

**Search for High-Mass Resonant Top-Antitop Pair Production
in Lepton+Jets Events in 8 TeV pp Collisions**

BY

Paul Jonathan Turner
B.S., Boise State University, 2009

THESIS

Submitted as partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Physics
in the Graduate College of the
University of Illinois at Chicago, 2015

Chicago, Illinois

Defense Committee:

Cecilia E. Gerber, Chair and Advisor
Mark R. Adams, Advisor
Richard J. Cavanaugh
Mikhail A. Stephanov
Dhiman Chakraborty, Northern Illinois University

This thesis is dedicated to my parents and family.

ACKNOWLEDGMENTS

I would like to thank my advisors, Dr. Cecilia Gerber and Dr. Mark Adams, for all the support they provided to me during my research. I would also like to thank all the professors and staff of the UIC physics department for providing a rich and fulfilling education during my doctoral studies. During my tenure I was granted many opportunities to expand my knowledge and experience different aspects of collaborative experimentation in high-energy particle physics. Specifically, I was able to take part in the NSF-funded Partners in International Research and Education program which allowed me to spend time at the Paul-Scherrer Institute and ETH Zürich in Switzerland researching and collaborating with the team behind the design of the pixel tracker used within CMS. During my time in Switzerland, Dr. Roland Horisberger, Dr. Tilman Rohe, and Dr. Andrey Starodumov provided invaluable assistance to the research projects I was working on.

Finally, I thank my family and friends for providing assistance and support during my studies.

PJT

TABLE OF CONTENTS

1	THEORETICAL INTRODUCTION	1
1.1	The Standard Model	1
1.2	Beyond the Standard Model	4
1.2.1	Topcolor	6
1.2.2	Kaluza-Klein excitation of gluons in Randall-Sundrum models	7
2	THE LARGE HADRON COLLIDER AND COMPACT MUON SOLENOID	9
2.1	The Large Hadron Collider	10
2.2	Compact Muon Solenoid Experiment	13
2.2.1	Layer 1: The Tracker	16
2.3	Layer 2: Electromagnetic Calorimeter	19
2.4	Layer 3: Hadronic Calorimeter	21
2.5	Layer 4: The Muon System	21
2.6	Trigger	23
3	OBJECT IDENTIFICATION	25
4	SIMULATED EVENTS	31
4.1	MC Simulation	31
4.1.1	Event Generators Used	32
4.1.2	MC Sample Cross-Sections	34
5	EVENT SELECTION	38
5.1	Preselection	40
5.2	Muon+Jets Channel	41
5.3	Electron+Jets Channel	42
6	CORRECTIONS AND EFFICIENCIES	44
6.1	Pileup Re-Weighting	44
6.2	Lepton Identification and Trigger Efficiencies	45
6.2.1	Electron Channel Trigger	52
6.3	Jet Energy Corrections	54
6.4	Jet Energy Resolution	56
6.5	B-Tagging Efficiency and Mistag Rate	56
6.6	Top-Tagging Efficiency and Mistag Rate	57

TABLE OF CONTENTS (continued)

7	EVENT RECONSTRUCTION AND FINAL SELECTION	62
7.1	Neutrino Reconstruction	62
7.2	Jet Assignment	64
7.3	Final Selection and Event Yields	66
8	STATISTICAL ANALYSIS	71
8.1	Systematic Uncertainties	74
8.1.1	Normalization Uncertainties	74
8.1.2	Shape Uncertainties	75
9	RESULTS	81
10	SUMMARY AND OUTLOOK	89
A	χ^2 CUT AND EVENT CATEGORIZATION OPTIMIZATION	94
B	KINEMATIC DISTRIBUTIONS	98
C	SYSTEMATIC UNCERTAINTY TEMPLATES	116
	CITED LITERATURE	127
	VITA	134

LIST OF TABLES

4.1	SM background process cross-sections used in the analysis. [73]	35
4.2	Theoretical cross-sections for the the signal models studied. Z' cross-sections are taken from Ref. [14] and g_{KK} cross-sections are taken from Ref. [20]. Cross-sections are multiplied by a factor $K = 1.3$ to account for higher-order corrections. [72]	36
4.3	$t\bar{t}$ and W+jets cross-sections used for Q^2 scale and matching systematics. . .	37
6.1	Overall efficiencies of the Ele30[...] and its logical OR with the PFJet320 trigger in the $e\mu$ control sample for different narrow- Z' mass hypotheses. The uncertainty is statistical only.	53
7.1	Number of observed and expected events in the muon and electron channels. The yields are modified by scale factors summarized in Tab. 7.2. The uncertainty on each background includes both the statistical error and the posterior uncertainties from Tab. 7.2.	68
7.2	Best-fit normalization scale factors for each background process obtained from a maximum-likelihood fit across all three event categories in both channels simultaneously. The top-tagging scale factor affects the normalization of $t\bar{t}$ background in the “1top” category, whereas the other scale factors apply to all categories equally. Also shown are the prior and posterior constraints on the rate uncertainty. These scale factors are used to normalize MC to data for all figures.	69
8.1	Table of systematic uncertainties considered in the statistical analysis. The uncertainties above the horizontal line are used for the determination of the background normalization and the CMS Top Tagger efficiency scale factor, as described in Sec. 6.6.	78
8.2	Effect of the systematic uncertainties on the normalization of background and signal samples in the electron channel for the “1top” (top), “0top1btag” (middle) and “0top0btag” (bottom) categories. The numbers correspond to the difference in percent with respect to the nominal yield. The upper (lower) value refers to the $+1\sigma$ (-1σ) variation of a given systematic.	79
8.3	Effect of the systematic uncertainties on the normalization of background and signal samples in the muon channel for the “1top” (top), “0top1btag” (middle) and “0top0btag” (bottom) category. The numbers correspond to the difference in percent with respect to the nominal yield. The upper (lower) value refers to the $+1\sigma$ (-1σ) variation of a given systematic.	80

LIST OF TABLES (continued)

9.1	Expected and observed 95% CL lower mass limits for the three models studied in this analysis. Limits are given for the electron channel, the muon channel, and the combination of the two.	82
9.2	Expected and observed 95% CL limits on the production cross-section times branching ratio of the narrow Z' model. Limits are given for the electron channel, the muon channel, and the combination of the two. Fig. 9.1 plots these numerical values against the theoretical cross-section for the model.	83
9.3	Expected and observed 95% CL limits on the production cross-section times branching ratio of the wide Z' model. Limits are given for the electron channel, the muon channel, and the combination of the two. Fig. 9.2 plots these numerical values against the theoretical cross-section for the model.	84
9.4	Expected and observed 95% CL limits on the production cross-section times branching ratio of the Kaluza-Klein gluon resonance model. Limits are given for the electron channel, the muon channel, and the combination of the two. Fig. 9.3 plots these numerical values against the theoretical cross-section for the model.	85
A.1	CMS Top Tagger efficiency for various signals and SM backgrounds. The efficiency is defined as the ratio of events containing a top-tagged jet which pass the selection cuts defined in Chap. 5.	96

LIST OF FIGURES

1.1	Diagram summarizing the interactions between elementary particles according to the Standard Model. Dark ovals represent types of particles and blue arcs represent interactions between them. An arc that links to a box is equivalent to a set of arcs that link every oval in the box. [6]	2
1.2	Diagram summarizing the Standard Model of particle physics. Shown are the elementary particles of the SM (the Higgs boson, the three generations of quarks and leptons, and the gauge bosons), including their names, masses, spins, charges, chiralities, and interactions with the strong, weak, and electromagnetic forces. The top shows the particle content before electroweak symmetry breaking and the bottom afterwards [8].	4
2.1	Layout of the accelerator complex at CERN [36].	11
2.2	Cross-section of a dipole magnet used in the LHC. The 2-in-1 structure of the beam pipe which contains counter-rotating beams can be seen [37].	12
2.3	Cutaway view of the CMS detector showing the outer four layers for muon detection interleaved with the iron return yoke of the magnet, the central calorimeters, and the central detector consisting of the tracking system. [39]	15
2.4	View of a transverse slice through CMS showing how different particles interact with the various subdetectors. [40]	16
2.5	Schematic cross-section of the CMS tracker in the r - z plane. Only the top half is shown, the bottom half being symmetric. The interaction point, labeled by a star, is the center of the Tracker. The green dashed lines separate the different parts of the Tracker. Strip tracker modules that only have a single layer are shown by thin black lines. Strip tracker modules consisting of two back-to-back rotated layers, which can provide some 3-D hit reconstruction, are shown by the thick blue lines. The innermost tracker, the pixel detector, is shown in red and provides true 3-D hit reconstruction. Reproduced from Ref. [42].	17
2.6	Layout of the CMS ECAL showing the arrangement of crystal modules, supermodules, endcaps, and the preshower. Reproduced from Ref. [38]	20
2.7	Layout of one quarter of the CMS HCAL showing the segmentation in the r - z plane for the barrel (HB), outer barrel (HO), and endcap (HE) detectors. Reproduced from Ref. [38]	22

LIST OF FIGURES (continued)

2.8	Layout of one quadrant of the CMS detector highlighting the muon system. The four barrel drift tube stations (MB1-MB4, green), and the endcap CSC/RPC stations (ME1-ME4, blue/red) are shown. Reproduced from Ref. [45]	23
4.1	Generated $t\bar{t}$ invariant mass for the different BSM signals in this analysis. Shown are (a) 1.5 TeV and (b) 3 TeV masses. Z' 1% (narrow) is shown in red, Z' 10% (wide) is shown in blue and the KK gluon is shown in black. . .	33
5.1	The difference between the expected decay topology for SM $t\bar{t}$ production (a) and high-mass resonant $t\bar{t}$ production (b) is shown. High-mass resonances result in a boosted topology where the decay products can merge together. This merging of final state objects must be taken into account when searching for these events.	39
6.1	Pileup distributions for data and $t\bar{t}$ MC before (a) and after (b) pileup reweighting. The distributions in (a) are taken at the generator level for MC and before any analysis cuts for data whereas the distributions for (b) are taken after the preselection described in Sec. 5.1 which accounts for the slight difference in shape of the data pileup distribution. The distributions have been scaled to the same area to highlight the shape difference.	45
6.2	Scale factors applied to MC samples for muon ID efficiency. The scale factor is given as a function muon p_T and η region. The scale factor is derived separately for the barrel ($ \eta < 0.9$), endcap ($1.2 < \eta < 2.1$) and overlap ($0.9 < \eta < 1.2$) regions as the muon detector uses different physical detectors systems for each region which affects the efficiency measurements.	46
6.3	Scale factors applied to MC samples for muon trigger efficiency. The scale factors are given as a factor of muon p_T and η region. The scale factor is derived separately for the barrel ($ \eta < 0.9$), endcap ($1.2 < \eta < 2.1$) and overlap ($0.9 < \eta < 1.2$) regions as the muon detector uses different physical detectors systems for each region which affects the efficiency measurements. .	47
6.4	Scale factors applied to MC samples for electron ID efficiency. The scale factor is given as a function electron p_T and η_{SC}	47
6.5	Data/MC plots for the $Z \rightarrow \ell\ell + \text{jets}$ control sample for muons and electrons (before any specific cut on the probe): $M_{\mu\mu}$ invariant mass (a), M_{ee} invariant mass (b), probe muon p_T (c), probe electron p_T (d), and number of AK5-jets with $p_T > 50$ GeV for muons (e) and electrons (f). The overall data/MC agreement is good except for the dielectron invariant mass. This is due to a known electron energy discrepancy for this kinematic quantity, but has no effect on the efficiency measurement. The gray band corresponds to the statistical uncertainty of the simulation. The bottom plot of each figure shows the ratio of data over background.	50

LIST OF FIGURES (continued)

6.6	Efficiency of the 2D-cut for data and MC events as a function of $\Delta R_{\min}(\ell, jets)$ in the $Z \rightarrow \mu\mu$ (a) and $Z \rightarrow ee$ (b) control samples. The scale factor is the ratio of these efficiencies and is plotted below each figure.	51
6.7	Data/MC comparisons of several kinematic variables in the $e\mu$ control sample used to measure the electron channel trigger efficiency. Shown are the number of primary vertices (a), number of AK5 jets with $p_T > 30$ GeV (b), the electron p_T (c) and the leading jet p_T (d). The Z' signal has been normalized to a cross section of 1pb and the sum of backgrounds has been normalized to data. The gray band corresponds to the statistical uncertainty of the simulated samples.	53
6.8	Data/MC efficiency ratio for the electron channel combined trigger plotted as a function of the leading AK5 jet p_T . The fit function (solid red line) used is $f(p_T) = D + \frac{C}{2} \left[1 + \operatorname{erf} \left(\frac{p_T - A}{\sqrt{2B}} \right) \right]$. The shaded band around the fit function corresponds to the $\pm 1\sigma$ statistical error band obtained by propagating the full correlation matrix of the fit result.	54
6.9	Data/MC comparisons of relevant kinematic variables in the W+jets control sample used to measure the CMS Top Tagging algorithm's mistag rate: p_T of the p_T leading topjet candidate (a), mass of the leading topjet candidate (b), p_T of the muon (c), and \cancel{E}_T (d). The sum of MC backgrounds is normalized to data. The gray band corresponds to the statistical uncertainty of the simulation.	59
7.1	Two examples of the reconstructed hadronic top mass, M_{had} , in correctly matched events where all three final state quarks, $q\bar{q}b$, are matched to a jet are shown. Both SM $t\bar{t}$ with $M_{t\bar{t}} < 1$ TeV (a) and a 1% width 2 TeV Z' signal MC (b) are shown with events separated by the number of reconstructed jets. The distributions are well modeled by Gaussian functions and the fit parameters are used in the χ^2 calculation.	65
7.2	Distributions of the χ^2 discriminator for the (a) electron and (b) muon channels after the kinematic selection described in Chap. 5. Also shown are the contributions from different $W+n$ parton samples for the (c) electron and (d) muon channels. The shoulder at ≈ 150 present in the muon channel is suppressed in the electron channel due to the contribution from the $W+1$ parton process being much smaller due to the triangle cut. The yields of the background processes are normalized to data using scale factors described in Chap. 8. The shaded band shows the change in the yield due to systematic uncertainties also described in Chap. 8.	67
7.3	Reconstructed $M_{t\bar{t}}$ distributions in the electron (left) and muon (right) channels. The three categories are "1top" on the top, "0top1btag" in the middle, and "0top0btag" on the bottom. The yields of the background processes are normalized to data using scale factors described in Tab. 7.2. The shaded band shows the change in the yield due to systematic uncertainties described in Sec. 8.1. The signal is normalized to a cross-section of 1 pb.	70

LIST OF FIGURES (continued)

9.1	Expected and observed 95% CL limits on the production cross-section times branching ratio of the narrow Z' model. Limits are given for the electron channel (a), the muon channel (b), and the combination of the two (c). The theoretical prediction for the cross-section of a Z' boson with a width of 1.2% are taken from Ref. [14] and multiplied by a K-factor of $K = 1.3$ to account for higher-order corrections. [72]	86
9.2	Expected and observed 95% CL limits on the production cross-section times branching ratio of the wide Z' model. Limits are given for the electron channel (a), the muon channel (b), and the combination of the two (c). The theoretical prediction for the cross-section of a Z' boson with a width of 10% are taken from Ref. [14] and multiplied by a K-factor of $K = 1.3$ to account for higher-order corrections. [72]	87
9.3	Expected and observed 95% CL limits on the production cross-section times branching ratio of the Kaluza-Klein gluon resonance model. Limits are given for the electron channel (a), the muon channel (b), and the combination of the two (c). The theoretical prediction for the cross-section of a Kaluza-Klein resonant gluon in the Randall-Sundrum model are taken from Ref. [20] and multiplied by a K-factor of $K = 1.3$ to account for higher-order corrections. [72]	88
10.1	Expected 95% CL Bayesian upper limits on production cross section times branching ratio for a narrow Z' boson. Individual analysis limits are shown separately along with the combined results.	91
10.2	Expected 95% CL Bayesian upper limits on production cross section times branching ratio for a wide Z' boson. Individual analysis limits are shown separately along with the combined results.	91
10.3	Expected 95% CL Bayesian upper limits on production cross section times branching ratio for a Kaluza-Klein gluon. Individual analysis limits are shown separately along with the combined results.	92
10.4	95% CL Bayesian upper limits on production cross section times branching ratio for a narrow Z' boson (a), wide Z' boson (b), and Kaluza-Klein gluon (c). The vertical dashed line represents the transition from the threshold lepton+jets analysis and the combined results. The theoretical prediction for the cross-section of a Z' boson with a width of 1.2% and 10% are taken from Ref. [14]. The theoretical prediction for a Kaluza-Klein resonant gluon in the Randall-Sundrum model are taken from Ref. [20]. The signal cross-sections are multiplied by a K-factor of $K = 1.3$ to account for higher-order corrections. [72]	93

LIST OF FIGURES (continued)

A.1	95% CL expected (Bayesian) limits on the production cross section times branching ratio of a narrow Z' resonance decaying to $t\bar{t}$ in the muon channel for the “3 categories” case described in the text: (a) using no N-subjettiness cut in the top-tag definition, (b) with a cut $\tau_{32} < 0.7$ in the top-tag definition. The corresponding limits for the settings used in Ref. [91], i.e. two categories based on the number of b -tags with $\chi^2 < 10$, are also shown (black line). The uncertainties on the normalization of each background template are the only systematics included in the limit calculation.	96
A.2	95% CL expected (Bayesian) limits on the production cross section times branching ratio of a narrow Z' resonance decaying to $t\bar{t}$ in the muon channel for the “4 categories” case described in the text: (a) using no N-subjettiness cut in the top-tag definition, (b) with a cut $\tau_{32} < 0.7$ in the top-tag definition. The corresponding limits for the settings used in [91], i.e. two categories based on the number of b -tags with $\chi^2 < 10$, are also shown (black line). The uncertainties on the normalization of each background template are the only systematics included in the limit calculation.	97
A.3	95% CL expected (Bayesian) limits on the production cross section times branching ratio of a narrow Z' resonance decaying to $t\bar{t}$ in the muon channel, comparing the “3 categories” and “4 categories” cases described in the text. For each categorization, a cut $\chi^2 < 50$ is used and results are shown for two different definitions of top-tagged jet (with and without the $\tau_{32} < 0.7$ cut). The uncertainties on the normalization of each background template are the only systematics included in the limit calculation.	97
B.1	Legend colors for plots in Appendix B	98
B.2	Data/MC comparison for the lepton p_T after all analysis cuts are applied. Events are split into the following categories: muon “1top” (a), electron “1top” (b), muon “0top1btag” (c), electron “0top1btag” (d), muon “0top0btag” (e), electron “0top0btag” (f). Each background process is corrected by scale factors from the maximum-likelihood fit to data as described in Chap. 8. The signal MC samples are normalized to a cross section of 1 pb. The shaded band shows the error associated with the background expectation and includes both the MC statistical uncertainty and the post-fit systematic uncertainty described in Chap. 8.	99
B.3	Data/MC comparison for the lepton η after all analysis cuts are applied. Events are split into the following categories: muon “1top” (a), electron “1top” (b), muon “0top1btag” (c), electron “0top1btag” (d), muon “0top0btag” (e), electron “0top0btag” (f). Each background process is corrected by scale factors from the maximum-likelihood fit to data as described in Chap. 8. The signal MC samples are normalized to a cross section of 1 pb. The shaded band shows the error associated with the background expectation and includes both the MC statistical uncertainty and the post-fit systematic uncertainty described in Chap. 8.	100

LIST OF FIGURES (continued)

B.4	Data/MC comparison for the minimum ΔR -distance between lepton and AK5-jets after all analysis cuts are applied. Events are split into the following categories: muon “1top” (a), electron “1top” (b), muon “0top1btag” (c), electron “0top1btag” (d), muon “0top0btag” (e), electron “0top0btag” (f). Each background process is corrected by scale factors from the maximum-likelihood fit to data as described in Chap. 8. The signal MC samples are normalized to a cross section of 1 pb. The shaded band shows the error associated with the background expectation and includes both the MC statistical uncertainty and the post-fit systematic uncertainty described in Chap. 8.	101
B.5	Data/MC comparison for the missing E_T after all analysis cuts are applied. Events are split into the following categories: muon “1top” (a), electron “1top” (b), muon “0top1btag” (c), electron “0top1btag” (d), muon “0top0btag” (e), electron “0top0btag” (f). Each background process is corrected by scale factors from the maximum-likelihood fit to data as described in Chap. 8. The signal MC samples are normalized to a cross section of 1 pb. The shaded band shows the error associated with the background expectation and includes both the MC statistical uncertainty and the post-fit systematic uncertainty described in Chap. 8.	102
B.6	Data/MC comparison for the number of AK5-jets with $p_T > 50$ GeV after all analysis cuts are applied. Events are split into the following categories: muon “1top” (a), electron “1top” (b), muon “0top1btag” (c), electron “0top1btag” (d), muon “0top0btag” (e), electron “0top0btag” (f). Each background process is corrected by scale factors from the maximum-likelihood fit to data as described in Chap. 8. The signal MC samples are normalized to a cross section of 1 pb. The shaded band shows the error associated with the background expectation and includes both the MC statistical uncertainty and the post-fit systematic uncertainty described in Chap. 8.	103
B.7	Data/MC comparison for the leading AK5-jet p_T after all analysis cuts are applied. Events are split into the following categories: muon “1top” (a), electron “1top” (b), muon “0top1btag” (c), electron “0top1btag” (d), muon “0top0btag” (e), electron “0top0btag” (f). Each background process is corrected by scale factors from the maximum-likelihood fit to data as described in Chap. 8. The signal MC samples are normalized to a cross section of 1 pb. The shaded band shows the error associated with the background expectation and includes both the MC statistical uncertainty and the post-fit systematic uncertainty described in Chap. 8.	104

LIST OF FIGURES (continued)

<p>B.8 Data/MC comparison for the leading AK5-jet η after all analysis cuts are applied. Events are split into the following categories: muon “1top” (a), electron “1top” (b), muon “0top1btag” (c), electron “0top1btag” (d), muon “0top0btag” (e), electron “0top0btag” (f). Each background process is corrected by scale factors from the maximum-likelihood fit to data as described in Chap. 8. The signal MC samples are normalized to a cross section of 1 pb. The shaded band shows the error associated with the background expectation and includes both the MC statistical uncertainty and the post-fit systematic uncertainty described in Chap. 8.</p>	105
<p>B.9 Data/MC comparison for the 2nd-leading AK5-jet p_T after all analysis cuts are applied. Events are split into the following categories: muon “1top” (a), electron “1top” (b), muon “0top1btag” (c), electron “0top1btag” (d), muon “0top0btag” (e), electron “0top0btag” (f). Each background process is corrected by scale factors from the maximum-likelihood fit to data as described in Chap. 8. The signal MC samples are normalized to a cross section of 1 pb. The shaded band shows the error associated with the background expectation and includes both the MC statistical uncertainty and the post-fit systematic uncertainty described in Chap. 8.</p>	106
<p>B.10 Data/MC comparison for the 2nd-leading AK5-jet η after all analysis cuts are applied. Events are split into the following categories: muon “1top” (a), electron “1top” (b), muon “0top1btag” (c), electron “0top1btag” (d), muon “0top0btag” (e), electron “0top0btag” (f). Each background process is corrected by scale factors from the maximum-likelihood fit to data as described in Chap. 8. The signal MC samples are normalized to a cross section of 1 pb. The shaded band shows the error associated with the background expectation and includes both the MC statistical uncertainty and the post-fit systematic uncertainty described in Chap. 8.</p>	107
<p>B.11 Data/MC comparison for kinematic variables of the “top-tagged” CA8-jet in the “1top” event category: top-tag jet p_T (muon channel) (a), top-tag jet p_T (electron channel) (b), top-tag jet η (muon channel) (c), top-tag jet η (electron channel) (d), top-tag jet mass (muon channel) (e), top-tag jet mass (electron channel) (f). Each background process is corrected by scale factors from the maximum-likelihood fit to data as described in Chap. 8. The signal MC samples are normalized to a cross section of 1 pb. The shaded band shows the error associated with the background expectation and includes both the MC statistical uncertainty and the post-fit systematic uncertainty described in Chap. 8.</p>	108

LIST OF FIGURES (continued)

<p>B.12 Data/MC comparison for kinematic variables of the “top-tagged” CA8-jet in the “1top” event category: top-tag jet groomed mass (muon channel) (a), top-tag jet groomed mass (electron channel) (b), top-tag jet minimum subjet-pairwise mass (muon channel) (c), top-tag jet minimum subjet-pairwise mass (electron channel) (d), top-tag jet τ_{32} variable (muon channel) (e), top-tag jet τ_{32} variable (electron channel) (f). Each background process is corrected by scale factors from the maximum-likelihood fit to data as described in Chap. 8. The signal MC samples are normalized to a cross section of 1 pb. The shaded band shows the error associated with the background expectation and includes both the MC statistical uncertainty and the post-fit systematic uncertainty described in Chap. 8.</p>	109
<p>B.13 Data/MC comparison for the hadronic-top p_T after all analysis cuts are applied. Events are split into the following categories: muon “1top” (a), electron “1top” (b), muon “0top1btag” (c), electron “0top1btag” (d), muon “0top0btag” (e), electron “0top0btag” (f). Each background process is corrected by scale factors from the maximum-likelihood fit to data as described in Chap. 8. The signal MC samples are normalized to a cross section of 1 pb. The shaded band shows the error associated with the background expectation and includes both the MC statistical uncertainty and the post-fit systematic uncertainty described in Chap. 8.</p>	110
<p>B.14 Data/MC comparison for the hadronic-top mass after all analysis cuts are applied. Events are split into the following categories: muon “1top” (a), electron “1top” (b), muon “0top1btag” (c), electron “0top1btag” (d), muon “0top0btag” (e), electron “0top0btag” (f). Each background process is corrected by scale factors from the maximum-likelihood fit to data as described in Chap. 8. The signal MC samples are normalized to a cross section of 1 pb. The shaded band shows the error associated with the background expectation and includes both the MC statistical uncertainty and the post-fit systematic uncertainty described in Chap. 8.</p>	111
<p>B.15 Data/MC comparison for the leptonic-top p_T after all analysis cuts are applied. Events are split into the following categories: muon “1top” (a), electron “1top” (b), muon “0top1btag” (c), electron “0top1btag” (d), muon “0top0btag” (e), electron “0top0btag” (f). Each background process is corrected by scale factors from the maximum-likelihood fit to data as described in Chap. 8. The signal MC samples are normalized to a cross section of 1 pb. The shaded band shows the error associated with the background expectation and includes both the MC statistical uncertainty and the post-fit systematic uncertainty described in Chap. 8.</p>	112

LIST OF FIGURES (continued)

<p>B.16 Data/MC comparison for the leptonic-top mass after all analysis cuts are applied. Events are split into the following categories: muon “1top” (a), electron “1top” (b), muon “0top1btag” (c), electron “0top1btag” (d), muon “0top0btag” (e), electron “0top0btag” (f). Each background process is corrected by scale factors from the maximum-likelihood fit to data as described in Chap. 8. The signal MC samples are normalized to a cross section of 1 pb. The shaded band shows the error associated with the background expectation and includes both the MC statistical uncertainty and the post-fit systematic uncertainty described in Chap. 8.</p>	113
<p>B.17 Data/MC comparison for the p_T of the $t\bar{t}$ pair after all analysis cuts are applied. Events are split into the following categories: muon “1top” (a), electron “1top” (b), muon “0top1btag” (c), electron “0top1btag” (d), muon “0top0btag” (e), electron “0top0btag” (f). Each background process is corrected by scale factors from the maximum-likelihood fit to data as described in Chap. 8. The signal MC samples are normalized to a cross section of 1 pb. The shaded band shows the error associated with the background expectation and includes both the MC statistical uncertainty and the post-fit systematic uncertainty described in Chap. 8.</p>	114
<p>B.18 Data/MC comparison for the χ^2 of the reconstructed $t\bar{t}$ pair after all analysis cuts are applied. Events are split into the following categories: muon “1top” (a), electron “1top” (b), muon “0top1btag” (c), electron “0top1btag” (d), muon “0top0btag” (e), electron “0top0btag” (f). Each background process is corrected by scale factors from the maximum-likelihood fit to data as described in Chap. 8. The signal MC samples are normalized to a cross section of 1 pb. The shaded band shows the error associated with the background expectation and includes both the MC statistical uncertainty and the post-fit systematic uncertainty described in Chap. 8.</p>	115
<p>C.1 $\pm 1\sigma$ systematic variation on the $M_{t\bar{t}}$ distribution due to the Jet Energy Scale: (left) total background, (right) narrow Z' signal with $M_{Z'} = 2$ TeV. Plots are shown for the three final categories in the muon channel: (from top to bottom) “1top”, “0top1btag”, and “0top0btag”. The nominal distribution (black line) is plotted with its statistical uncertainty, the $\pm 1\sigma$ templates account for differences in both shape and normalization with respect to the nominal sample.</p>	117
<p>C.2 $\pm 1\sigma$ systematic variation on the $M_{t\bar{t}}$ distribution due to the Jet Energy Resolution: (left) total background, (right) narrow Z' signal with $M_{Z'} = 2$ TeV. Plots are shown for the three final categories in the muon channel: (from top to bottom) “1top”, “0top1btag”, and “0top0btag”. The nominal distribution (black line) is plotted with its statistical uncertainty, the $\pm 1\sigma$ templates account for differences in both shape and normalization with respect to the nominal sample.</p>	118

LIST OF FIGURES (continued)

C.3	$\pm 1\sigma$ systematic variation on the $M_{t\bar{t}}$ distribution due to the b -tagging SF: (left) total background, (right) narrow Z' signal with $M_{Z'} = 2$ TeV. Plots are shown for the three final categories in the muon channel: (from top to bottom) “1top”, “0top1btag”, and “0top0btag”. The nominal distribution (black line) is plotted with its statistical uncertainty, the $\pm 1\sigma$ templates account for differences in both shape and normalization with respect to the nominal sample.	119
C.4	$\pm 1\sigma$ systematic variation on the $M_{t\bar{t}}$ distribution due to the b -mistag SF: (left) total background, (right) narrow Z' signal with $M_{Z'} = 2$ TeV. Plots are shown for the three final categories in the muon channel: (from top to bottom) “1top”, “0top1btag”, and “0top0btag”. The nominal distribution (black line) is plotted with its statistical uncertainty, the $\pm 1\sigma$ templates account for differences in both shape and normalization with respect to the nominal sample.	120
C.5	Effect on the $M_{t\bar{t}}$ distribution due to a 20% systematic variation of the top-tagging SF: (left) total background, (right) narrow Z' signal with $M_{Z'} = 2$ TeV. Plots are shown for the three final categories in the muon channel: (from top to bottom) “1top”, “0top1btag”, and “0top0btag”. The nominal distribution (black line) is plotted with its statistical uncertainty, the $\pm 1\sigma$ templates account for differences in both shape and normalization with respect to the nominal sample.	121
C.6	$\pm 1\sigma$ systematic variation on the $M_{t\bar{t}}$ distribution due to the top-mistag SF: (left) total background, (right) narrow Z' signal with $M_{Z'} = 2$ TeV. Plots are shown for the three final categories in the muon channel: (from top to bottom) “1top”, “0top1btag”, and “0top0btag”. The nominal distribution (black line) is plotted with its statistical uncertainty, the $\pm 1\sigma$ templates account for differences in both shape and normalization with respect to the nominal sample.	122
C.7	$\pm 1\sigma$ systematic variation on the $M_{t\bar{t}}$ distribution due to the top-mistag SF: (left) SM $t\bar{t}$, (right) SM $W + \text{jets}$. Plots are shown for the three final categories in the muon channel: (from top to bottom) “1top”, “0top1btag”, and “0top0btag”. The nominal distribution (black line) is plotted with its statistical uncertainty, the $\pm 1\sigma$ templates account for differences in both shape and normalization with respect to the nominal sample.	123
C.8	$\pm 1\sigma$ systematic variation on the $M_{t\bar{t}}$ distribution due to the Parton Distribution Functions: (left) total background, (right) narrow Z' signal with $M_{Z'} = 2$ TeV. Plots are shown for the three final categories in the muon channel: (from top to bottom) “1top”, “0top1btag”, and “0top0btag”. The nominal distribution (black line) is plotted with its statistical uncertainty, the $\pm 1\sigma$ templates account for differences in both shape and normalization with respect to the nominal sample.	124

LIST OF FIGURES (continued)

C.9	$\pm 1\sigma$ systematic variation on the $M_{t\bar{t}}$ distribution for the $t\bar{t}$ background, due to the choice of the renormalization scale. Plots are shown for the three final categories in the muon channel: (from top to bottom) “1top”, “0top1btag”, and “0top0btag”. The nominal distribution (black line) is plotted with its statistical uncertainty, the $\pm 1\sigma$ templates account for differences in both shape and normalization with respect to the nominal sample.	125
C.10	$\pm 1\sigma$ systematic variation on the $M_{t\bar{t}}$ distribution for the $W + \text{jets}$ background, due to the choice of (left) the matching scale and (right) the renormalization scale. Plots are shown for the three final categories in the muon channel: (from top to bottom) “1top”, “0top1btag”, and “0top0btag”. The nominal distribution (black line) is plotted with its statistical uncertainty, the $\pm 1\sigma$ templates account for differences in both shape and normalization with respect to the nominal sample.	126

LIST OF ABBREVIATIONS

AK	Anti- κ_T
BSM	Beyond the Standard Model
BDT	Boosted Decision Tree
CA	Cambridge-Aacheen
CERN	European Organization for Nuclear Research
CL	Confidence Level
CMS	Compact Muon Solenoid
CSC	Cathode Strip Chamber
CSV	Combined Secondary Vertex
CSVL	Combined Secondary Vertex, Loose Working Point
CSVM	Combined Secondary Vertex, Medium Working Point
CSVT	Combined Secondary Vertex, Tight Working Point
DT	Drift Tube
ECAL	Electromagnetic Calorimeter
\cancel{E}_T	Missing Transverse Momentum
FSR	Final State Radiation
GSF	Gaussian-Sum Filter
HB	Hadronic Calorimeter Barrel
HCAL	Hadronic Calorimeter
HE	Hadronic Calorimeter Endcap
HLT	High Level Trigger
HO	Hadronic Calorimeter Outer Barrel
ISR	Initial State Radiation
JEC	Jet Energy Correction
JER	Jet Energy Resolution
KK	Kaluza-Klein

LIST OF ABBREVIATIONS (continued)

L1	Level-1 Trigger
LHC	Large Hadron Collider
LO	Leading Order
MVA	Multivariate Analysis
NLO	Next-to Leading Order
NNLO	Next-to-next-to Leading Order
PDF	Parton Distribution Function
PF	Particle Flow
pp	proton-proton
PS	Proton Synchrotron
PSB	Proton Synchrotron Booster
p_T	Transverse Momentum
QCD	Quantum Chromodynamics
ROC	Read Out Chip
RPC	Resistive Plate Chamber
SC	Supercluster
SM	Standard Model
SPS	Super Proton Synchrotron
TEC	Tracker End Cap
TIB	Tracker Inner Barrel
TID	Tracker Inner Disk
TOB	Tracker Outer Barrel

SUMMARY

In this thesis we present a search for boosted top-antitop quark pairs, consistent with heavy resonance decay, produced in $\sqrt{s}=8$ TeV proton-proton collisions at the Large Hadron Collider recorded by the Compact Muon Solenoid Experiment. Data samples corresponding to 19.7 fb^{-1} of integrated luminosity were analyzed by selecting events containing one electron or muon and at least two high transverse momentum jets consistent with the semileptonic decay of the top-antitop quark pair. The highly boosted topology of heavy resonance decay into top-antitop quark pairs requires a dedicated event selection, including the use of new top tagging algorithms to select events with boosted hadronic top quark decays by studying the jet substructure. The invariant mass of the top-antitop quark pair is reconstructed using a χ^2 approach, and we look for excess above the Standard Model background predictions for evidence of undiscovered new heavy resonances. No such evidence is found, and we use a Bayesian statistical analysis to set model-independent 95% Confidence Level upper limits on the production cross-section times branching ratio for narrow 1% width and wide 10% width resonances. In addition, we place limits on two benchmark models that predict top-antitop quark resonant production including a leptophobic Topcolor Z' and a Kaluza-Klein excitation of a gluon in a Randall-Sundrum model. We then compare these limits to the searches for resonant top-antitop quark pair production done using the fully-leptonic and all-hadronic decay modes of the top-antitop quark pair and present a combined result where all decay channels are used to produce the strongest limits on resonant top-antitop quark pair production to date.

