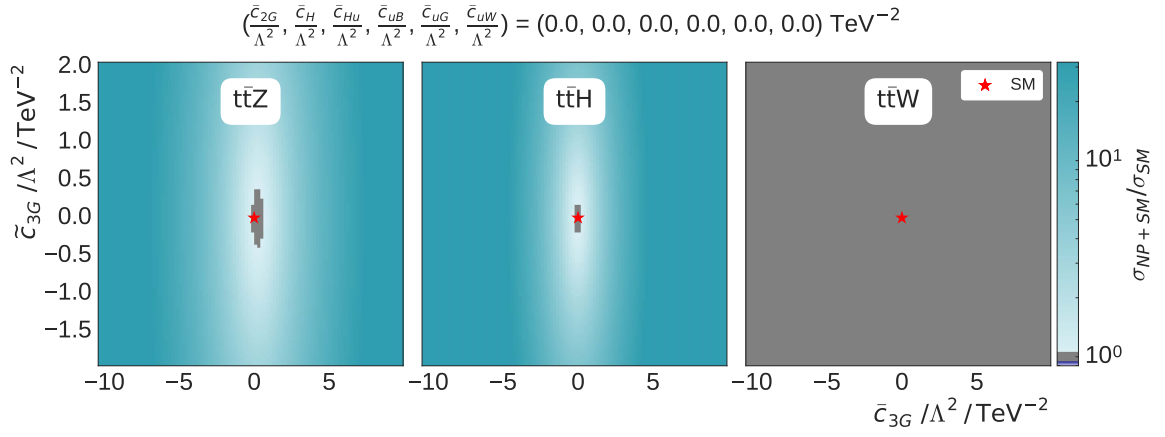


Figure B.13. Signal scaling shown in the $\bar{c}_{3G}$, $\tilde{c}_{3G}$ plane with all other coefficients fixed to zero (a) or their best-fit values (b) for $t\bar{t}Z$ (left), $t\bar{t}H$ (center), and $t\bar{t}W$ (right). The color represents the scaling ($\sigma_{\text{NP} + \text{SM}}/\sigma_{\text{SM}}$) due to NP effects. The star represents the SM point in which all $c_i = 0$. The negative log likelihood is shown in (c). The best fit is represented by a cross. The 68 % and 95 % CL contours are shown with dashed and solid lines, respectively.

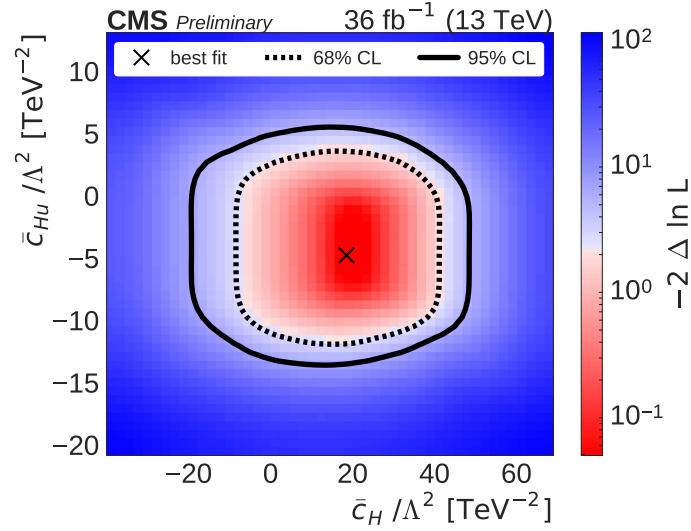
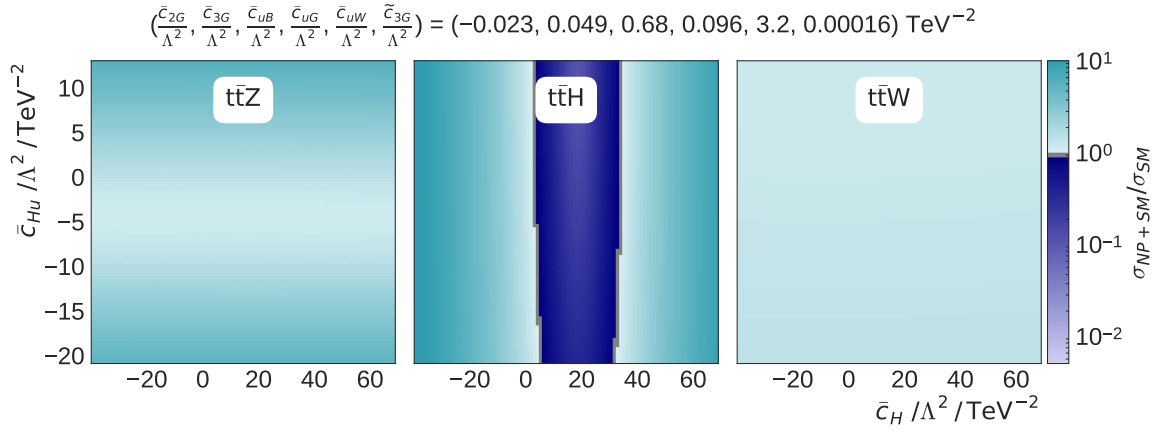
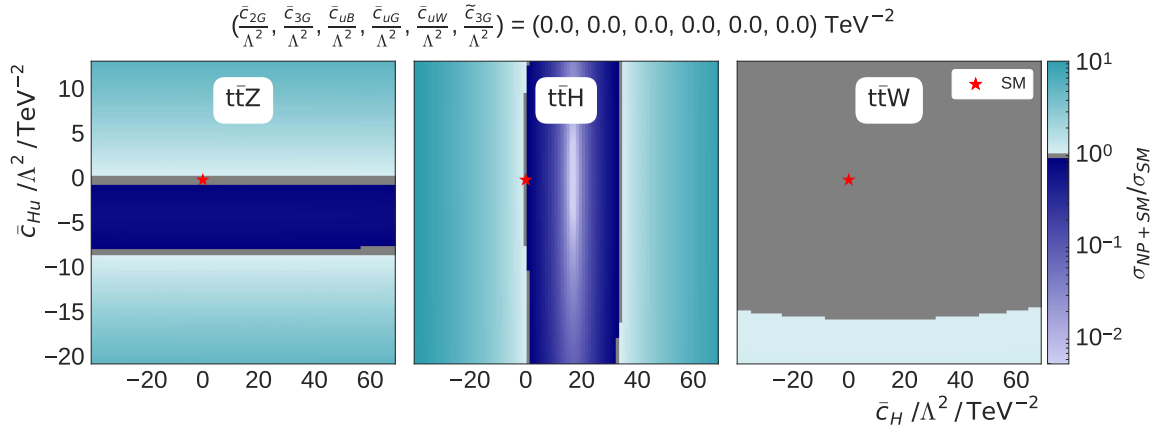


Figure B.14. Signal scaling shown in the $\bar{c}_{Hu}$, $\bar{c}_H$ plane with all other coefficients fixed to zero (a) or their best-fit values (b) for $t\bar{t}Z$ (left), $t\bar{t}H$ (center), and $t\bar{t}W$ (right). The color represents the scaling ($\sigma_{NP+SM}/\sigma_{SM}$) due to NP effects. The star represents the SM point in which all $c_i=0$. The negative log likelihood is shown in (c). The best fit is represented by a cross. The 68 % and 95 % CL contours are shown with dashed and solid lines, respectively.

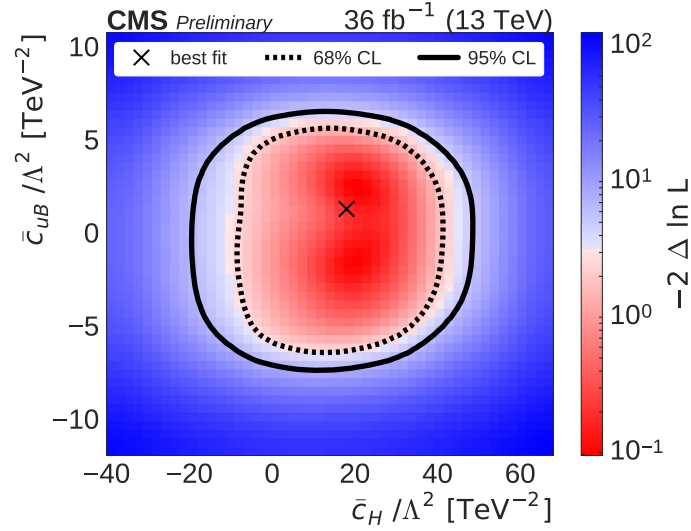
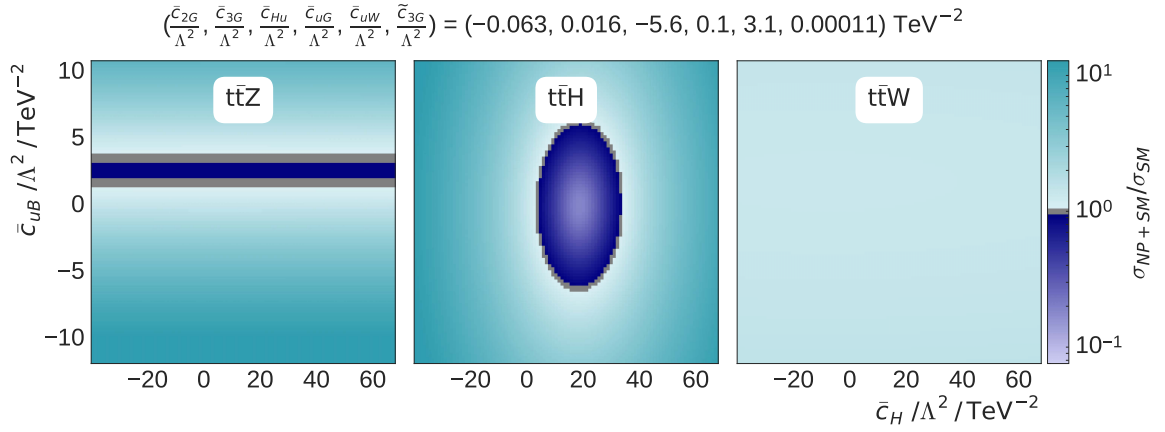
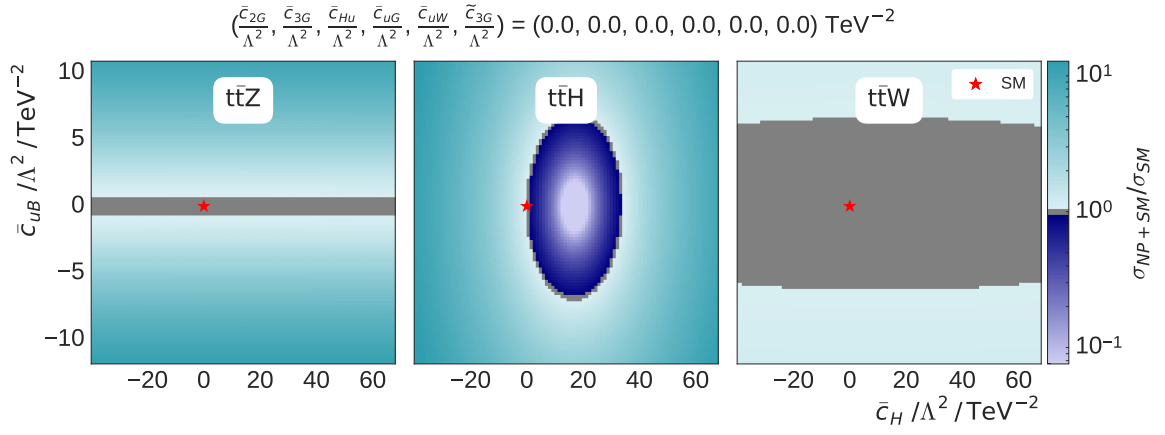


Figure B.15. Signal scaling shown in the $\bar{c}_H, \bar{c}_{uB}$ plane with all other coefficients fixed to zero (a) or their best-fit values (b) for $t\bar{t}Z$ (left), $t\bar{t}H$ (center), and $t\bar{t}W$ (right). The color represents the scaling ($\sigma_{NP+SM}/\sigma_{SM}$) due to NP effects. The star represents the SM point in which all $c_i=0$. The negative log likelihood is shown in (c). The best fit is represented by a cross. The 68 % and 95 % CL contours are shown with dashed and solid lines, respectively.