1. THEORETICAL INTRODUCTION

1.1 The Standard Model

The Standard Model (SM) of particle physics was developed throughout the latter half of the 20th century and describes all the subatomic particles known to exist and their interactions with each other. The theory describes three of the four known fundamental forces: electromagnetic, weak, and strong nuclear interactions. The current formulation was finalized when the existence of quarks was experimentally confirmed in the mid-1970s. Since then a large number of experiments have tested and confirmed the predictions of the SM, including most recently the discovery of the Higgs boson in 2012 at the Large Hadron Collider [1,2], the tau neutrino at the DONUT experiment [3], and the top quark at the Tevatron in 1995 [4,5]. Figure 1.1 shows the complete set of particles and their interactions within the SM.

Quantum field theory, in which particles are treated as excited states of an underlying physical field, provides the mathematical framework for the SM. Field theories are usually constructed by finding the most general renormalizable Lagrangian that is invariant under the symmetries of a system. Symmetries of the Lagrangian are very important since Noether's theorem [7] states that each symmetry has a corresponding conserved quantity. Translational and rotational symmetry combined with inertial reference frame invariance make up what is known as the global Poincaré symmetry, postulated for all relativistic quantum field theories, which corresponds to conservation of energy, momentum, and angular momentum. In addition to the global Poincaré symmetry, the local $SU(3)_C \times SU(2)_L \times U(1)_Y$ internal gauge symmetry defines the SM.

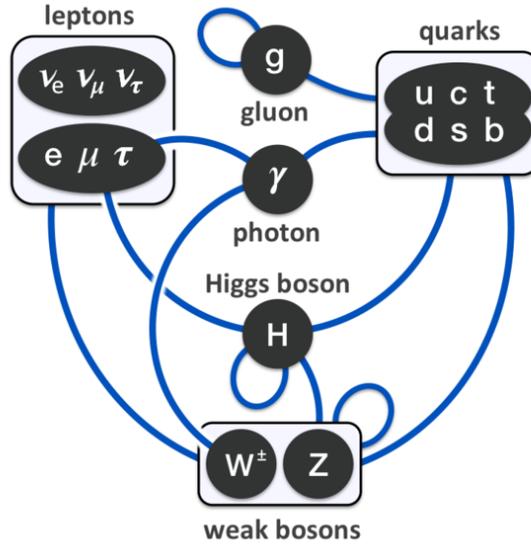


Figure 1.1: Diagram summarizing the interactions between elementary particles according to the Standard Model. Dark ovals represent types of particles and blue arcs represent interactions between them. An arc that links to a box is equivalent to a set of arcs that link every oval in the box. [6]

The quantum chromodynamics (QCD) sector, governed by the $SU(3)_C$ symmetry, defines the interactions between gluons and quarks resulting in the strong force. Quarks and gluons make up hadrons such as the proton, neutron, and pion. The gluon is the force carrier of the theory. Color charge is the conserved quantum number of the theory, and is analogous to standard electromagnetic charge. Only colored particles participate in strong interactions. Three color charges and three anti-color charges exist: red, green, blue, anti-red, anti-green, and anti-blue. Quarks carry color, antiquarks carry anti-color, and gluons carry one color and one anti-color. Leptons do not carry color and therefore do not participate in the strong interaction. The theory is a non-abelian gauge theory and has two important, but peculiar properties. The first peculiar property of QCD is known as asymptotic freedom in which the strength of the interaction between quarks and gluons diminishes as the energy of the interaction increases. Asymptotic freedom is important for calculations in QCD as it allows the theory to be treated perturbatively.

The second peculiar property is known as confinement, which means the force between quarks does not diminish as their separation increases. Instead, as two quarks are separated, the energy contained in the gluon field increases until there is enough energy to create another quark-antiquark pair which creates two hadrons where there was one before. Although confinement is not analytically proven, it is widely believed to be true due to the failure to observe free quarks. There exists a rule that states all observable particles must be colorless, meaning they are made up of equal amounts of color and anti-color or have equal amounts of all three colors or anti-colors. This rule makes sense in the context of confinement and gives rise to the possible colorless combinations of $q\bar{q}$ (meson), qqq (baryon), and $\bar{q}\bar{q}\bar{q}$ (antibaryon). A fourth category of $q\bar{q}q\bar{q}$ is theoretically possible but hasn't been observed experimentally yet.

The $SU(2)_L \times U(1)_Y$ symmetry generates the electromagnetic and weak forces. It is described by the Glashow-Salam-Wienberg model, which is a unified electroweak theory. In the SM, the Higgs mechanism causes spontaneous breaking of $SU(2)_L \times U(1)_Y$ to $U(1)_{em}$, which generates the W^\pm and Z^0 bosons, and the photon. This spontaneous symmetry breaking only occurs at low energies. The $SU(2)_L$ symmetry gives rise to the conserved quantum number called weak isospin, L , and the $U(1)_Y$ symmetry gives rise to the conserved quantum number called hypercharge, Y . The $U(1)_{em}$ symmetry gives rise to the well-known conserved quantum number known as electric charge, Q . These quantum numbers are connected by the Gell-Mann-Nishijima formula:

$$Q = L_3 + \frac{1}{2}Y$$

where L_3 is the third component of the isospin. The Higgs mechanism also leads to mass terms for the weak gauge bosons and the masses of the quarks and leptons. Figure 1.2 shows a diagram that summarizes the SM in both the high-energy (unbroken) symmetric phase and the low-energy (broken) phase.

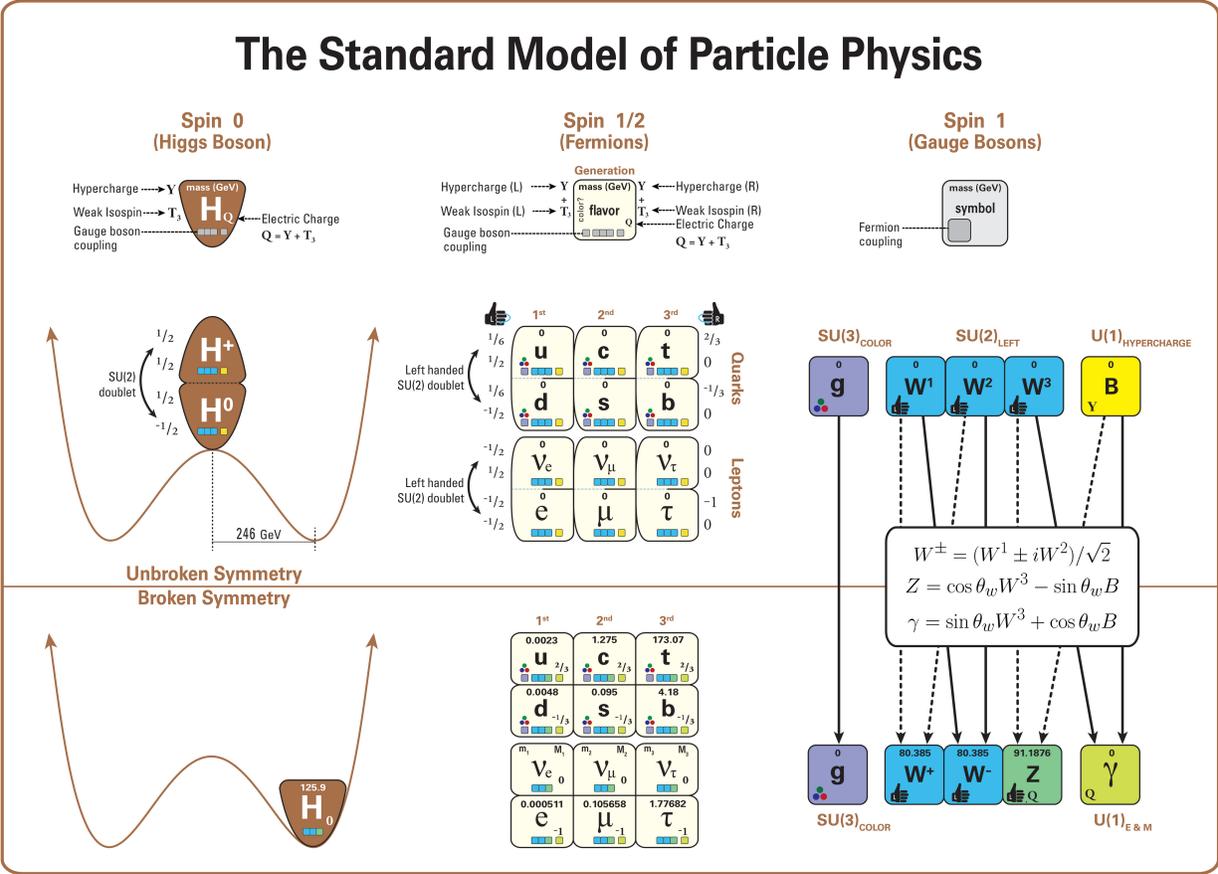


Figure 1.2: Diagram summarizing the Standard Model of particle physics. Shown are the elementary particles of the SM (the Higgs boson, the three generations of quarks and leptons, and the gauge bosons), including their names, masses, spins, charges, chiralities, and interactions with the strong, weak, and electromagnetic forces. The top shows the particle content before electroweak symmetry breaking and the bottom afterwards [8].

1.2 Beyond the Standard Model

Despite the huge successes of the SM, there are a number of inadequacies that motivate research to extend the SM into a Unified field theory of everything. These inadequacies include:

- The SM does not attempt to explain gravitation. The canonical theory of gravitation, general relativity, cannot be described consistently in terms of a quantum field theory.

- The Higgs mechanism gives rise to the hierarchy problem, which questions why the weak force is 10^{32} times stronger than gravity and why the Higgs boson is much lighter than the Planck mass.
- The SM cannot explain the observed amount of cold dark matter observed by cosmologists and gives contributions to dark energy which are many orders of magnitude too large.
- The asymmetry of matter and anti-matter left over from the big bang is thought to be explained by new physics near the electroweak scale.
- The SM requires 19 numerical constants whose values are unrelated and arbitrary, which some consider to be inelegant. An additional 7 or 8 constants are believed to be required to explain neutrino masses, which are massless in the SM.

New physics theories that go beyond the SM (BSM) are proposed that address these unanswered questions.

The top quark plays a special role in BSM theories because it is the heaviest known fundamental particle with a mass close to the electroweak scale. It is closely connected to the hierarchy problem, where the largest corrections to the Higgs boson mass are due to top quark loops, as it has a strong Yukawa coupling to the Higgs potential. Therefore studies of top quark production may provide valuable insight into electroweak symmetry breaking.

Additionally, many BSM theoretical models predict the production of new heavy resonances which preferentially decay to top quark pairs, $t\bar{t}$. Such models include axigluons [9,10], colorons [11–14], extended gauge theories with massive color-singlet Z -like bosons [15–17] and models in which a pseudoscalar Higgs boson may couple strongly to top quarks [18]. Furthermore, various extensions of the Randall-Sundrum model [19] with extra dimensions predict Kaluza-Klein (KK) excitations of gluons g_{KK} [20] or gravitons [21], both of which can have enhanced couplings to $t\bar{t}$ pairs. All of these models predict an additional resonant component to SM $t\bar{t}$ production and a model-independent search for BSM physics can be done by searching for peaks in the $t\bar{t}$ invariant mass spectrum, $M_{t\bar{t}}$.

Analyses at the Tevatron have searched for heavy $t\bar{t}$ resonances up to ~ 900 GeV [22–27] and previous searches at the LHC have set sub-picobarn limits on the production cross-section time branching ratio in the mass range of 1-3 TeV [28–33] in proton-proton collisions with $\sqrt{s} \leq 8$ TeV.

This thesis presents a model-independent search for $Z' \rightarrow t\bar{t} \rightarrow W^+bW^-\bar{b}$ production, where one of the W bosons decays leptonically into a muon or electron plus a neutrino and the other hadronically into $q\bar{q}$. The symbol Z' is used throughout this thesis to denote a generic resonance without reference to any model, unless otherwise noted. The search is based on proton-proton (pp) collision data collected by the Compact Muon Solenoid experiment at $\sqrt{s} = 8$ TeV, corresponding to an integrated luminosity of 19.7fb^{-1} . This analysis is an update to a previous analysis performed on the same dataset which was published in Physical Review Letters in 2013 [33]. This updated analysis leverages a new triggering strategy in the electron channel and the use of boosted hadronically decaying top quark tagging through the analysis of jet substructure. With no significant excess of $t\bar{t}$ production observed, we place 95% CL Bayesian upper limits on the production cross-section time branching ratio for narrow 1% width and wide 10% width resonances. In addition to model-independent limits, we look for evidence of two benchmark models which are widely used in other boosted $t\bar{t}$ resonance searches: Topcolor and Kaluza-Klein excitations of gluons in the Randall-Sundrum RS1 warped extra dimensions model.

1.2.1 Topcolor

Topcolor is a theory that proposes a solution to the question of why the top quark mass is so much larger than the first two generations of quarks. Additionally, Topcolor can be incorporated into supersymmetric theories to solve their flavor changing neutral current problem [34]. Topcolor proposes that the symmetry of the SM is extended to include an additional $SU(3) \times U(1)$ symmetry that is broken. The full symmetry proposed by Topcolor is $SU(3)_1 \times SU(3)_2 \times SU(2)_L \times U(1)_{Y_1} \times U(1)_{Y_2}$ which is broken to $SU(3)_{QCD} \times U(1)_{EM}$ [13]. The

gauge fields associated with $SU(3)_1$ and $U(1)_{Y_1}$ couple preferentially to the third generation of quarks while the gauge fields associated with $SU(3)_2$ and $U(1)_{Y_2}$ couple preferentially to the first and second generations. The symmetry breaking to $SU(3)_{QCD} \times U(1)_{EM}$ yields nine additional massive bosons: eight massive “top-gluons” and one Z' boson. Under Topcolor, the assumption is that these “top-gluons” couple preferentially to third generations quarks and a bosonic top quark condensate is formed by bound top-antitop quark pairs [35]. The top quark condensate generates mass terms analogous to the SM Higgs Mechanism, even for the top quark itself, with the bulk of the top quark mass coming from the top quark condensate while the SM Higgs Mechanism generates the masses of the other two generations. This offers an explanation to why the top quark mass is so vastly different than the other quarks. Since the top and bottom quarks do not both have an enhanced mass due to this mechanism, the condensate of bottom quarks which would also form from the interactions with the “top-gluons” must be broken somehow. The Z' boson is assumed to couple asymmetrically to the top and bottom quarks in such a way that breaks the bottom quark condensate. With the Z' boson coupling strongly to top quarks, its decay into resonant top-antitop pairs promises a means of discovery. In this thesis a leptophobic Z' model [14], where the couplings to leptons are highly suppressed and the coupling to top quarks is greatly enhanced, is used as a benchmark for narrow 1.2% width and wide 10% width $t\bar{t}$ resonances.

1.2.2 Kaluza-Klein excitation of gluons in Randall-Sundrum models

The central idea behind Kaluza-Klein excitations of gluons in extended Randall-Sundrum models is that the four-dimensional universe we live in is actually embedded in a higher-dimensional space as a “brane”. Kaluza-Klein theories attempt to unify Einstein’s general relativity with electromagnetism by extending the four dimensions of space-time to five and re-deriving Einstein’s field equations in this new five-dimensional space. Klein proposed that the fifth dimension is “compactified”, being curled up in a very small circle such that a particle moving along this dimension would return to where it began. The radius of this

circle is said to be the size of the dimension. In a simplistic interpretation of the consequences of this extra compactified dimension, one might expect to have standing waves analogous to the quantum mechanic particle-in-a-box. If the extra dimension has a radius of r , then the invariant mass of such standing waves would be $M_n = nh/rc$. The set of possible mass values for these resonances is called the Kaluza-Klein tower. In the Randall-Sundrum RS1 model [19], the higher-dimensional universe is a five-dimensional anti-de Sitter space with warped geometry described by the metric:

$$ds^2 = \frac{1}{k^2 y^2} (dy^2 + \eta_{\mu\nu} dx^\mu dx^\nu)$$

where k is a constant, $\eta_{\mu\nu}$ has the “-+++” metric signature, dx is the standard four-dimensional metric, and y is the coordinate of the extra dimension. The boundaries at $y = 1/k$ and $y = 1/Wk$, with k on the order of the Planck scale and Wk around the TeV scale, form two branes that bound the extra dimension. The brane at $y = 1/k$ is said to have positive brane energy and is called the Planck brane while the brane located at $y = 1/Wk$ is called the TeV brane and has negative brane energy. All SM particles are strongly localized on the TeV brane. This model has two possible explanations for the weakness of gravity. With the warping of space-time only occurring along the fifth dimension, the proposed graviton mediator particle would be strongly localized at the Planck brane and therefore gravity would be much weaker on the TeV brane. Also, objects moving from the Planck brane to the TeV brane through the fifth dimension would be growing, becoming lighter, and moving slower through time. The details of these calculations are beyond the scope of this thesis, but an important consequence of them is that excited Kaluza-Klein states are allowed to decay into SM particles. In particular, the first excited state of the gluon, g_{KK} , is localized near the TeV brane and has an enhanced coupling to SM top quarks which would manifest as resonant $t\bar{t}$ production over the SM prediction, making it an ideal search candidate for verification of the RS1 model [20].

2. THE LARGE HADRON COLLIDER AND COMPACT MUON SOLENOID

The Large Hadron Collider (LHC), located at and operated by the European Organization for Nuclear Research (CERN), is the largest and most complex experimental facility that has been built thus far. It is also the largest and most powerful hadron collider in the world with a circumference of 27 km and a design beam energy of 7 TeV. Located under the Franco-Swiss border near Geneva, Switzerland, the LHC produces counter-rotating beams of protons that collide with each other at several interaction points that house different experiments. Currently, six different experimental detectors share four interaction points along the LHC beam: ATLAS, CMS, LHCb, ALICE, TOTEM, and LHCf.

The motivation of the LHC is to study the predictions of the SM and shed light on various alternatives and extensions to it. New discoveries, such as supersymmetry or extra dimensions, could pave the way toward a unified theory of particle physics and expand our knowledge of the universe. These discoveries are thought to manifest at high energies in the TeV scale, and as such hadron colliders are well suited to study these new energy domains. Most previous hadron colliding accelerators, such as the Tevatron, used proton-antiproton collisions which provided a higher reaction rate for some processes. However, this advantage over proton-proton collisions decreases as the collision energy increases toward the energy regime of the LHC. In addition, the rate at which antiprotons can be created puts a severe limitation on the luminosity of a proton-antiproton collider. Therefore, the LHC was designed to have an energy seven times higher than previous hadron colliding accelerators.

This opens the door to studying a wide range of physics, but requires a very careful design of the accelerator complex and the detectors.

2.1 The Large Hadron Collider

The LHC was built from 1998 to 2008 with the help of thousands of scientists and engineers from hundreds of countries, universities and laboratories. The first recorded collisions took place on November 23, 2009 at a center of mass energy of 900 GeV. Since then, both the energy and intensity of the collisions have increased with 8 TeV center of mass energy collisions taking place during 2012. Data recorded by the various experiments has been produced at unprecedented rates and is analyzed by a grid-based computer network called the LHC Computing grid, which by 2012 was the world's largest computing grid.

Figure 2.1 shows the overall layout of the LHC accelerator complex. The main operation mode of the LHC is proton-proton collisions, but it can also accommodate heavy ion collisions such as lead-lead and proton-lead. For proton-proton collisions everything starts with a small bottle of hydrogen being fed into a Duoplasmatron which strips the hydrogen of its electrons producing protons that feed into the first accelerator, Linac2. Linac2 is a linear accelerator that boosts protons to 50 MeV before injecting them into the Proton Synchrotron Booster (PSB), which further boosts them to 1.4 GeV. Following the PSB, they are injected into the Proton Synchrotron (PS) which boosts them to 25 GeV. Next, the Super Proton Synchrotron (SPS) boosts them to 450 GeV before they are transferred into the two beam pipes of the LHC, with one beam rotating clockwise and the other counterclockwise. The LHC accelerates the beams to their final energy. Figure 2.2 shows a cross-section of one of the 1,232 dipole magnets that are used to keep the beams on their circular path showing this 2-in-1 beam structure. In addition, 392 quadrupole magnets are used to keep the beam focused.

The proton beams are designed to have a bunch structure, rather than continuous beams, with up to 2,808 bunches of up to 115 billion protons in each bunch. The bunch structure ensures that interactions between the two beams occur at discrete intervals, which immensely

CERN's Accelerator Complex

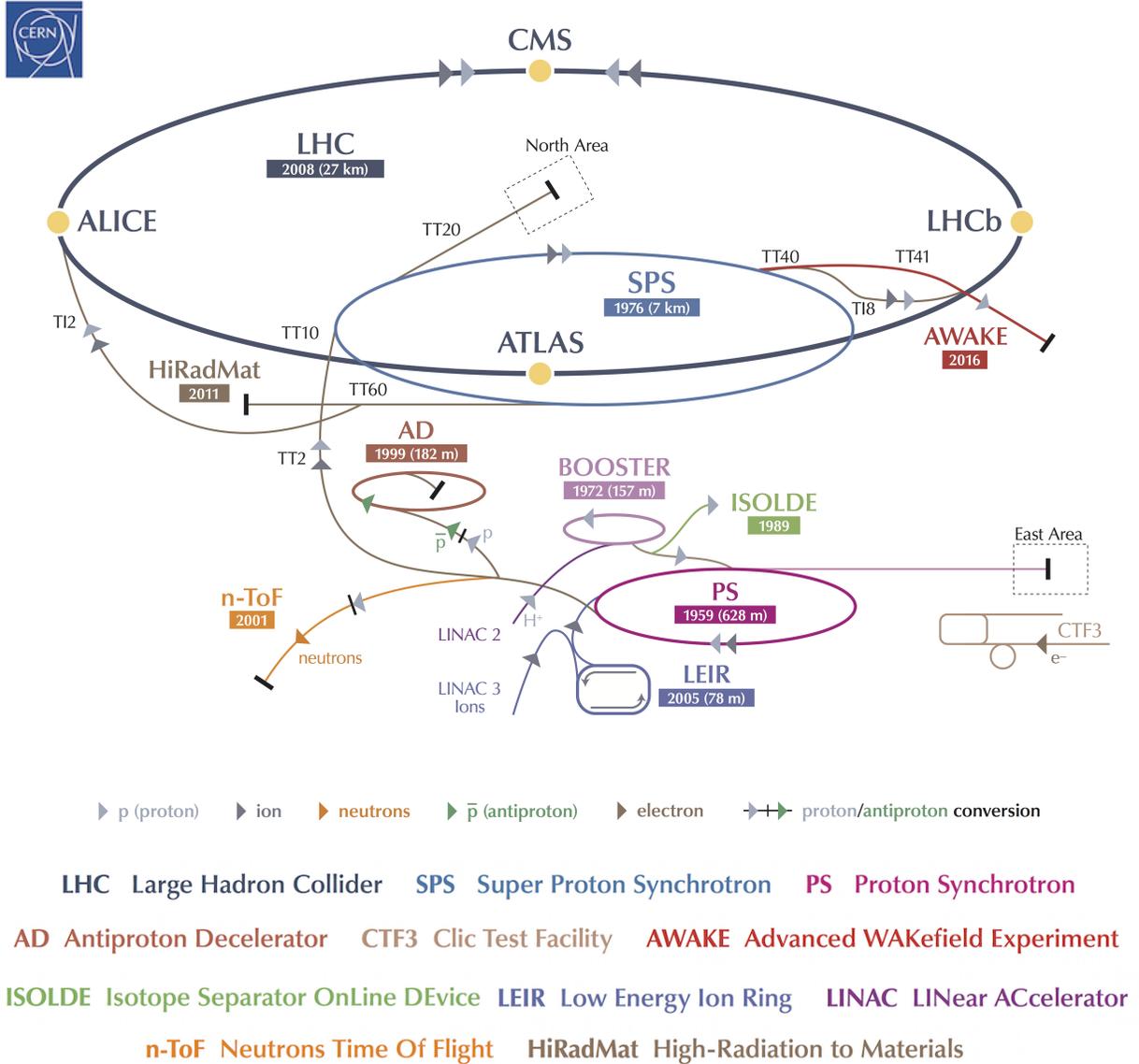


Figure 2.1: Layout of the accelerator complex at CERN [36].

LHC DIPOLE : STANDARD CROSS-SECTION

CERN AC/DI/MM - HE107 - 30 04 1999

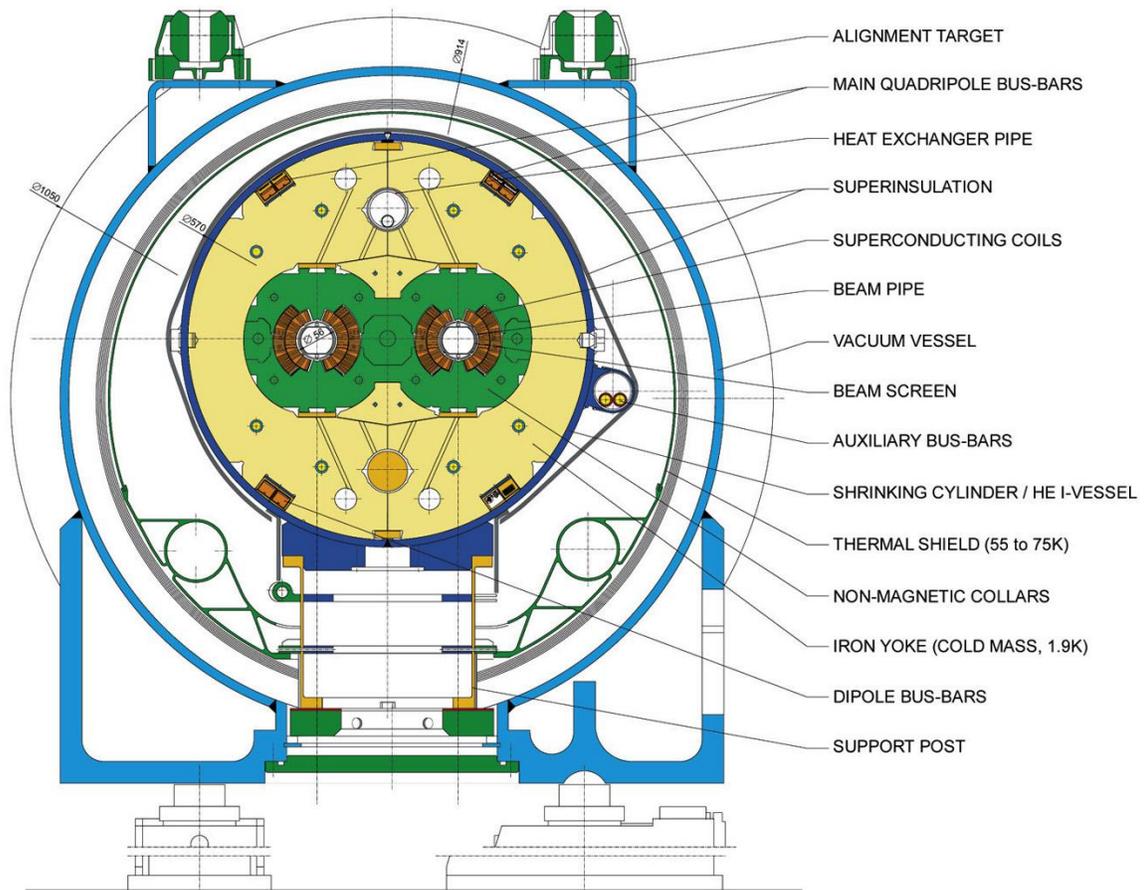


Figure 2.2: Cross-section of a dipole magnet used in the LHC. The 2-in-1 structure of the beam pipe which contains counter-rotating beams can be seen [37].

simplifies detector design. During the 2012 data taking run only 1,380 bunches were used which corresponds to a bunch crossing interval of 50 ns, twice that of the design limit of 25 ns. Even with the very large number of protons per bunch, the probability of a hard proton-proton collision is very small due to their small size. Every effort is used to squeeze the bunches into a very small transverse size before the interaction points in order to maximize the probability of a collision. This probability is also known as the instantaneous luminosity, \mathcal{L} , and is determined by the bunch revolution frequency, f , the number of bunches, b , the number of protons per bunch, $n_{1,2}$, and the transverse size of the bunches at the interaction point, $\sigma_{x,y}$. The LHC has a design luminosity of $10^{34}\text{cm}^{-2}\text{s}^{-1}$. Equation 2.1 shows how to calculate the instantaneous luminosity.

$$\mathcal{L} = f \cdot b \cdot \frac{n_1 n_2}{4\pi\sigma_x\sigma_y} \quad (2.1)$$

2.2 Compact Muon Solenoid Experiment

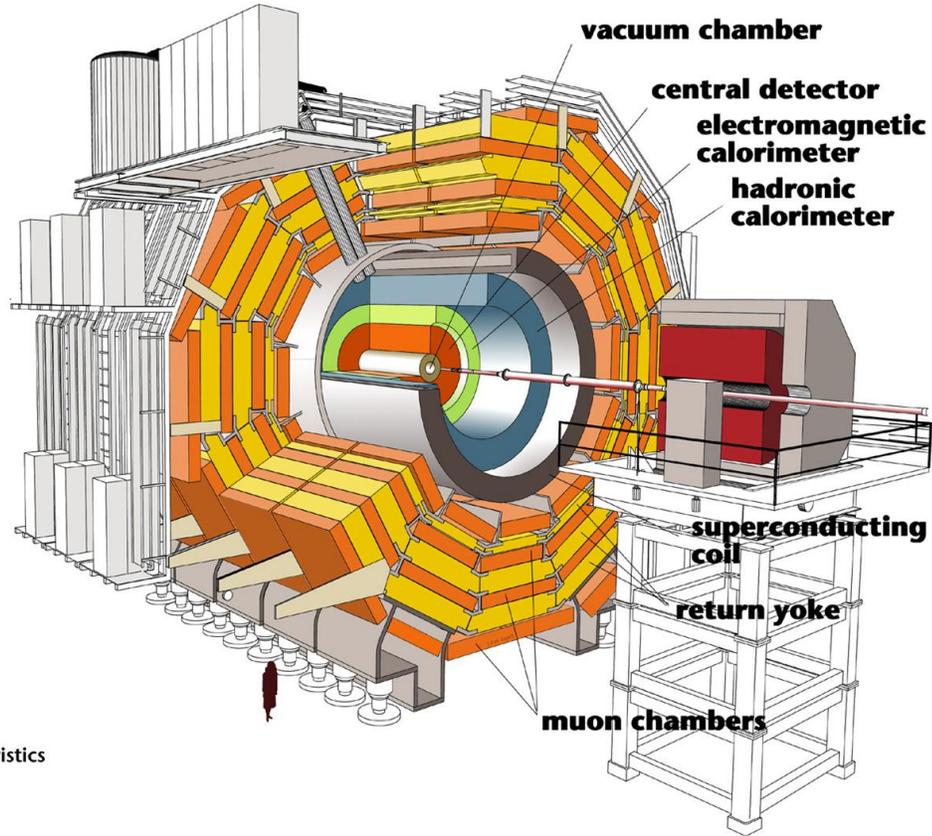
The Compact Muon Solenoid (CMS) detector is a general-purpose particle physics detector built on the LHC and is extensively described in Ref. [38]. A concise description of the most relevant aspects of the detector is given in this section. The goals of the CMS experiment are to explore physics at the TeV scale, to look for evidence of BSM physics, to study heavy ion collisions, and to study the properties of the recently discovered Higgs boson, among others. The CMS detector is built around a superconducting solenoid magnet that generates a magnetic field of 4 Tesla. It consists of several hermetic subdetectors in a “shell” layout as seen in Fig. 2.3. The detector requirements needed to satisfy the physics goals of the experiment can be summarized as:

- Good muon identification and momentum resolution, including good dimuon mass resolution ($\approx 1\%$ at 100 GeV) and the ability to determine muon charge, even for very high energy muons that have very low curvature tracks.

- Good charged-particle momentum resolution and reconstruction efficiency, including efficient triggering and offline tagging of heavy flavor decays.
- Good electromagnetic energy resolution, including diphoton and dielectron mass resolution.
- Good missing-transverse-energy and dijet-mass resolution.

The CMS detector has a total length of 22 m and a radius of 7.5 m. The central section of the detector is referred to as the barrel and the ends as the endcaps. CMS uses a Cartesian coordinate system with the origin at the center of the detector near the interaction point. Because of the cylindrical symmetry around the beam axis, it is chosen as the z -axis such that conversion to cylindrical coordinates is straightforward. With the x -axis pointing toward the center of the LHC ring and the y -axis points upwards towards the surface of the earth, the cylindrical coordinates of (r, ϕ, θ) are defined in the usual sense. The distance from the beam axis is r , the azimuthal angle with respect to the x -axis is ϕ , and the polar angle with respect to the z -axis is θ . The rapidity of high energy particles, $y = \frac{1}{2} \ln\left(\frac{E+p_z}{E-p_z}\right)$, is a useful Lorentz invariant under z -axis boosts. However, the quantity requires energy and momentum measurements and CMS usually substitutes a good approximation called pseudorapidity, $\eta = -\ln \tan \frac{\theta}{2}$, defined only using the polar angle θ . The spatial separation of two particles is usually denoted by $\Delta R = \sqrt{(\Delta\phi)^2 + (\Delta\eta)^2}$.

At the heart of CMS is a 13 m long, 6 m inner diameter superconducting solenoid magnet. The magnet provides a 4 Tesla axial magnetic field which produces a large bending power on charged particles. This bending power allows momentum measurements on charged particles by looking at the radius of curvature of the particle track. The bore of the magnet is large enough to accommodate the inner tracker and calorimetry. Fig. 2.4 shows how various types of particles traverse and interact with the various subdetectors. A detailed description of each layer of subdetectors is given in the next sections.



Detector characteristics

Width: 22m
Diameter: 15m
Weight: 14'500t

Figure 2.3: Cutaway view of the CMS detector showing the outer four layers for muon detection interleaved with the iron return yoke of the magnet, the central calorimeters, and the central detector consisting of the tracking system. [39]

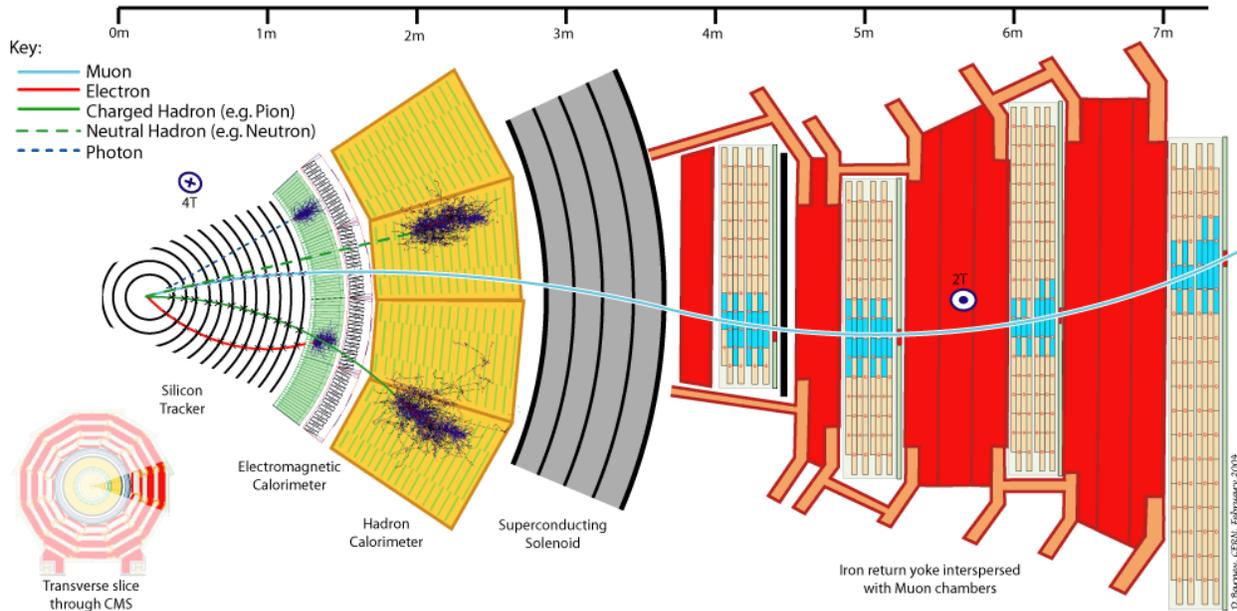


Figure 2.4: View of a transverse slice through CMS showing how different particles interact with the various subdetectors. [40]

2.2.1 Layer 1: The Tracker

The innermost layer of the CMS detector is the Tracker. The Tracker’s main purpose is to reconstruct charged particle tracks to a high precision in order to aide in the reconstruction of a collision event. The tracking volume is a 5.8 m long, 2.6 m diameter cylinder consisting of an all silicon detector. CMS uses 10 layers of silicon microstrip detectors and 3 layers of silicon pixel detectors in the barrel to deal with high track multiplicities. In total, 1,440 pixel modules and over 15,000 strip detector modules are used with a total active area of $\approx 200 \text{ m}^2$. This makes it the largest silicon tracker ever built. Figure 2.5 shows the layout of the Tracker.

The Tracker provides efficient and precise charged particle trajectory reconstruction in a robust manner. This allows the reconstruction of vertices, including secondary vertices from heavy flavor decays, and momentum measurements based on track curvature. A nominal momentum resolution of 0.7% for particles with transverse momentum of 1 GeV and 5.0%

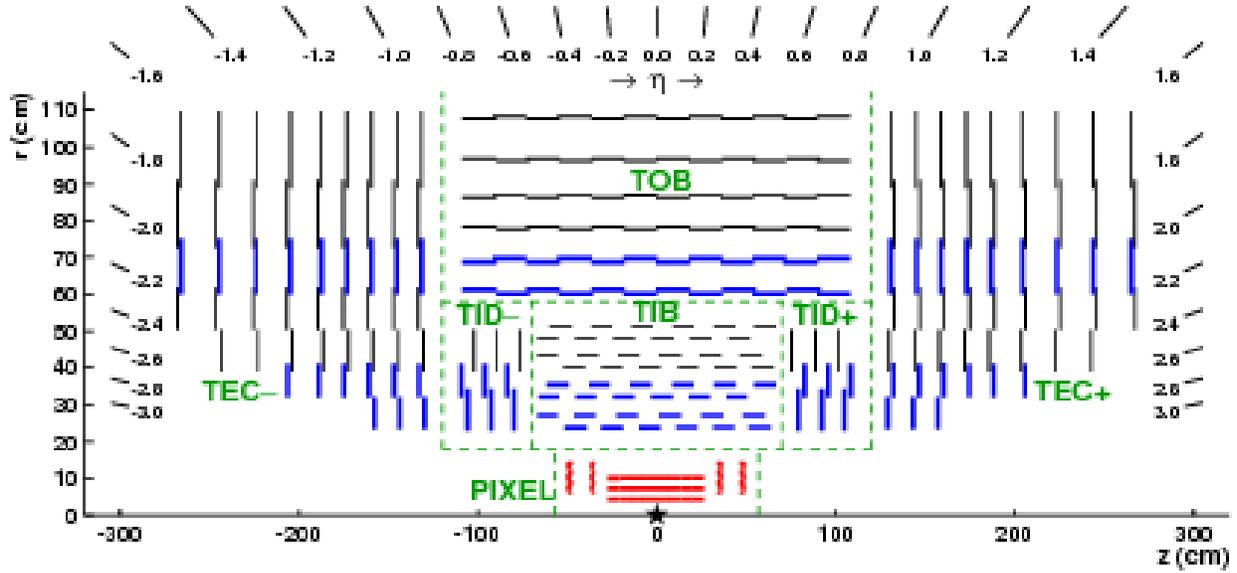


Figure 2.5: Schematic cross-section of the CMS tracker in the r - z plane. Only the top half is shown, the bottom half being symmetric. The interaction point, labeled by a star, is the center of the Tracker. The green dashed lines separate the different parts of the Tracker. Strip tracker modules that only have a single layer are shown by thin black lines. Strip tracker modules consisting of two back-to-back rotated layers, which can provide some 3-D hit reconstruction, are shown by the thick blue lines. The innermost tracker, the pixel detector, is shown in red and provides true 3-D hit reconstruction. Reproduced from Ref. [42].

for particles with transverse momentum of 1000 GeV in the central region can be attained.

In addition the impact parameter resolution is $10 \mu\text{m}$ for high momentum tracks. [41]

Pixel Detector

The pixel detector consists of 65 million pixels arranged in 3 barrel layers and 2 endcap disks. Each pixel is $100 \times 150 \mu\text{m}^2$ and is read out by individual electronics which contain programmable zero-suppression thresholds. The design of the pixel detector is modular, with each module comprising of up to 16 readout chips (ROC), and each ROC consisting of 52×80 pixels. The barrel region uses full modules of 2×8 ROCs and half modules of 1×8 ROCs, while the endcap disks use seven different module arrangement ranging from 1×2 to 2×5 ROCs. Each ROC contains one set of data storage buffers in the periphery of the chip that hold the information from the individual pixels that register a hit until it

receives a signal to read out the data. This periphery is very important for the operation of the pixel detector since there is no possible way to readout 65 million channels for every bunch crossing.

The 3 barrel layers are located at $r = 4.4$ cm, $r = 7.3$ cm, and $r = 10.2$ cm with a length of 53 cm. The endcaps are placed at $z = \pm 34.5$ cm and $z = \pm 46.5$ cm and extend from $r = 6$ cm to $r = 16$ cm. The design is such that any minimally ionizing charged particle with $|\eta| < 2.2$ will leave at least two hits in the pixel detector.

Silicon Microstrip Detector

The silicon microstrip detector consists of single- and double-layer modules that provide either 2-D or 3-D hit measurements. The double-layer modules consist of two strip detectors rotated with respect to each other and are called stereo modules. The stereo modules can provide 3-D hit measurements for low track multiplicity, but become ambiguous when more than one track traverses a module. The pixel detector provides true 3-D hit measurements and is needed to untangle the ambiguity of the hits in the microstrip detector. The Tracker Inner Barrel (TIB), Tracker Inner Disk (TID), and the inner 4 rings of the Tracker End Caps (TEC) are made with thin silicon sensors of $320 \mu\text{m}$ wafer thickness. The Tracker Outer Barrel (TOB) and outer 6 rings of the TEC are made with thicker $500 \mu\text{m}$ sensors. In the TIB, the strip pitch ranges from $80 \mu\text{m}$ to $120 \mu\text{m}$ with the first two layers having stereo modules. The TOB uses strips with a pitch between $120 \mu\text{m}$ and $180 \mu\text{m}$ and has two stereo module layers.

Overall, 15 different sensor geometries are used with typical dimensions of $6 \times 12 \text{ cm}^2$ for the inner barrel and $10 \times 9 \text{ cm}^2$ for the outer barrel. The total number of sensors is 24,244 with an active area of 198 m^2 consisting of 9.3 million strips.

2.3 Layer 2: Electromagnetic Calorimeter

The electromagnetic calorimeter (ECAL) of CMS is a homogeneous calorimeter made of scintillating crystals. Lead tungstate (PbWO_4) is chosen for its high density and optical qualities. It is divided into a central barrel ($|\eta| < 1.4442$) consisting of 61,200 crystals and two endcaps ($1.556 < |\eta| < 3.0$) each with 7,324 crystals. There exists an uninstrumented gap in the region $1.4442 < |\eta| < 1.556$. A preshower detector is placed in front of each endcap which assists in π_0 detection. The high density of lead tungstate means the radiation length, X_0 , is 0.89 cm. The scintillation decay time of the crystals is very fast, with about 80% of the light being emitted within the LHC design bunch crossing of 25 ns. The scintillation light is detected by avalanche photodiodes in the barrel region and vacuum phototriodes in the endcaps. Each crystal in the barrel has dimensions of 22 mm \times 22 mm \times 230 mm and points toward the interaction point as can be seen in Fig. 2.6. The endcaps use crystals with dimensions of 28.6 mm \times 28.6 mm \times 220 mm.

The preshower is a sampling calorimeter that is 20 cm thick and consists of two lead layers and two silicon strip sensor layers. The first layer is a $2X_0$ thick layer of lead followed by a silicon sensing layer. Then a $1X_0$ thick lead layer and another silicon sensing layer follows. This design ensures 90% of incident photons produce electron-positron pairs before reaching the second sensing layer. This improves the correct identification of electrons, positrons, and neutral pions and their position resolution.

The overall energy resolution of the ECAL can be quantified by

$$\left(\frac{\sigma_E}{E}\right)^2 = \left(\frac{S}{\sqrt{E}}\right)^2 + \left(\frac{N}{E}\right)^2 + C^2 \quad (2.2)$$

where the first term corresponds to the stochastic term, the second term corresponds to the noise, and the last term is the constant term, with E in GeV. The value of S and C are determined uniquely by the material choice of the ECAL with the values for lead tungstate being $S = 2.8\%$ and $C = 0.3\%$. The value of N is determined by test beam measurements

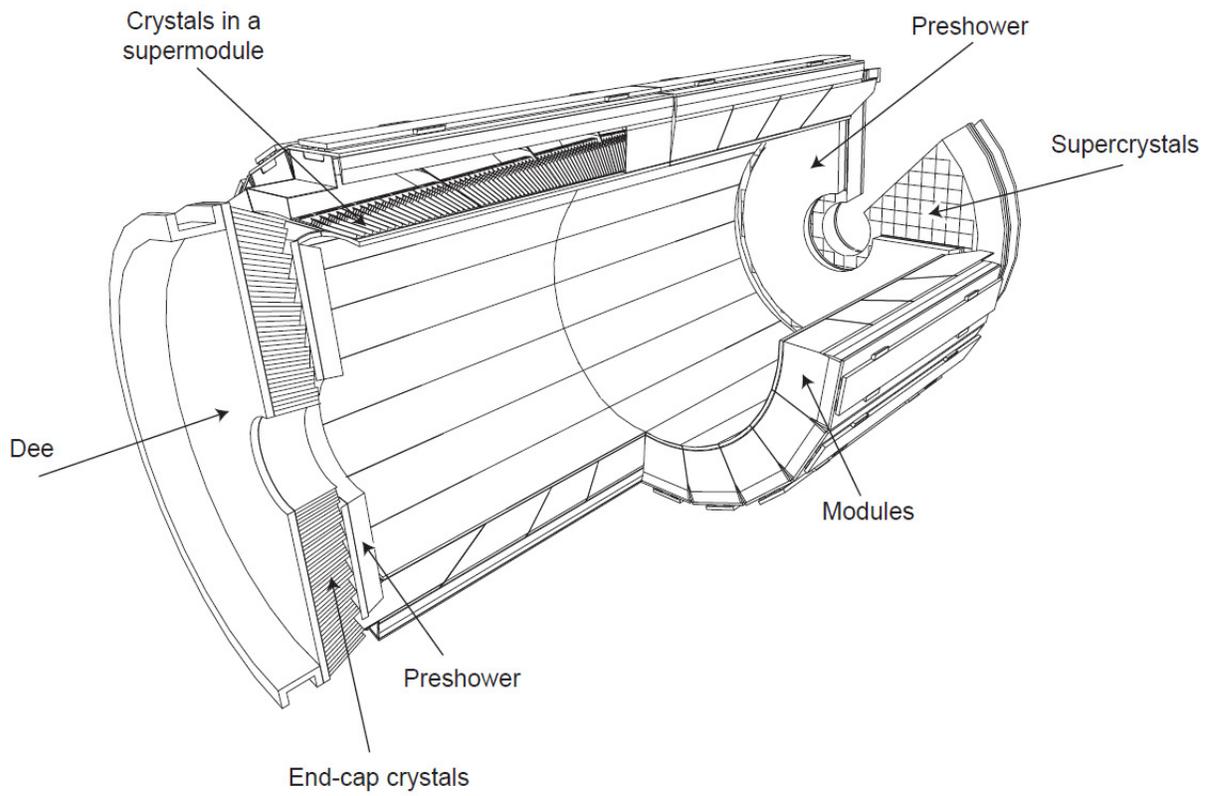


Figure 2.6: Layout of the CMS ECAL showing the arrangement of crystal modules, supermodules, endcaps, and the preshower. Reproduced from Ref. [38]

and is on the order of 12%. The energy of electrons with E_T of about 45 GeV from $Z \rightarrow e^+e^-$ are measured in the central $|\eta| < 0.8$ region with a resolution better than 2%, and is between 2% and 5% elsewhere. This resolution improves to 1.5% for low-bremsstrahlung electrons which deposit most of their energy within a 3×3 array of crystals [43].

2.4 Layer 3: Hadronic Calorimeter

The hadronic calorimeter (HCAL) of CMS is a sampling calorimeter consisting of iron absorbers and scintillating detection layers. In the region $|\eta| < 1.74$, the individual cells have a width of 0.087 in both η and ϕ . In the central region of the HCAL, these larger cells correspond to a 5×5 array of ECAL crystals, which are more finely segmented. Outside the central region, the ECAL arrays corresponding to a single HCAL cell contain fewer crystals. The HCAL and ECAL cells together form calorimeter towers projecting radially outwards from the interaction point, with the total size of the towers increasing in the endcap regions. This segmentation and layout of the HCAL can be seen in Fig. 2.7. Jet energies are measured using a combination of the HCAL and ECAL with a resolution $\Delta E/E \approx 100\%/\sqrt{E[\text{GeV}]} \oplus 5\%$ [43].

2.5 Layer 4: The Muon System

The Muon system of the CMS detector consists of a barrel and endcap system embedded in the iron return yoke of the magnet shown in Fig. 2.8. Three different detector technologies are used: drift tubes (DT), cathode strip chambers (CSC) and resistive plate chambers (RPC). DT technology provides the highest precision measurements on the track position, but are relatively slow. As such they are only used in the barrel region where muon rate is expected to be low, less than 10 particles per second per square cm. The rates in the endcap are higher so faster CSCs are used. RPCs complement both the barrel and the endcap by providing very fast signals, albeit at lower resolutions. Muons are measured in the pseudorapidity range $|\eta| < 2.4$. The p_T resolution in the barrel is better than 10% for

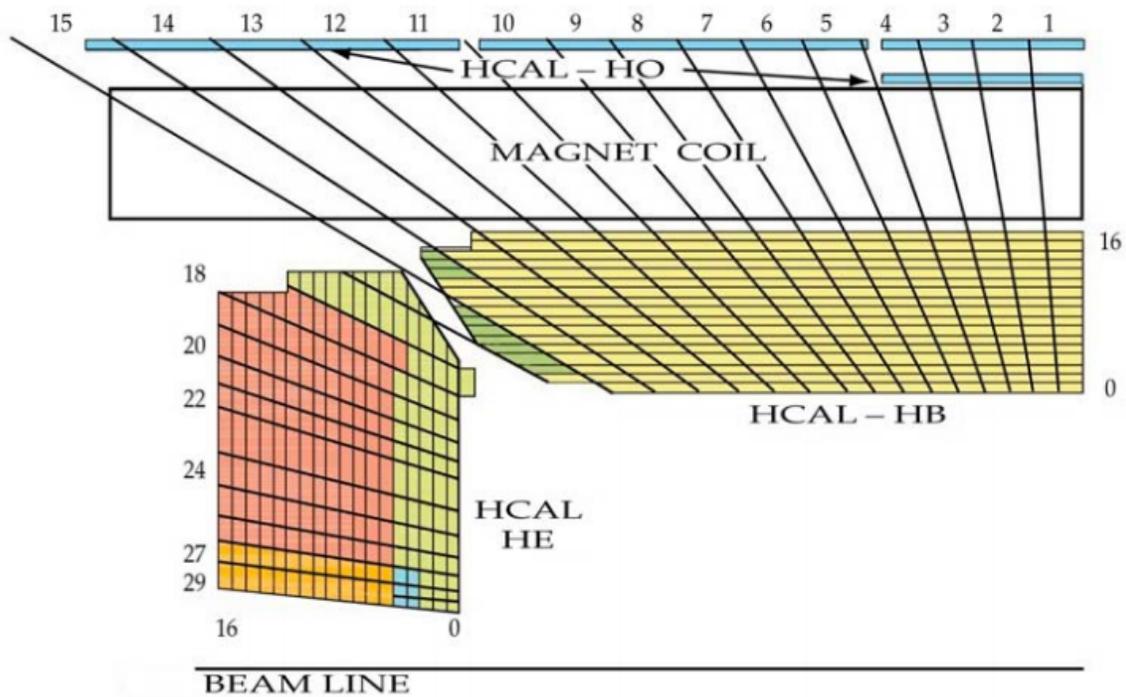


Figure 2.7: Layout of one quarter of the CMS HCAL showing the segmentation in the r - z plane for the barrel (HB), outer barrel (HO), and endcap (HE) detectors. Reproduced from Ref. [38]

muons with p_T up to 1 TeV. Combining tracks measured in the muon system to those in the inner Tracker results in a muon p_T resolution of 1.3-2.0% in the barrel and less than 6% in the endcaps for muons with $20\text{GeV} < p_T < 100\text{ GeV}$ [44].

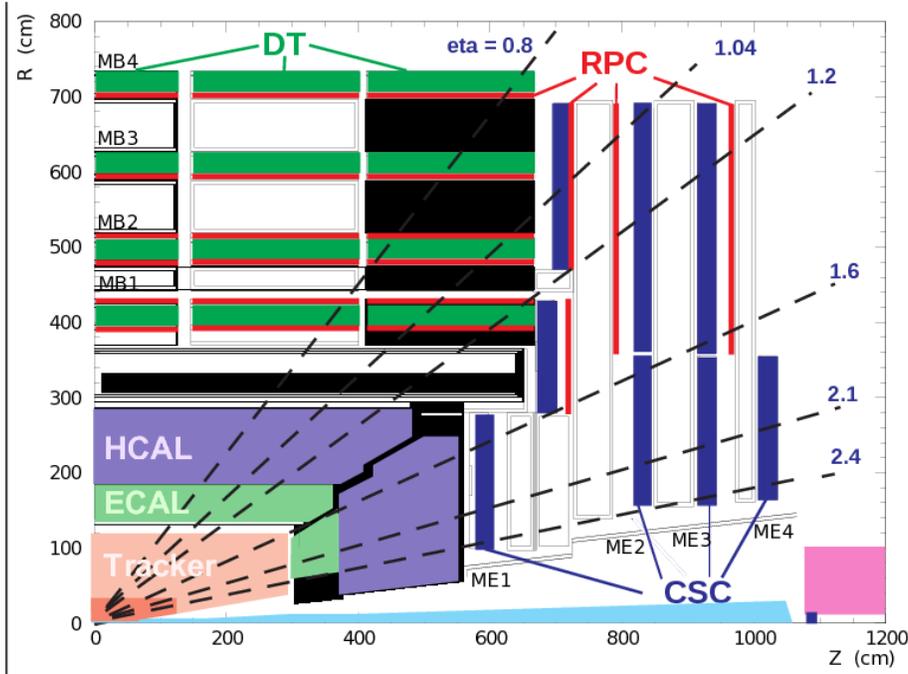


Figure 2.8: Layout of one quadrant of the CMS detector highlighting the muon system. The four barrel drift tube stations (MB1-MB4, green), and the endcap CSC/RPC stations (ME1-ME4, blue/red) are shown. Reproduced from Ref. [45]

2.6 Trigger

CMS employs a multi-level triggering system whose purpose is to reduce event rate to a manageable level for data recording. Current technology and storage limitations require the event recording rate to be on the order of a few hundred events per second. With a bunch crossing of 40MHz, event filtering must be employed to select approximately 2-3 per million bunch crossings to save. This filtering is done in two steps called Level-1 (L1) Trigger and High-Level Trigger (HLT). The L1 Trigger is implemented in hardware using field-programmable gate arrays and other custom hardware and reduces the rate from 40 MHz to ≈ 100 kHz. It uses fast signals from the muon and calorimetry systems to make

decisions while the rest of the event data is pipelined in the various subdetectors waiting for a L1 Trigger to read out. Once the subdetectors receive a L1 Trigger they send their pipelined data out to the data acquisition system to be processed by the HLT. The HLT has access to all the information collected in each subdetector for each event triggered by the L1 Trigger and uses this data to reconstruct the event and make decisions on whether to save the event or not. The HLT is implemented in software and runs on approximately one thousand commercial processors. The HLT selects events to save at a rate of $\approx 100\text{Hz}$ based on various interesting physics phenomena such as high p_T leptons, large momentum imbalances, and very high energy events.

3. OBJECT IDENTIFICATION

CMS uses a particle-flow [46] (PF) event reconstruction to identify all stable particles resulting from a collision. A large number of particles are produced with each collision and each one interacts with the various subdetectors within CMS leaving deposits of energy. The CMS PF algorithm optimally combines the information from the various subdetectors to build a list of stable particles that include muons, electrons, photons, charged hadrons, and neutral hadrons. This information includes charged particle tracks from the tracking system and energy deposits in both the ECAL and HCAL, utilizing the excellent granularity of these systems to disentangle overlapping particles.

The design of CMS makes it ideal for a PF event reconstruction. The Tracker measures the direction and energy of all charged particles without affecting their energies significantly before all the other subdetectors. Charged particle tracks can be efficiently reconstructed in $|\eta| < 2.6$ with a small fake rate down to a transverse momentum (p_T) of 150 MeV. Muons can be identified from other particles because they are the only particle that can traverse the entire detector and leave tracks in the Muon system. Muons detected by the Muon system alone are called Global Muons. Muons are reconstructed using a combination of the Tracker and the Muon system with a very large efficiency as PF Muons. The ECAL surrounding the tracker measures and reconstructs photons with an excellent energy resolution. The granularity of the ECAL and the high magnetic field allows photons to be distinguished from charged particle energy deposits such as electrons. Electrons are reconstructed by a combination of the track they produce in the Tracker and the energy deposits left in the ECAL. Due to their small mass, electrons tend to radiate Bremsstrahlung photons and lose

energy before being absorbed in the ECAL. Electron tracks are refit using a Gaussian-Sum Filter (GSF) [47] to try to follow their trajectories to the ECAL and accurately measure their energy which includes the energy of the Bremsstrahlung photons. Charged and neutral hadrons deposit most of their energy in the HCAL. The combination information from the ECAL and HCAL allows a hadron energy resolution of $\approx 10\%$ at 100 GeV. This resolution allows energy deposits from neutral hadrons to be separated from energy deposits of charged hadrons by comparing the energy of the calorimetry cells to the energy of the charged hadrons as measured by the Tracker.

The CMS PF algorithm identifies all PF candidates in the event as either electrons, muons, photons, charged hadrons, or neutral hadrons. Since multiple proton-proton collisions can occur in the same bunch crossing the excellent tracking resolution of the Tracker is used to identify the separate proton-proton collision points, or vertices. Vertices are reconstructed by clustering tracks with a deterministic annealing algorithm. [48] The candidate vertices produced by deterministic annealing that have at least two tracks assigned to them are fed into an adaptive vertex fitter [49] which determines the best estimate of the vertex position parameters. In each vertex fit the tracks are assigned a weight between zero and one, w_i , that reflects the likelihood that track truly belongs to the vertex. The weights are close to one for tracks coinciding with the vertex and become very small as tracks are more than a few standard deviations away. Each vertex must satisfy a minimal set of quality criteria including a cut on the number of degrees of freedom in the adaptive vertex fit, $n_{\text{dof}} \geq 4$, where n_{dof} is defined as:

$$n_{\text{dof}} = -3 + 2 \sum_{i=1}^{\text{\#tracks}} w_i. \quad (3.1)$$

Each vertex candidate must also satisfy $\sqrt{x^2 + y^2} < 2$ cm, $|z| < 24$ cm. The primary vertex of an event is defined as the vertex with the highest $\sum_{\text{tracks}} p_T^2$. The efficiency for primary vertex reconstruction is estimated to be nearly 100% when the number of tracks used to reconstruct it is greater than two. [42] All charged hadrons not associated to the primary

vertex in the event are labeled “pileup” and ignored in the subsequent steps. PF candidates not identified as isolated leptons or pileup are clustered into jets.

Specifically, the definitions for various physics objects used in this analysis are as follows:

- Muon candidates used in this analysis must satisfy the following requirements:
 - candidate reconstructed both as a Global Muon from the Muon system and as a PF Muon;
 - normalized χ^2 of the Global Muon track fit $\chi^2/N_{\text{DOF}} < 10$;
 - the Global Muon track fit must include a hit in at least 1 muon chamber;
 - at least 2 muon stations have segments matched to the Global Muon;
 - longitudinal distance of muon inner track with respect to the primary vertex $|\Delta z| < 0.5$ cm;
 - transverse impact parameter of muon inner track with respect to the primary vertex $|d_{xy}| < 0.2$ cm;
 - at least 1 pixel hit found for the inner track;
 - more than 5 tracker layers with hits

No isolation requirements are applied to the muon candidates as muons from the decay of high- p_T top quarks tend to be close to the b-jet and not well isolated.

- Electron candidates are identified by several discriminating variables in a boosted decision tree (BDT) multivariate analysis (MVA) trained on electron candidates that fire an electron trigger. These variables include observables from the Tracker, the calorimeters, and the comparison of Tracker and ECAL measurements. The Tracker improves the separation of electrons and charged hadrons by comparing the track fit with the GSF and without. Since electrons are expected to deposit their energy in several ECAL cells, called a supercluster (SC), the calorimeters can distinguish electrons based on transverse shape of electromagnetic showers. The fraction of energy deposited in the HCAL, which is expected to be small for electrons, can also be used. Comparing measurements between the Tracker and the ECAL, such as the track momentum and ECAL

energy, also provides discriminating power to find electrons. Electron candidates used in this analysis must satisfy the following MVA-based identification criteria [50]:

- pass the photon conversion veto;
- number of tracker layers before first hit belonging to the track (missing hits) = 0;
- lower cut on the electron Triggering-MVA discriminator (mvaTrigV0):
 - * mvaTrigV0 > 0.94 if $|\eta_{\text{SC}}| < 0.8$;
 - * mvaTrigV0 > 0.85 if $0.8 < |\eta_{\text{SC}}| < 1.479$;
 - * mvaTrigV0 > 0.92 if $1.479 < |\eta_{\text{SC}}| < 2.5$

where η_{SC} is the pseudorapidity of the electron SC.

As with the muon candidates, no isolation requirements are made on electron candidates.

- Jets are clustered from PF candidates not identified as leptons or pileup. Two clustering algorithms are used in this analysis. The anti- κ_T (AK) algorithm [51] with cone-parameter $R = 0.5$ is used to reconstruct AK5 jets. The Cambridge-Aachen (CA) algorithm [52] with a larger cone-radius of $R = 0.8$ is used to reconstruct CA8 jets. Each jet candidate is required to pass the following minimal jet quality criteria:
 - number of constituent particles > 1;
 - fraction of jet energy coming from electrons < 0.99;
 - fraction of jet energy coming from neutral hadrons < 0.99;
 - fraction of jet energy coming from photons < 0.99;
 - if $|\eta| < 2.4$, charged hadron energy fraction > 0;
 - if $|\eta| < 2.4$, charged multiplicity > 0
- b-tagged jets are identified using the Combined Secondary Vertex (CSV) algorithm [53] on AK5 jets. The CSV algorithm searches for jets originating from secondary vertices with displaced tracks that are consistent with heavy flavor decays. The CSV algorithm uses a likelihood ratio technique to combine several low correlation discriminating variables into a single discriminator which a cut is performed on. It categorizes jets

into three vertex categories: jets with a real secondary vertex determined by the vertex finding algorithms, jets with a “psuedo”-vertex made of tracks more than 2σ away from the jet vertex, and a “no vertex” category for jets not falling into either of the previous categories. The discriminating variables used include:

- Vertex category;
- Flight distance significance in the transverse plane of the detector;
- Vertex mass;
- Number of tracks at the vertex;
- Ratio of energy carried by tracks at the vertex with respect to the whole jet;
- η of tracks at the vertex with respect to the jet axis;
- 2D impact parameter significance of the first track at the vertex that raises the invariant mass of the vertex above 1.5 GeV;
- Number of tracks in the jet;
- 3D impact parameter significance for each track in the jet

Only the last two variables are used for jets in the “no vertex” category. The combined discriminator is valued between zero and one, corresponding to the probability of the jet originating from a heavy flavor decay. We utilize the medium and loose working points, which have cuts on the discriminator value of > 0.679 and > 0.244 respectively, with an efficiency/mistag rate of 70%/1.5% and 85%/10%.

- top-tagged jets are identified using the CMS Top Tagging algorithm [54, 55] on CA8 jets. The CMS Top Tagging algorithm uses jet substructure and jet mass observables to identify hadronically decaying top quarks. The algorithm iteratively decomposes a CA8 jet by reversing the clustering done by the CA algorithm in order to find sufficiently hard and well separated subjets. In the analysis top-tagged CA8 jets are used to reconstruct the hadronic leg of the $t\bar{t}$ decay. Pruning [56, 57] is used to groom the jet mass and improve the mass resolution. The decomposition of the CA8 jets goes as follows:

- Two subclusters are selected from the reversal of the pairwise clustering algorithm;
- If the subclusters satisfy $\sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} > 0.4 - A \times p_T^{cluster}$ the decomposition continues, where $A = 0.0004$ is the optimal value used. If this is not satisfied, the decomposition fails;
- If the subclusters satisfy $p_T^{cluster} > \delta_p \times p_T^{hardjet}$ then the decomposition succeeds, where $\delta_p = 0.05$ is the optimal value used;
- The decomposition is repeated on each passing subcluster until both subclusters pass, both fail, or the subcluster consists of a single constituent.

The decomposition produces either 1,2,3, or 4 subjets depending on whether the decomposition steps succeed. After the decomposition a CA8 jet candidate is said to be top-tagged if it passes the following selection:

- $N_{\text{subjets}} \geq 3$, where N_{subjets} is number of subjets reconstructed by the CMS Top Tagging algorithm;
- $140 \text{ GeV} < m_{\text{jet}} < 250 \text{ GeV}$, where m_{jet} is the ungroomed mass of the CA8 jet;
- $m_{\text{min}} > 50 \text{ GeV}$, where m_{min} is the minimum-pairwise mass between the subjets;
- $\tau_{32} \equiv \tau_3/\tau_2 < 0.7$, where τ_N is a jet-shape variable, known as N -subjettiness, designed to determine the consistency of the jet substructure with the decay of N quarks. For details on τ_N and its definition see Ref. [58];

Kinematic distributions for these top-tag variables can be seen in Appendix B, Fig. B.11 and Fig. B.12.

- Missing transverse energy, \cancel{E}_T , is defined as the negative vector sum of all PF objects. Fully corrected AK5 jets are used in the calculation.

4. SIMULATED EVENTS

This analysis searched for high-mass resonances decaying to $t\bar{t}$ in proton-proton collision data collected by the CMS experiment in 2012. Data samples corresponding to 19.7 fb^{-1} of integrated luminosity at a center-of-mass energy of $\sqrt{s} = 8 \text{ TeV}$ were analyzed. MC event generators were used to produce simulated samples of both SM processes and BSM signals. All simulated samples were generated with a center-of-mass energy of $\sqrt{s} = 8 \text{ TeV}$ and include the simulation of additional inelastic proton-proton interactions within the same bunch crossing (“in-time pileup”) and the simulation of additional contribution in the signal read-out from the previous and next bunch crossing (“out-of-time pileup”). The samples were generated with a wide pileup distribution which is suitable for re-weighting to the actual pileup conditions determined after data taking, which were unknown at the time of sample generation. All simulated samples underwent a detailed simulation of particle propagation through the CMS apparatus and subsequent detector response with the GEANT4 v9.2 [59] toolkit. Both data samples and simulated samples were reconstructed using the same software. Simulated samples have various corrections applied to them to match data as detailed in Chap. 6.

4.1 MC Simulation

MC event simulation consists of many interconnected pieces working together to produce a whole picture of a physics process. Several software libraries exist that generate randomized high-energy particle physics events such as those produced within CMS. A typical event generator simulates not only the main (hard) process, but initial-state composition

and substructure (through parton distribution functions or PDFs), initial-state radiation showers (ISR), final-state radiation showers (FSR), and hadronization. The hard process is described by the tree-level perturbative quantum field theory calculation, usually through the use of Feynman diagrams. Despite its simple structure the observed process usually contains many modifications like loop diagram corrections or ISR/FSR from photon and gluon bremsstrahlung. These modifications complicate the calculation of event cross-section and sometimes analytical methods are not enough. The cross-sections are calculated using Markov-Chain integration instead. [60] Furthermore, the non-perturbative nature of QCD bound states makes hadronization and shower modeling important. The final-state particles from the event generators are fed into the GEANT4 v9.2 [59] detector simulation, allowing a precise prediction for the analysis.

4.1.1 Event Generators Used

The primary background MC, $t\bar{t}$, was generated with the next-to-leading-order generator POWHEG v1.0 [61–64] interfaced to PYTHIA v6.4 [65] for the showering. Higher-order parton radiations (ISR/FSR) are calculated for up to three extra partons at tree-level. Two additional samples for high mass $t\bar{t}$ pairs were produced where $700 \text{ GeV} < M_{t\bar{t}} < 1000 \text{ GeV}$ and $1000 \text{ GeV} < M_{t\bar{t}}$ since the inclusive $t\bar{t}$ sample produces very low statistics in this mass range and our signals are predicted to lie there.

W and Z boson production in association with jets were generated with MADGRAPH v5.1. [66] Exclusive samples with 1, 2, 3 or 4 additional partons generated in the matrix element were used. Single top quark production was simulated using POWHEG [64] and the diboson processes WW , WZ , and ZZ were simulated using PYTHIA v6.2. [67]

Two generic high-mass resonances were simulated, a Z' boson and a Kaluza-Klein excitation of a gluon. The Z' signal model was generated with MADGRAPH v4.4. [68] The simplified model used assumes the same left- and right-handed coupling to fermions as the SM Z boson. Higher-order parton radiations with up to three extra partons are calculated

as with $t\bar{t}$. Various masses were simulated with the width set to 1% (narrow) and 10% (wide) of the Z' mass. The Z' boson decays to $t\bar{t}$ in all generated events.

PYTHIA was used to model the parton showering and fragmentation for all the MADGRAPH samples. Parton showers describe the collinear and soft radiation from partons using a Markov chain technique based on Sudakov form factors. When using parton showering along with extra partons in the matrix element generator one must be careful to avoid double counting of partons. The MLM algorithm [69] is used to avoid overlapping between phase-space descriptions given the matrix element generator MADGRAPH and the showering program PYTHIA.

Kaluza-Klein excitation of a gluon (KK gluon) also results in resonant $t\bar{t}$ production. PYTHIA v8 [70] was used to simulate these resonances. KK gluons produce resonances which are much wider compared to the Z' model which can be seen in Fig. 4.1.