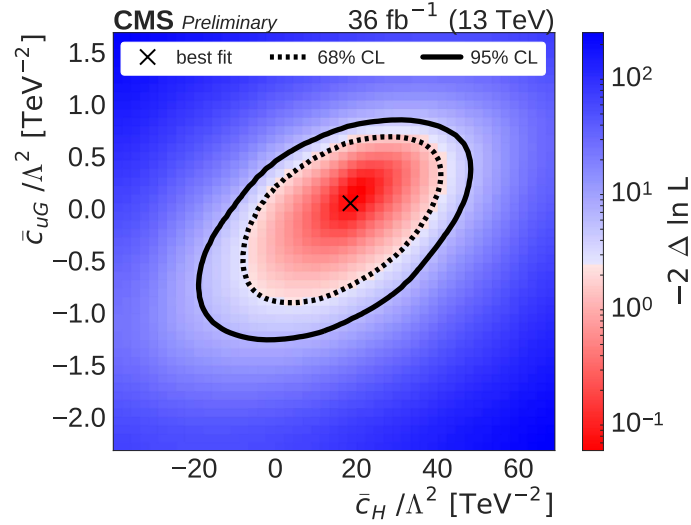
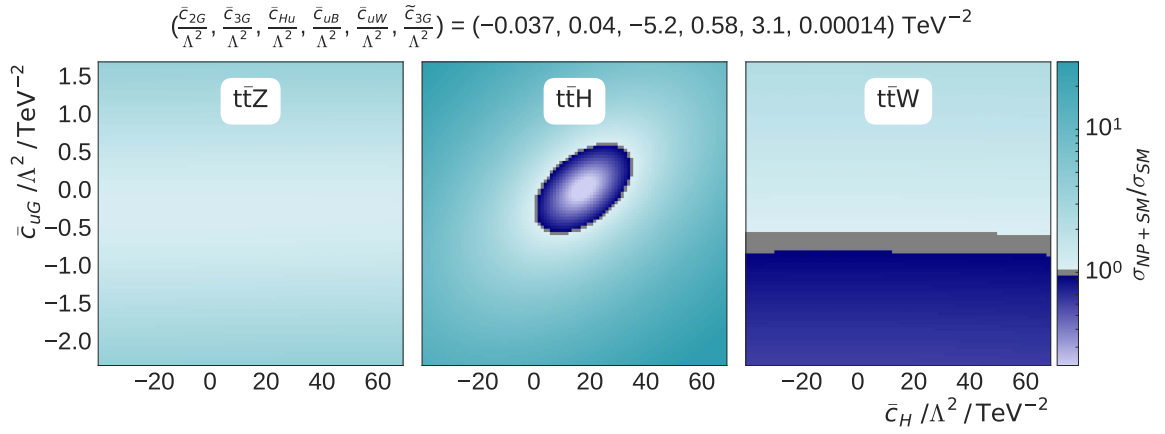
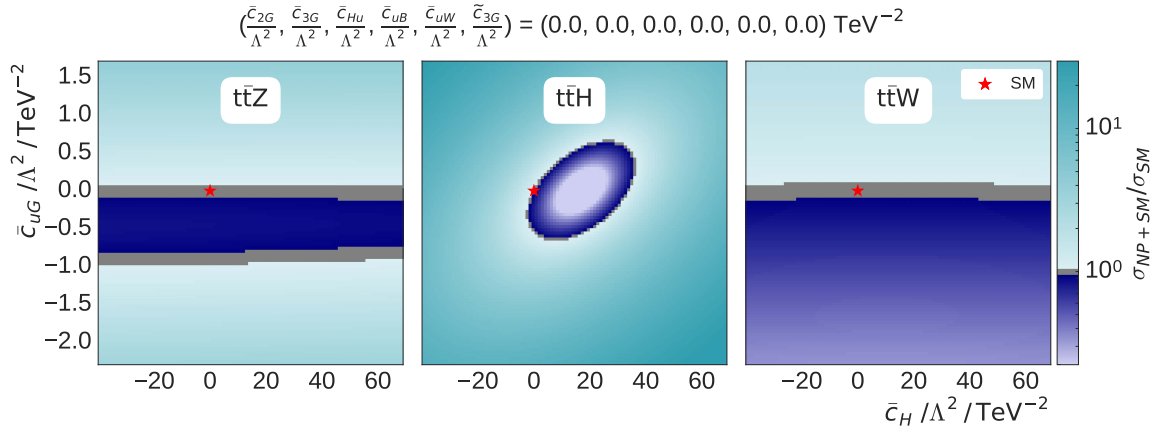


Figure B.16. Signal scaling shown in the $\bar{c}_H, \bar{c}_{uG}$ plane with all other coefficients fixed to zero (a) or their best-fit values (b) for $t\bar{t}Z$ (left), $t\bar{t}H$ (center), and $t\bar{t}W$ (right). The color represents the scaling ($\sigma_{NP+SM}/\sigma_{SM}$) due to NP effects. The star represents the SM point in which all $c_i = 0$. The negative log likelihood is shown in (c). The best fit is represented by a cross. The 68 % and 95 % CL contours are shown with dashed and solid lines, respectively.

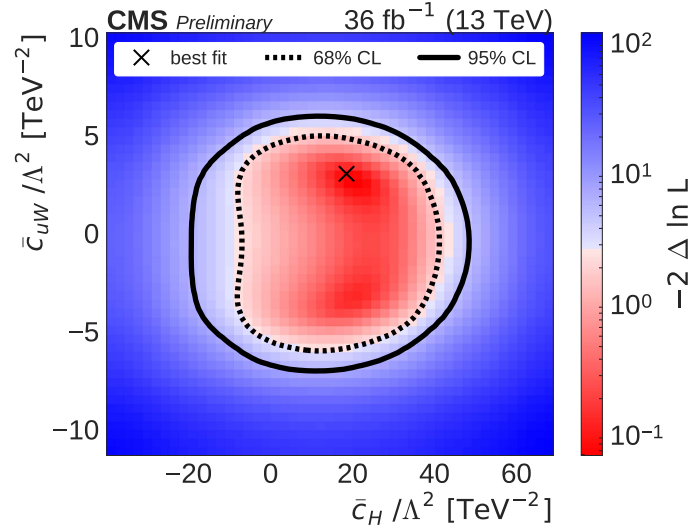
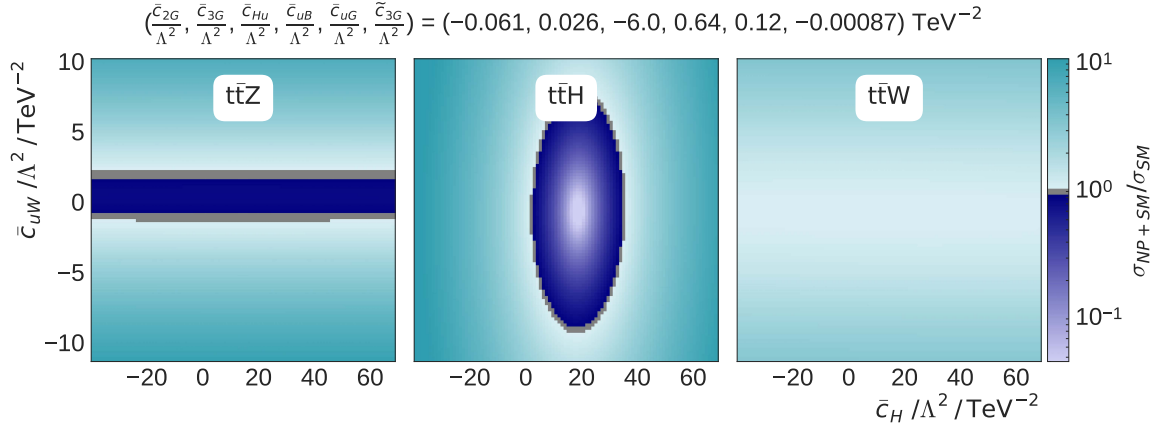
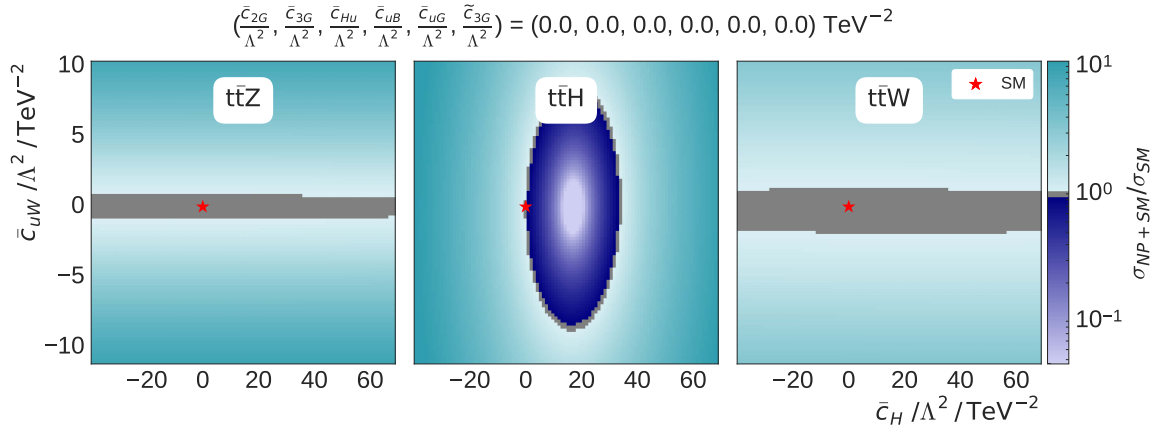


Figure B.17. Signal scaling shown in the $\bar{c}_H, \bar{c}_{uW}$ plane with all other coefficients fixed to zero (a) or their best-fit values (b) for $t\bar{t}Z$ (left), $t\bar{t}H$ (center), and $t\bar{t}W$ (right). The color represents the scaling ($\sigma_{NP+SM}/\sigma_{SM}$) due to NP effects. The star represents the SM point in which all $c_i = 0$. The negative log likelihood is shown in (c). The best fit is represented by a cross. The 68 % and 95 % CL contours are shown with dashed and solid lines, respectively.