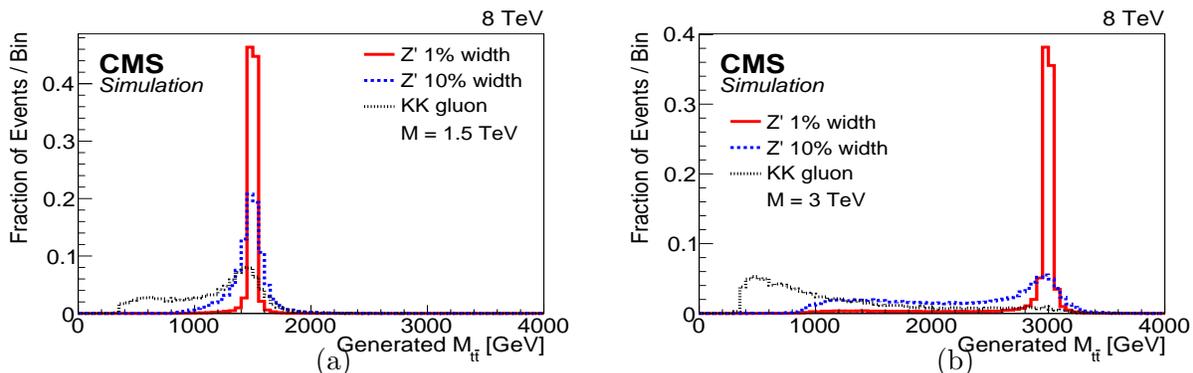


Figure 4.1: Generated $t\bar{t}$ invariant mass for the different BSM signals in this analysis. Shown are (a) 1.5 TeV and (b) 3 TeV masses. Z' 1% (narrow) is shown in red, Z' 10% (wide) is shown in blue and the KK gluon is shown in black.

Samples produced with MADGRAPH and the PYTHIA use the CTEQ6L PDF set. [71] The POWHEG produced $t\bar{t}$ process uses the CT10 PDF set and the single top process uses the CTEQ6M PDF set.

4.1.2 MC Sample Cross-Sections

All MC samples used in the analysis and their cross-section are listed in Tab. 4.1. Tab. 4.2 details the MC signal samples for the narrow Z' , wide Z' and KK gluon models. These tables include the cross section times branching ratio to $t\bar{t}$ for each hypothesis from theoretical calculations. [14, 20] The cross-sections are multiplied by a factor $K = 1.3$ to account for higher-order corrections. [72]

Additional MC samples were generated to study the effects of theoretical uncertainties on the $t\bar{t}$ and W +jets processes. Missing higher order effects in the $t\bar{t}$ and W +jets samples are estimated by varying the renormalization and factorization scales during MC generation. These additional MC samples, labeled scaleup and scaledown, were generated with the Q^2 scale varied by factors of two and one half respectively. Additionally, the jet matching threshold in the W +jets samples is varied by factors of two and one half to estimate the uncertainty on extra hard parton radiation. These samples are labeled matchingup and matchingdown. The cross-sections for these additional MC samples for the scale and matching uncertainties are detailed in Tab. 4.3.

Backgrounds

Process	σ (pb)	
$t\bar{t}$	245.8	(NNLO)
$t\bar{t}$, $700 < M_{t\bar{t}} < 1000$	18.19	(NNLO)
$t\bar{t}$, $1000 < M_{t\bar{t}}$	3.44	(NNLO)
W+1jet	6663	(NNLO)
W+2jets	2159	(NNLO)
W+3jets	640	(NNLO)
W+4jets	264	(NNLO)
single top, s-channel	3.79	(approx. NNLO)
single top, t-channel	56.4	(approx. NNLO)
single top, tW-channel	11.1	(approx. NNLO)
single antitop, s-channel	1.76	(approx. NNLO)
single antitop, t-channel	30.7	(approx. NNLO)
single antitop, tW-channel	11.1	(approx. NNLO)
Z+1jet	666	(NNLO)
Z+2jets	215	(NNLO)
Z+3jets	60.7	(NNLO)
Z+4jets	27.4	(NNLO)
WW	54.8	(NLO)
WZ	33.2	(NLO)
ZZ	8.1	(NLO)

Table 4.1: SM background process cross-sections used in the analysis. [73]

Signals

Process	$\sigma \cdot \text{BR}$ (pb)	
1% width Z' (M=500 GeV)	17.82	(LO)
1% width Z' (M=750 GeV)	4.31	(LO)
1% width Z' (M=1000 GeV)	1.24	(LO)
1% width Z' (M=1250 GeV)	0.411	(LO)
1% width Z' (M=1500 GeV)	0.160	(LO)
1% width Z' (M=2000 GeV)	$2.75 \cdot 10^{-2}$	(LO)
1% width Z' (M=3000 GeV)	$1.16 \cdot 10^{-3}$	(LO)
10% width Z' (M=500 GeV)	145.06	(LO)
10% width Z' (M=750 GeV)	33.29	(LO)
10% width Z' (M=1000 GeV)	9.84	(LO)
10% width Z' (M=1250 GeV)	3.37	(LO)
10% width Z' (M=1500 GeV)	1.28	(LO)
10% width Z' (M=2000 GeV)	0.218	(LO)
10% width Z' (M=3000 GeV)	$8.59 \cdot 10^{-3}$	(LO)
g_{KK} (M=1000 GeV)	6.30	(LO)
g_{KK} (M=1300 GeV)	1.82	(LO)
g_{KK} (M=1600 GeV)	0.623	(LO)
g_{KK} (M=1900 GeV)	0.245	(LO)
g_{KK} (M=2200 GeV)	0.108	(LO)
g_{KK} (M=2500 GeV)	0.053	(LO)
g_{KK} (M=2800 GeV)	0.0288	(LO)
g_{KK} (M=3100 GeV)	0.0169	(LO)

Table 4.2: Theoretical cross-sections for the the signal models studied. Z' cross-sections are taken from Ref. [14] and g_{KK} cross-sections are taken from Ref. [20]. Cross-sections are multiplied by a factor $K = 1.3$ to account for higher-order corrections. [72]

Backgrounds (scale and matching systematics)

Process	σ (pb)	
$t\bar{t}$ scaleup	245.8	(NNLO)
$t\bar{t}$, $700 < M_{t\bar{t}} < 1000$ scaleup	17.0	(NNLO)
$t\bar{t}$, $1000 < M_{t\bar{t}}$ scaleup	3.20	(NNLO)
$t\bar{t}$ scaledown	245.8	(NNLO)
$t\bar{t}$, $700 < M_{t\bar{t}} < 1000$ scaledown	19.2	(NNLO)
$t\bar{t}$, $1000 < M_{t\bar{t}}$ scaledown	3.93	(NNLO)
W+1jet scaleup	5626	(NNLO)
W+2jets scaleup	1641	(NNLO)
W+3jets scaleup	398.5	(NNLO)
W+4jets scaleup	140.7	(NNLO)
W+1jet scaledown	7772	(NNLO)
W+2jets scaledown	2816	(NNLO)
W+3jets scaledown	1160	(NNLO)
W+4jets scaledown	585.7	(NNLO)
W+1jet matchingup	4086	(NNLO)
W+2jets matchingup	1015	(NNLO)
W+3jets matchingup	245.2	(NNLO)
W+4jets matching up	70.6	(NNLO)
W+1jet matchingdown	8338	(NNLO)
W+2jets matchingdown	2906	(NNLO)
W+3jets matchingdown	985.4	(NNLO)
W+4jets matchingdown	518.0	(NNLO)

Table 4.3: $t\bar{t}$ and W+jets cross-sections used for Q^2 scale and matching systematics.

5. EVENT SELECTION

We perform a model independent search for high-mass resonances, denoted by Z' , decaying to a top-antitop quark pair, $t\bar{t}$. The model independence of the search is ensured by using selection criteria tailored to generic $t\bar{t}$ resonant production without any model specific search parameters. Both top quarks decay to a W boson and a bottom quark, $Z' \rightarrow t\bar{t} \rightarrow W^+bW^-\bar{b}$, with the bottom quark hadronizing and manifesting as a jet in the event. The W boson decays into a fermion-antifermion pair, $(67.60 \pm 0.27)\%$ of the time into a quark-antiquark pair (hadronic decay) and $(32.40 \pm 0.27)\%$ of the time into a lepton-neutrino pair (leptonic decay). [74] We focus only on the semileptonic decay of the $t\bar{t}$ system where one of the W bosons decays leptonically into a muon or electron and a neutrino and the other decays hadronically. The all hadronic and dileptonic $t\bar{t}$ decay modes, where both W bosons decay either hadronically or leptonically, are studied by other groups within the CMS collaboration. The high mass resonances we search for result in highly Lorentz boosted top quarks which causes the decay products of the top quark to be confined to a smaller area inside the detector. SM $t\bar{t}$ production, on the other hand, usually results in a well separated topology where the lepton and all four final state quarks are reconstructed as separate objects. Fig. 5.1 shows the expected decay topology for a typical signal event and a SM $t\bar{t}$ pair for comparison. As the mass of the resonance increases, the resulting boost gets larger and the ability to distinguish individual decay products becomes impossible as they begin to overlap with each other. For large enough boost the whole top quark decay merges into a single “fat” jet.

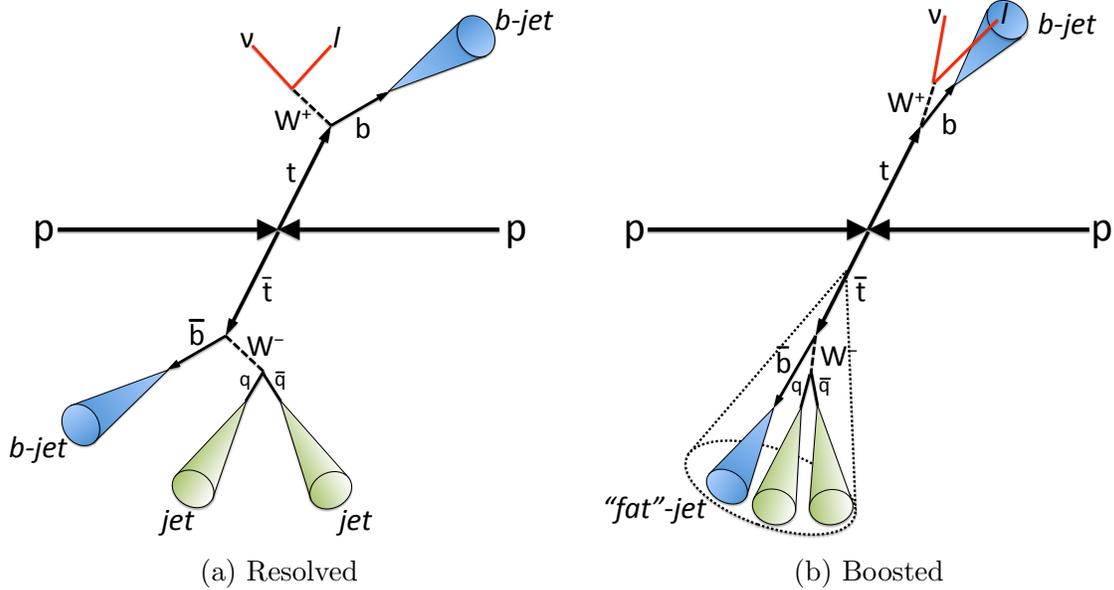


Figure 5.1: The difference between the expected decay topology for SM $t\bar{t}$ production (a) and high-mass resonant $t\bar{t}$ production (b) is shown. High-mass resonances result in a boosted topology where the decay products can merge together. This merging of final state objects must be taken into account when searching for these events.

We search for and select events consistent with a boosted $t\bar{t}$ decay coming from a high-mass resonance, while rejecting events from SM background processes. Background processes with decay topologies that are significantly different from our signal can usually be reduced through carefully selected kinematic cuts. Some backgrounds, such as SM $t\bar{t}$ production, have identical decay products to our signal and are therefore called irreducible background. We base our event selection criteria on the expected boosted $t\bar{t}$ decay topology where we expect the hadronically decaying top quark to be merged into as little as one jet, and the leptonically decaying top quark to produce a non-isolated lepton. Non-isolated leptons are usually associated with multijet QCD production and we utilize some specialized selection cuts to reduce this background. We expect the final state objects to be boosted as well, so we search for leptons and jets with high transverse momentum. We also expect a significant amount of \cancel{E}_T from the undetectable neutrino. The high p_T lepton present in our signal provides a nice handle to trigger events on, although we must ensure the trigger does not

require an isolated lepton as this would severely degrade the efficiency of the trigger for high mass resonances with boosted decay topologies.

Therefore we search for events containing a high p_T lepton that can be non-isolated, \cancel{E}_T from the undetected neutrino, and at least two high p_T jets. In addition a top-quark tagging algorithm is utilized to identify events containing a “fat” jet consistent with a boosted hadronically decaying top quark. This algorithm is discussed in more detail in Sec. 6.6.

5.1 Preselection

In order to reduce the dataset to a reasonable size a preselection step is used. Noise-suppressing filters and trigger requirements are applied at this stage. This preselection also separates the events based on the lepton flavor of the leptonically decaying top quark into two statistically independent channels. Only the electron+jets channel and the muon+jets channel are considered.

Events recorded by the CMS detector are sometimes rejected due to known issues and limitations of the various sub detectors. To account for these conditions, the collected data must undergo the following additional selection criteria compared to the MC simulations.

1. HCAL instrumentation issues sometimes cause anomalous noise not related to the electronic readout. A dedicated filter is applied that rejects events where this is present.
2. Events in which a charged particle moving nearly parallel to the beam axis causing a large number of hits in the silicon strip detector are called “beam scraping” events. These events are filtered out by their low fraction of high quality reconstructed tracks.
3. Events taken when the detector was not fully operational or that did not pass a data certification procedure are filtered out based on certified good luminosity sections.

The electron+jets and muon+jets channels are statistically independent after the preselection and are processed independently afterwards. To allow a combination with the all-hadronic channel in which both top quarks decay fully hadronically we veto events containing two or more CA8 “top-tagged” jets that fall into their signal region. This veto has a

minimal effect on our signal. The dileptonic search is automatically statistically independent to our search since we require exactly one lepton. The selection criteria for each channel are detailed in the following sections.

5.2 Muon+Jets Channel

The following selection criteria must be fulfilled for events in the muon+jets channel.

1. Events are selected by and must pass the single-muon HLT trigger, `HLT_Mu40_eta2p1_v*` (hereafter called “Mu40”). This trigger is unprescaled and does not apply isolation requirements on the muon. It triggers on events containing a muon with $p_T > 40$ GeV within $|\eta| < 2.1$.
2. At least one good primary vertex is required. The primary vertex must have $|z| < 24$ cm, $\sqrt{x^2 + y^2} < 2$ cm, and $N_{dof} \geq 4$ where N_{dof} is the weighted number of tracks used in the reconstruction of the primary vertex.
3. Exactly one muon candidate, as defined in Chap. 3, with $p_T > 45$ GeV and $|\eta| < 2.1$ is required. A veto on additional leptons passing the same selection criteria is performed to ensure statistical independence between channels.
4. At least two AK5 jets with $p_T > 50$ GeV and $|\eta| < 2.4$ are required.
5. The leading AK5 jet must have $p_T > 150$ GeV.
6. In lieu of lepton isolation, a 2D-cut is made between the muon and j , the nearest AK5 jet with $p_T > 25$ GeV. The event is accepted if either $\Delta R(\mu, j) > 0.5$ or $p_T^{rel}(\mu, j) > 25$ GeV. p_T^{rel} is defined as the component of the lepton’s momentum orthogonal to the jet axis and is given by $p_T^{rel} = |p_l| \sin(\alpha)$, where α is the angle between the muon and j .
7. $H_T^{lep} > 150$ GeV, where $H_T^{lep} \equiv \cancel{E}_T + p_T^\mu$.
8. $\cancel{E}_T > 50$ GeV.

5.3 Electron+Jets Channel

The selection criteria for the electron+jets channel is very similar to those of the muon+jets channel, with the addition of a topological cut to filter out the multijet QCD background that isn't present in the muon+jets channel. A slightly more complex triggering strategy is used where we select events that pass a logical OR of two triggers. The single electron `HLT_Ele30_CaloIdVT_TrkIdT_PFnPUJet100_PFnPUJet25_v*` (hereafter called “Ele30[...]”) trigger is the primary trigger required. This trigger looks for events containing an electron candidate with $p_T > 30$ GeV, and two jets with $p_T > 25$ GeV with the leading jet $p_T > 100$ GeV. The logic of this trigger removes jets near the electron from consideration in the jet requirements, which causes the trigger to become inefficient for high mass Z' signals and their boosted decay topology. Effectively this trigger becomes an electron plus three jets trigger instead of an electron plus two jets trigger since there is usually a jet close to the electron which gets ignored. To mitigate this inefficiency we use the logical OR with an additional single-jet HLT trigger `HLT_PFJet320_v*` (hereafter called “PFJet320”). This strategy recovers the lost signal events. A study of the efficiency of this combined trigger in data and simulated events is presented in Sec. 6.2.1. Since these two triggers record data events into separate datasets a careful combination of the datasets must be done to avoid double counting of data events. The PFJet320 trigger records data events into the JetHT dataset while the Ele30[...] trigger records data events into the ElectronHad dataset. We look for data events in the ElectronHad dataset using the electron trigger, and then add events from the JetHT dataset that pass the jet trigger but fail the electron trigger. In this way we ensure we do not double count events in both datasets. For simulated samples a simple logical OR is used between the two triggers.

The following selection criteria must be fulfilled for events in the electron+jets channel.

1. Either Ele30[...] or PFJet320 must be passed. These triggers are unrescaled and do not apply isolation requirements on the electron.

2. At least one good primary vertex as defined in Sec. 5.2
3. Exactly one electron candidate, as defined in Chap. 3, with $p_T > 35$ GeV and $|\eta| < 2.5$ is required. A veto on additional leptons passing the same selection criteria is performed to ensure statistical independence between channels.
4. At least two AK5 jets with $p_T > 50$ GeV and $|\eta| < 2.4$ are required.
5. The leading AK5 jet must have $p_T > 150$ GeV.
6. A 2D-cut is made between the electron and j , the nearest AK5 jet with $p_T > 25$ GeV. The event is accepted if either $\Delta R(e, j) > 0.5$ or $p_T^{rel}(e, j) > 25$ GeV. p_T^{rel} is the same as defined in Sec. 5.2.
7. $H_T^{lep} > 150$ GeV, where $H_T^{lep} \equiv \cancel{E}_T + p_T^e$.
8. $\cancel{E}_T > 50$ GeV.
9. Triangular cut [75] which defines topological requirements to ensure \cancel{E}_T does not point along the transverse direction of the electron (e) or the leading AK5 jet (j):

$$-\frac{1.5}{75 \text{ GeV}} \cancel{E}_T + 1.5 < \Delta\phi\{(e \text{ or } j), \cancel{E}_T\} < \frac{1.5}{75 \text{ GeV}} \cancel{E}_T + 1.5$$

10. The reconstructed leptonic top quark transverse momentum must be $p_T^{t,lep} > 140$ GeV. $p_T^{t,lep}$ is defined in Sec. 7.2.

6. CORRECTIONS AND EFFICIENCIES

A number of corrections are needed in order to more accurately simulate data taking conditions at the CMS experiment and to correct for biases that may arise from selection cuts. These corrections include pileup, lepton selection, trigger efficiencies, jet energy corrections (JEC), jet energy resolution (JER), b-tagging efficiency, and top-tagging efficiency.

To correct for a discrepancy between data and simulated samples for a specific efficiency a scale factor is used. Scale factors are derived from the ratio of efficiency in data to the efficiency in MC as shown in Equation 6.1. These scale factors are applied to MC samples and correct for the difference in efficiency between data and simulation.

$$SF = \frac{\epsilon_{data}}{\epsilon_{MC}} \quad (6.1)$$

6.1 Pileup Re-Weighting

As discussed in Chap. 4, the MC simulations are generated with a wide range of pileup conditions since the true pileup conditions were unknown at the time of sample generation. Pileup interactions affect the detector response and resolution and as such the simulated samples must be re-weighted so the events in simulation are comparable to those in data. Each MC sample is re-weighted such that the number of pileup interactions matches data as shown in Fig. 6.1. The average pileup during the 2012 data taking period had a mean number of interactions per bunch crossing of $\langle \mu \rangle = 21$.

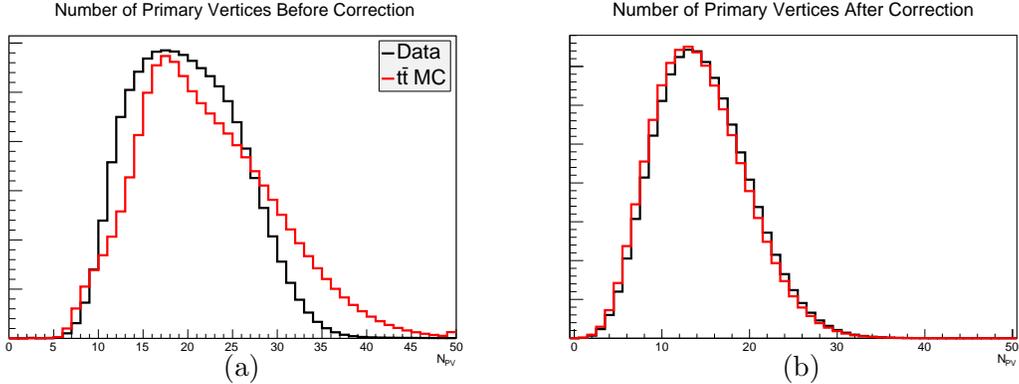


Figure 6.1: Pileup distributions for data and $t\bar{t}$ MC before (a) and after (b) pileup reweighting. The distributions in (a) are taken at the generator level for MC and before any analysis cuts for data whereas the distributions for (b) are taken after the preselection described in Sec. 5.1 which accounts for the slight difference in shape of the data pileup distribution. The distributions have been scaled to the same area to highlight the shape difference.

6.2 Lepton Identification and Trigger Efficiencies

The use of $Z \rightarrow \ell\ell$ dilepton resonances to measure lepton efficiencies has been well established by the high energy physics community. [76] Events containing a Z peak can be used to measure lepton identification and triggering efficiencies using a tag-and-probe method due to the very clean signal. The Z resonance is reconstructed as a pair of leptons with one leg passing very stringent quality control cuts (“tag”) and one passing a loose identification (“probe”). Probes that pass whatever efficiency is to be measured are called “passing probes” and probes that fail are called “failing probes”. The tag+passing probes and tag+failing probes events are fit separately with signal+background models, where the signal model is the line shape expected from a Z resonance. The efficiency is computed from the ratio of the passing probe signal model to the sum of passing and failing probe signal model.

Lepton identification and trigger efficiency scale factors have been measured centrally by CMS using a tag-and-probe method in $Z \rightarrow \ell\ell$ events. These results can be used in the analysis since the “probe” leptons have much looser selection requirements including

no requirement of isolation. In the muon channel we use the scale factors from Ref. [77] to correct for muon identification and trigger efficiency differences in all of our simulated samples. These scale factors are given as a function of muon η and p_T and are very close to unity. Figures 6.2–6.3 show the scale factors applied to MC.

In the electron channel we use the scale factors from Ref. [50] to correct for electron identification efficiencies in all of our simulated samples. These scale factors are given as a function of electron p_T and supercluster pseudorapidity η_{SC} and are also very close to unity. Fig. 6.4 shows the scale factor applied to MC.

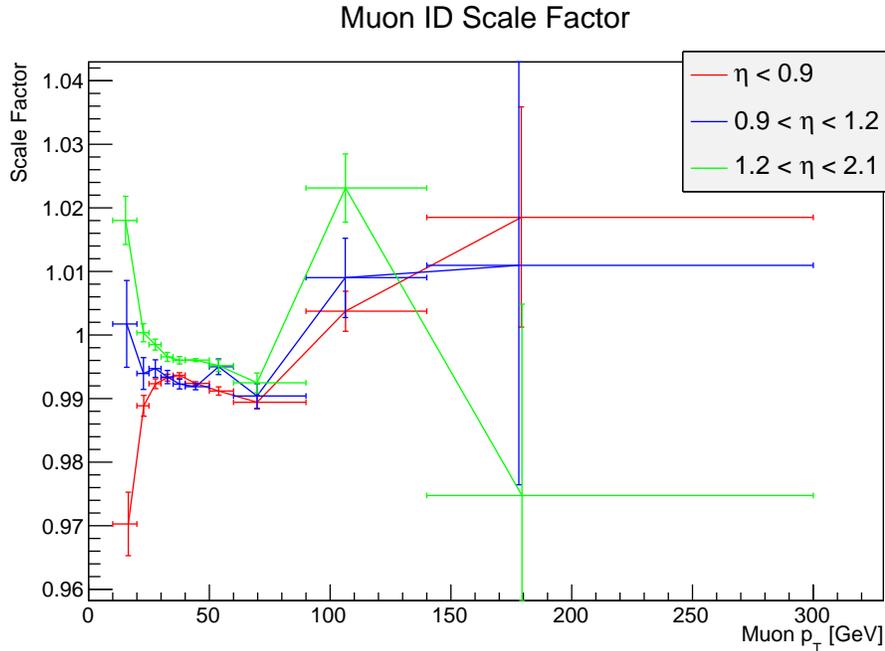


Figure 6.2: Scale factors applied to MC samples for muon ID efficiency. The scale factor is given as a function muon p_T and η region. The scale factor is derived separately for the barrel ($|\eta| < 0.9$), endcap ($1.2 < |\eta| < 2.1$) and overlap ($0.9 < |\eta| < 1.2$) regions as the muon detector uses different physical detectors systems for each region which affects the efficiency measurements.

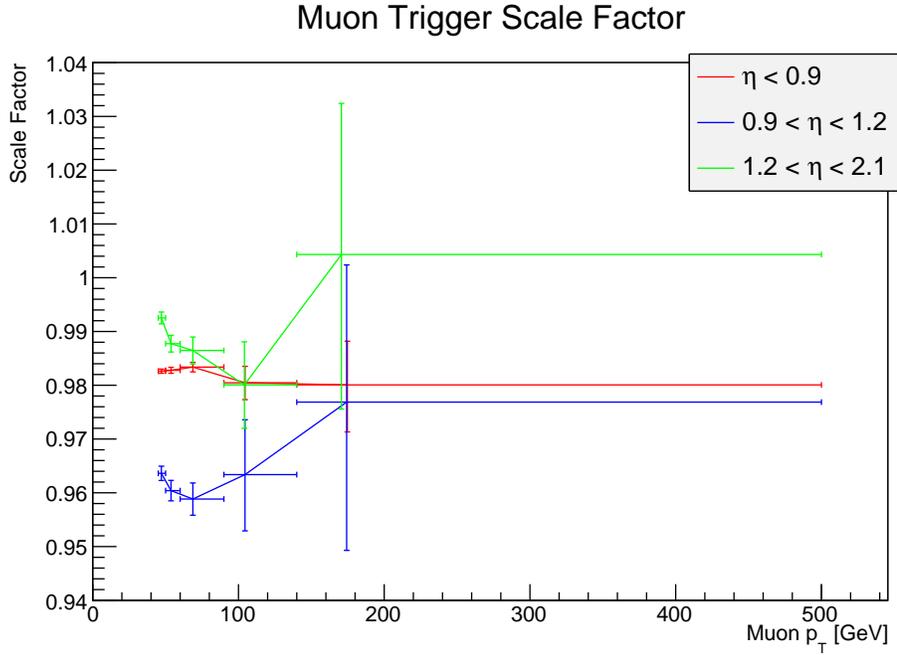


Figure 6.3: Scale factors applied to MC samples for muon trigger efficiency. The scale factors are given as a factor of muon p_T and η region. The scale factor is derived separately for the barrel ($|\eta| < 0.9$), endcap ($1.2 < |\eta| < 2.1$) and overlap ($0.9 < |\eta| < 1.2$) regions as the muon detector uses different physical detectors systems for each region which affects the efficiency measurements.

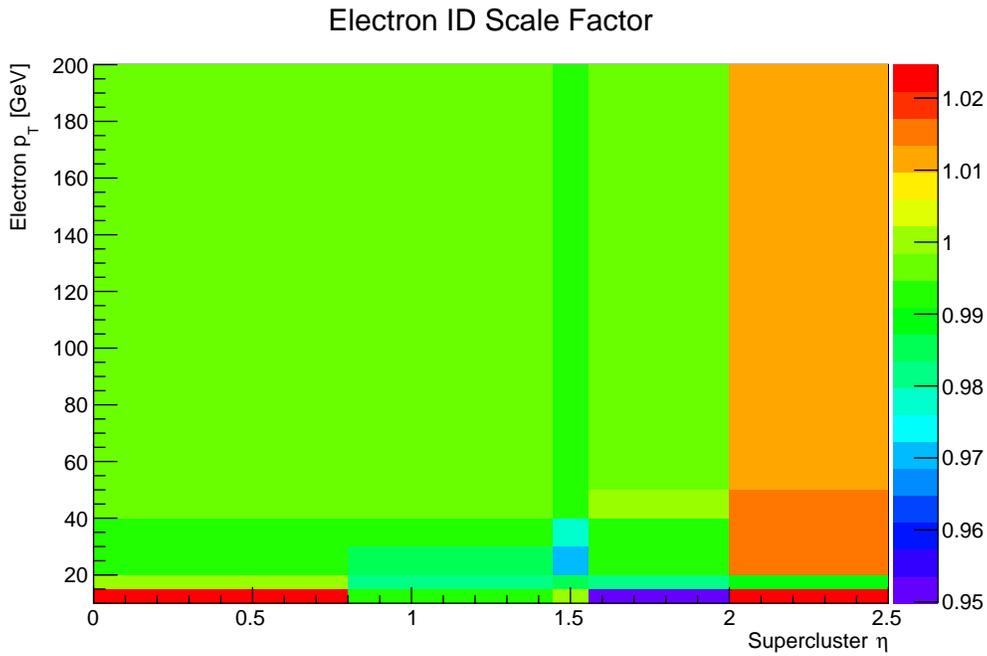


Figure 6.4: Scale factors applied to MC samples for electron ID efficiency. The scale factor is given as a function electron p_T and η_{SC} .

Due to the boosted nature of the event topology in this analysis it is important to not apply isolation requirements on the lepton candidate. The so-called 2D-cut replaces the isolation requirement to some extent and reduces mostly QCD without compromising the signal acceptance. To measure the efficiency of this 2D-cut we use a tag-and-probe method in a $Z \rightarrow \ell\ell + \text{jets}$ control sample. The kinematic cuts that define this control sample mimic the cuts used in the analysis and use the same object reconstruction, with the goal of producing a sample similar to our signal region. The full selection requirements are as follows:

1. Events are selected with the following triggers:

HLT_IsoMu24_eta2p1_v* (muon);

HLT_Ele27_WP80_v* (electron)

2. exactly 2 opposite-sign same-flavor leptons and veto on other leptons
3. at least 2 AK5-jets with $p_T > 50$ GeV and $|\eta| < 2.4$
4. at least 1 AK5-jet with $p_T > 150$ GeV and $|\eta| < 2.4$
5. $H_T^{lep} \equiv p_T^{\ell 1} + p_T^{\ell 2} > 150$ GeV
6. one lepton has to be the “tag”, defined as the highest- p_T candidate satisfying the following requirements
 - $p_T^{\text{tag}} > 50$ GeV
 - minimum $\Delta R_{(\text{tag}, \text{jets})} > 0.5$, calculated with respect to AK5-jets with $p_T > 25$ GeV
 - relative isolation cut
 - muon: $I_{\text{rel}}^{\mu} < 0.12$ (cone size $R = 0.4$)
 - electron: $I_{\text{rel}}^e < 0.1$ (cone size $R = 0.3$)

The other lepton is the “probe”.

The tag-and-probe fit specifics used are as follows:

- Muon ($\ell = \mu$):
 - dilepton invariant mass interval: [70,130] GeV
 - signal PDF: sum of two voigtians
 - background PDF: exponential
- Electron ($\ell = e$):
 - di-electron invariant mass interval: [60,120] GeV
 - signal PDF: Breit-Wigner convoluted with crystal-ball
 - background PDF: exponential

Fig. 6.5 shows control plots for the selection defined above. For both muons and electrons the sample is fully dominated by Z+jets production, and thus is the only MC sample used to measure the efficiency of the 2D-cut. Using this approach the ratio of the 2D-cut efficiency in data and MC was measured to be flat and unitary as can be seen in Fig. 6.6. As such, no specific scale factor was applied in the analysis.

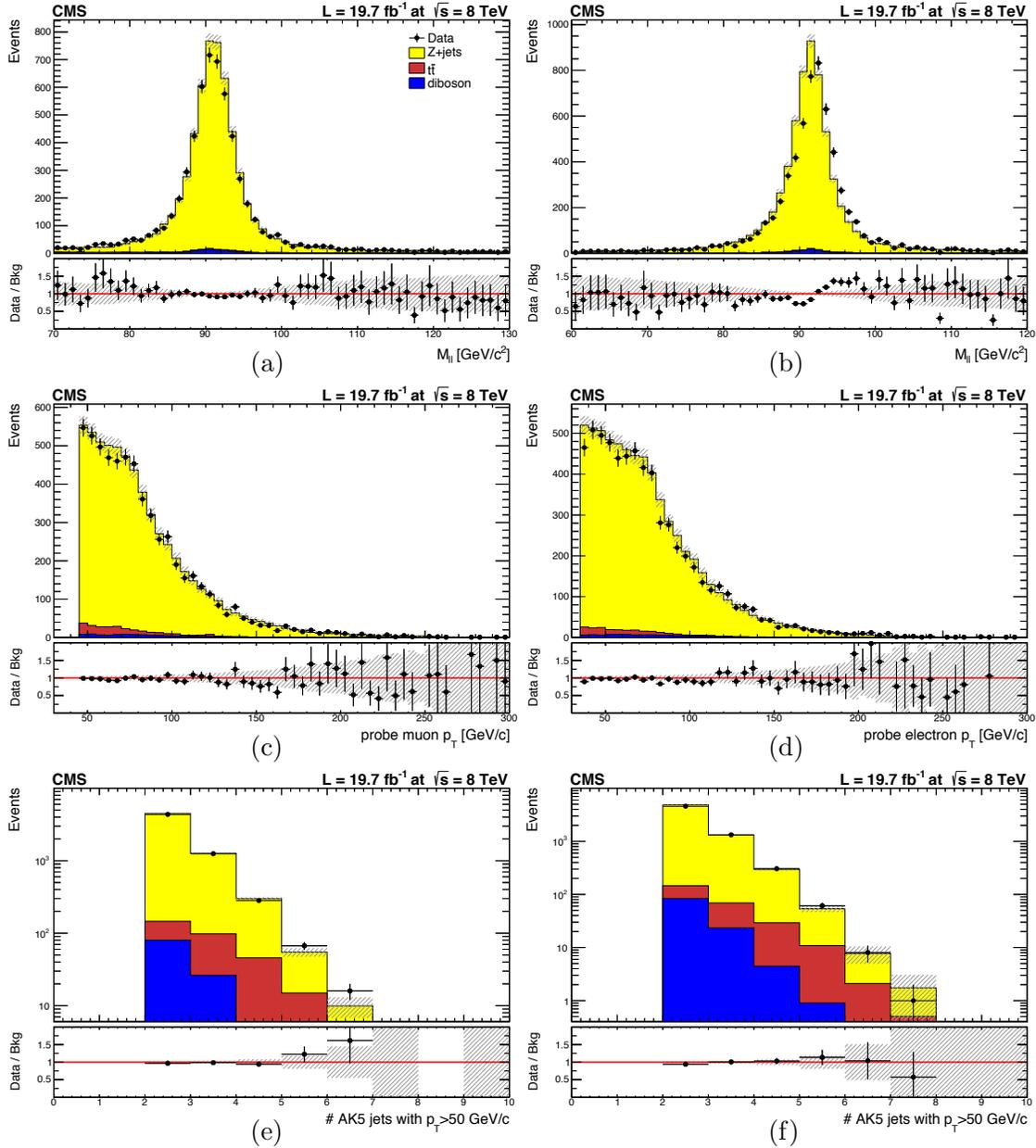


Figure 6.5: Data/MC plots for the $Z \rightarrow \ell\ell + \text{jets}$ control sample for muons and electrons (before any specific cut on the probe): $M_{\mu\mu}$ invariant mass (a), M_{ee} invariant mass (b), probe muon p_T (c), probe electron p_T (d), and number of AK5-jets with $p_T > 50$ GeV for muons (e) and electrons (f). The overall data/MC agreement is good except for the dielectron invariant mass. This is due to a known electron energy discrepancy for this kinematic quantity, but has no effect on the efficiency measurement. The gray band corresponds to the statistical uncertainty of the simulation. The bottom plot of each figure shows the ratio of data over background.