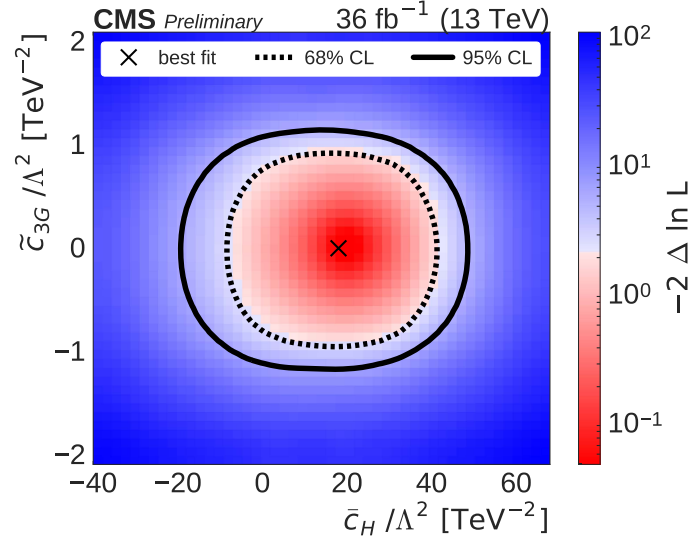
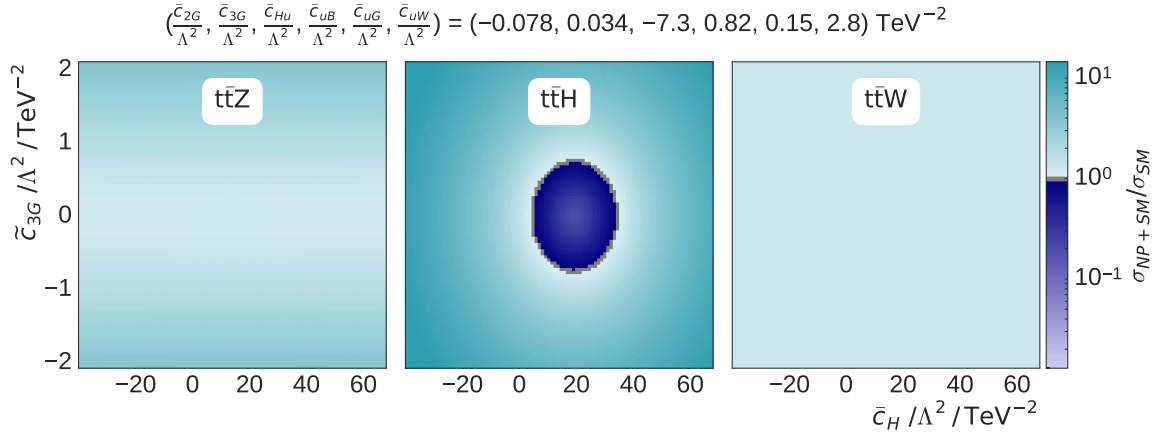
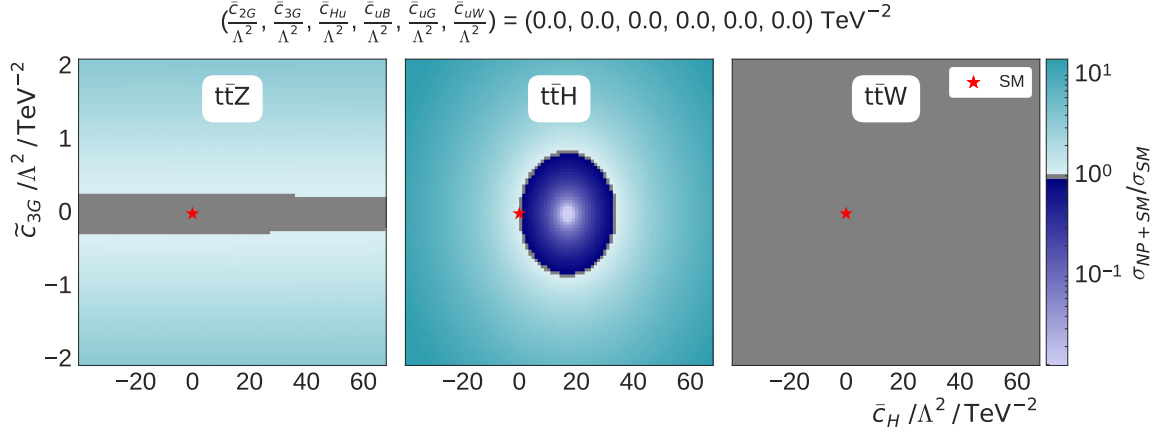


Figure B.18. Signal scaling shown in the $\bar{c}_H$, $\tilde{c}_{3G}$ plane with all other coefficients fixed to zero (a) or their best-fit values (b) for $t\bar{t}Z$ (left), $t\bar{t}H$ (center), and $t\bar{t}W$ (right). The color represents the scaling ($\sigma_{\text{NP} + \text{SM}}/\sigma_{\text{SM}}$) due to NP effects. The star represents the SM point in which all $c_i = 0$. The negative log likelihood is shown in (c). The best fit is represented by a cross. The 68 % and 95 % CL contours are shown with dashed and solid lines, respectively.

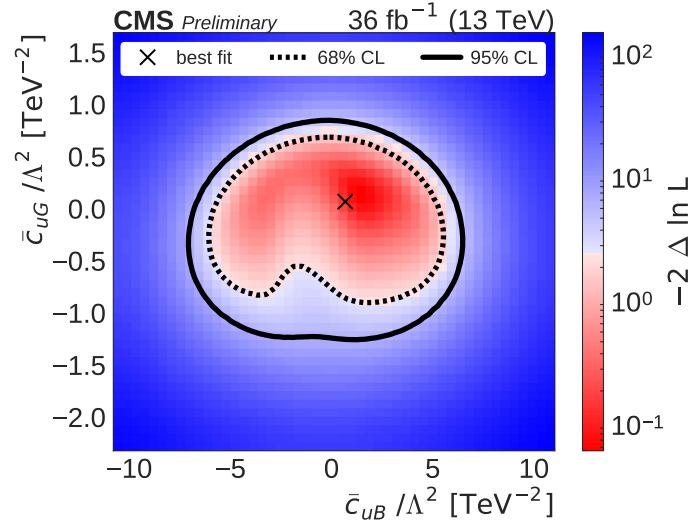
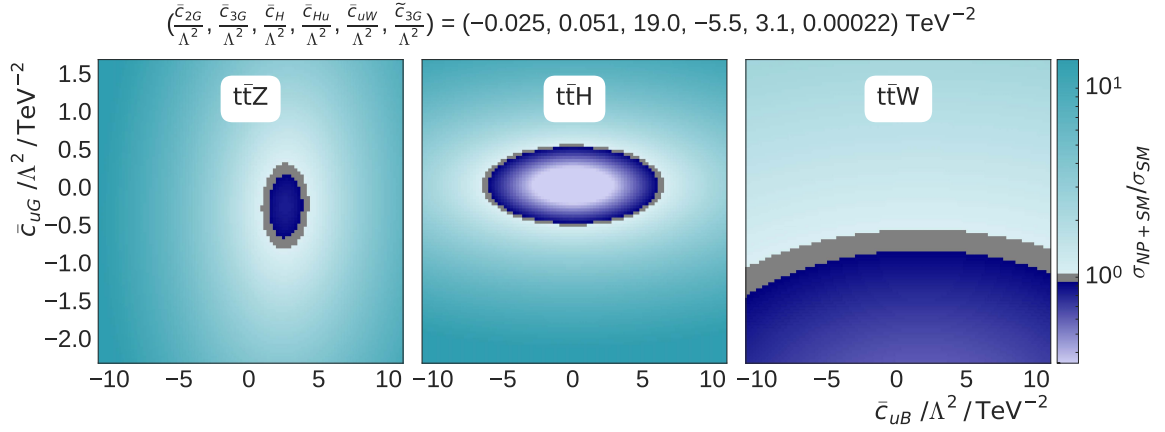
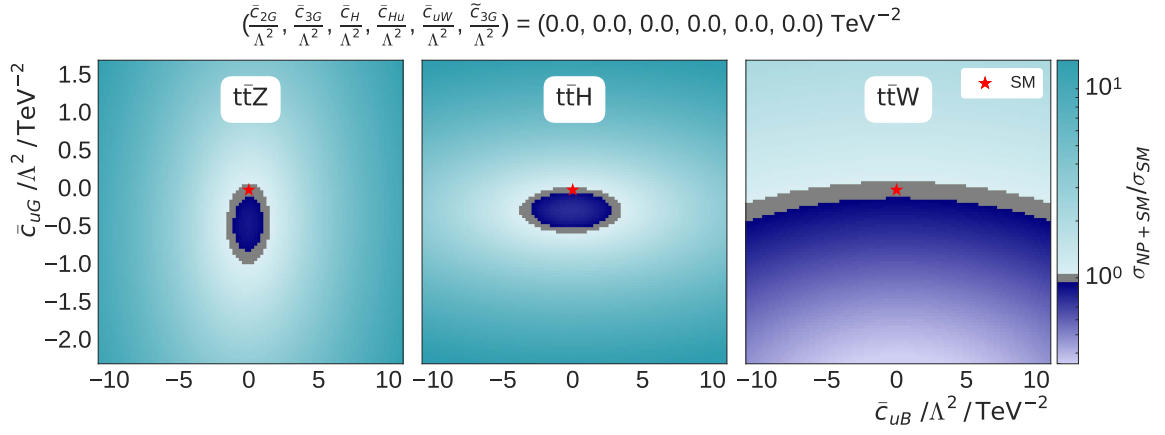


Figure B.19. Signal scaling shown in the $\bar{c}_{uB}$, $\bar{c}_{uG}$ plane with all other coefficients fixed to zero (a) or their best-fit values (b) for $t\bar{t}Z$ (left), $t\bar{t}H$ (center), and $t\bar{t}W$ (right). The color represents the scaling ($\sigma_{\text{NP}+\text{SM}}/\sigma_{\text{SM}}$) due to NP effects. The star represents the SM point in which all $c_i = 0$. The negative log likelihood is shown in (c). The best fit is represented by a cross. The 68 % and 95 % CL contours are shown with dashed and solid lines, respectively.