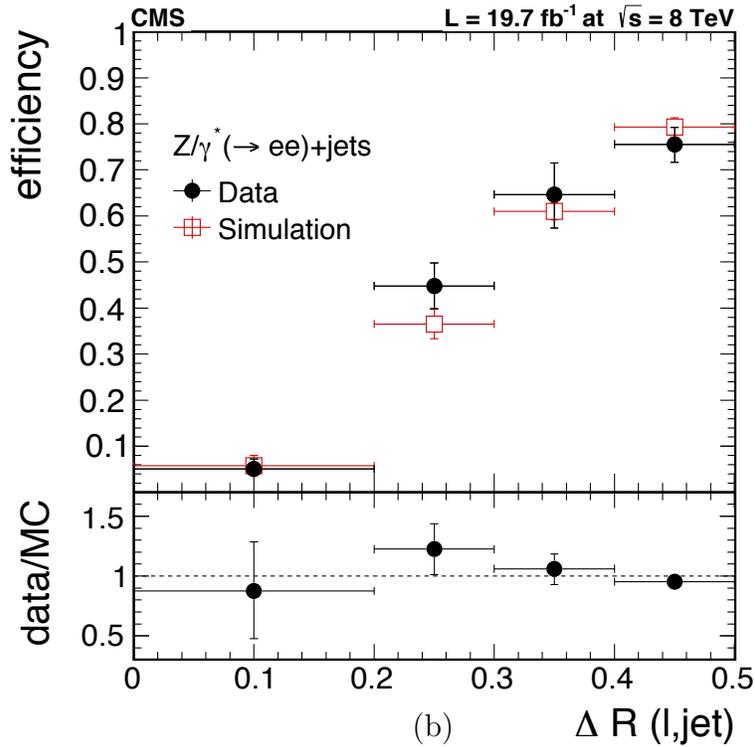
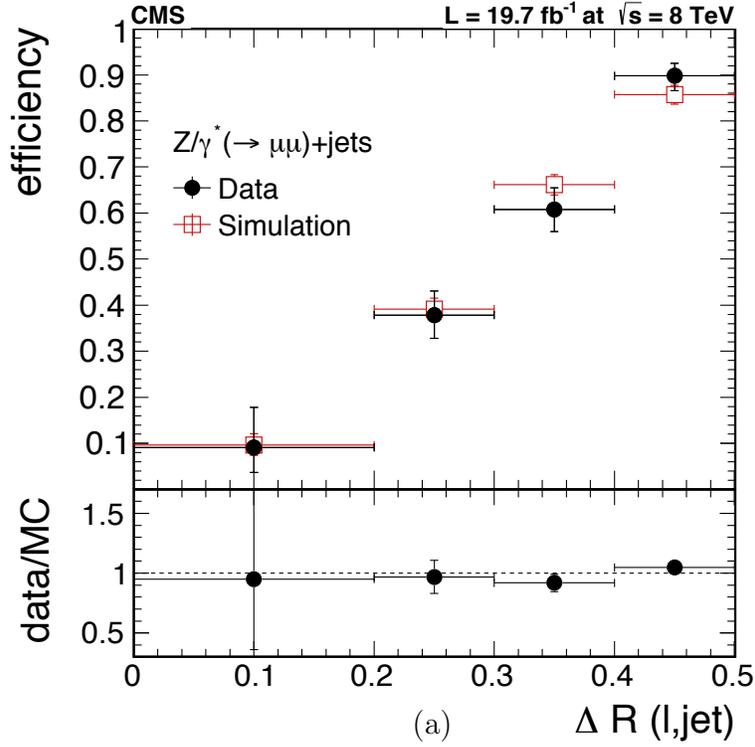


Figure 6.6: Efficiency of the 2D-cut for data and MC events as a function of $\Delta R_{\min}(\ell, jets)$ in the $Z \rightarrow \mu\mu$ (a) and $Z \rightarrow ee$ (b) control samples. The scale factor is the ratio of these efficiencies and is plotted below each figure.

6.2.1 Electron Channel Trigger

For the electron channel we combined two triggers and as such had to measure the trigger efficiency ourselves. The performance of the combined trigger was studied using a tag-and-probe approach using a $e\mu$ control sample which is dominated by $t\bar{t}$ production. The choice to use dilepton $t\bar{t}$ events instead of $Z \rightarrow \ell\ell$ for the tag-and-probe study was because $Z \rightarrow \ell\ell$ events contain mostly isolated leptons and we wanted to have a sample close to our signal region. The scale factor is derived separately for the barrel ($\eta < 0.9$), endcap ($1.2 < \eta < 2.1$) and overlap ($0.9 < \eta < 1.2$) regions as the muon detector uses different physical detectors systems for each region which affects the efficiency measurements.. To obtain the control sample we start with the muon selection described in Sec. 5.2 and remove the veto on additional leptons. We then require an electron matching the kinematic cuts for the electron in the electron channel in Sec. 5.3. A final requirement enforcing the two leptons to be oppositely charged is then applied. The sample is dominated by $t\bar{t}$ production with small contributions from W/Z+jets and single top processes. Kinematic plots showing the data/MC agreement in this control samples are shown in Fig. 6.7 for the number of primary vertices, the number of AK5 jets with $p_T > 30$ GeV, the electron p_T and the leading jet p_T .

The efficiency of the combined trigger is given by the fraction of events in this $e\mu$ samples that fire either Ele30[...] or PFJet320. The overall efficiency of the combined trigger is $90.0 \pm 0.7\%$ in data and $95.1 \pm 0.2\%$ in the $t\bar{t}$ MC. The trigger efficiency is flat in all kinematic variables except the leading jet p_T , where a turn on occurs above 320 GeV due to the inclusion the single-jet trigger. Therefore we characterize the scale factor as a function of leading jet p_T which is shown in Fig. 6.8.

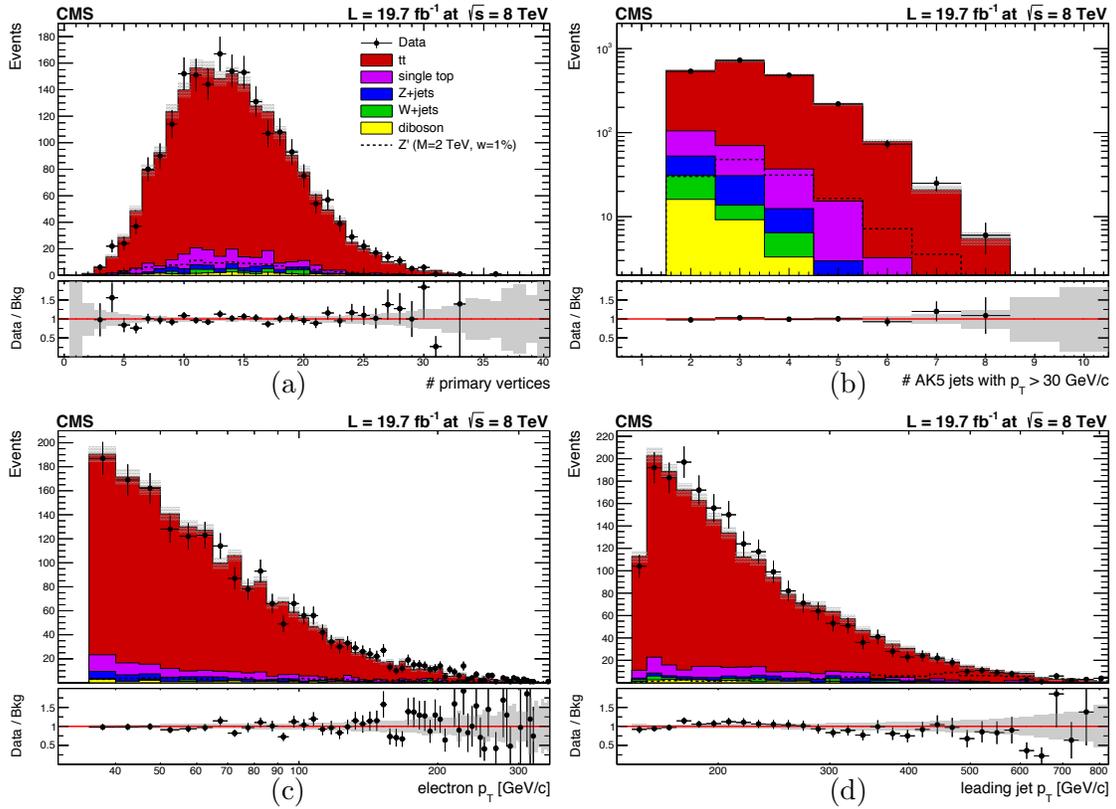


Figure 6.7: Data/MC comparisons of several kinematic variables in the $e\mu$ control sample used to measure the electron channel trigger efficiency. Shown are the number of primary vertices (a), number of AK5 jets with $p_T > 30$ GeV (b), the electron p_T (c) and the leading jet p_T (d). The Z' signal has been normalized to a cross section of 1pb and the sum of backgrounds has been normalized to data. The gray band corresponds to the statistical uncertainty of the simulated samples.

Table 6.1: Overall efficiencies of the Ele30[...] and its logical OR with the PFJet320 trigger in the $e\mu$ control sample for different narrow- Z' mass hypotheses. The uncertainty is statistical only.

$M_{Z'} (w = 1\%)$	Ele30[...]	Ele30[...] PFJet320
1 TeV	$94.1^{+0.9}_{-1.0}$	$95.8^{+0.8}_{-0.9}$
1.25 TeV	$92.4^{+0.9}_{-1.0}$	$97.6^{+0.5}_{-0.6}$
1.5 TeV	$90.5^{+0.6}_{-0.6}$	$98.5^{+0.2}_{-0.3}$
2 TeV	$86.9^{+1.0}_{-1.0}$	$99.1^{+0.3}_{-0.3}$
3 TeV	$83.6^{+1.6}_{-1.7}$	$99.3^{+0.3}_{-0.5}$

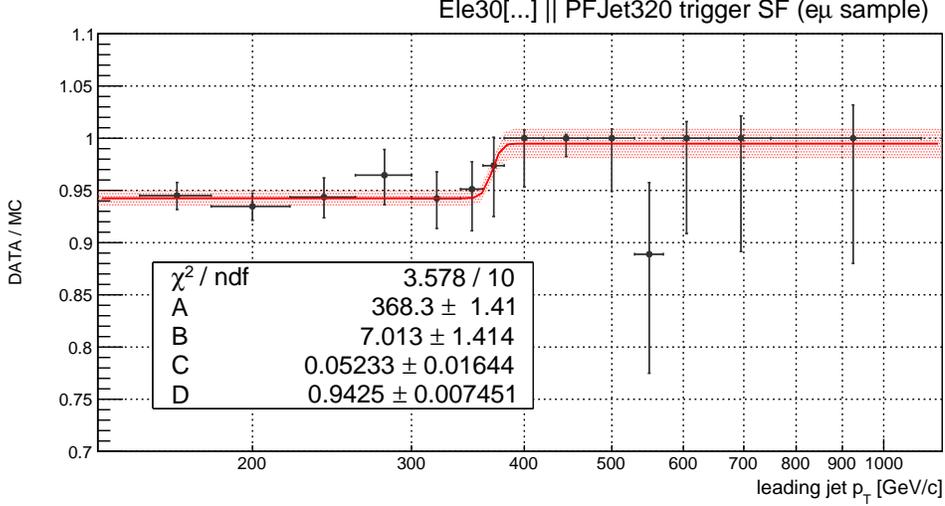


Figure 6.8: Data/MC efficiency ratio for the electron channel combined trigger plotted as a function of the leading AK5 jet p_T . The fit function (solid red line) used is $f(p_T) = D + \frac{C}{2} \left[1 + \text{erf} \left(\frac{p_T - A}{\sqrt{2}B} \right) \right]$. The shaded band around the fit function corresponds to the $\pm 1\sigma$ statistical error band obtained by propagating the full correlation matrix of the fit result.

We parameterize the combined trigger scale factor as a constant plus an error function, with the constant modeling the Ele30[...] trigger efficiency scale factor and the error function modeling the turn on of the PFJet320 trigger efficiency scale factor. The parameterized scale factor is of the form:

$$f(p_T) = D + \frac{C}{2} \left[1 + \text{erf} \left(\frac{p_T - A}{\sqrt{2}B} \right) \right]$$

The effect on the inclusion of the PFJet320 trigger on the overall trigger efficiency in several signal samples is shown in Tab. 6.1. It is clear the combined trigger gives a significant performance boost in comparison to the Ele30[...] trigger alone.

6.3 Jet Energy Corrections

Jet energy is measured with a combination of the ECAL and HCAL sub-detectors, however the detector response is non-linear. As such it is not straightforward to translate the measured jet energy to energy of the parton producing the jet. The jet energy corrections are a set of tools used to map the measured jet energy deposition to the parton energy. A

factorized set of corrections are applied, with each level of correction taking care of a different effect. At each level of correction a scale factor is applied to the jet four-momentum, and the corrections are applied sequentially with the output of one level being the input of the subsequent level. The following correction levels are applied to each jet candidate in both data and simulation:

- **L1 Pile Up:** The L1 correction removes the energy coming from pile up events. This correction in principle removes dependence on luminosity from the event and the subsequent corrections are applied to a luminosity independent sample.
- **L2 Relative Jet Correction:** The L2 correction ensures a uniformity in η for jet response by correcting jets to those in the central region ($|\eta| < 1.3$).
- **L3 Absolute Jet Correction:** The L3 correction ensures a uniform jet response in p_T . After the L3 correction the corrected jet p_T is equal on average to the parton p_T .
- **L2L3 Residual:** The CMS jet energy response simulation was seen to be very successful but the comparison between data and MC still has small differences, up to 10%, depending on η . The L2L3 residual calibration is applied to data only to account for these known discrepancies.

The jet energy corrections are derived separately for AK5 jets and CA8 jets.

For AK5 jets a jet-lepton cleaning procedure is needed due to the non-isolated nature of boosted leptonic top decays. Since we do not require isolated leptons in the analysis, it is probable that leptons may be clustered into jets. To avoid a double-counting of energy the lepton four-momentum is subtracted from the jet four-momentum if the lepton is within $\Delta R < 0.5$ of the jet. This subtraction is not performed if the lepton candidate is isolated as it wasn't included in the jet clustering in the first place. The subtraction is performed on the raw jet energy, where the jet energy corrections of the jet are first reversed and then the lepton four-momentum is subtracted. Jet energy corrections are then reapplied to the now cleaned jet energy.

6.4 Jet Energy Resolution

Measurements of dijet energy resolution have shown that JER in data is worse than in MC. Therefore jets in MC must be smeared so that the resolution matches that measured in data. This smearing is done by scaling the reconstructed jet energy (*recopt*) by a factor, f , dependent on the generated jet energy (*genpt*) and some factor α :

$$f = \max(0, 1 + \alpha * (\text{recopt} - \text{genpt})/\text{recopt})$$

The α factors are such that the MC rescaled dijet energy resolution matches the measurement in data.

6.5 B-Tagging Efficiency and Mistag Rate

We use b-tagging to identify events containing AK5 jets with b-flavored content. Data and MC events are tagged using the Combined Secondary Vertex algorithm at the medium operating point (CSVM). The efficiency of this algorithm has been extensively studied in events with muons in jets dominated by $b\bar{b}$ decay [78]. MC events are weighted to correct for any discrepancy between the efficiency of b-tagging a jet in data and simulated events. For an event with N_T tagged and $N_{\bar{T}}$ untagged AK5 jets the event weight is given by

$$w = \prod_{i=1}^{N_T} SF_T(f_i, p_{T,i}) \times \prod_{i=1}^{N_{\bar{T}}} SF_{\bar{T}}(f_i, p_{T,i}) \quad (6.2)$$

The scale factors for tagged jets, SF_T , and for untagged jets, $SF_{\bar{T}}$, are defined as:

$$SF_T(f, p_T) = \frac{\epsilon_f^{\text{data}}(p_T)}{\epsilon_f^{\text{MC}}(p_T)} \quad SF_{\bar{T}}(f, p_T) = \frac{1 - \epsilon_f^{\text{data}}(p_T)}{1 - \epsilon_f^{\text{MC}}(p_T)} \quad (6.3)$$

The scale factors and efficiencies depend on flavor of the jet, f . This equation encapsulates both the b-tagging efficiency scale factors, $SF_T(b)$, and the b-tagging mistag rate scale

factors, $SF_T(l)$, simply by taking into account the true jet flavor, f . The jet flavor can be b for b-flavored jets, c for c-flavored jets, or l for light-flavored jets including jets originating from gluons. The scale factors used in this analysis are taken from Ref. [78]. The MC efficiencies are measured in the final event selection defined in Chap. 5.

The scale factors for untagged jets, $SF_{\bar{T}}$, are derived from the scale factors for tagged jets, SF_T , in Equation 6.4.

$$SF_{\bar{T}}(f = \{b, c, l\}, p_T) = \frac{1 - SF_T(f, p_T)\epsilon_f^{\text{MC}}(p_T)}{1 - \epsilon_f^{\text{MC}}(p_T)} \quad (6.4)$$

Although the number of events with tagged jets does not change, their contribution to the yields is modified by the event weight when allowing the scale factors to vary within their errors. The use of event weights thus allows for effective bin migration between the categories containing b-tags and those without.

6.6 Top-Tagging Efficiency and Mistag Rate

We use the CMS Top Tagging algorithm [54, 55] to identify events with boosted hadronically decaying top quarks through jet substructure and jet mass observables. The efficiency of this algorithm has been measured in both data and MC [79], however the scale factors derived from this measurement cannot be applied to this analysis due to significant overlap between the measured sample and our signal region. Additionally, the scale factors were derived for a version of the algorithm that uses a different τ_{32} cut than the one used in this analysis. Since no suitable scale factors were available, and producing a control sample outside the signal region of this analysis was unfeasible, we performed an *in situ* measurement of the top-tagging efficiency scale factor. This was done by setting the scale factor to 1.0 with a flat 20% error and calculating event weights based on the same method for b-tagging in Equation 6.2. The scale factor was determined to be 0.90 ± 0.05 by the maximum-likelihood fit described in Chap. 8.

To study the mistag rate of the CMS Top Tagging algorithm a sample outside our signal region that is enriched with W+jets events is selected. The kinematic selection is similar to that of the Z' analysis with the exception of a veto on events that have an AK5 jet tagged as a b-jet at the CSV-loose operating point which removes most of the $t\bar{t}$ from the sample. The event is required to contain a CA8 jet with $p_T > 400$ GeV as a candidate for the mistagged top jet. To ensure we are outside of the signal region we invert the χ^2 criterion on the leptonic leg of the $t\bar{t}$ decay to ensure the probe jet is not biased.

The selection cuts are:

1. Require the event to pass the unrescaled, single muon Mu40 trigger with no isolation requirement on the muon
2. At least one good primary vertex
3. At least two AK5 jets with $p_T > 50$ GeV and $|\eta| < 2.4$
4. At least one CA8 jet with $p_T > 400$ GeV as the topjet candidate
5. $\chi_{\text{lep}}^2 > 50$
6. Veto on additional leptons: veto events with an electron, second muon.
7. Lepton 2D-cut: $\Delta R(\mu, \text{closest jet}) > 0.5$ or $p_T^{\text{rel}}(\mu, \text{closest jet}) > 25$ GeV, measured w.r.t. the closest jet with $p_T > 25$ GeV
8. $H_T^{\text{lep}} > 150$ GeV, where $H_T^{\text{lep}} = \cancel{E}_T + p_T^\mu$
9. $\cancel{E}_T > 20$ GeV.
10. No AK5 jets b-tagged at the CSV-loose working point.

The resulting control sample is dominated by W+jets events (85%) with a small residual $t\bar{t}$ contribution of 3.5%. Fig. 6.9 shows plots for the control sample for relevant kinematic variables. The fraction of events in the control sample that contain a jet tagged by the CMS Top Tagging algorithm is small, with 73% coming from W+jets and 26% coming from $t\bar{t}$. The mistag rate efficiency in MC is defined as the rate of events containing a jet tagged by the CMS Top Tagging algorithm as shown in Equation 6.5. The mistag rate efficiency measurement in data is corrected by the $t\bar{t}$ MC prediction in Equation 6.6 since 26% of

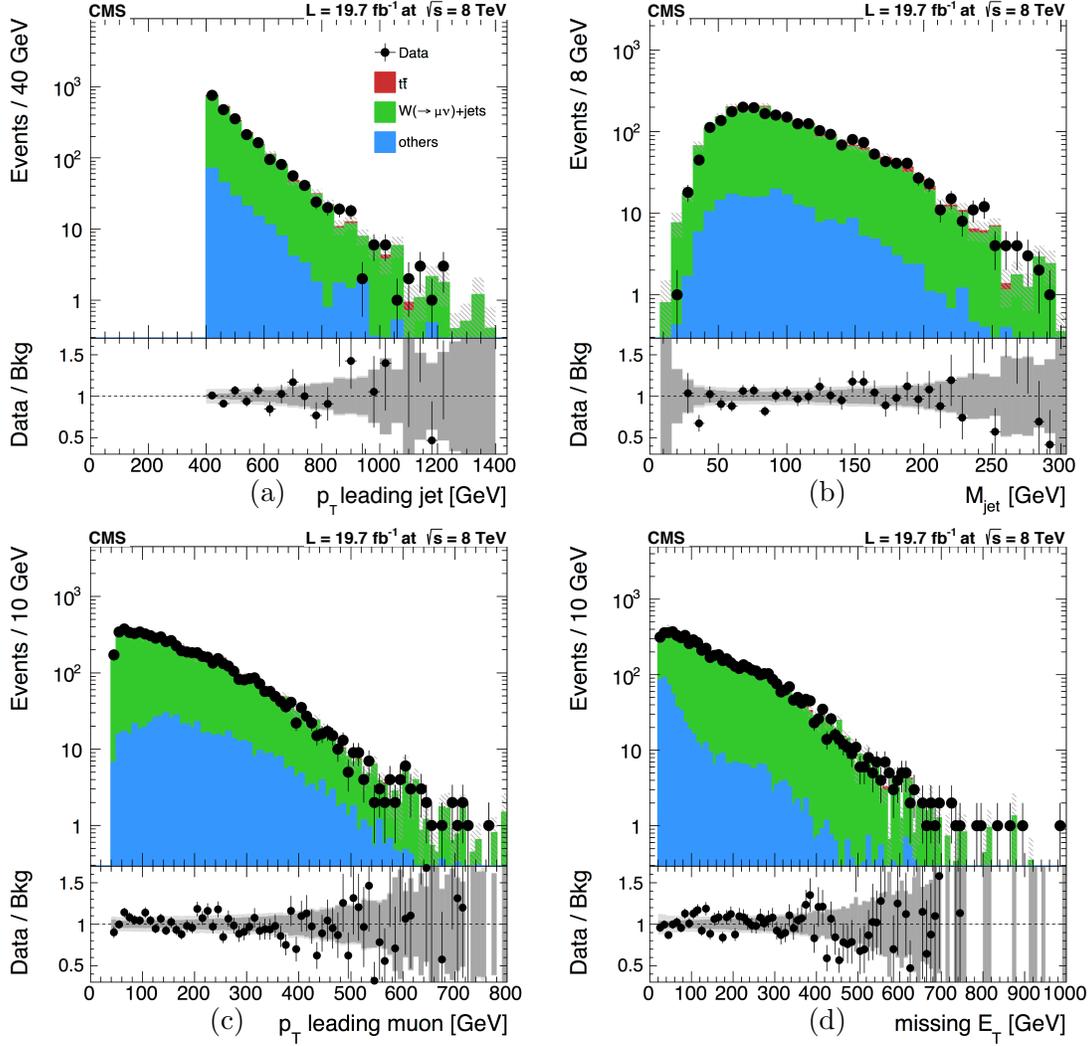


Figure 6.9: Data/MC comparisons of relevant kinematic variables in the W+jets control sample used to measure the CMS Top Tagging algorithm’s mistag rate: p_T of the p_T leading topjet candidate (a), mass of the leading topjet candidate (b), p_T of the muon (c), and \cancel{E}_T (d). The sum of MC backgrounds is normalized to data. The gray band corresponds to the statistical uncertainty of the simulation.

events that are top-tagged are expected to come from $t\bar{t}$ contamination. The ratio of these mistag rates gives the mistag scale factor for the CMS Top Tagging algorithm.

$$\varepsilon_{mistag}^{MC} \equiv \frac{N_{pass,tagged}^{MC}}{N_{pass}^{MC}} \quad (6.5)$$

$$\epsilon_{mistag}^{data} \equiv \frac{N_{pass,tagged}^{data} - N_{pass,tagged}^{t\bar{t}MC}}{N_{pass}^{data} - N_{pass}^{t\bar{t}MC}} \quad (6.6)$$

This mistag rate in both data and MC is very small, so the measured mistag efficiencies have large statistical error. Parameterizing the scale factors as a function of the kinematic variables becomes impossible due to the low statistics, so we adopt a flat scale factor. The mistag rate obtained in data is 0.0127 ± 0.003 , and the one for MC is 0.0152 ± 0.001 . This results in a scale factor consistent with unity, with a value of

$$SF_{mistag} = 0.83 \pm 0.21.$$

The event weight is determined by the top-tagging efficiency, the top-tagging efficiency scale factor, the top-tagging mistag rate and the top-tagging mistag rate scale factor through Equation 6.7.

$$w = \prod_{i=1}^{N_{tagged}} SF(f_i) \times \prod_{j=1}^{N_{untagged}} \frac{1 - SF(f_j)\epsilon^{MC}(f_j)}{1 - \epsilon^{MC}(f_j)} \quad (6.7)$$

The first product loops over all CA8 jets in the event that are tagged by the CMS Top Tagging algorithm, i , while the second product loops over the untagged CA8 jets. Only CA8 jets with $p_T > 400$ GeV are considered for both loops. If the jet can be matched to a hadronically decaying top quark at the generator level then both the scale factor, SF , and the efficiency, ϵ^{MC} , are set to the top tagging efficiency and related scale factor. If the jet is not matched to a hadronically decaying top quark then these values are set to those of the mistag rate and related scale factor. The mistag rate in MC is taken as 0.0152 with its related scale factor of 0.83, as measured in the W+jets control sample. The top-tagging efficiency is taken as the fraction of $t\bar{t}$ events in the W+jets control sample that are top-tagged, 0.199, with a scale factor of 1.0. Since the scale factor is set to 1.0, the untagged scale factor derived from this efficiency is also 1.0. Only when the scale factor is varied by $\pm 20\%$

to generate templates to measure the scale factor *in situ* during the limit setting procedure does the choice in efficiency affect the untagged jets, but this effect is minimal.

Although the number of events with tagged jets does not change, their contribution to the yields is modified by the event weight when allowing the scale factors to vary within their errors. The use of event weights thus allows for effective bin migration between the categories containing top-tags and those without.

7. EVENT RECONSTRUCTION AND FINAL SELECTION

In order to compare predicted event signals to signals recorded by the CMS detector we must use a method of reconstructing the components of the $t\bar{t}$ decay from the event objects. After the kinematic selections detailed in Chap. 5 we are left with events that contain a number of reconstructed physics objects that are compatible with the semileptonic decay products of a $t\bar{t}$ pair, namely a charged lepton, \cancel{E}_T , and at least 2 high- p_T jets. To reconstruct the $t\bar{t}$ system we must separately reconstruct both the hadronic top, t_{had} , and leptonic top, t_{lep} . To fully reconstruct the leptonic leg of the decay we must reconstruct the neutrino momentum. Once the neutrino momentum is found the jets in the event must be assigned to either t_{lep} , t_{had} , or neither. The jet assignment is done using a χ^2 approach where all possibilities are considered and the assignment that yields the reconstructed top mass nearest the true top mass is taken. A final cut to reduce W +jets background is made and the events are split into exclusive categories.

7.1 Neutrino Reconstruction

On the leptonic leg of the $t\bar{t}$ system we have a leptonically decaying W boson which decays into a lepton and a neutrino. The lepton leaves a signature in the detector and we derive its momentum and energy from this, but the neutrino leaves the detector without a trace. We need the neutrino momentum in order to reconstruct t_{lep} .

In this analysis we interpret the momentum imbalance of the event, \cancel{E}_T , as the transverse component of the neutrino momentum. In order to solve for the longitudinal component of the neutrino momentum we assume the W boson was on shell. In that case we can write the

4-vector momentum conservation:

$$\mathbf{P}_W = \mathbf{P}_l + \mathbf{P}_\nu, \quad \mathbf{P}^W = \mathbf{P}^l + \mathbf{P}^\nu \quad (7.1)$$

Multiplying these equations together we get:

$$\mathbf{P}_W \cdot \mathbf{P}^W = (\mathbf{P}_l + \mathbf{P}_\nu) \cdot (\mathbf{P}^l + \mathbf{P}^\nu) = \mathbf{P}_l \cdot \mathbf{P}^l + \mathbf{P}_\nu \cdot \mathbf{P}^\nu + 2\mathbf{P}_l \cdot \mathbf{P}^\nu \quad (7.2)$$

Which simplifies to:

$$M_W^2 = M_l^2 + M_\nu^2 + 2(E_l E_\nu - \vec{p}_l \cdot \vec{p}_\nu) \quad (7.3)$$

The lepton and neutrino mass terms are negligibly small compared to their momentum terms, and can be discarded. The neutrino energy, E_ν , can be rewritten in terms of its momentum as $E_\nu = \sqrt{p_{T,\nu}^2 + p_{z,\nu}^2}$. Rearranging the terms we have:

$$\begin{aligned} \frac{M_W^2}{2} &= E_l \sqrt{p_{T,\nu}^2 + p_{z,\nu}^2} - (p_{x,l} p_{x,\nu} + p_{y,l} p_{y,\nu} + p_{z,l} p_{z,\nu}) \\ &= E_l \sqrt{p_{T,\nu}^2 + p_{z,\nu}^2} - (p_{T,l} p_{T,\nu} \cos(\Delta\phi_{l,\nu}) + p_{z,l} p_{z,\nu}) \end{aligned} \quad (7.4)$$

Let $\alpha = \frac{M_W^2}{2} + p_{T,l} p_{T,\nu} \cos(\Delta\phi_{l,\nu})$ and rearrange the terms:

$$\alpha + p_{z,l} p_{z,\nu} = E_l \sqrt{p_{T,\nu}^2 + p_{z,\nu}^2} \quad (7.5)$$

Square both sides:

$$\alpha^2 + p_{z,l}^2 p_{z,\nu}^2 + 2\alpha p_{z,l} p_{z,\nu} = E_l^2 (p_{T,\nu}^2 + p_{z,\nu}^2) \quad (7.6)$$

This is a quadratic equation that can be solved for $p_{z,\nu}$, with the solution being:

$$p_{z,\nu} = \frac{\alpha p_{z,l}}{p_{T,l}^2} \pm \sqrt{\frac{\alpha^2 p_{z,l}^2}{p_{T,l}^4} - \frac{E_l^2 p_{T,\nu}^2 - \alpha^2}{p_{T,l}^2}} \quad (7.7)$$

The longitudinal component of the neutrino momentum has either zero, one, or two real solutions. If no real solutions are possible, we take the real part of the imaginary solution. For events with two real solutions we consider both, effectively doubling the number of reconstruction hypotheses.

7.2 Jet Assignment

The reconstruction of \mathbf{P}_{lep} and \mathbf{P}_{had} , the 4-vectors of t_{lep} and t_{had} , is done using a χ^2 approach. The objective is to assign each object in the event to either t_{lep} , t_{had} , or neither of the two. Because of the boosted topology expected in the event we do not assume that we can find a match for each parton. Therefore the hadronic side of the decay, t_{had} , and its decay products, $q\bar{q}b$, will be reconstructed as at least one jet. A special case is made for events containing a CA8 top-tagged jet. In these events the hadronic leg of the event t_{had} is set to the top-tagged jet, and all jets with $\Delta R < 1.3$ from the top-tagged jet are removed from the consideration of jet assignment to t_{lep} . The leptonic side of the decay, t_{lep} , will have one b quark which reconstructs as a single jet, one lepton, and one neutrino. We reconstruct t_{lep} with exactly one jet, the lepton, and the neutrino momentum solutions. A list of reconstruction hypothesis is built where exactly one jet is assigned to t_{lep} , and at least one jet is assigned to t_{had} . This list contains $N_{\nu sol} \times N_{jet} \times (2^{(N_{jet}-1)} - 1)$ possible object assignments for the $t\bar{t}$ system with \mathbf{P}_{lep} and \mathbf{P}_{had} given by the sum of the 4-vectors of the objects assigned to t_{lep} and t_{had} .

We eventually select one hypothesis per event with the criterion for selection being based on the idea that the reconstructed top quark masses should be close to the true top quark mass. For each hypothesis the invariant mass of t_{lep} and t_{had} , M_{lep} and M_{had} , is calculated and the quantity χ^2 is calculated. We select the hypothesis that minimizes the χ^2 quantity.

$$\chi^2 = \left[\frac{M_{lep} - \bar{M}_{lep}}{\sigma_{M_{lep}}} \right]^2 + \left[\frac{M_{had} - \bar{M}_{had}}{\sigma_{M_{had}}} \right]^2 \quad (7.8)$$

The parameters $\bar{M}_{lep}, \bar{M}_{had}, \sigma_{M_{lep}}$, and $\sigma_{M_{had}}$ are chosen such that χ^2 is small for the correct hypothesis. They are derived from a subset of MC events in which all 4 quarks from the $t\bar{t}$ system decay are matched to a reconstructed jet in the event, using $\Delta R(parton, jet) < 0.3$ as the matching criterion, and the $t\bar{t}$ system is reconstructed from these jets with the correct assignments. Multiple partons can be matched to a single jet, allowing for the boosted topology of our signal. About 5-6% of the SM $t\bar{t}$ MC events are “matchable” after the selection cuts in both channels.

The M_{lep} and M_{had} distributions for “matchable” events are fitted with a Gaussian to find the values for $\bar{M}_{lep}, \bar{M}_{had}, \sigma_{M_{lep}}$, and $\sigma_{M_{had}}$. Examples of the distributions used for the fit are shown in Fig. 7.1. The mean and width of the distributions is found to be largely independent of the Z' mass and the number of reconstructed jets assigned to the hadronic leg of the decay. In this analysis the values of $\bar{M}_{lep} = 174$ GeV, $\sigma_{M_{lep}} = 18$ GeV, $\bar{M}_{had} = 181$ GeV, and $\sigma_{M_{had}} = 15$ GeV are used in the χ^2 calculation.