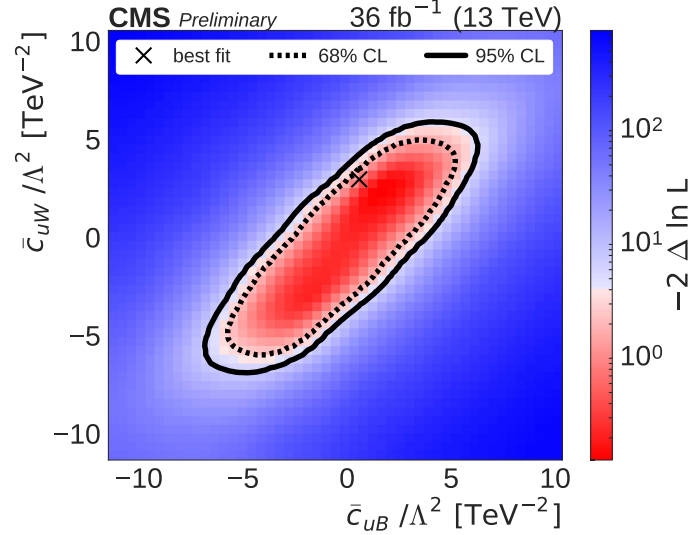
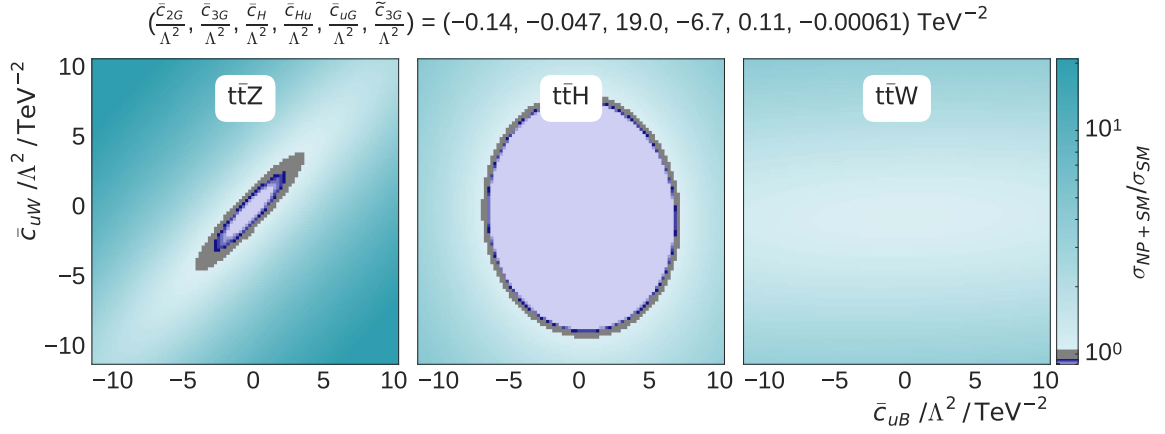
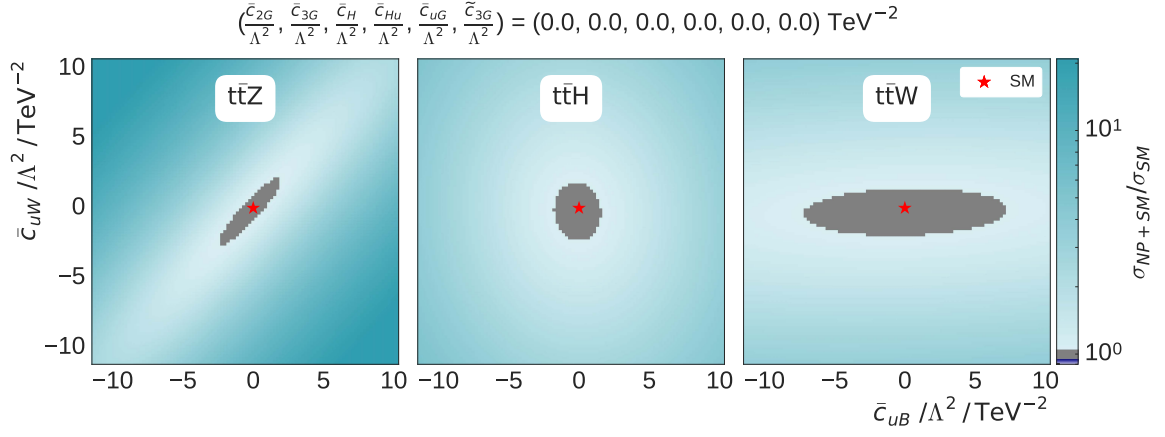


Figure B.20. Signal scaling shown in the $\tilde{c}_{uB}$, $\tilde{c}_{uW}$ plane with all other coefficients fixed to zero (a) or their best-fit values (b) for $t\bar{t}Z$ (left), $t\bar{t}H$ (center), and $t\bar{t}W$ (right). The color represents the scaling ($\sigma_{NP+SM}/\sigma_{SM}$) due to NP effects. The star represents the SM point in which all $c_i=0$. The negative log likelihood is shown in (c). The best fit is represented by a cross. The 68 % and 95 % CL contours are shown with dashed and solid lines, respectively.

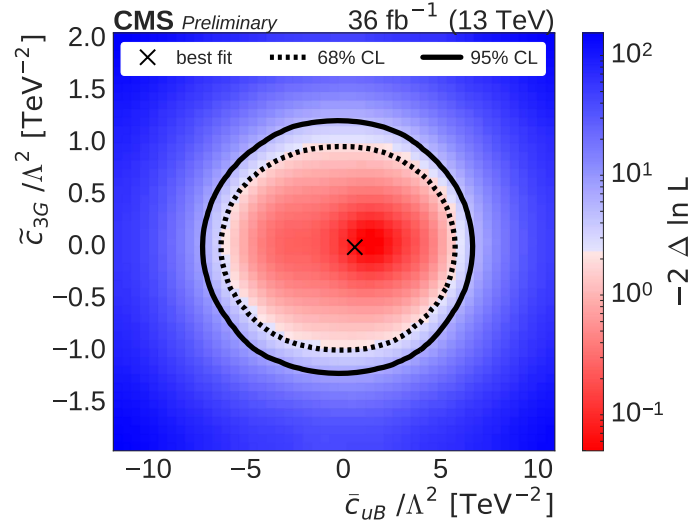
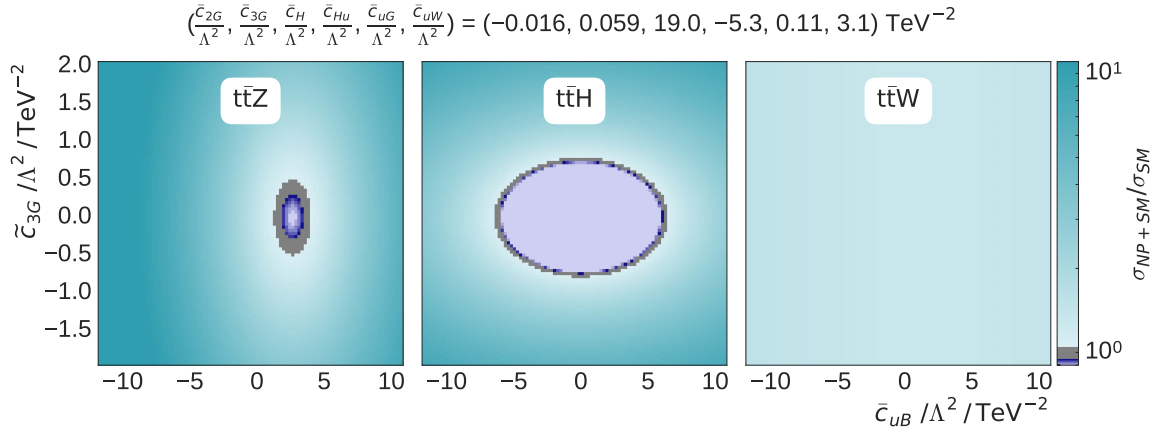
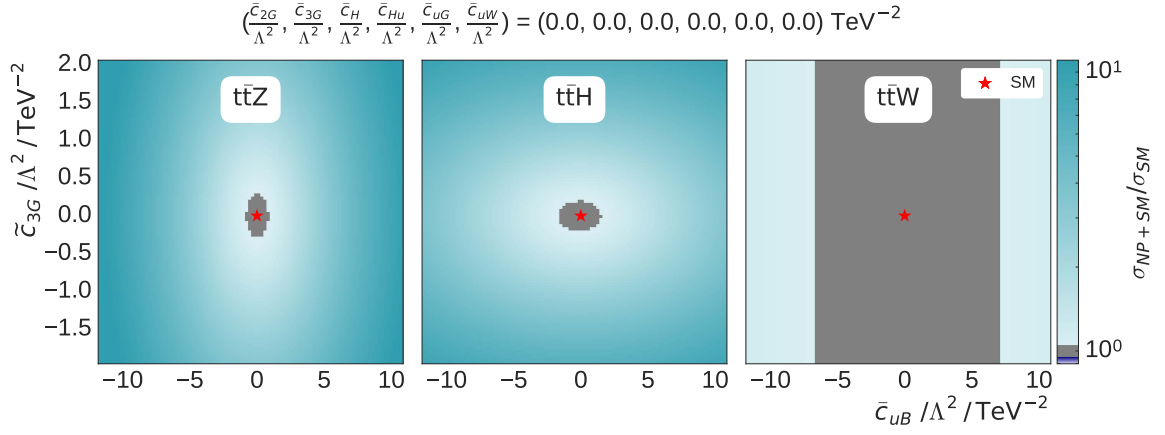


Figure B.21. Signal scaling shown in the $\bar{c}_{uB}$, $\bar{c}_{3G}$ plane with all other coefficients fixed to zero (a) or their best-fit values (b) for $t\bar{t}Z$ (left), $t\bar{t}H$ (center), and $t\bar{t}W$ (right). The color represents the scaling ($\sigma_{\text{NP} + \text{SM}}/\sigma_{\text{SM}}$) due to NP effects. The star represents the SM point in which all $c_i = 0$. The negative log likelihood is shown in (c). The best fit is represented by a cross. The 68 % and 95 % CL contours are shown with dashed and solid lines, respectively.

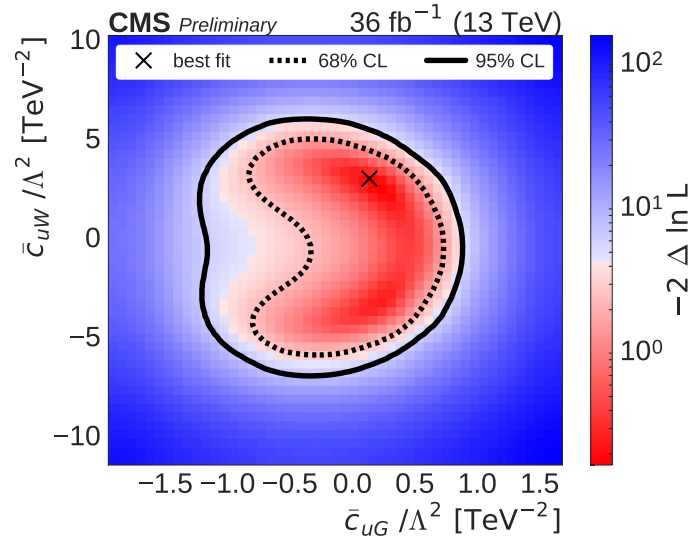
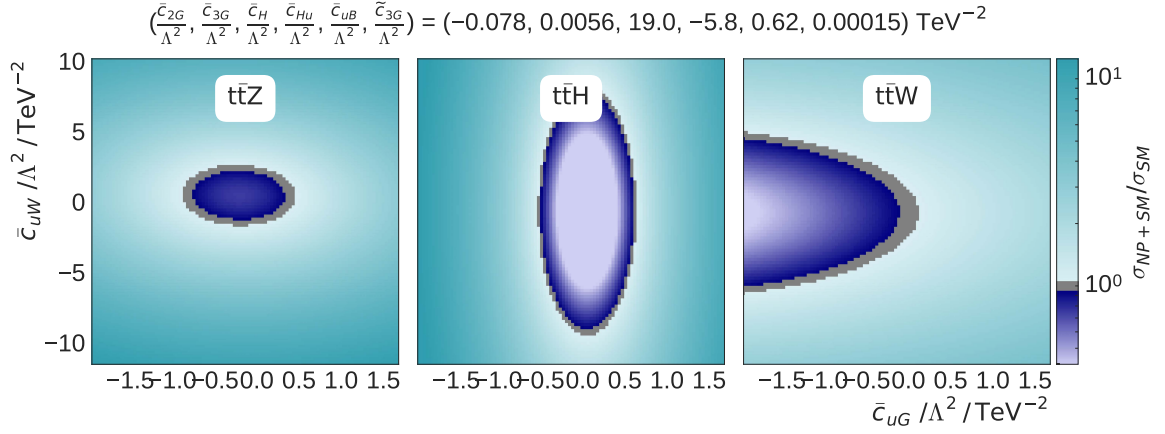
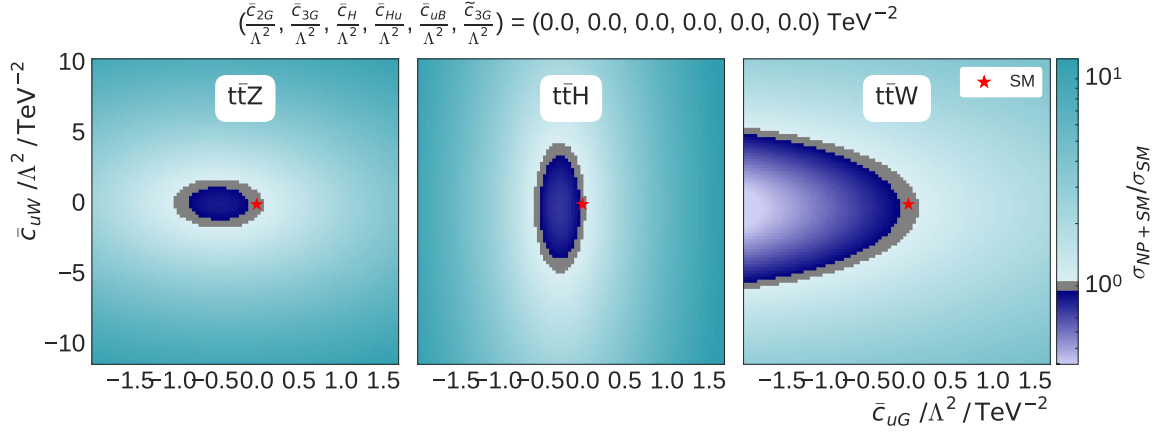


Figure B.22. Signal scaling shown in the $\tilde{c}_{uG}$, $\tilde{c}_{uW}$ plane with all other coefficients fixed to zero (a) or their best-fit values (b) for $t\bar{t}Z$ (left), $t\bar{t}H$ (center), and $t\bar{t}W$ (right). The color represents the scaling ($\sigma_{\text{NP} + \text{SM}}/\sigma_{\text{SM}}$) due to NP effects. The star represents the SM point in which all $c_i = 0$. The negative log likelihood is shown in (c). The best fit is represented by a cross. The 68 % and 95 % CL contours are shown with dashed and solid lines, respectively.

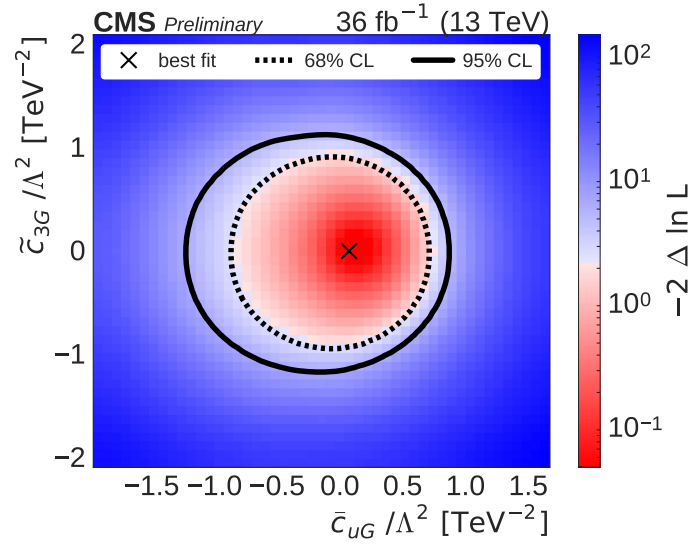
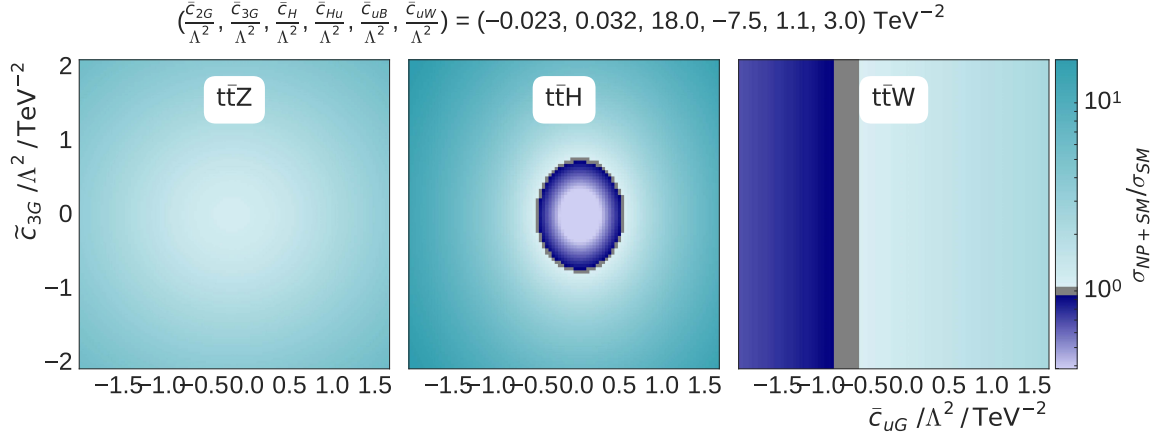
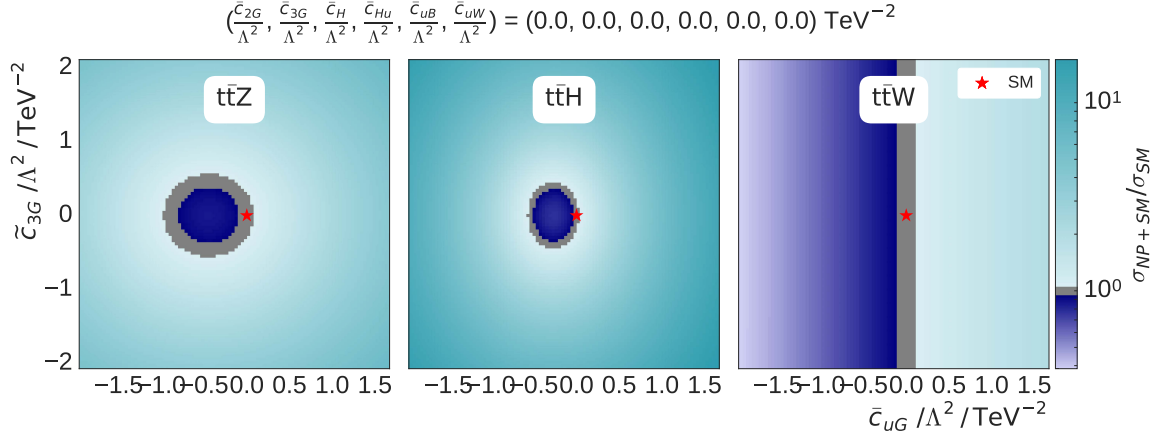


Figure B.23. Signal scaling shown in the $\tilde{c}_{uG}$, $\tilde{c}_{3G}$ plane with all other coefficients fixed to zero (a) or their best-fit values (b) for $t\bar{t}Z$ (left), $t\bar{t}H$ (center), and $t\bar{t}W$ (right). The color represents the scaling ($\sigma_{NP+SM}/\sigma_{SM}$) due to NP effects. The star represents the SM point in which all $c_i=0$. The negative log likelihood is shown in (c). The best fit is represented by a cross. The 68 % and 95 % CL contours are shown with dashed and solid lines, respectively.