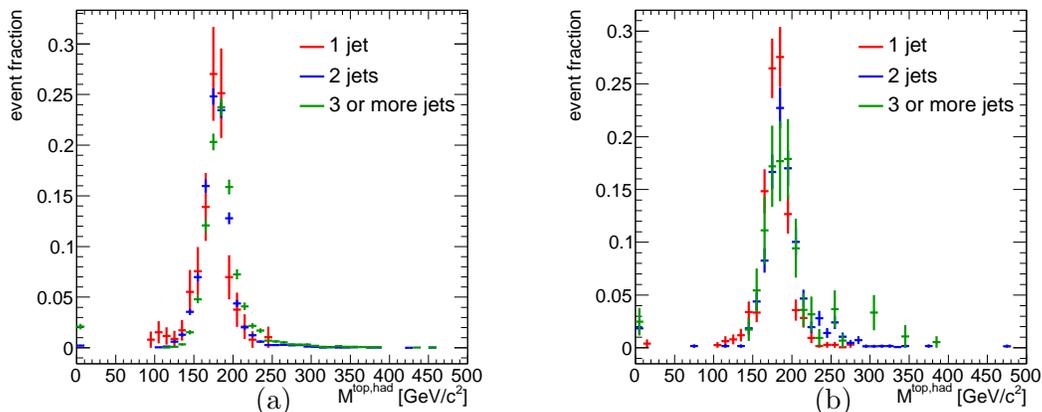


Figure 7.1: Two examples of the reconstructed hadronic top mass, M_{had} , in correctly matched events where all three final state quarks, $q\bar{q}b$, are matched to a jet are shown. Both SM $t\bar{t}$ with $M_{t\bar{t}} < 1$ TeV (a) and a 1% width 2 TeV Z' signal MC (b) are shown with events separated by the number of reconstructed jets. The distributions are well modeled by Gaussian functions and the fit parameters are used in the χ^2 calculation.

Figure 7.2 shows the χ^2 distribution for data and the MC prediction for the electron and muon channels as well as the contribution to the χ^2 distribution from different $W+n$ parton samples that pass the event selection. As expected, χ^2 is small for signal and $t\bar{t}$ events and

larger for the dominant background of W +jets events. The peak around $\chi^2 = 110$ is due to events reconstructing t_{lep} around the top mass and incorrectly reconstructing t_{had} . The shoulder around $\chi^2 = 150$ is primarily from W +1 parton events where the second selected jet is expected to come from soft radiation. In these events both t_{had} and t_{lep} are not able to be reconstructed, resulting in the large χ^2 value. The triangular cut described in Sec. 5.3 removes a significant portion of the W +1 parton events so this feature is less pronounced in that channel.

7.3 Final Selection and Event Yields

Events are required to have a minimum χ^2 value smaller than 50, which reduces contributions from W +jet background processes and enhances the sensitivity of the analysis. The value of χ^2 to cut at was determined by studying the expected limits for several different cut values as shown in Appendix A. As Fig. 7.2 and Fig. B.18 show, almost all the signal sample lies below $\chi^2 = 50$. It was found that the sensitivity of the analysis increases monotonically as the χ^2 cut increases until the signal depletes at $\chi^2 = 50$.

Furthermore, events are separated into three categories based on the event contents. Events that contain a CA8 top-tagged jet are separated into a top-tag category, “1top”. The signal purity in this category is very high, and it contributes to the majority of the analysis sensitivity. Events without a CA8 top-tagged jet are separated according to the number of b-tagged jets. Events with no b-tags are put in the “0top0btag” category, and events with one or more are put in the “0top1btag” category. The separation of events based on b-tagging yields categories with different contributions from W +jets background. This provides a better handle on constraining the systematic uncertainties. The optimal choice of event categorization was also studied at the same time as the χ^2 optimization with the results presented in Appendix A.

The number of events left after the final selection for both the electron+jets and muon+jets channel are summarized in Tab. 7.1. Z' samples are normalized to a cross-section of 1 pb.

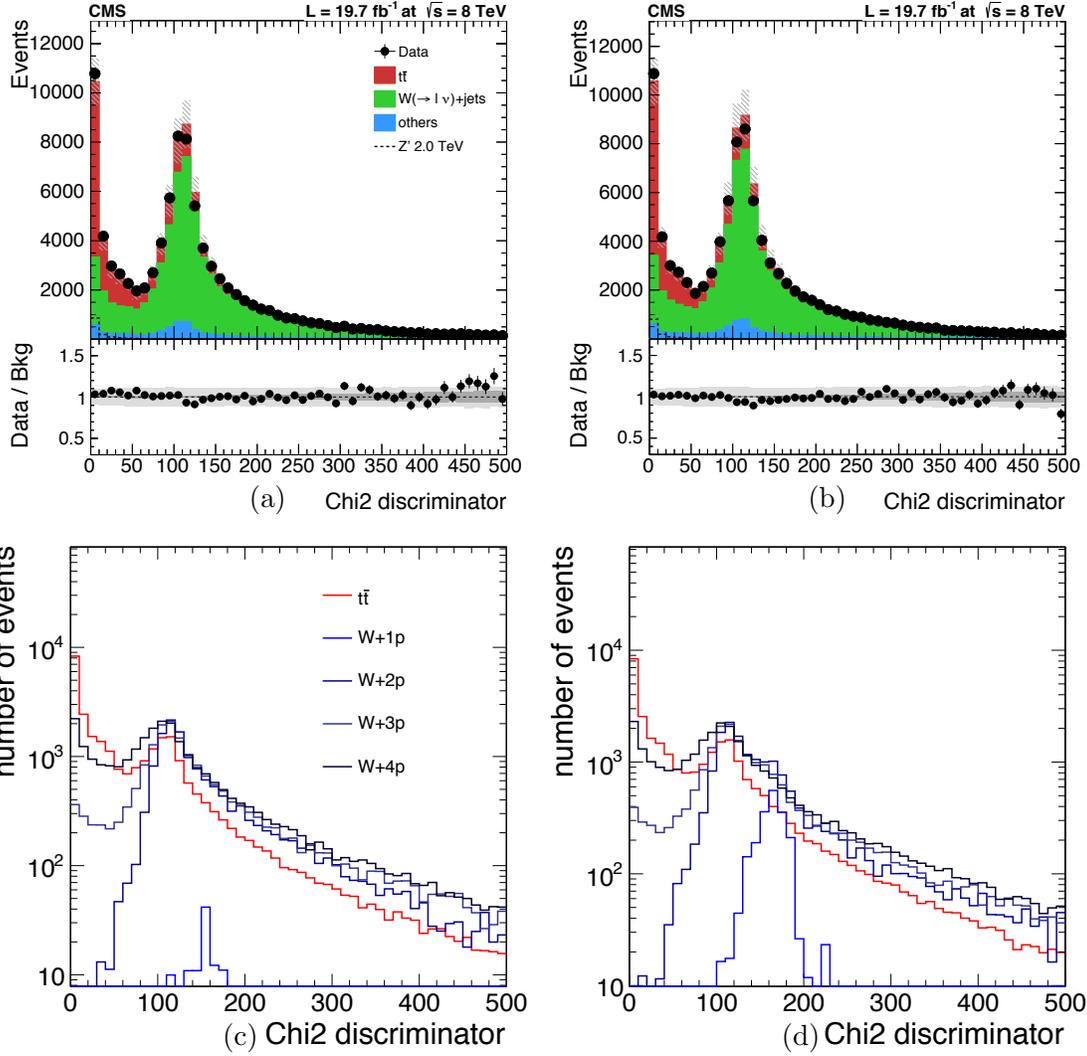


Figure 7.2: Distributions of the χ^2 discriminator for the (a) electron and (b) muon channels after the kinematic selection described in Chap. 5. Also shown are the contributions from different $W+n$ parton samples for the (c) electron and (d) muon channels. The shoulder at ≈ 150 present in the muon channel is suppressed in the electron channel due to the contribution from the $W+1$ parton process being much smaller due to the triangle cut. The yields of the background processes are normalized to data using scale factors described in Chap. 8. The shaded band shows the change in the yield due to systematic uncertainties also described in Chap. 8.

muon channel

sample	0-toptag – 0-btag	0-toptag – 1-btag	1-toptag
Z' ($M = 1$ TeV, $w = 1\%$)	85.9	437.3	81.6
Z' ($M = 2$ TeV, $w = 1\%$)	223.0	467.3	432.4
Z' ($M = 3$ TeV, $w = 1\%$)	328.6	447.8	381.7
Z' ($M = 1$ TeV, $w = 10\%$)	88.0	404.6	80.5
Z' ($M = 2$ TeV, $w = 10\%$)	183.3	437.5	350.2
Z' ($M = 3$ TeV, $w = 10\%$)	189.7	364.2	258.6
g_{KK} ($M = 1$ TeV)	82.5	348.8	69.2
g_{KK} ($M = 2$ TeV)	143.8	357.0	255.5
g_{KK} ($M = 3$ TeV)	114.1	280.7	181.1
$t\bar{t}$	3099.8 ± 144.8	10856.4 ± 501.3	440.1 ± 29.2
W + jets (+light)	5379.6 ± 343.1	227.1 ± 18.3	36.4 ± 4.6
W + jets (+c)	1560.5 ± 375.9	437.7 ± 106.4	16.9 ± 5.1
W + jets (+b)	118.5 ± 28.4	280.8 ± 65.6	3.2 ± 1.4
single top	258.1 ± 65.6	762.4 ± 191.5	9.5 ± 3.3
Z + jets	304.9 ± 149.5	55.4 ± 27.2	2.2 ± 1.2
diboson	135.0 ± 27.3	27.5 ± 5.8	0.9 ± 0.4
Total Background	10856.4 ± 555.1	12647.3 ± 552.0	509.2 ± 30.3
DATA	10099	12510	493

electron channel

sample	0-toptag – 0-btag	0-toptag – 1-btag	1-toptag
Z' ($M = 1$ TeV, $w = 1\%$)	104.2	462.1	90.6
Z' ($M = 2$ TeV, $w = 1\%$)	224.9	495.3	460.2
Z' ($M = 3$ TeV, $w = 1\%$)	286.7	448.0	343.2
Z' ($M = 1$ TeV, $w = 10\%$)	90.7	420.4	84.4
Z' ($M = 2$ TeV, $w = 10\%$)	178.7	461.1	366.2
Z' ($M = 3$ TeV, $w = 10\%$)	169.5	412.7	264.4
g_{KK} ($M = 1$ TeV)	93.6	346.8	65.4
g_{KK} ($M = 2$ TeV)	137.1	367.5	268.3
g_{KK} ($M = 3$ TeV)	111.4	294.6	178.5
$t\bar{t}$	2918.0 ± 136.4	10586.8 ± 488.8	447.3 ± 29.7
W + jets (+light)	5268.0 ± 335.8	243.8 ± 19.4	37.5 ± 4.6
W + jets (+c)	1483.7 ± 357.4	398.4 ± 96.8	12.1 ± 3.9
W + jets (+b)	128.1 ± 30.5	289.2 ± 67.5	3.5 ± 1.4
single top	264.6 ± 67.2	723.8 ± 181.9	10.6 ± 3.5
Z + jets	244.2 ± 119.8	45.5 ± 22.4	1.9 ± 1.0
diboson	124.1 ± 25.1	29.7 ± 6.2	1.1 ± 0.4
Total Background	10430.7 ± 528.7	12317.1 ± 535.6	514.1 ± 30.6
DATA	10204	12157	465

Table 7.1: Number of observed and expected events in the muon and electron channels. The yields are modified by scale factors summarized in Tab. 7.2. The uncertainty on each background includes both the statistical error and the posterior uncertainties from Tab. 7.2.

process	best-fit value	prior uncertainty [%]	posterior uncertainty [%]
$t\bar{t}$	0.92	15	4.2
W +jets (light flavour)	0.98	9	6.3
W +jets (c flavour)	1.26	23	24
W +jets (b flavour)	0.99	23	23
single top	1.09	23	25
Z +jets	0.87	50	49
diboson	1.02	20	20
top-tagging scale factor	0.90	unconstrained	5

Table 7.2: Best-fit normalization scale factors for each background process obtained from a maximum-likelihood fit across all three event categories in both channels simultaneously. The top-tagging scale factor affects the normalization of $t\bar{t}$ background in the “1top” category, whereas the other scale factors apply to all categories equally. Also shown are the prior and posterior constraints on the rate uncertainty. These scale factors are used to normalize MC to data for all figures.

The total yield of MC samples are normalized to data using scale factors derived in a maximum likelihood fit to $M_{t\bar{t}}$ across both channels simultaneously. This maximum likelihood fit is the same that will be discussed in the next chapter with only rate uncertainties considered. Table 7.2 summarizes these normalization scale factors. Figure 7.3 shows the invariant mass of the reconstructed $t\bar{t}$ system for both data and MC in the three categories. In addition several kinematic distributions are plotted in Appendix B. There is excellent agreement between the MC and data, with no significant deviations from the SM predictions observed.

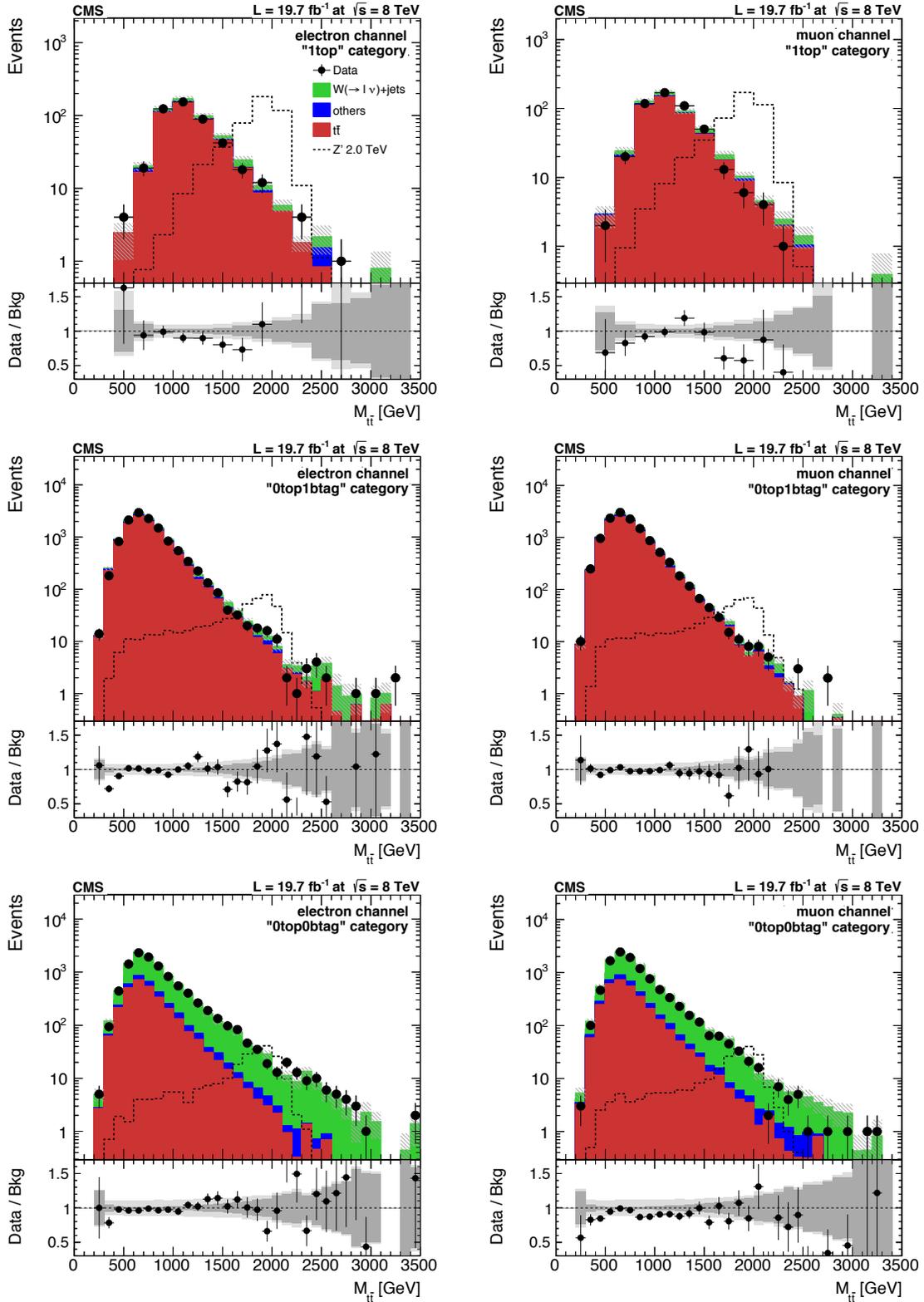


Figure 7.3: Reconstructed $M_{t\bar{t}}$ distributions in the electron (left) and muon (right) channels. The three categories are “1top” on the top, “0top1btag” in the middle, and “0top0btag” on the bottom. The yields of the background processes are normalized to data using scale factors described in Tab. 7.2. The shaded band shows the change in the yield due to systematic uncertainties described in Sec. 8.1. The signal is normalized to a cross-section of 1 pb.

8. STATISTICAL ANALYSIS

We rely on a Bayesian statistical method to extract the 95% confidence level (CL) limits on the $Z' \rightarrow t\bar{t}$ cross-section. The invariant mass distribution of the reconstructed $t\bar{t}$ system, $M_{t\bar{t}}$, is used in a template based statistical evaluation. We use the `theta` software [80] to build and evaluate the statistical model.

The method used to extract the limits on the signal models is a binned-likelihood statistical model which is based on Poisson statistics. Each region, or bin, in the reconstructed $M_{t\bar{t}}$ distribution can be treated as a separate counting experiment. Events measured in data either fall into one bin or another, and the average number of events that fall into a certain bin is known from the SM MC prediction. This expected number of events can be modified by the inclusion of a signal process. The probability of an event falling into a bin is proportional to the bin size, and the probability of events falling into a bin with zero width is zero. Therefore each bin can be treated as a Poisson experiment with the probability of observing n events in a bin with a predicted event count of μ is given by the Poisson formula:

$$P(n; \mu) = \frac{e^{-\mu} \cdot \mu^n}{n!} \tag{8.1}$$

The likelihood function is then given by the product of the Poisson probabilities in each bin, i .

$$L(\mathbf{n}|\boldsymbol{\mu}) = \prod_i \frac{e^{-\mu_i} \cdot \mu_i^{n_i}}{n_i!} \tag{8.2}$$

The vectors \mathbf{n} and $\boldsymbol{\mu}$ are the observed and predicted event yields in all the bins. The index, i , enumerates all the statistically independent counting experiments, which in our case are the

bins in the 6 $M_{i\bar{i}}$ distributions for the 2 channels and 3 categories. The predicted event yields for each bin, μ_i , are given by the sum of the different physics processes, called templates. This yield is given by

$$\mu_i = \sum_k \beta_k \cdot T_{k,i} , \quad (8.3)$$

where k denotes the different physics processes and $T_{k,i}$ is the content of the i -th bin in the k -th process. The term β_k scales the overall normalization of the k -th template and correspond to normalization uncertainties. Table 7.2 summarizes the log-normal priors used for these normalization uncertainties. Since the signal samples are scaled to a cross-section of 1 pb, $\beta_{Z'}$ and β_{kk} can be interpreted as the cross-section for the Z' and kk -gluon signal samples in units of pb. μ_i can be further modified by systematic uncertainties with the inclusion of nuisance parameters, θ_j .

$$\mu_i = \sum_k \prod_j \lambda_{ijk}(\theta_j) \cdot \beta_k \cdot T_{k,i} \quad (8.4)$$

The event yields are modified by $\lambda_{ijk}(\theta_j)$ which allow the effects of the j -th systematic uncertainty to be accounted for. This treatment of the effect of the systematic uncertainties requires shifted templates, T_k^\pm , to be produced for each process, T_k . These shifted templates reflect the effect of the $\pm 1\sigma$ variation of the systematic uncertainty. The nuisance parameters follow normal distributions with a mean of 0 and a variance of 1, $\theta_j = \mathcal{N}(0, 1)$, and are used to interpolate between the templates. We use a quadratically matched asymmetric log-normal distribution for the parameterization of $\lambda_{ijk}(\theta_j)$.

$$\lambda_{ijk}(\theta_j) = \exp\left[\frac{1}{2}(\sigma_{ijk}^+ + \sigma_{ijk}^-)\theta_j + \frac{1}{2}(\sigma_{ijk}^+ - \sigma_{ijk}^-)\theta_j^2\right] \quad (8.5)$$

$$\sigma_{ijk}^\pm = \pm \ln(T_{ijk}^\pm/T_{ik}) \quad (8.6)$$

The parameters σ_{ijk}^\pm are derived from the shifted templates for each uncertainty j . This parameterization reduces to a scale factor of T_{ijk}^\pm/T_{ik} when the associated nuisance parameter $\theta_j = \pm 1$, scaling T_k to T_{jk}^\pm .

The nuisance parameter associated with the top-tagging efficiency scale factor uncertainty is treated differently than the rest and given a flat prior instead of Gaussian one. There is no reason to believe the top-tagging efficiency scale factor would be much larger or smaller than one, so we set the scale factor to one with a 20% uncertainty. The flat prior allows the top-tagging efficiency scale factor to float freely between 0.8 – 1.2, and the scale factor to be measured *in situ* during the limit setting procedure.

The ‘‘Barlow-Beeston lite’’ [81] method is used to account for the statistical uncertainty due to the finite size of MC samples. Barlow and Beeston [82] developed the original method which was modified to use only one additional nuisance parameter for each bin. This additional nuisance parameter, ν_i , has a Gaussian prior with a mean of zero and a width that corresponds to the statistical uncertainty on the MC prediction in the bin. To improve background estimation and stability in the numerical maximization of the likelihood we rebin the tails of the $M_{t\bar{t}}$ distributions. This rebinning is done in such a way that the statistical uncertainty from the sum of the SM MC samples in each bin is less than 30%.

Too many nuisance parameters make the numerical maximization of the likelihood function unstable, so we use the ‘‘profile’’-likelihood function, L_p , where the newly introduced nuisance parameters have been ‘‘maximized out’’ analytically.

$$L_p(\beta_k, \theta_j | \text{data}) := \max_{\nu_i} L(\beta_k, \theta_j, \nu_i | \text{data}). \quad (8.7)$$

The Bayesian 95% confidence level upper limit is the value $\hat{\beta}_{signal}$ which satisfies:

$$\int_0^{\hat{\beta}_{signal}} d\beta_{signal} \int d(\beta_k, \theta_j) L_p(\beta_{signal}, \beta_k, \theta_j) \pi(\beta_{signal}, \beta_k, \theta_j) = 0.95 \quad (8.8)$$

The integral in Equation 8.8 is evaluated with a Markov chain Monte Carlo method.

To determine the expected upper limits for the signal cross-sections we generate several toy experiments with no signal. Each toy experiment is generated by first generating a set of random numbers for $\theta_j \propto \mathcal{N}(0, 1)$. These values are used to calculate μ_i . Next, a set of random numbers is generated that are used as the counts of the simulated experiments using the Poisson distribution: $n_i = \text{Pois}(\mu_i)$. The expected limit for each signal model is determined by the mean upper limit for the toy experiments. The median central 68% and 95% of the upper limits on each signal model for the toy experiments define the $\pm 1\sigma$ and $\pm 2\sigma$ bands on the expected limit. The observed limit is obtained by evaluating Equation 8.8 for data.

8.1 Systematic Uncertainties

We include several sources of systematic uncertainty in the statistical model described in the previous section. These uncertainties affect both the overall normalization of the various background processes and the shape of the reconstructed $M_{t\bar{t}}$ distribution in the form of shifted templates. The uncertainties included are described below and summarized in Tab. 8.1.

8.1.1 Normalization Uncertainties

A 15% uncertainty on the $t\bar{t}$ cross-section is used [83]. W +jets production cross-sections have a 9% uncertainty applied for light flavor jets [84] and a 23% uncertainty applied for heavy flavor jets [85]. The Z +jets cross-section has been measured to an accuracy of 7% and 9% for two and three additional jets [86], but the measurement was done in a significantly

different phase-space than our analysis. We therefore take a conservative uncertainty of 50% on the Z +jets cross-section. For single-top processes we use the uncertainty measured in the tW -channel at 23% [87] since this channel has the largest contribution. WW , WZ , and ZZ diboson production cross-section uncertainty is taken to be 20% [88, 89].

In addition to cross-section uncertainties we apply flat normalization uncertainties to both the luminosity and the electron trigger efficiency. We vary the luminosity by 2.6% [90]. From the results of the trigger efficiency study in Sec. 6.2.1 we apply a 1% uncertainty to the electron channel.

8.1.2 Shape Uncertainties

The following systematic uncertainties affect both the normalization and the shape of the $M_{t\bar{t}}$ distributions. The shifts due to these uncertainties are used in the templates used in the statistical analysis.

- **Muon ID and Trigger:** As described in Sec. 6.2 we apply muon ID and trigger efficiency scale factors. Templates for each MC sample are generated by varying these scale factors by $\pm 1\sigma$. Systematic uncertainties of the tag-and-probe method which affect the normalization by 0.5% and 0.2% are taken into account.
- **Electron ID:** As described in Sec. 6.2 we apply electron ID scale factors. Templates for each MC sample are generated by varying these scale factors by $\pm 1\sigma$.
- **Jet Energy Correction (JEC):** The uncertainties in the JECs described in Sec. 6.3 are on the order of a few percent. Templates for each MC sample are generated by varying the JECs by $\pm 1\sigma$ for both the AK5 and CA8 jets simultaneously. The variations are propagated to \cancel{E}_T .
- **Jet Energy Resolution (JER):** The JER correction described in Sec. 6.4 is varied by $\pm 1\sigma$ simultaneously for AK5 and CA8 jets to generate the templates for each MC sample. The variations are propagated to \cancel{E}_T .

- **b-tagging Efficiency:** As described in Sec. 6.5 we apply b-tagging efficiency scale factors. These scale factors are varied by $\pm 1\sigma$ to generate the templates. These scale factors are measured for jets with p_T up to 800 GeV. For jets above this energy we use the scale factor at the value of $p_T = 800$ GeV with twice the quoted uncertainty. The uncertainties of SF_b and SF_c are treated as fully correlated with SF_c being twice the uncertainty of SF_b .
- **b-tagging Mistag Rate:** As described in Sec. 6.5 we apply b-tagging mistag rate scale factors. The uncertainty on SF_l is taken as uncorrelated to SF_b and SF_c . This scale factor is measured for jets with p_T up to 800 GeV. For jets above this energy we use the scale factor at the value of $p_T = 800$ GeV with double the quoted uncertainty. The templates are generated by varying this scale factor by $\pm 1\sigma$.
- **Top-Tagging Efficiency:** As described in Sec. 6.6 we set the top-tagging efficiency scale factor to unity due to this scale factor being unknown for our signal region. Templates are generated by setting this scale factor to 0.8 and 1.2 and using a flat prior, with the goal of using the interpolated templates to measure this scale factor *in situ* during the limit setting procedure.
- **Top-Tag Mistag Rate:** As described in Sec. 6.6 we measure the mistag rate of the CMS Top Tagging algorithm in a side-band region dominated by W +Jets events. The scale factor obtained is 0.83 ± 0.21 . We vary this scale factor by $\pm 1\sigma$ to obtain templates .
- **Pileup:** As described in Sec. 6.1 each MC samples is reweighted to match the instantaneous luminosity profile in data. The luminosity profile in data uses a minimum-bias cross-section of 69.4 mb. We vary the minimum-bias cross-section by $\pm 5\%$ to generate the templates.
- **Parton Distribution Functions (PDFs):** To evaluate the uncertainty associated with the PDFs used in MC event generation we reweight events according to the uncertainties of the eigenvectors of the PDF set used. We generate the new event

weights with the shifts produced by individual eigenvectors being added in quadrature for each bin. The resulting uncertainty is taken to be fully correlated across both channels.

- **Scale Uncertainty:** The Q^2 scale is varied to estimate the effect of higher order corrections in W/Z +jets and $t\bar{t}$ samples. The samples were generated with varied scales ($\mu = Q/2$ and $\mu = 2Q$). Due to limited statistics in the Q^2 systematic samples and in order to not underestimate the error, the largest variation in each bin is taken as a symmetric uncertainty for that bin. Due to low statistics in W +heavy flavor jets samples this variation is evaluated over all bins before the χ^2 cut and taken as a flat uncertainty.
- **Matching Uncertainty:** The jet matching threshold is varied in the W/Z +jets samples to estimate the effect of extra hard parton radiation. Due to low statistics in W +heavy flavor jets samples this variation is evaluated over all bins before the χ^2 cut and taken as a flat uncertainty.

The effect on the overall normalizations in each category across the two channels for each of these uncertainties is shown in Tables 8.2–8.3. Appendix C shows the effect on the shape of the $M_{t\bar{t}}$ distributions due to these uncertainties.

Source of systematic uncertainty	Uncertainty	Type
$t\bar{t}$ cross section	15%	Normalization
Single top cross section	23%	Normalization
diboson cross section	20%	Normalization
W +light jets cross section	9%	Normalization
W +heavy jets cross section	23%	Normalization
Z +jets cross section	50%	Normalization
Luminosity	2.6%	Normalization
electron trigger	1%	Normalization
muon trigger and id	$\pm 1\sigma(p_T, \eta)$	Normalization & Shape
electron id	$\pm 1\sigma(p_T, \eta)$	Normalization & Shape
Jet Energy Scale	$\pm 1\sigma(p_T, \eta)$	Normalization & Shape
Jet Energy Resolution	$\pm 1\sigma(\eta)$	Normalization & Shape
b-tagging	$\pm 1\sigma(p_T, \eta)$	Normalization & Shape
b-tagging mistag rate	$\pm 1\sigma(p_T, \eta)$	Normalization & Shape
top-tagging	unconstrained	Normalization & Shape
top-tagging mistag rate	$\pm 25\%$	Normalization & Shape
Pileup	$\pm 1\sigma$	Normalization & Shape
PDFs	CTEQ6 (CT10) set	Normalization & Shape
Scale ($Q^2 = M_t^2 + \sum p_{T,\text{jet}}^2$) for $t\bar{t}$	$4Q^2$ and $0.25Q^2$	Normalization & Shape
Scale ($Q^2 = M_V^2 + \sum p_{T,\text{jet}}^2$) for W/Z +jets	$4Q^2$ and $0.25Q^2$	Normalization & Shape
Matching for W/Z +jets	2 and $0.5 \times$ default	Normalization & Shape

Table 8.1: Table of systematic uncertainties considered in the statistical analysis. The uncertainties above the horizontal line are used for the determination of the background normalization and the CMS Top Tagger efficiency scale factor, as described in Sec. 6.6.

“1top”

Systematic	Background		Narrow Z' signal		
	$t\bar{t}$	$W + Jets$	$M = 1$ TeV	$M = 2$ TeV	$M = 3$ TeV
PU	-0.30 +0.26	+2.00 -2.62	+0.06 -0.27	-0.48 +0.45	-0.91 +0.93
elec-ID	+0.04 -0.04	+0.04 -0.04	+0.04 -0.04	+0.04 -0.04	+0.05 -0.05
JEC	+2.50 -2.45	+4.30 -3.58	+2.54 -2.11	-0.27 -0.19	-0.94 +0.49
JER	+2.10 -2.32	+2.63 -1.78	+1.79 -0.78	+1.55 -2.10	+1.24 -1.10
b -tag SF	-0.04 +0.04	+0.07 -0.07	+0.05 -0.05	-0.97 +0.98	-2.11 +2.11
b -mistag SF	-0.04 +0.04	-0.10 +0.09	-0.01 +0.01	-0.03 +0.03	+0.08 -0.08
top-mistag SF	+1.54 -1.55	+24.92 -24.95	+0.17 -0.17	+0.67 -0.67	+1.12 -1.12
PDF	+16.12 -13.09	+10.40 -9.28	+1.60 -1.68	+1.51 -1.53	+5.85 -5.66
Q^2 scale ($t\bar{t}$)	-13.59 +15.03				
Q^2 scale ($W + jets$)		-41.23 +41.23			
matching scale ($W + jets$)		-0.98 +0.98			

“0top1btag”

Systematic	Background		Narrow Z' signal		
	$t\bar{t}$	$W + Jets$	$M = 1$ TeV	$M = 2$ TeV	$M = 3$ TeV
PU	+0.30 -0.26	+0.66 -0.59	+0.45 -0.42	-0.01 +0.10	+0.46 -0.28
elec-ID	+0.04 -0.04	+0.05 -0.05	+0.04 -0.04	+0.04 -0.04	+0.04 -0.04
JEC	+6.09 -5.92	+7.24 -6.14	+2.24 -2.83	+1.44 -1.19	+1.15 -1.80
JER	+2.42 -2.37	+3.43 -1.87	+1.07 -0.92	+1.31 -1.59	+0.70 -2.07
b -tag SF	+1.30 -1.33	+2.13 -2.14	+1.20 -1.25	+2.09 -2.21	+2.41 -2.62
b -mistag SF	+0.14 -0.14	+3.77 -3.79	+0.05 -0.05	+0.16 -0.17	+0.15 -0.15
top-mistag SF	-0.03 +0.03	-0.06 +0.06	-0.01 +0.01	-0.08 +0.08	-0.12 +0.12
PDF	+10.98 -9.62	+5.92 -5.31	+0.84 -0.78	+2.22 -2.23	+6.09 -5.80
Q^2 scale ($t\bar{t}$)	-9.00 +9.03				
Q^2 scale ($W + jets$)		-42.82 +40.48			
matching scale ($W + jets$)		-1.67 +0.65			

“0top0btag”

Systematic	Background		Narrow Z' signal		
	$t\bar{t}$	$W + Jets$	$M = 1$ TeV	$M = 2$ TeV	$M = 3$ TeV
PU	+0.31 -0.28	+0.50 -0.46	+0.14 +0.12	+0.60 -0.66	+0.86 -0.98
elec-ID	+0.05 -0.05	+0.05 -0.05	+0.05 -0.05	+0.04 -0.04	+0.04 -0.04
JEC	+5.45 -4.72	+6.47 -6.00	+1.98 -1.99	+0.77 -1.34	+0.72 -0.75
JER	+2.25 -2.28	+2.11 -2.07	+1.82 -1.39	+1.80 -2.79	+1.28 -1.11
b -tag SF	-4.58 +4.69	-0.31 +0.31	-6.25 +6.47	-7.33 +7.63	-8.64 +9.05
b -mistag SF	-0.43 +0.43	-0.60 +0.61	-0.30 +0.31	-0.29 +0.29	-0.31 +0.31
top-mistag SF	-0.03 +0.03	-0.06 +0.06	-0.01 +0.01	-0.07 +0.07	-0.09 +0.09
PDF	+11.62 -9.95	+4.13 -3.76	+2.04 -1.80	+2.98 -2.86	+7.28 -6.57
Q^2 scale ($t\bar{t}$)	-11.29 +11.41				
Q^2 scale ($W + jets$)		-48.31 +48.24			
matching scale ($W + jets$)		+2.93 -3.01			

Table 8.2: Effect of the systematic uncertainties on the normalization of background and signal samples in the electron channel for the “1top” (top), “0top1btag” (middle) and “0top0btag” (bottom) categories. The numbers correspond to the difference in percent with respect to the nominal yield. The upper (lower) value refers to the $+1\sigma$ (-1σ) variation of a given systematic.

“1top”

Systematic	Background		Narrow Z' signal		
	$t\bar{t}$	$W + Jets$	$M = 1$ TeV	$M = 2$ TeV	$M = 3$ TeV
PU	-0.32 +0.19	-0.67 +0.53	-1.23 +1.11	-0.48 +0.37	-1.66 +1.62
muon-ID+trigger	+2.03 -2.04	+2.05 -2.05	+1.71 -1.72	+2.31 -2.33	+2.71 -2.73
JEC	+2.54 -3.74	+0.20 -4.37	+2.71 -2.87	-1.20 -1.39	-1.44 +0.33
JER	+0.88 -0.71	-1.08 -0.91	+0.49 +0.76	+0.28 -0.67	+0.07 +0.31
b -tag SF	-0.11 +0.11	+0.23 -0.23	+0.14 -0.14	-0.81 +0.81	-1.60 +1.60
b -mistag SF	-0.06 +0.06	+0.20 -0.21	-0.05 +0.05	+0.04 -0.04	-0.01 +0.01
top-mistag SF	+1.35 -1.35	+24.93 -24.95	+0.18 -0.18	+0.68 -0.68	+1.07 -1.07
PDF	+16.43 -13.05	+9.86 -8.74	+1.30 -1.40	+1.73 -1.64	+6.06 -5.82
Q^2 scale ($t\bar{t}$)	-14.76 +15.79				
Q^2 scale ($W + jets$)		-40.71 +40.71			
matching scale ($W + jets$)		-0.38 +0.38			

“0top1btag”

Systematic	Background		Narrow Z' signal		
	$t\bar{t}$	$W + Jets$	$M = 1$ TeV	$M = 2$ TeV	$M = 3$ TeV
PU	+0.06 -0.07	+0.04 -0.16	-0.65 +0.66	+0.09 -0.18	+0.13 +0.02
muon-ID+trigger	+1.41 -1.41	+1.63 -1.63	+1.61 -1.61	+2.22 -2.23	+2.51 -2.53
JEC	+5.39 -5.89	+6.84 -6.69	+1.77 -2.69	+0.72 -2.06	-1.18 -3.17
JER	+1.31 -1.16	+0.74 -0.38	+0.09 -0.20	+0.04 -0.38	+0.38 -0.78
b -tag SF	+1.32 -1.35	+2.17 -2.17	+1.34 -1.38	+2.39 -2.52	+3.02 -3.27
b -mistag SF	+0.15 -0.15	+3.71 -3.73	+0.01 -0.01	+0.14 -0.14	+0.40 -0.40
top-mistag SF	-0.03 +0.03	-0.05 +0.05	-0.01 +0.01	-0.08 +0.08	-0.12 +0.12
PDF	+10.58 -9.37	+5.70 -5.15	+0.98 -0.94	+2.02 -1.95	+6.77 -6.26
Q^2 scale ($t\bar{t}$)	-6.81 +6.84				
Q^2 scale ($W + jets$)		-43.13 +42.64			
matching scale ($W + jets$)		+1.30 -1.48			

“0top0btag”

Systematic	Background		Narrow Z' signal		
	$t\bar{t}$	$W + Jets$	$M = 1$ TeV	$M = 2$ TeV	$M = 3$ TeV
PU	+0.17 -0.18	+0.18 -0.15	+0.68 -0.72	+0.21 -0.10	+0.10 -0.20
muon-ID+trigger	+1.42 -1.43	+1.59 -1.59	+1.67 -1.68	+2.30 -2.31	+2.63 -2.65
JEC	+4.66 -5.52	+6.90 -6.64	+1.78 -2.60	-0.35 -2.11	+0.50 -1.78
JER	+1.45 -1.27	+1.41 -1.22	+0.73 +0.49	+0.35 -1.22	+0.17 -0.22
b -tag SF	-4.52 +4.64	-0.31 +0.31	-5.95 +6.15	-7.14 +7.42	-7.92 +8.26
b -mistag SF	-0.44 +0.44	-0.61 +0.61	-0.32 +0.32	-0.27 +0.27	-0.27 +0.27
top-mistag SF	-0.03 +0.03	-0.06 +0.06	-0.02 +0.02	-0.06 +0.06	-0.09 +0.09
PDF	+11.36 -9.73	+4.23 -3.87	+1.62 -1.86	+2.98 -2.84	+6.84 -6.22
Q^2 scale ($t\bar{t}$)	-9.81 +9.95				
Q^2 scale ($W + jets$)		-47.42 +47.36			
matching scale ($W + jets$)		+2.58 -2.58			

Table 8.3: Effect of the systematic uncertainties on the normalization of background and signal samples in the muon channel for the “1top” (top), “0top1btag” (middle) and “0top0btag” (bottom) category. The numbers correspond to the difference in percent with respect to the nominal yield. The upper (lower) value refers to the $+1\sigma$ (-1σ) variation of a given systematic.

9. RESULTS

The statistical analysis described in Chap. 8 is used to derive the 95% CL upper limits on the cross-section times branching ratio for heavy resonances decaying to a $t\bar{t}$ pair. The measurement is performed by analyzing the distributions for the reconstructed invariant mass, $M_{t\bar{t}}$. As any theoretical model that predicts resonant $t\bar{t}$ production will produce “bumps” in the $M_{t\bar{t}}$ distribution these limits are said to be model-independent. The narrow and wide Z' models used to generate our signal model can be used generically for any model producing a 1% or 10% width resonance decaying to $t\bar{t}$. The kk -gluon signal model produces a wider resonance with more off-shell masses, but can also be used for resonant models that produce similar $M_{t\bar{t}}$ distributions.

In addition to the model-independent limits on resonance production cross-section times branching ratio, we place limits on the models used to generate our signal samples. These limits are extracted by comparing the theoretically predicted signal cross-section to the observed limit as a function of resonance mass. Resonant masses which have a higher predicted cross-section than the excluded limit are ruled out at 95% CL.

For each of the three benchmark models considered limits are set for both the electron and muon channel as well as their combination. We scale the background yields using the results of the maximum-likelihood fit across both channels according to the scale factors described in Tab. 7.2. We use the systematic uncertainties described in Sec. 8.1 when calculating the likelihood and setting the limits.

The exclusion limits on production cross-section times branching ratio for the three signal models are summarized in Tables 9.2–9.4. For a narrow 1% width 2 TeV resonance we

establish a limit of 0.0172 pb for cross-section times branching ratio with the combination of the electron and muon channels. For a wider 10% width 2 TeV resonance we establish a limit of 0.0299 pb. For a Kaluza-Klein gluon resonance 2 TeV resonance we establish a limit of 0.0625 pb.

These limits are visualized in Fig. 9.1 for the narrow Z' resonance, Fig. 9.2 for the wide Z' resonance, and Fig. 9.3 for the Kaluza-Klein gluon resonance. The dashed lines in these figures indicate the theoretical cross-section for their respective models and the intersection with the observed limit gives the 95% CL mass exclusion limits. The narrow Z' model is excluded below 2.29 TeV with the expected exclusion limit being 2.21 TeV. The wide Z' model is excluded below 2.75 TeV with the expected exclusion limit being 2.64 TeV. Finally, the Kaluza-Klein gluon resonance model is excluded below 2.61 TeV with the expected exclusion being 2.52 TeV. Table 9.1 summarizes the mass exclusion limits for the individual channels and their combination.

signal model	mass limits [TeV]					
	electron channel		muon channel		combination	
	obs.	exp.	obs.	exp.	obs.	exp
narrow Z' (1% width)	2.13	2.07	2.11	2.09	2.29	2.21
wide Z' (10% width)	2.59	2.51	2.64	2.54	2.75	2.64
KK gluon	2.42	2.34	2.48	2.37	2.61	2.52

Table 9.1: Expected and observed 95% CL lower mass limits for the three models studied in this analysis. Limits are given for the electron channel, the muon channel, and the combination of the two.

electron channel

$M_{Z'}$ (TeV)	Expected (pb)	Exp. $\pm 1\sigma$ (pb)	Exp. $\pm 2\sigma$ (pb)	Observed (pb)
0.5	26.2	42 – 15.9	71.8 – 11.1	16.3
0.75	1.1	1.54 – 0.784	2.15 – 0.572	1.46
1	0.326	0.453 – 0.23	0.635 – 0.168	0.299
1.25	0.156	0.226 – 0.111	0.316 – 0.0824	0.266
1.5	0.0789	0.111 – 0.0538	0.149 – 0.0384	0.0615
2	0.0294	0.0441 – 0.02	0.069 – 0.0144	0.0247
3	0.0197	0.0302 – 0.0134	0.0443 – 0.00989	0.0159

muon channel

$M_{Z'}$ (TeV)	Expected (pb)	Exp. $\pm 1\sigma$ (pb)	Exp. $\pm 2\sigma$ (pb)	Observed (pb)
0.5	23.7	38.6 – 14.7	66.2 – 10.3	13.9
0.75	1.27	1.85 – 0.907	2.62 – 0.634	1.86
1	0.345	0.489 – 0.244	0.692 – 0.176	0.319
1.25	0.161	0.229 – 0.114	0.323 – 0.0806	0.295
1.5	0.0838	0.119 – 0.0587	0.176 – 0.0424	0.112
2	0.0279	0.0426 – 0.0195	0.0627 – 0.0134	0.0282
3	0.0149	0.0229 – 0.00969	0.0346 – 0.00693	0.00945

combination

$M_{Z'}$ (TeV)	Expected (pb)	Exp. $\pm 1\sigma$ (pb)	Exp. $\pm 2\sigma$ (pb)	Observed (pb)
0.5	16.7	27.5 – 10.7	51.9 – 6.87	10.5
0.75	0.75	1.1 – 0.541	1.62 – 0.404	1.11
1	0.245	0.35 – 0.171	0.517 – 0.128	0.197
1.25	0.114	0.172 – 0.0822	0.254 – 0.061	0.242
1.5	0.058	0.0848 – 0.0409	0.128 – 0.0285	0.0602
2	0.0198	0.03 – 0.0131	0.0513 – 0.00876	0.0172
3	0.0117	0.0183 – 0.0075	0.0348 – 0.00519	0.00766

Table 9.2: Expected and observed 95% CL limits on the production cross-section times branching ratio of the narrow Z' model. Limits are given for the electron channel, the muon channel, and the combination of the two. Fig. 9.1 plots these numerical values against the theoretical cross-section for the model.

electron channel

$M_{Z'}$ (TeV)	Expected (pb)	Exp. $\pm 1\sigma$ (pb)	Exp. $\pm 2\sigma$ (pb)	Observed (pb)
0.5	26.6	47.2 – 15.3	78.9 – 9.72	25.4
0.75	1.51	2.15 – 1.09	2.97 – 0.781	2.26
1	0.489	0.704 – 0.354	0.939 – 0.252	0.495
1.25	0.228	0.317 – 0.166	0.435 – 0.12	0.446
1.5	0.125	0.182 – 0.0853	0.245 – 0.063	0.115
2	0.0528	0.0782 – 0.0361	0.11 – 0.0263	0.0417
3	0.0543	0.0809 – 0.0358	0.116 – 0.0256	0.0416

muon channel

$M_{Z'}$ (TeV)	Expected (pb)	Exp. $\pm 1\sigma$ (pb)	Exp. $\pm 2\sigma$ (pb)	Observed (pb)
0.5	17.6	31.2 – 10.8	52.1 – 7.22	9.36
0.75	1.58	2.23 – 1.11	3.25 – 0.788	1.98
1	0.477	0.661 – 0.327	0.929 – 0.251	0.477
1.25	0.258	0.372 – 0.183	0.513 – 0.127	0.5
1.5	0.125	0.183 – 0.0862	0.273 – 0.0608	0.187
2	0.0508	0.0769 – 0.0349	0.12 – 0.0242	0.0492
3	0.0463	0.0718 – 0.0313	0.111 – 0.0216	0.0291

combination

$M_{Z'}$ (TeV)	Expected (pb)	Exp. $\pm 1\sigma$ (pb)	Exp. $\pm 2\sigma$ (pb)	Observed (pb)
0.5	12.7	23.3 – 7.68	50.7 – 4.92	10.2
0.75	0.986	1.46 – 0.713	2.23 – 0.523	1.43
1	0.357	0.502 – 0.244	0.754 – 0.175	0.314
1.25	0.184	0.273 – 0.125	0.446 – 0.095	0.419
1.5	0.0931	0.135 – 0.0648	0.206 – 0.045	0.113
2	0.036	0.054 – 0.0237	0.0858 – 0.0157	0.0299
3	0.0342	0.0533 – 0.0219	0.0898 – 0.0151	0.0236

Table 9.3: Expected and observed 95% CL limits on the production cross-section times branching ratio of the wide Z' model. Limits are given for the electron channel, the muon channel, and the combination of the two. Fig. 9.2 plots these numerical values against the theoretical cross-section for the model.

electron channel

$M_{g_{KK}}$ (TeV)	Expected (pb)	Exp. $\pm 1\sigma$ (pb)	Exp. $\pm 2\sigma$ (pb)	Observed (pb)
0.7	3.08	4.82 – 2.08	7.51 – 1.47	6.14
1	0.764	1.09 – 0.528	1.52 – 0.376	0.751
1.2	0.469	0.679 – 0.328	0.959 – 0.243	0.831
1.4	0.283	0.42 – 0.194	0.588 – 0.14	0.49
1.5	0.235	0.35 – 0.163	0.521 – 0.118	0.28
1.6	0.182	0.268 – 0.129	0.408 – 0.0892	0.156
1.8	0.134	0.2 – 0.0933	0.299 – 0.0669	0.116
2	0.107	0.162 – 0.0726	0.245 – 0.0505	0.0942
2.5	0.0976	0.148 – 0.0654	0.226 – 0.0442	0.0825
3	0.136	0.21 – 0.0854	0.311 – 0.0615	0.0955

muon channel

$M_{g_{KK}}$ (TeV)	Expected (pb)	Exp. $\pm 1\sigma$ (pb)	Exp. $\pm 2\sigma$ (pb)	Observed (pb)
0.7	3.48	5.12 – 2.39	8.13 – 1.64	5.91
1	0.748	1.11 – 0.519	1.56 – 0.381	0.842
1.2	0.504	0.728 – 0.349	1.04 – 0.245	0.942
1.4	0.288	0.417 – 0.198	0.615 – 0.131	0.55
1.5	0.225	0.326 – 0.156	0.482 – 0.105	0.365
1.6	0.183	0.265 – 0.124	0.375 – 0.0875	0.211
1.8	0.123	0.188 – 0.0847	0.276 – 0.0589	0.11
2	0.0991	0.151 – 0.0678	0.228 – 0.0478	0.0921
2.5	0.0924	0.141 – 0.0591	0.211 – 0.0389	0.0721
3	0.112	0.18 – 0.0711	0.269 – 0.0497	0.0787

combination

$M_{g_{KK}}$ (TeV)	Expected (pb)	Exp. $\pm 1\sigma$ (pb)	Exp. $\pm 2\sigma$ (pb)	Observed (pb)
0.7	2.03	3.17 – 1.36	5.56 – 1.06	4.55
1	0.556	0.819 – 0.376	1.27 – 0.273	0.509
1.2	0.374	0.553 – 0.249	0.898 – 0.182	0.812
1.4	0.221	0.326 – 0.144	0.504 – 0.109	0.454
1.5	0.173	0.258 – 0.118	0.425 – 0.0799	0.266
1.6	0.133	0.198 – 0.0909	0.3 – 0.0648	0.123
1.8	0.0916	0.137 – 0.0607	0.222 – 0.0417	0.0744
2	0.0725	0.112 – 0.0486	0.173 – 0.0325	0.0625
2.5	0.0656	0.101 – 0.0425	0.173 – 0.0295	0.0548
3	0.0838	0.132 – 0.0545	0.2 – 0.0349	0.0586

Table 9.4: Expected and observed 95% CL limits on the production cross-section times branching ratio of the Kaluza-Klein gluon resonance model. Limits are given for the electron channel, the muon channel, and the combination of the two. Fig. 9.3 plots these numerical values against the theoretical cross-section for the model.

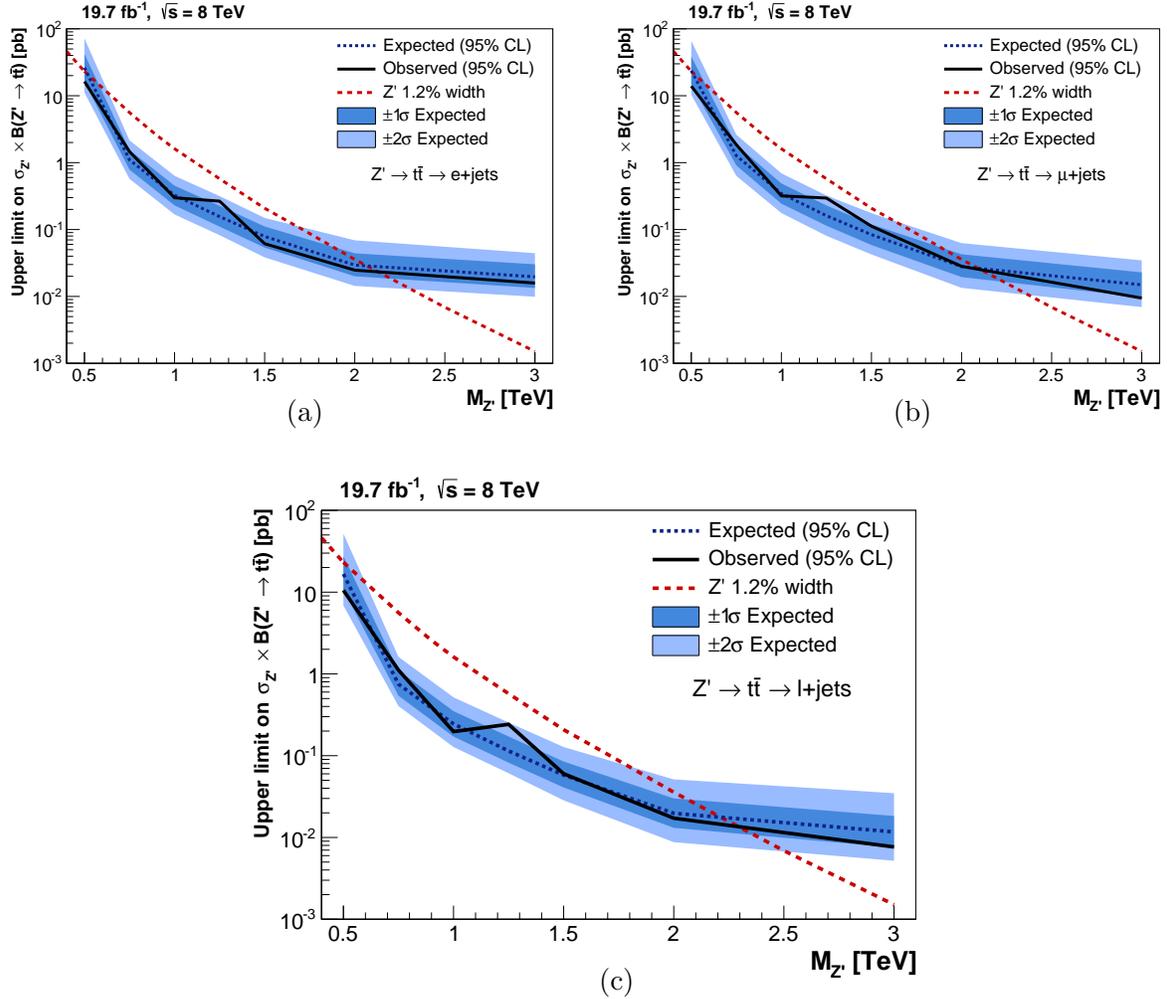


Figure 9.1: Expected and observed 95% CL limits on the production cross-section times branching ratio of the narrow Z' model. Limits are given for the electron channel (a), the muon channel (b), and the combination of the two (c). The theoretical prediction for the cross-section of a Z' boson with a width of 1.2% are taken from Ref. [14] and multiplied by a K-factor of $K = 1.3$ to account for higher-order corrections. [72]

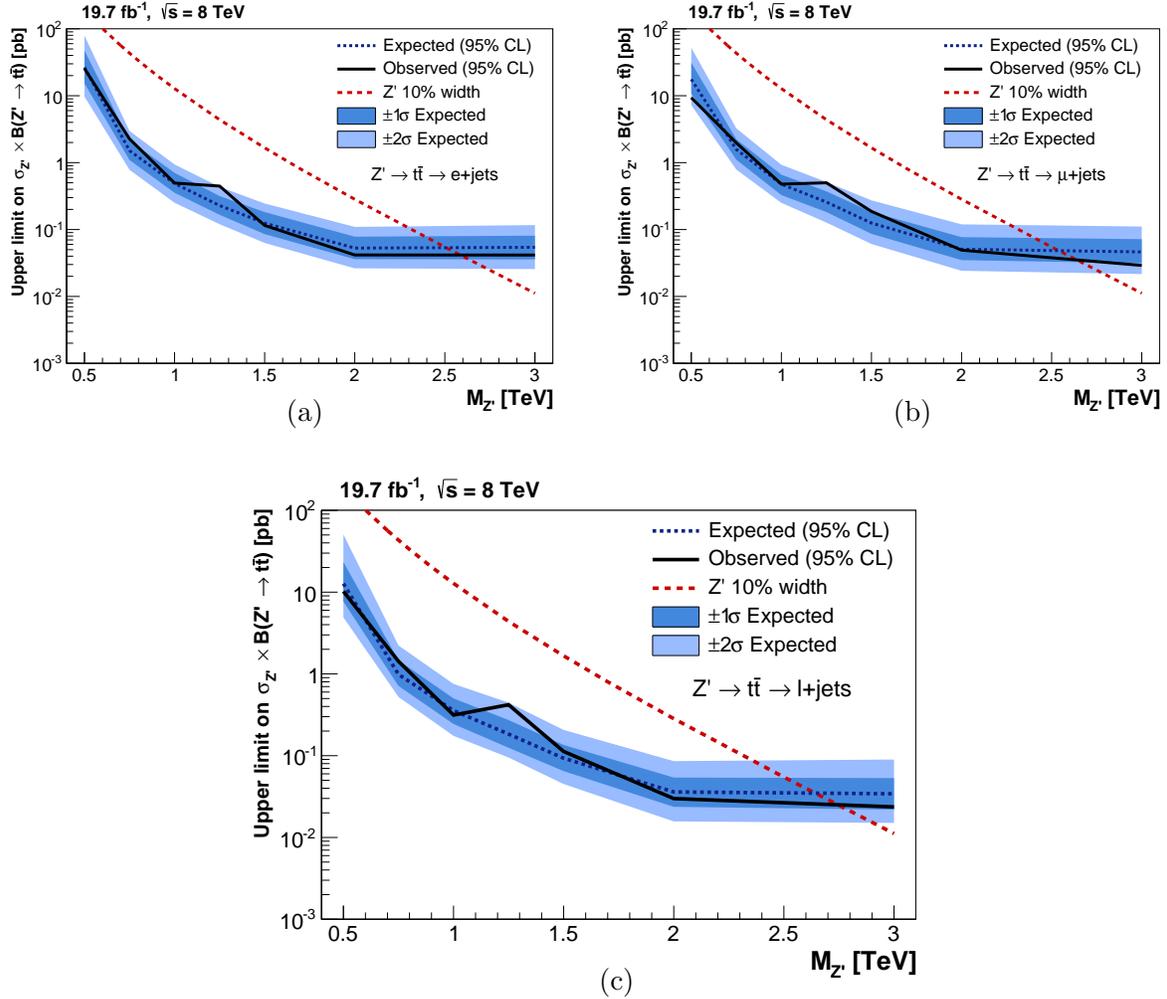


Figure 9.2: Expected and observed 95% CL limits on the production cross-section times branching ratio of the wide Z' model. Limits are given for the electron channel (a), the muon channel (b), and the combination of the two (c). The theoretical prediction for the cross-section of a Z' boson with a width of 10% are taken from Ref. [14] and multiplied by a K-factor of $K = 1.3$ to account for higher-order corrections. [72]

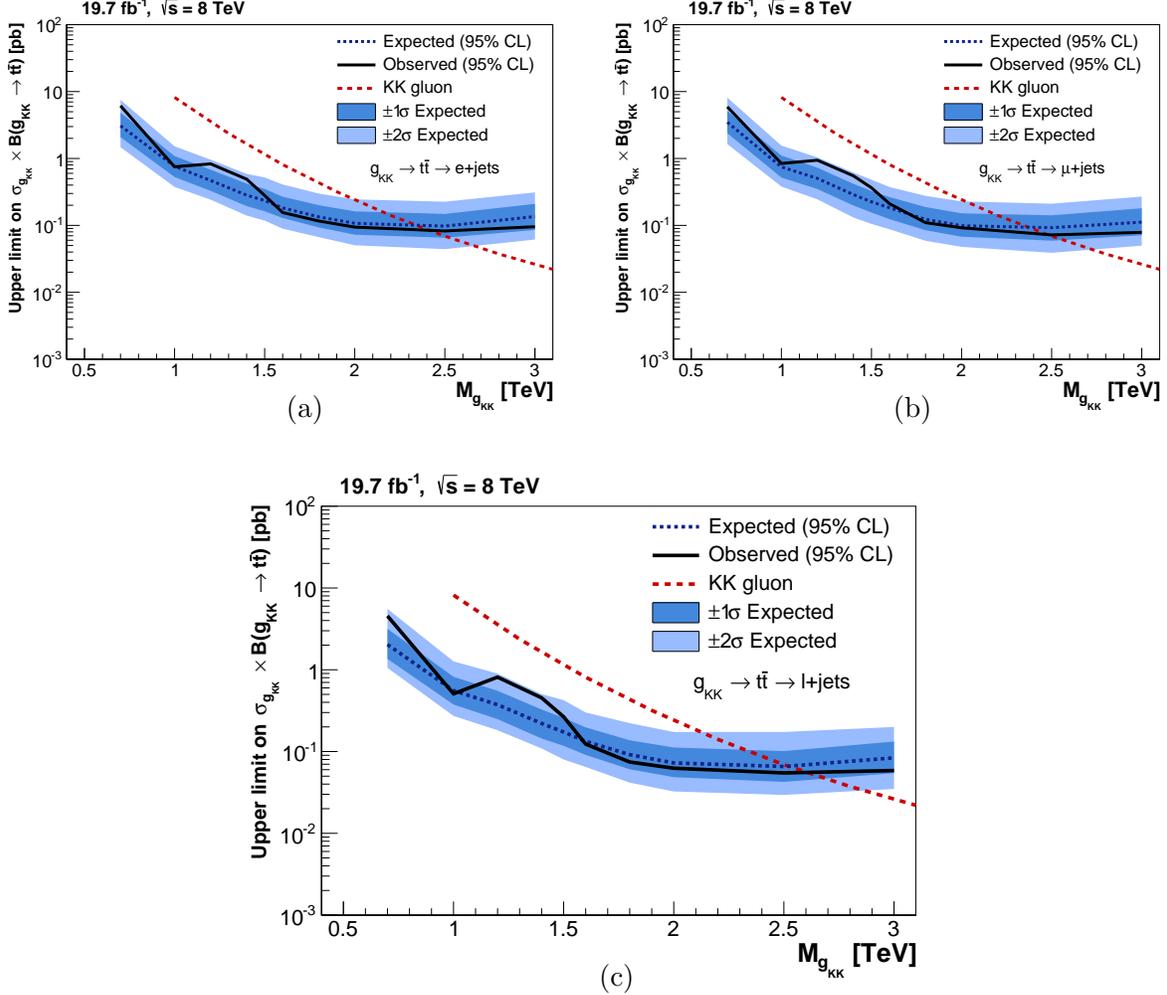


Figure 9.3: Expected and observed 95% CL limits on the production cross-section times branching ratio of the Kaluza-Klein gluon resonance model. Limits are given for the electron channel (a), the muon channel (b), and the combination of the two (c). The theoretical prediction for the cross-section of a Kaluza-Klein resonant gluon in the Randall-Sundrum model are taken from Ref. [20] and multiplied by a K-factor of $K = 1.3$ to account for higher-order corrections. [72]

10. SUMMARY AND OUTLOOK

This thesis has presented the results from a model-independent search for the production of heavy resonances decaying to $t\bar{t}$ pairs in the semileptonic channel. The data sample corresponds to an integrated luminosity of 19.7 fb^{-1} of 8 TeV center of mass energy proton-proton collisions recorded by CMS in 2012. We searched for three signal models: a Z' boson with a 1% and 10% relative width [14] and a Kaluza-Klein resonant gluon [20] in a Randall-Sundrum model [19]. No evidence for such massive new states were found after analyzing events containing an electron or muon and jets. Therefore, limits were placed on the production cross-section times branching ratio using a Bayesian statistical model at 95% CL. We exclude narrow Z' resonances below 2.29 TeV, wide Z' resonances below 2.75 TeV, and Kaluza-Klein gluons below 2.61 TeV.

The boosted lepton+jets search for high-mass resonances presented in this thesis significantly improves upon previously published results [91]. The improvements are due to the top-tagging made possible by studying jet substructure and the inclusion of a single-jet trigger in the electron channel. The improvements on the expected 95% CL upper limits on the production cross-section times branching ratio are between 30%-40% for resonances with masses between 1-2 TeV.

The object definitions of electrons, muons, and top-tag jets in Chap. 3 were chosen in collaboration with other groups at CMS who were performing parallel analyses of the 8 TeV dataset. These groups studied the same Z' and Kaluza-Klein gluon signals we did, but with different decay topologies for the $t\bar{t}$ pair. The dilepton analysis group searched for high-mass resonances decaying to $t\bar{t}$, with both top quarks decaying leptonically. The all-hadronic anal-

ysis group searched for high-mass resonant $t\bar{t}$, with both top quarks decaying hadronically. An additional group studied the threshold production of Z' and Kaluza-Klein gluons which lack the signature boosted topology. Because of the shared object definitions, these analyses are statistically independent from one another. Therefore it is straightforward to combine these separate searches together to obtain the strongest exclusion limits possible. The combination of analyses is done using the same statistical techniques described in Chap. 8. The nuisance parameters that drive identical systematic uncertainties in the different analyses are fully correlated in this combined statistical analysis.

Figure 10.1 shows the expected 95% CL Bayesian upper limits on the production cross-section time branching ratio for a narrow Z' signal for each analysis independently and their statistical combination. Similar plots for the wide Z' signal and Kaluza-Klein gluon can be seen in Figures 10.2–10.3. These plots show that the lepton+jets analysis presented in this thesis provide the best exclusion limits for all signal models. However, the combination provides a significant improvement over this analysis alone, especially in the low mass region where the threshold lepton+jets analysis provides better exclusion limits. Figure 10.4 shows the observed limit for all three signal models along with the expected limit and its ± 1 and ± 2 bands. The combined analysis excludes narrow Z' resonances below 2.4 TeV, wide Z' resonances below 2.9 TeV, and Kaluza-Klein gluons below 2.8 TeV. These results provide the most stringent constraints on resonant $t\bar{t}$ production to date.

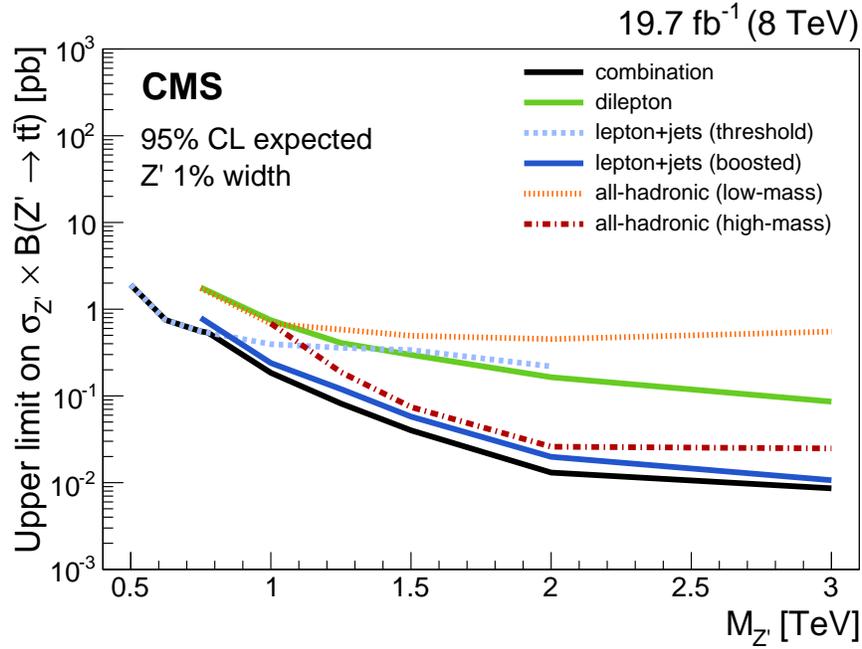


Figure 10.1: Expected 95% CL Bayesian upper limits on production cross section times branching ratio for a narrow Z' boson. Individual analysis limits are shown separately along with the combined results.

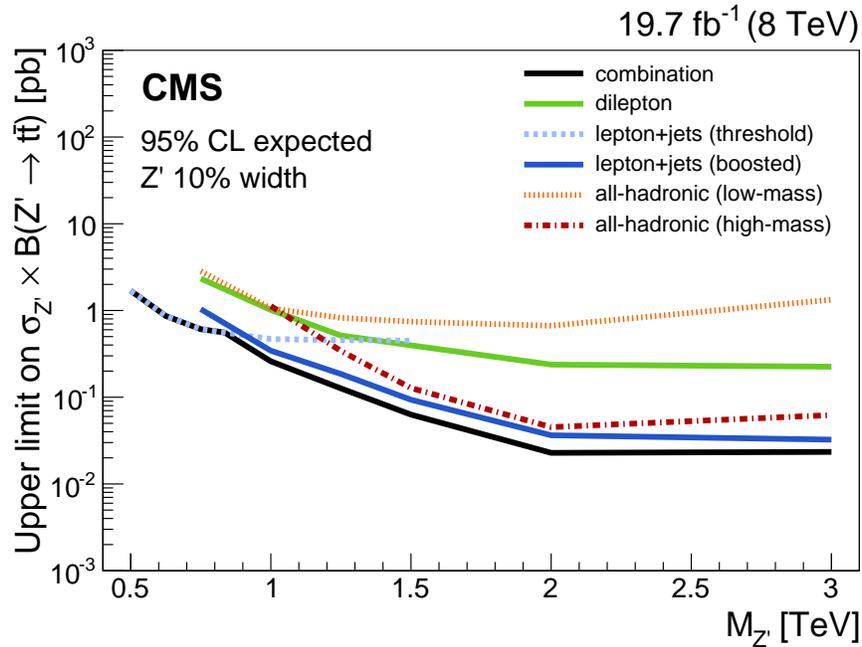


Figure 10.2: Expected 95% CL Bayesian upper limits on production cross section times branching ratio for a wide Z' boson. Individual analysis limits are shown separately along with the combined results.

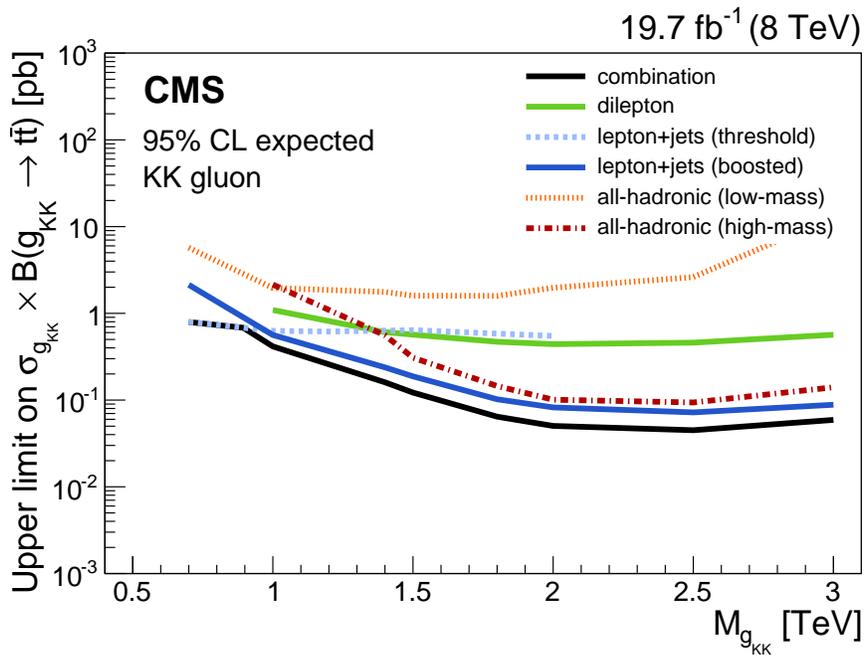
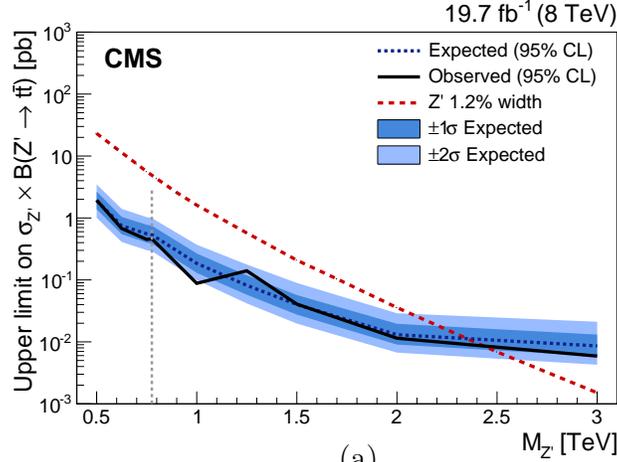
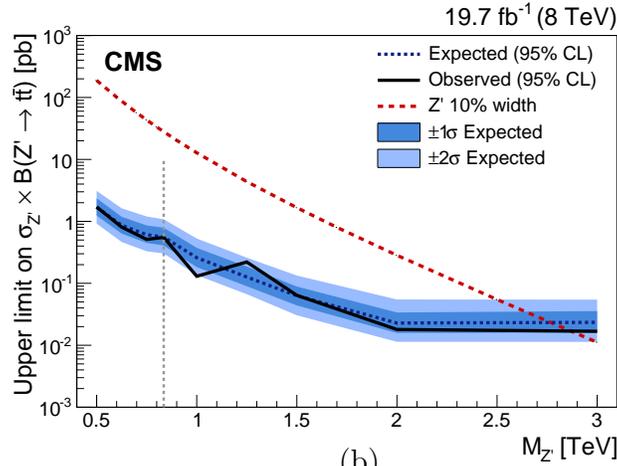


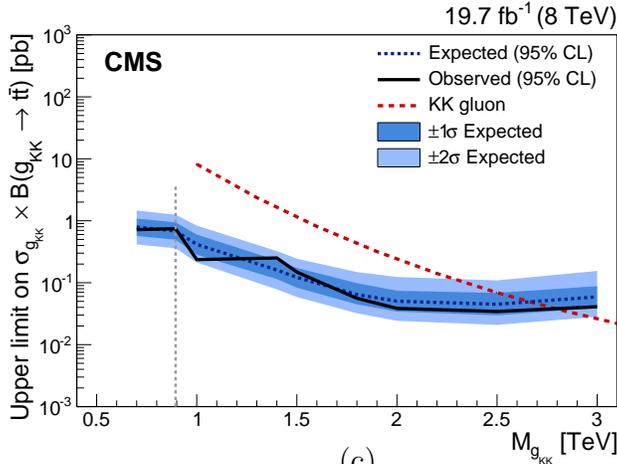
Figure 10.3: Expected 95% CL Bayesian upper limits on production cross section times branching ratio for a Kaluza-Klein gluon. Individual analysis limits are shown separately along with the combined results.