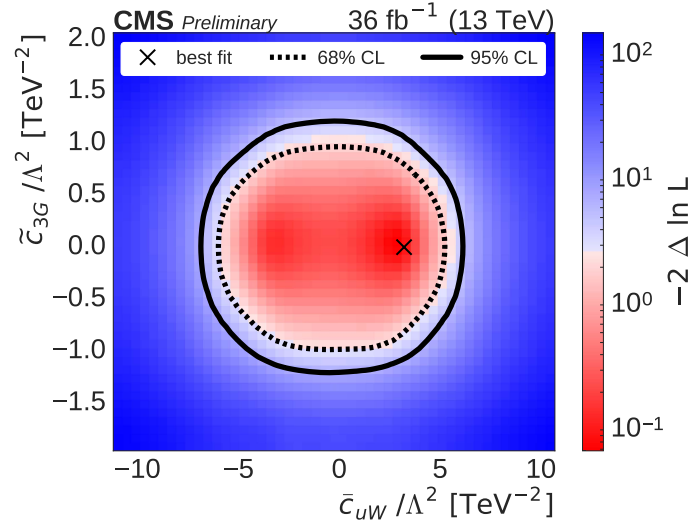
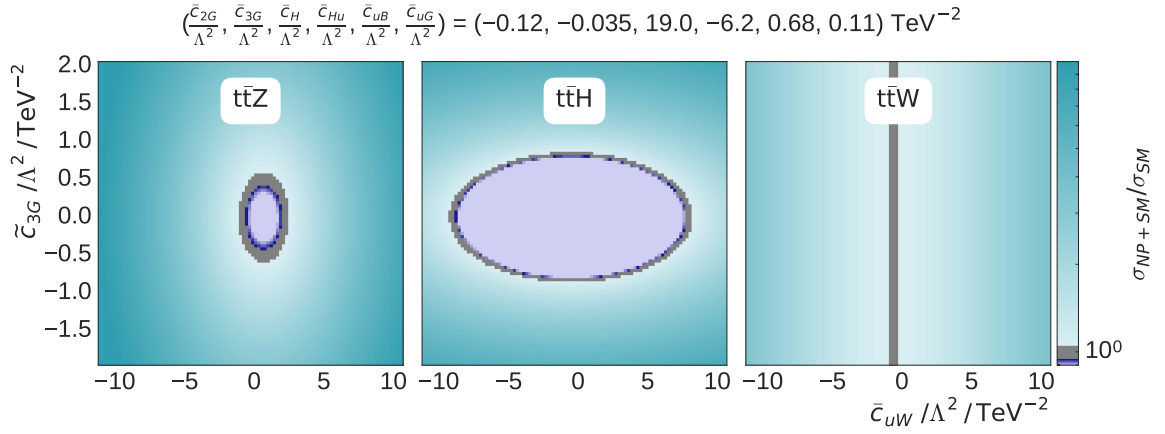
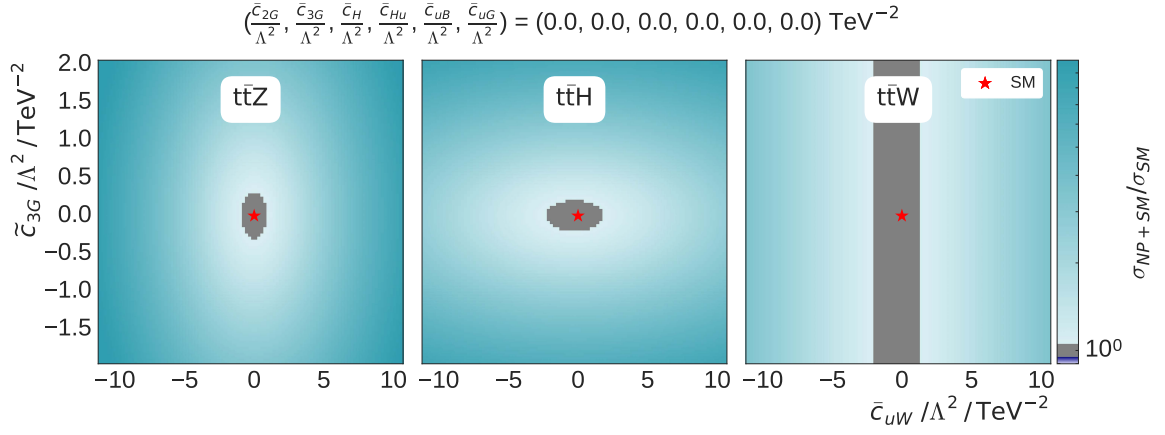


Figure B.24. Signal scaling shown in the $\tilde{c}_{uW}$, $\tilde{c}_{3G}$ plane with all other coefficients fixed to zero (a) or their best-fit values (b) for $t\bar{t}Z$ (left), $t\bar{t}H$ (center), and $t\bar{t}W$ (right). The color represents the scaling ($\sigma_{NP+SM}/\sigma_{SM}$) due to NP effects. The star represents the SM point in which all $c_i=0$. The negative log likelihood is shown in (c). The best fit is represented by a cross. The 68 % and 95 % CL contours are shown with dashed and solid lines, respectively.

BIBLIOGRAPHY

1. V. Khachatryan et al. “Observation of top quark pairs produced in association with a vector boson in pp collisions at $\sqrt{s}=8$ TeV”. In: *JHEP* 01 (2016), p. 096. doi : 10 . 1007/ JHEP01 (2016) 096.
2. A. M. Sirunyan et al. “Measurement of the cross section for top quark pair production in association with a W or Z boson in proton-proton collisions at $\sqrt{s}=13$ TeV”. 2017.
3. G. Arnison et al. “Experimental observation of isolated large transverse energy electrons with associated missing energy at $s=540$ GeV”. In: *Physics Letters B* 122.1 (1983), pp. 103–116. doi : 10 . 1016/0370–2693(83)91177–2.
4. M. Banner et al. “Observation of single isolated electrons of high transverse momentum in events with missing transverse energy at the CERN pp collider”. In: *Physics Letters B* 122.5 (1983), pp. 476–485. doi : 10 . 1016/0370–2693(83)91605–2.
5. R. D. Klauber. *Student friendly quantum field theory: basic principles and quantum electrodynamics*. Sandtrove Press, 2013.
6. F. Wilczek. “Nobel Lecture: Asymptotic freedom: From paradox to paradigm”. In: *Rev. Mod. Phys.* 77 (3 Sept. 2005), pp. 857–870. doi : 10 . 1103/RevModPhys . 77 . 857.
7. S. L. Glashow. “Partial-symmetries of weak interactions”. In: *Nuclear Physics* 22.4 (1961), pp. 579–588. doi : 10 . 1016/0029–5582(61)90469–2.
8. A. Salam. “Weak and Electromagnetic Interactions”. In: *Conf. Proc.* C680519 (1968), pp. 367–377.
9. S. Weinberg. “A Model of Leptons”. In: *Phys. Rev. Lett.* 19 (21 Nov. 1967), pp. 1264–1266. doi : 10 . 1103/PhysRevLett . 19 . 1264.
10. D. J. Griffiths. *Introduction to elementary particles; 2nd rev. version*. Physics textbook. New York, NY: Wiley, 2008.
11. A. Romanino. “The Standard model of particle physics”. In: *Physics of elementary particles and astrophysics. Proceedings, International Baikal Summer School, Bolshie Koty, Russia, July 23-30, 2009*. 2009. www.slac.stanford.edu/econf/C0907232/pdf/001.pdf.

12. W. Cottingham and D. Greenwood. *An Introduction to the Standard Model of Particle Physics*. Cambridge University Press, 1998.
13. J. Goldstone. “Field theories with « Superconductor » solutions”. In: *Il Nuovo Cimento (1955-1965)* 19.1 (Jan. 1, 1961), pp. 154–164. doi : 10 . 1007/BF02812722.
14. S. Chatrchyan et al. “Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC”. In: *Phys. Lett. B* 716 (2012), p. 30. doi : 10 . 1016/j . physletb . 2012 . 08 . 021.
15. M. E. Peskin and D. V. Schroeder. *An Introduction to quantum field theory*. Reading, USA: Addison-Wesley, 1995.
16. D. Hanneke, S. Fogwell, and G. Gabrielse. “New Measurement of the Electron Magnetic Moment and the Fine Structure Constant”. In: *Phys. Rev. Lett.* 100 (12 Mar. 2008), p. 120801. doi : 10 . 1103/PhysRevLett . 100 . 120801.
17. P. Cladé et al. “Determination of the Fine Structure Constant Based on Bloch Oscillations of Ultracold Atoms in a Vertical Optical Lattice”. In: *Phys. Rev. Lett.* 96 (3 Jan. 2006), p. 033001. doi : 10 . 1103/PhysRevLett . 96 . 033001.
18. Planck Collaboration et al. “Planck 2013 results. I. Overview of products and scientific results”. In: *AAP* 571, A1 (Nov. 2014), A1. doi : 10 . 1051/0004-6361/201321529.
19. F. Abe et al. “Observation of Top Quark Production in $\bar{p}p$ Collisions with the Collider Detector at Fermilab”. In: *Phys. Rev. Lett.* 74 (14 Apr. 1995), pp. 2626–2631. doi : 10 . 1103/PhysRevLett . 74 . 2626.
20. S. Abachi et al. “Search for High Mass Top Quark Production in $p\bar{p}$ Collisions at $\sqrt{s} = 1.8$ TeV”. In: *Phys. Rev. Lett.* 74 (13 Mar. 1995), pp. 2422–2426. doi : 10 . 1103/PhysRevLett . 74 . 2422.
21. CMS Collaboration. “Precise determination of the mass of the Higgs boson and tests of compatibility of its couplings with the standard model predictions using proton collisions at 7 and 8 TeV”. In: *Eur. Phys. J. C* 75.CMS-HIG-14-009. CERN-PH-EP-2014-288. CMS-HIG-14-009 (Dec. 2014), 212. 75 p. cds . cern . ch/record/1979247.
22. V. A. Novikov et al. “Theory of Z boson decays”. In: *Reports on Progress in Physics* 62.9 (1999), p. 1275. stacks . iop . org/0034-4885/62/i=9/a=201.
23. The CMS Collaboration. “Search for the associated production of the Higgs boson with a top-quark pair”. In: *Journal of High Energy Physics* 2014.9 (Sept. 16, 2014), p. 87. doi : 10 . 1007/JHEP09(2014)087.

24. *Search for the associated production of a Higgs boson with a top quark pair in final states with a τ lepton at $\sqrt{s}=13$ TeV*. Tech. rep. CMS-PAS-HIG-17-003. Geneva: CERN, 2017. cds.cern.ch/record/2257067.
25. *Search for Higgs boson production in association with top quarks in multilepton final states at $\sqrt{s}=13$ TeV*. Tech. rep. CMS-PAS-HIG-17-004. Geneva: CERN, 2017. cds.cern.ch/record/2256103.
26. S. Chatrchyan et al. “Measurement of associated production of vector bosons and $t\bar{t}$ in pp collisions at $\sqrt{s}=7$ TeV”. In: *Phys. Rev. Lett.* 110 (2013), p. 172002. doi : 10.1103/PhysRevLett.110.172002.
27. V. Khachatryan et al. “Measurement of top quark-antiquark pair production in association with a W or Z boson in pp collisions at $\sqrt{s}=8$ TeV”. In: *Eur. Phys. J.C* 74.9 (2014), p. 3060. doi : 10.1140/epjc/s10052-014-3060-7.
28. *Evidence for the associated production of a vector boson (W, Z) and top quark pair in the dilepton and trilepton channels in pp collision data at $\sqrt{s}=8$ TeV collected by the ATLAS detector at the LHC*. Tech. rep. ATLAS-CONF-2014-038. Geneva: CERN, July 2014. cds.cern.ch/record/1735215.
29. M. Maggiore. *A modern introduction to quantum field theory*. Oxford Master Series in Statistical, Computational, and Theoretical Physics. Oxford: Oxford Univ., 2005.
30. S. Weinberg. “Baryon- and Lepton-Nonconserving Processes”. In: *Phys. Rev. Lett.* 43 (21 Nov. 1979), pp. 1566–1570. doi : 10.1103/PhysRevLett.43.1566.
31. B. Grzadkowski et al. “Dimension-Six Terms in the Standard Model Lagrangian”. In: *JHEP* 10 (2010), p. 085. doi : 10.1007/JHEP10(2010)085.
32. A. Alloul, B. Fuks, and V. Sanz. “Phenomenology of the Higgs effective Lagrangian via FeynRules”. In: *Journal of High Energy Physics* 2014.4 (Apr. 2014), p. 110. doi : 10.1007/JHEP04(2014)110.
33. C. Degrande et al. “Effective field theory: A modern approach to anomalous couplings”. In: *Annals of Physics* 335 (2013), pp. 21–32. doi : 10.1016/j.aop.2013.04.016.
34. Forthommel. *Map of the CERN accelerator complex*. 2011. commons.wikimedia.org/wiki/File:Cern-accelerator-complex.svg.
35. CERN Bulletin. *CERN outreach*. 2014. cds.cern.ch/record/1706606/files/Diagram_image.png.
36. L. Evans and P. Bryant. “LHC Machine”. In: *Journal of Instrumentation* 3.08 (2008),

S08001. stacks.iop.org/1748-0221/3/i=08/a=S08001.

37. G. P. Willering et al. “Performance of CERN LHC Main Dipole Magnets on the Test Bench From 2008 to 2016”. In: *IEEE Trans. Appl. Supercond.* 27.4 (2017), p. 4002705. doi:10.1109/TASC.2016.2633423.
38. P. Lebrun and L. Tavian. “Cooling with Superfluid Helium”. In: *CERN Accelerator School: Course on Superconductivity for Accelerators* (2014), 453–476. 24 p. cds.cern.ch/record/1974065. arxiv:1501.07156.
39. A. Kumar and A. Sharma. *Search for Dark Matter with LHC proton Beam Dump*. 2016. arxiv:1608.08818.
40. M. Bajko et al. *Report of the Task Force on the Incident of 19th September 2008 at the LHC*. Tech. rep. LHC-PROJECT-Report-1168. CERN-LHC-PROJECT-Report-1168. Geneva: CERN, Mar. 2009. cds.cern.ch/record/1168025.
41. W. Herr and B. Muratori. *Concept of luminosity*. 2006. cds.cern.ch/record/941318.
42. The CMS collaboration. *CMS Luminosity Public Results*. 2017. twiki.cern.ch/twiki/bin/view/CMSPublic/LumiPublicResults.
43. CMS collaboration. *CMS Physics: Technical Design Report Volume 1: Detector Performance and Software*. Tech. rep. Geneva, 2006. cds.cern.ch/record/922757.
44. D. Barney and E. Quigg. *Interactive Slice of the CMS detector*. cms-docdb.cern.ch/cgi-bin/PublicDocDB/ShowDocument?docid=4172.
45. CMS Collaboration. “The CMS experiment at the CERN LHC”. In: *Journal of Instrumentation* 3.08 (2008), S08004. stacks.iop.org/1748-0221/3/i=08/a=S08004.
46. S. Chatrchyan et al. “Description and performance of track and primary-vertex reconstruction with the CMS tracker”. In: *JINST* 9.10 (2014), P10009. doi:10.1088/1748-0221/9/10/P10009.
47. V. Karimäki et al. *The CMS tracker system project: Technical Design Report*. Tech. rep. Geneva, 1997. cds.cern.ch/record/368412.
48. *The CMS tracker: addendum to the Technical Design Report*. Tech. rep. Geneva, 2000. cds.cern.ch/record/490194.
49. V. Khachatryan et al. “CMS Tracking Performance Results from early LHC Operation”. In: *Eur. Phys. J. C* 70 (2010), pp. 1165–1192. doi:10.1140/epjc/s10052-010-1491-

3.

50. A. M. Sirunyan et al. “Particle-flow reconstruction and global event description with the CMS detector”. In: (2017). arxiv:1706.04965.
51. T. Ma and P. Dressendorfer. *Ionizing Radiation Effects in MOS Devices and Circuits*. A Wiley-Interscience publication. Wiley, 1989.
52. A. Saha. “Phase 1 upgrade of the CMS pixel detector”. In: *Journal of Instrumentation* 12.02 (2017), p. C02033. stacks.iop.org/1748-0221/12/i=02/a=C02033.
53. *The CMS muon project: Technical Design Report*. Tech. rep. Geneva, 1997. cds.cern.ch/record/343814.
54. L. Guiducci. *CMS muon system towards LHC Run 2 and beyond*. Tech. rep. CMS-CR-2014-333. Geneva: CERN, Oct. 2014. cds.cern.ch/record/1966038.
55. S. Chatrchyan et al. “The performance of the CMS muon detector in proton-proton collisions at $\sqrt{s} = 7$ TeV at the LHC”. In: *JINST* 8 (2013), P11002. doi:10.1088/1748-0221/8/11/P11002.
56. A. Tapper and D. Acosta. *CMS Technical Design Report for the Level-1 Trigger Upgrade*. Tech. rep. CERN-LHCC-2013-011. CMS-TDR-12. June 2013. cds.cern.ch/record/1556311.
57. V. Khachatryan et al. “The CMS trigger system”. In: *Journal of Instrumentation* 12.01 (2017), P01020. stacks.iop.org/1748-0221/12/i=01/a=P01020.
58. R. Fruhwirth. “Application of Kalman filtering to track and vertex fitting”. In: *Nucl. Instrum. Meth. A* 262 (1987), pp. 444–450. doi:10.1016/0168-9002(87)90887-4.
59. W. Adam et al. “Reconstruction of electrons with the Gaussian-sum filter in the CMS tracker at the LHC”. In: *Journal of Physics G: Nuclear and Particle Physics* 31.9 (2005), N9. stacks.iop.org/0954-3899/31/i=9/a=N01.
60. M. Cacciari, G. P. Salam, and G. Soyez. “The anti- k_t jet clustering algorithm”. In: *Journal of High Energy Physics* 2008.04 (2008), p. 063. stacks.iop.org/1126-6708/2008/i=04/a=063.
61. CMS Collaboration. “Identification of b-quark jets with the CMS experiment”. In: *Journal of Instrumentation* 8.04 (2013), P04013. stacks.iop.org/1748-0221/8/i=04/a=P04013.
62. *Identification of b quark jets at the CMS Experiment in the LHC Run 2*. Tech. rep.

- CMS-PAS-BTV-15-001. Geneva: CERN, 2016. cds.cern.ch/record/2138504.
63. CMS collaboration. “Jet energy scale and resolution in the CMS experiment in pp collisions at 8 TeV”. In: *Journal of Instrumentation* 12.02 (2017), P02014. stacks.iop.org/1748-0221/12/i=02/a=P02014.
 64. *Pileup Removal Algorithms*. Tech. rep. CMS-PAS-JME-14-001. Geneva: CERN, 2014. cds.cern.ch/record/1751454.
 65. *Pileup Jet Identification*. Tech. rep. CMS-PAS-JME-13-005. Geneva: CERN, 2013. cds.cern.ch/record/1581583.
 66. A. Brinkerhoff. *Observation of Top Quark Pairs Produced in Association with a W or Z Boson in PP Collisions at the LHC*. 2015. curate.nd.edu/show/3b59183390q.
 67. J. Alwall et al. “MADGRAPH 5: Going Beyond”. In: *JHEP* 06 (2011), p. 128. doi : 10.1007/JHEP06(2011)128.
 68. T. Sjöstrand, S. Mrenna, and P. Z. Skands. “PYTHIA 6.4 physics and manual”. In: *JHEP* 05 (2006), p. 026. doi : 10.1088/1126-6708/2006/05/026.
 69. S. Agostinelli et al. “GEANT4 – a simulation toolkit”. In: *Nucl. Instrum. Meth. A* 506 (2003), p. 250. doi : 10.1016/S0168-9002(03)01368-8.
 70. J. Pumplin, D. R. Stump, J. Huston, H. L. Lai, P. Nadolsky, and W. K. Tung. “New generation of parton distributions with uncertainties from global QCD analysis”. In: *JHEP* 07 (2002), p. 012. doi : 10.1088/1126-6708/2002/07/012.
 71. CMS Collaboration. *Calibration of the Combined Secondary Vertex b-Tagging discriminant using dileptonic ttbar and Drell-Yan events*. CMS Analysis Note AN-13-130. 2013.
 72. S. Chatrchyan et al. “Measurement of differential top-quark pair production cross sections in pp collisions at $\sqrt{s} = 7$ TeV”. In: *Eur. Phys. J. C* 73 (2013), p. 2339. doi : 10.1140/epjc/s10052-013-2339-4.
 73. B. P. Roe et al. “Boosted decision trees as an alternative to artificial neural networks for particle identification”. In: *Nucl. Instrum. Meth. A* 543 (2005), p. 577. doi : 10.1016/j.nima.2004.12.018.
 74. A. Hoecker et al. “TMVA: Toolkit for Multivariate Data Analysis”. In: *PoS ACAT* (2007), p. 040. [arxiv:physics/0703039](https://arxiv.org/abs/physics/0703039).
 75. LHC Higgs Combination Group. *Procedure for the LHC Higgs boson search combination*

in Summer 2011. Tech. rep. CMS-NOTE-2011-005, ATL-PHYS-PUB-2011-11. 2011. cds.cern.ch/record/1379837.

76. Particle Data Group, J. Beringer, et al. “Review of Particle Physics”. In: *Phys. Rev. D* 86 (2012), p. 010001. doi : 10.1103/PhysRevD.86.010001.
77. G. Cowan et al. “Asymptotic formulae for likelihood-based tests of new physics”. In: *The European Physical Journal C* 71.2 (Feb. 9, 2011), p. 1554. doi : 10.1140/epjc/s10052-011-1554-0.
78. R. Barlow. “Extended maximum likelihood”. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 297.3 (1990), pp. 496–506. doi : 10.1016/0168-9002(90)91334-8.
79. CMS Collaboration. *CMS luminosity based on pixel cluster counting – Summer 2013 update*. CMS Physics Analysis Summary CMS-PAS-LUM-13-001. 2013. cds.cern.ch/record/1598864.
80. CMS Collaboration. “Measurement of the inelastic proton-proton cross section at $\sqrt{s}=7\text{TeV}$ ”. In: *Phys. Lett. B* 722 (2013), p. 5. doi : 10.1016/j.physletb.2013.03.024.
81. S. Chatrchyan et al. “Determination of jet energy calibration and transverse momentum resolution in CMS”. In: *JINST* 6 (2011), P11002. doi : 10.1088/1748-0221/6/11/P11002.
82. S. Alekhin et al. *The PDF4LHC Working Group interim report*. Tech. rep. 2011. arxiv : 1101.0536.
83. M. Botje et al. *The PDF4LHC Working Group interim recommendations*. Tech. rep. 2011. arxiv : 1101.0538.
84. M. V. Garzelli et al. “ $t\bar{t}W^\pm$ and $t\bar{t}Z$ hadroproduction at NLO accuracy in QCD with parton shower and hadronization effects”. In: *JHEP* 11 (2012), p. 056. doi : 10.1007/JHEP11(2012)056.
85. R. D. Ball et al. “Parton distributions for the LHC Run II”. In: *JHEP* 04 (2015), p. 040. doi : 10.1007/JHEP04(2015)040.
86. T. Sjöstrand, S. Mrenna, and P. Z. Skands. “A brief introduction to PYTHIA 8.1”. In: *Comput. Phys. Commun.* 178 (2008), p. 852. doi : 10.1016/j.cpc.2008.01.036.
87. T. Sjostrand et al. “An Introduction to PYTHIA 8.2”. In: *Comput. Phys. Commun.* 191 (2015), pp. 159–177. doi : 10.1016/j.cpc.2015.01.024.

88. P. Skands, S. Carrazza, and J. Rojo. “Tuning PYTHIA 8.1: the Monash 2013 tune”. In: *Eur. Phys. J. C* 74 (2014), p. 3024. doi : 10.1140/epjc/s10052-014-3024-y.
89. V. Khachatryan et al. “Event generator tunes obtained from underlying event and multiparton scattering measurements”. In: *Eur. Phys. J. C* 76 (2016), p. 155. doi : 10.1140/epjc/s10052-016-3988-x.
90. H. B. Hartanto et al. “Higgs boson production in association with top quarks in the POWHEG BOX”. In: *Phys. Rev. D* 91 (2015), p. 094003. doi : 10.1103/PhysRevD.91.094003.
91. T. Melia et al. “ W^+W^- , WZ and ZZ production in the POWHEG BOX”. In: *JHEP* 11 (2011), p. 078. doi : 10.1007/JHEP11(2011)078.
92. P. Nason and G. Zanderighi. “ W^+W^- , WZ and ZZ production in the POWHEG-BOX-V2”. In: *Eur. Phys. J. C* 74 (2014), p. 2702. doi : 10.1140/epjc/s10052-013-2702-5.
93. J. Alwall et al. “The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations”. In: *JHEP* 07 (2014), p. 079. doi : 10.1007/JHEP07(2014)079.
94. S. Alioli et al. “A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX”. In: *JHEP* 06 (2010), p. 043. doi : 10.1007/JHEP06(2010)043.
95. J. M. Campbell and R. K. Ellis. “MCFM for the Tevatron and the LHC”. In: *Nucl. Phys. Proc. Suppl.* 205-206 (2010), p. 10. doi : 10.1016/j.nuclphysbps.2010.08.011.
96. F. Cascioli et al. “ZZ production at hadron colliders in NNLO QCD”. In: *Phys. Lett. B* 735 (2014), p. 311. doi : 10.1016/j.physletb.2014.06.056.
97. F. Caola et al. “QCD corrections to ZZ production in gluon fusion at the LHC”. In: *Phys. Rev. D* 92 (2015), p. 094028. doi : 10.1103/PhysRevD.92.094028.
98. CMS Collaboration. “Identification of b-quark jets with the CMS experiment”. In: *JINST* 8 (2013), p. 04013. doi : 10.1088/1748-0221/8/04/P04013.
99. CMS Collaboration. *CMS Luminosity Measurement for the 2016 Data Taking Period*. CMS Physics Analysis Summary CMS-PAS-LUM-17-001. 2017. cds . cern . ch / record/2138682.
100. M. Aaboud et al. “Measurement of the Inelastic Proton-Proton Cross Section at $\sqrt{s}=13$ TeV with the ATLAS Detector at the LHC”. In: *Phys. Rev. Lett.* 117 (2016), p. 182002. doi : 10.1103/PhysRevLett.117.182002.

101. S. Chatrchyan et al. “Performance of CMS muon reconstruction in pp collision events at $\sqrt{s}=7\text{TeV}$ ”. In: *JINST* 7 (2012), P10002. doi : 10.1088/1748-0221/7/10/P10002.
102. V. Khachatryan et al. “Performance of electron reconstruction and selection with the CMS detector in proton-proton collisions at $\sqrt{s}=8\text{TeV}$ ”. In: *JINST* 10 (2015), P06005. doi : 10.1088/1748-0221/10/06/P06005.
103. CMS Collaboration. “Performance of the CMS missing transverse momentum reconstruction in pp data at $\sqrt{s}=8\text{TeV}$ ”. In: *JINST* 10 (2015), P02006. doi : 10.1088/1748-0221/10/02/P02006.
104. CMS Collaboration. *Performance of missing energy reconstruction in 13 TeV pp collision data using the CMS detector*. CMS Physics Analysis Summary CMS-PAS-JME-16-004. 2016. cds.cern.ch/record/2205284.
105. D. de Florian et al. *Handbook of LHC Higgs Cross Sections: 4. Deciphering the Nature of the Higgs Sector*. Tech. rep. 2016. doi : 10.23731/CYRM-2017-002.
106. R. Röntsch and M. Schulze. “Constraining couplings of the top quarks to the Z boson in $t\bar{t}b\bar{b}+Z$ production at the LHC”. In: (Apr. 2014), p. 25. arxiv.org/abs/1404.1005. arxiv:1404.1005.
107. A. Alloul et al. “FeynRules 2.0 — A complete toolbox for tree-level phenomenology”. In: *Computer Physics Communications* 185.8 (2014), pp. 2250–2300. doi : https://doi.org/10.1016/j.cpc.2014.04.012.
108. J. Alwall et al. “MadGraph 5: going beyond”. In: *Journal of High Energy Physics* 2011.6 (June 2011), p. 128. doi : 10.1007/JHEP06(2011)128.
109. ATLAS and CMS Collaborations. *Procedure for the LHC Higgs boson search combination in summer 2011*. ATL-PHYS-PUB-2011-011, CMS NOTE-2011/005. 2011. cdsweb.cern.ch/record/1379837.
110. A. Martin. University of Notre Dame, Indiana, USA, 2017.
111. B. Cousins. “Probability Density Functions for Positive Nuisance Parameters”. 2010. www.physics.ucla.edu/~cousins/stats/cousins_lognormal_prior.pdf.
112. Python Core Team. “Python 2.7: A dynamic, open source programming language”. 2010.
113. S. van der Walt, S. C. Colbert, and G. Varoquaux. “The NumPy Array: A Structure for Efficient Numerical Computation”. In: *Computing in Science Engineering* 13.2 (Mar. 2011), pp. 22–30. doi : 10.1109/MCSE.2011.37.

114. O. Mattelaer. “On the maximal use of Monte Carlo samples: re-weighting events at NLO accuracy”. In: *The European Physical Journal C* 76.12 (Dec. 5, 2016), p. 674. doi : 10.1140/epjc/s10052-016-4533-7.
115. J. S. Gainer et al. “Exploring theory space with Monte Carlo reweighting”. In: *Journal of High Energy Physics* 10 (Oct. 13, 2014), p. 78. doi : 10.1007/JHEP10(2014)078.
116. D. Skeeahan et al. “Opportunistic High Energy Physics Computing in User Space with Parrot”. In: *IEEE Symposium on Cluster, Cloud, and Grid Computing (CCGrid)*. 2014.
117. A. Woodard et al. “Exploiting volatile opportunistic computing resources with Lobster”. In: *Journal of Physics: Conference Series* 664.3 (2015), p. 032035. doi : 10.1088/1742-6596/664/3/032035.
118. M. Mascheroni et al. “CMS distributed data analysis with CRAB3”. In: *Journal of Physics: Conference Series* 664.6 (2015), p. 062038. doi : 10.1088/1742-6596/664/6/062038.
119. A. Woodard et al. “Scaling Data Intensive Physics Applications to 10k Cores on Non-dedicated Clusters with Lobster”. In: *2015 IEEE International Conference on Cluster Computing, CLUSTER 2015, Chicago, IL, USA, September 8-11, 2015*. 2015, pp. 322–331. doi : 10.1109/CLUSTER.2015.53.
120. M. Wolf et al. “Scaling Up a CMS Tier-3 Site with Campus Resources and a 100 Gb/s Network Connection: What Could Go Wrong? (poster)”. In: *Journal of Physics: Conference Series* 898.8 (2017), p. 082041. doi : 10.1088/1742-6596/898/8/082041.
121. M. Wolf et al. “Opportunistic Computing with Lobster: Lessons Learned from Scaling up to 25k Non-Dedicated Cores (slides)”. In: *Journal of Physics: Conference Series* 898.5 (2017), p. 052036. doi : 10.1088/1742-6596/898/5/052036.
122. M. Albrecht, D. Rajan, and D. Thain. “Making work queue cluster-friendly for data intensive scientific applications”. In: *2013 IEEE International Conference on Cluster Computing (CLUSTER)*. Sept. 2013, pp. 1–8. doi : 10.1109/CLUSTER.2013.6702628.
123. H. Meng et al. “Conducting reproducible research with Umbrella: Tracking, creating, and preserving execution environments”. In: *12th IEEE International Conference on e-Science*. Baltimore, MD, USA, Oct. 2016, pp. 91–100. doi : 10.1109/eScience.2016.7870889.
124. H. Meng et al. “A case study in preserving a high energy physics application with Parrot”. In: *Journal of Physics: Conference Series* 664.3 (2015), p. 032022. doi : 10.1088/1742-6596/664/3/032022.

125. H. Meng et al. “Conducting reproducible research with Umbrella: Tracking, creating, and preserving execution environments”. In: *2016 IEEE 12th International Conference on e-Science (e-Science)*. Oct. 2016, pp. 91–100. doi : 10 . 1109 / eScience . 2016 . 7870889.
126. M. Albrecht et al. “Makeflow: A Portable Abstraction for Data Intensive Computing on Clusters, Clouds, and Grids”. In: *Proceedings of the 1st ACM SIGMOD Workshop on Scalable Workflow Execution Engines and Technologies*. SWEET ’12. Scottsdale, Arizona, USA: ACM, 2012, 1:1–1:13. doi : 10 . 1145 / 2443416 . 2443417.
127. A. Woodard and A. Wightman. *annawoodard/NPfitProduction: v1.0*. Mar. 2018. doi : 10 . 5281 / zenodo . 1197392.
128. A. Woodard. *annawoodard/NPfit: v1.1*. Mar. 2018. doi : 10 . 5281 / zenodo . 1197390.
129. J. Ellis. “TikZ-Feynman: Feynman diagrams with TikZ”. In: *Comput. Phys. Commun.* 210 (2017), pp. 103–123. doi : 10 . 1016 / j . cpc . 2016 . 08 . 019.