(a)



(b)



(c)

Figure 10.4: 95% CL Bayesian upper limits on production cross section times branching ratio for a narrow Z' boson (a), wide Z' boson (b), and Kaluza-Klein gluon (c). The vertical dashed line represents the transition from the threshold lepton+jets analysis and the combined results. The theoretical prediction for the cross-section of a Z' boson with a width of 1.2% and 10% are taken from Ref. [14]. The theoretical prediction for a Kaluza-Klein resonant gluon in the Randall-Sundrum model are taken from Ref. [20]. The signal cross-sections are multiplied by a K-factor of $K = 1.3$ to account for higher-order corrections. [72]

A. χ^2 CUT AND EVENT CATEGORIZATION OPTIMIZATION

We performed optimization studies on event selection and categorization detailed in Chap. 7. In particular we studied the optimization of event categorization and χ^2 cuts. The previous version of this analysis [91] used a χ^2 cut of 10 and separated events into 2 categories. Events with an AK5 jet b-tagged at the CSVT working point were placed into a “BTag” category and the others into a “NoBTag” category.

In this updated analysis we wanted to include at least one category containing events tagged by the CMS Top Tagging algorithm to improve upon previously published results. This algorithm is moderately efficient at tagging the boosted hadronic top decays in Z' and Kaluza-Klein gluon signals while rejecting events from SM backgrounds. By including a category of “top-tag” events we create a signal-rich region which drastically improves previously published limits. Tab. A.1 shows the efficiency of this algorithm on various signals and SM backgrounds.

The definition of a top-tagged jet in Chap. 3 includes a cut on N -subjettiness, τ_{32} , which is supposed to increase the top-tagging efficiency and provide more discriminating power. An additional definition of top-tagged jets was studied where we remove this requirement. We also studied the effect of including subjet b-tagging where CA8 subjets are b-tagged. The categorization and χ^2 cut were optimized based on the expected sensitivity by calculating the 95% CL expected limits for narrow Z' signals in the muon channel. The only systematic uncertainties included were the cross-section rate uncertainties. The following categorizations were studied:

- **3 categories:**

“0 top-tag + 0 b -tag” (0top0b)

“0 top-tag + ≥ 1 b -tag” (0top1b)

“1 top-tag” (1top).

- **4 categories:**

“0 top-tag + 0 b -tag” (0top0b)

“0 top-tag + ≥ 1 b -tag” (0top1b)

“1 top-tag + 0 sum- b -tag” (1top0sb)

“0 top-tag + ≥ 1 sum- b -tag” (1top1sb)

The variable “sum- b -tag” corresponds to the sum of the number of CSVM-tagged AK5 jets and the number of CSVM-tagged subjects of the top-tag jet.

In each case we consider both definitions of a top-tagged jet, with and without the $\tau_{32} < 0.7$ cut, and three different χ^2 cuts at 30, 50, and 60.

Fig. A.1 shows the expected limits for the “3 categories” case. Fig. A.2 shows the expected limits for the “4 categories” case. In both figures the expected limits from the “2 categories” case used in the previous analysis are also shown for comparison. In each case, a clear improvement over the previously published results is shown. In particular, slightly better results are achieved when using the “3 categories” case with a χ^2 cut of 50. Fig. A.3 shows a comparison of the expected limits between the “3 categories” and “4 categories” cases when the χ^2 cut is set to 50. Both categorizations have slightly improved limits when including the τ_{32} cut in the top-tag definition. Because of low MC statistics in categories containing a top-tag the choice to use “3 categories” was made. A χ^2 cut of 50 was chosen as the optimal value and the definition of a top-tagged jet should include the N -subjettiness cut based on these studies.

Table A.1: CMS Top Tagger efficiency for various signals and SM backgrounds. The efficiency is defined as the ratio of events containing a top-tagged jet which pass the selection cuts defined in Chap. 5.

Sample	Efficiency
Z' ($M = 1 \text{ TeV}/c^2$, $w = 1\%$)	0.151
Z' ($M = 2 \text{ TeV}/c^2$, $w = 1\%$)	0.415
Z' ($M = 3 \text{ TeV}/c^2$, $w = 1\%$)	0.341
Z' ($M = 4 \text{ TeV}/c^2$, $w = 1\%$)	0.274
g_{KK} ($M = 1 \text{ TeV}/c^2$)	0.129
g_{KK} ($M = 2 \text{ TeV}/c^2$)	0.347
g_{KK} ($M = 3 \text{ TeV}/c^2$)	0.305
g_{KK} ($M = 4 \text{ TeV}/c^2$)	0.280
$t\bar{t}$	0.0355
W+jets	0.00677

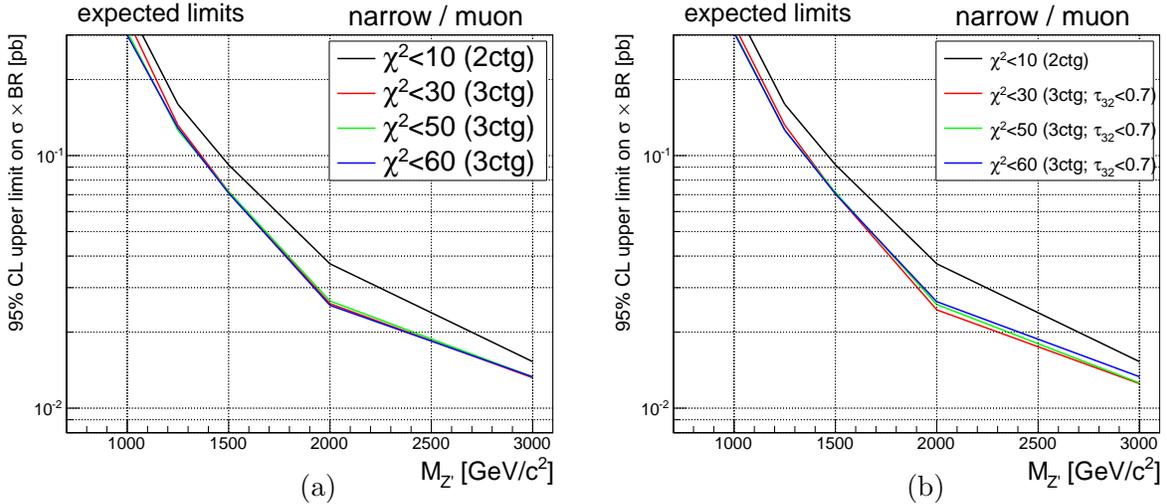


Figure A.1: 95% CL expected (Bayesian) limits on the production cross section times branching ratio of a narrow Z' resonance decaying to $t\bar{t}$ in the muon channel for the “3 categories” case described in the text: (a) using no N-subjettiness cut in the top-tag definition, (b) with a cut $\tau_{32} < 0.7$ in the top-tag definition. The corresponding limits for the settings used in Ref. [91], i.e. two categories based on the number of b -tags with $\chi^2 < 10$, are also shown (black line). The uncertainties on the normalization of each background template are the only systematics included in the limit calculation.

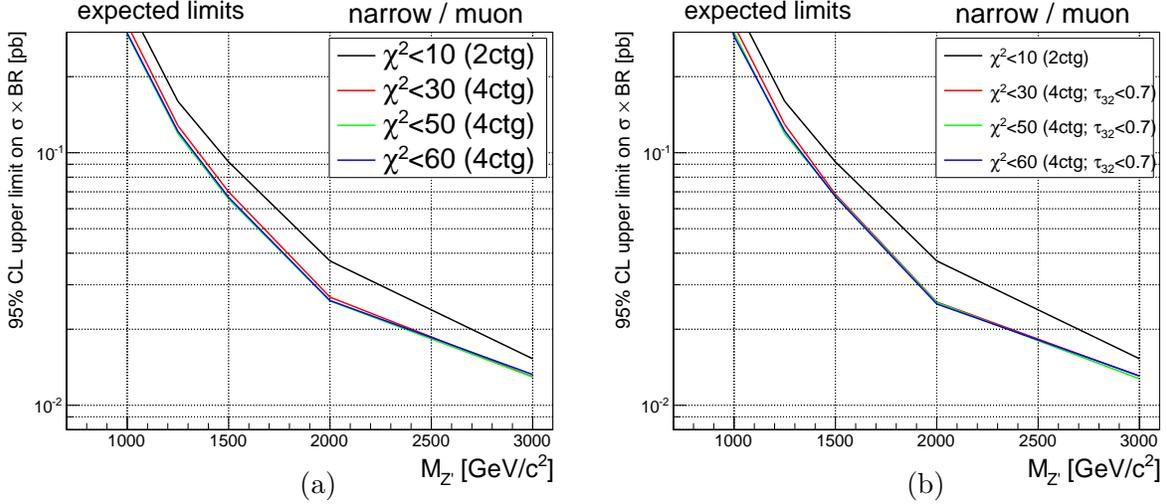


Figure A.2: 95% CL expected (Bayesian) limits on the production cross section times branching ratio of a narrow Z' resonance decaying to $t\bar{t}$ in the muon channel for the “4 categories” case described in the text: (a) using no N-subjettiness cut in the top-tag definition, (b) with a cut $\tau_{32} < 0.7$ in the top-tag definition. The corresponding limits for the settings used in [91], i.e. two categories based on the number of b -tags with $\chi^2 < 10$, are also shown (black line). The uncertainties on the normalization of each background template are the only systematics included in the limit calculation.

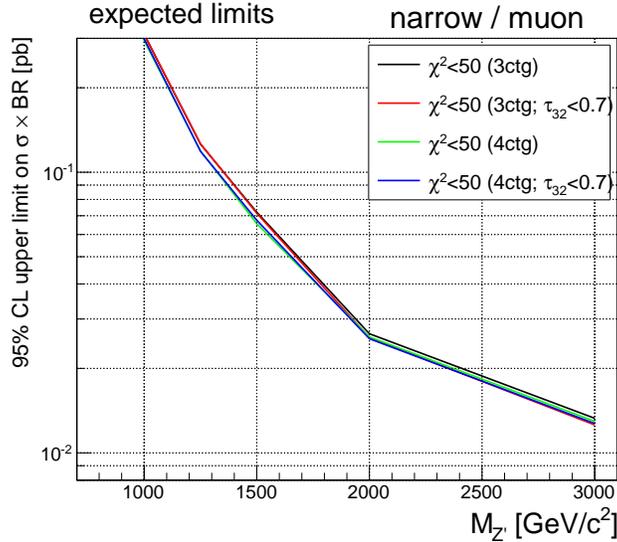


Figure A.3: 95% CL expected (Bayesian) limits on the production cross section times branching ratio of a narrow Z' resonance decaying to $t\bar{t}$ in the muon channel, comparing the “3 categories” and “4 categories” cases described in the text. For each categorization, a cut $\chi^2 < 50$ is used and results are shown for two different definitions of top-tagged jet (with and without the $\tau_{32} < 0.7$ cut). The uncertainties on the normalization of each background template are the only systematics included in the limit calculation.

B. KINEMATIC DISTRIBUTIONS

Figures B.2–B.17 show several distributions of kinematic quantities relevant to the analysis. Both data and MC simulation are plotted for comparison. The samples are split between the two channels and three categories described in Chap. 7.

The yields of the MC samples are corrected using scale factors derived in a maximum-likelihood fit to data across all channels and categories simultaneously. This fit and the resulting scale factors are described in detail in Chap. 8. The uncertainty bands in the plots include both the MC statistical uncertainty and the post-fit systematic uncertainty.

All signal samples have been normalized to a cross section of 1 pb in all of the plots. The plot colors all follow the legend shown in Fig. B.1.

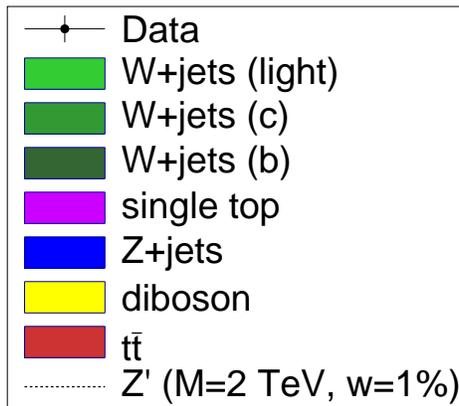


Figure B.1: Legend colors for plots in Appendix B

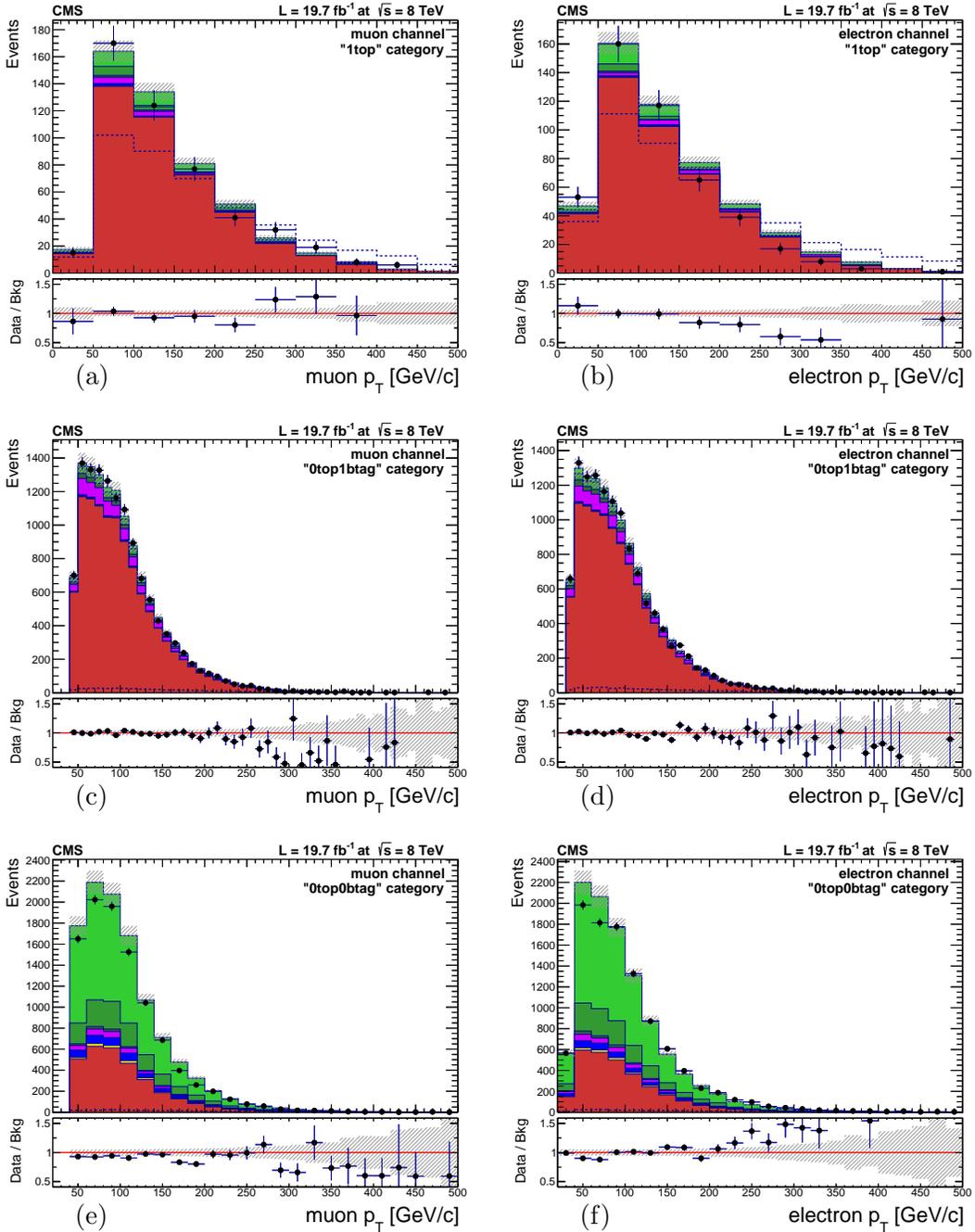


Figure B.2: Data/MC comparison for the lepton p_T after all analysis cuts are applied. Events are split into the following categories: muon "1top" (a), electron "1top" (b), muon "0top1btag" (c), electron "0top1btag" (d), muon "0top0btag" (e), electron "0top0btag" (f). Each background process is corrected by scale factors from the maximum-likelihood fit to data as described in Chap. 8. The signal MC samples are normalized to a cross section of 1 pb. The shaded band shows the error associated with the background expectation and includes both the MC statistical uncertainty and the post-fit systematic uncertainty described in Chap. 8.

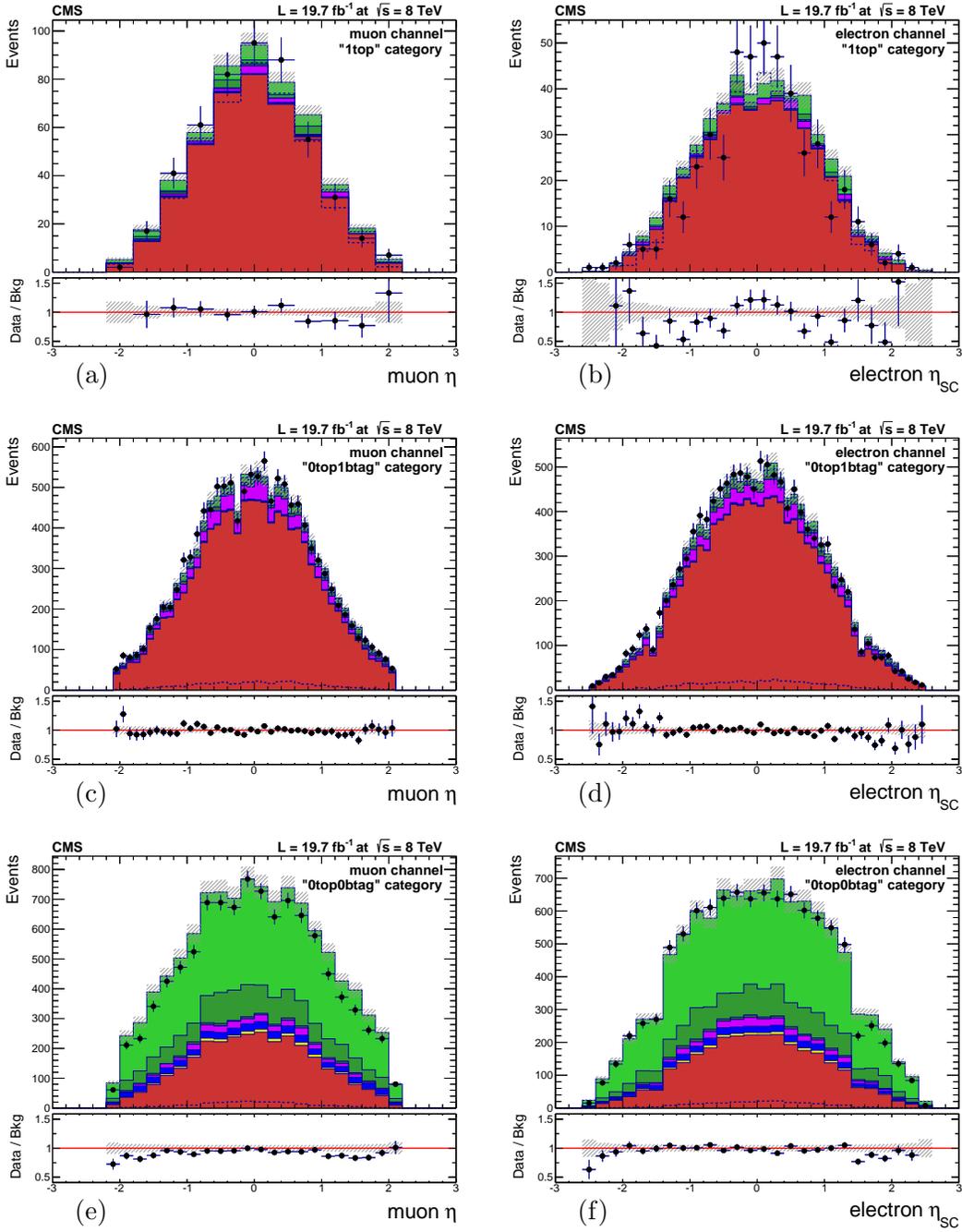


Figure B.3: Data/MC comparison for the lepton η after all analysis cuts are applied. Events are split into the following categories: muon “1top” (a), electron “1top” (b), muon “0top1btag” (c), electron “0top1btag” (d), muon “0top0btag” (e), electron “0top0btag” (f). Each background process is corrected by scale factors from the maximum-likelihood fit to data as described in Chap. 8. The signal MC samples are normalized to a cross section of 1 pb. The shaded band shows the error associated with the background expectation and includes both the MC statistical uncertainty and the post-fit systematic uncertainty described in Chap. 8.

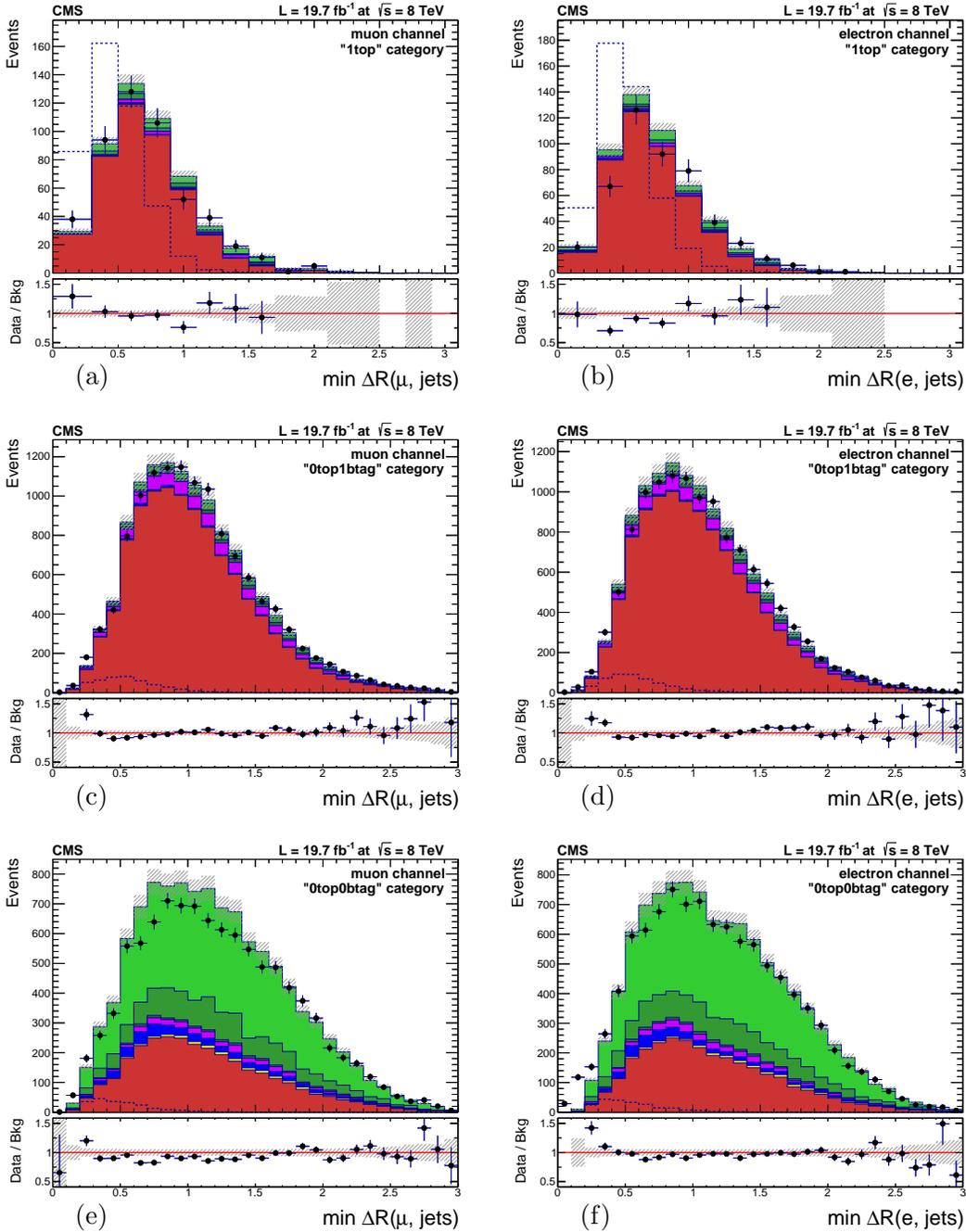


Figure B.4: Data/MC comparison for the minimum ΔR -distance between lepton and AK5-jets after all analysis cuts are applied. Events are split into the following categories: muon “1top” (a), electron “1top” (b), muon “0top1btag” (c), electron “0top1btag” (d), muon “0top0btag” (e), electron “0top0btag” (f). Each background process is corrected by scale factors from the maximum-likelihood fit to data as described in Chap. 8. The signal MC samples are normalized to a cross section of 1 pb. The shaded band shows the error associated with the background expectation and includes both the MC statistical uncertainty and the post-fit systematic uncertainty described in Chap. 8.

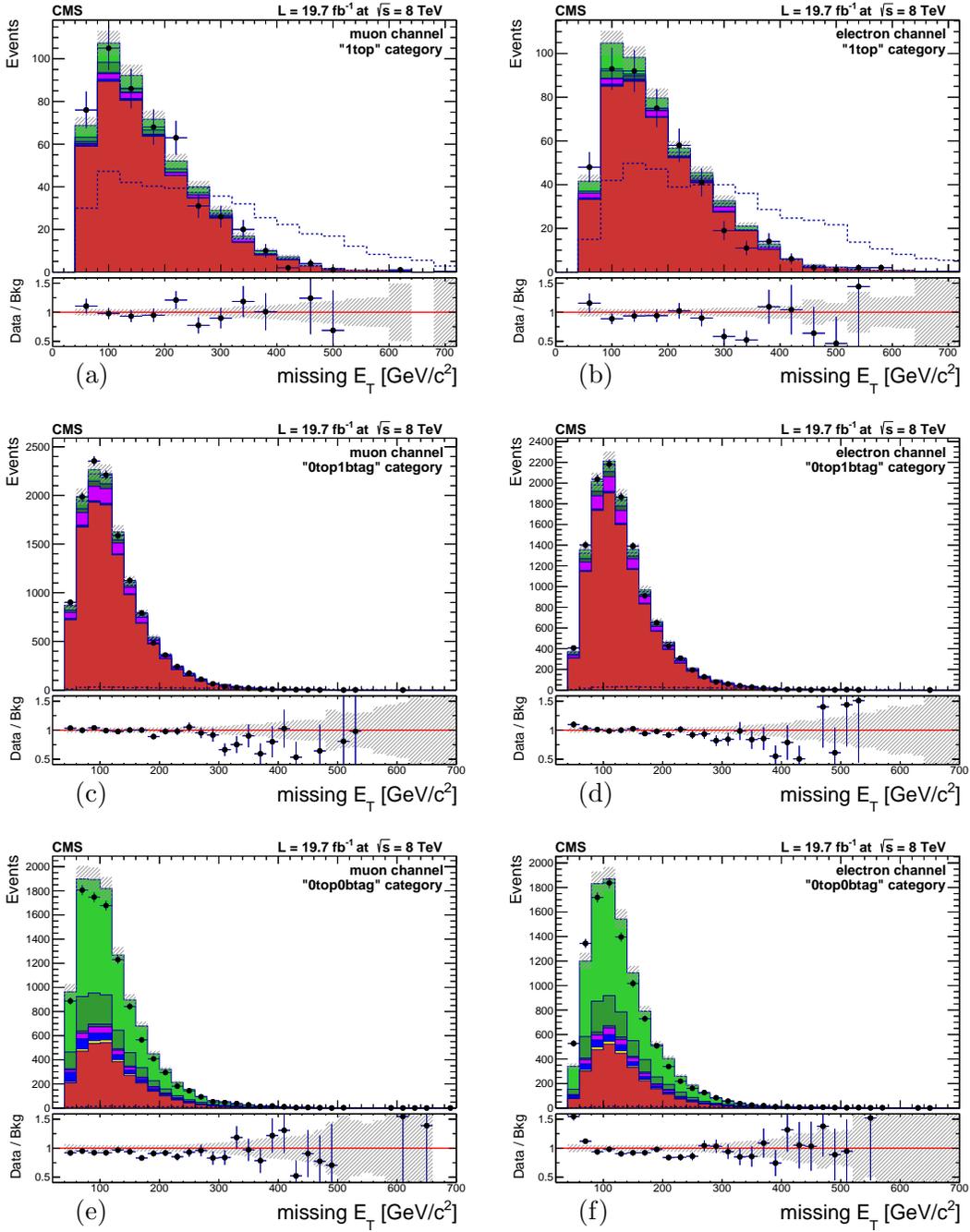


Figure B.5: Data/MC comparison for the missing E_T after all analysis cuts are applied. Events are split into the following categories: muon “1top” (a), electron “1top” (b), muon “0top1btag” (c), electron “0top1btag” (d), muon “0top0btag” (e), electron “0top0btag” (f). Each background process is corrected by scale factors from the maximum-likelihood fit to data as described in Chap. 8. The signal MC samples are normalized to a cross section of 1 pb. The shaded band shows the error associated with the background expectation and includes both the MC statistical uncertainty and the post-fit systematic uncertainty described in Chap. 8.

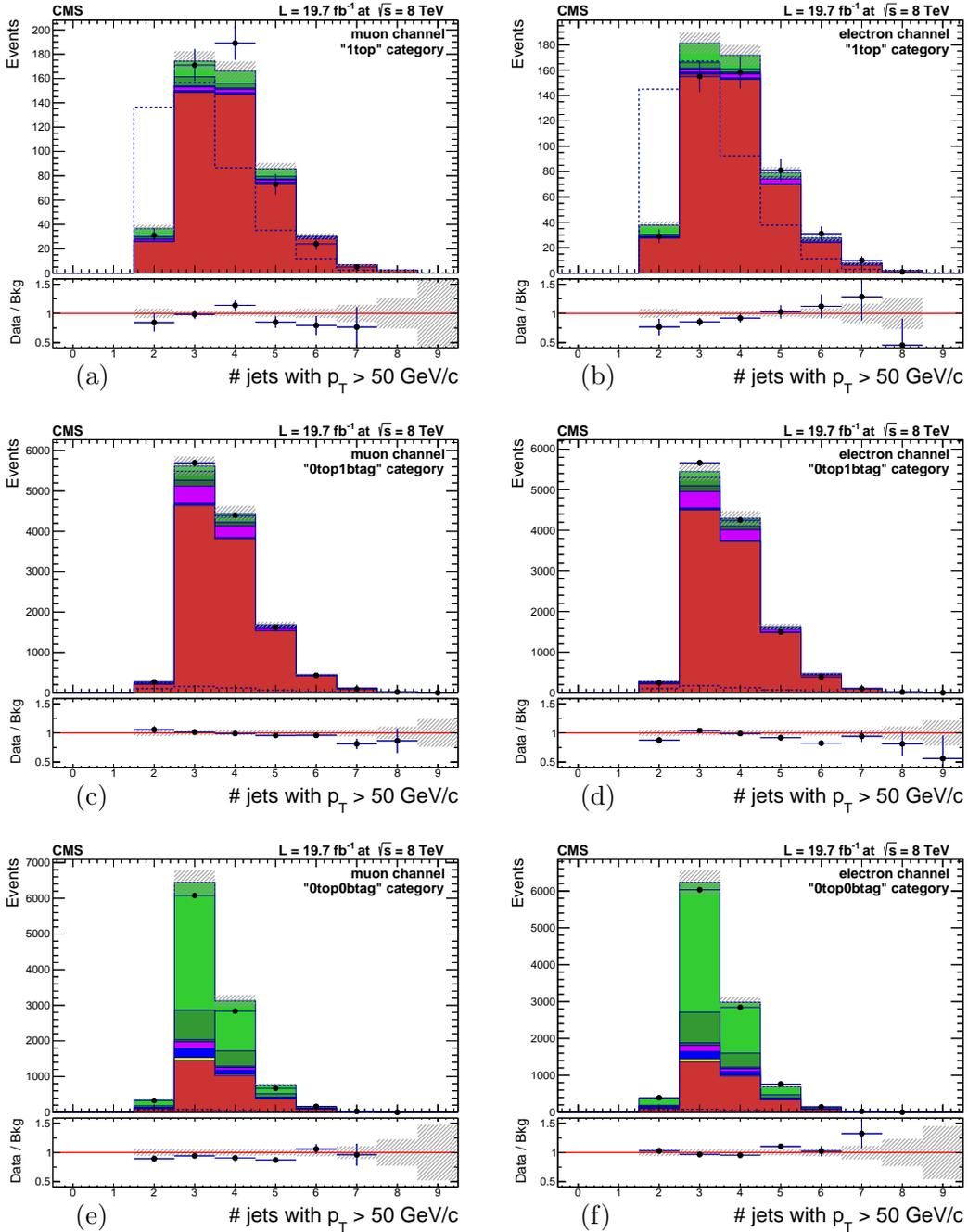


Figure B.6: Data/MC comparison for the number of AK5-jets with $p_T > 50$ GeV after all analysis cuts are applied. Events are split into the following categories: muon “1top” (a), electron “1top” (b), muon “0top1btag” (c), electron “0top1btag” (d), muon “0top0btag” (e), electron “0top0btag” (f). Each background process is corrected by scale factors from the maximum-likelihood fit to data as described in Chap. 8. The signal MC samples are normalized to a cross section of 1 pb. The shaded band shows the error associated with the background expectation and includes both the MC statistical uncertainty and the post-fit systematic uncertainty described in Chap. 8.

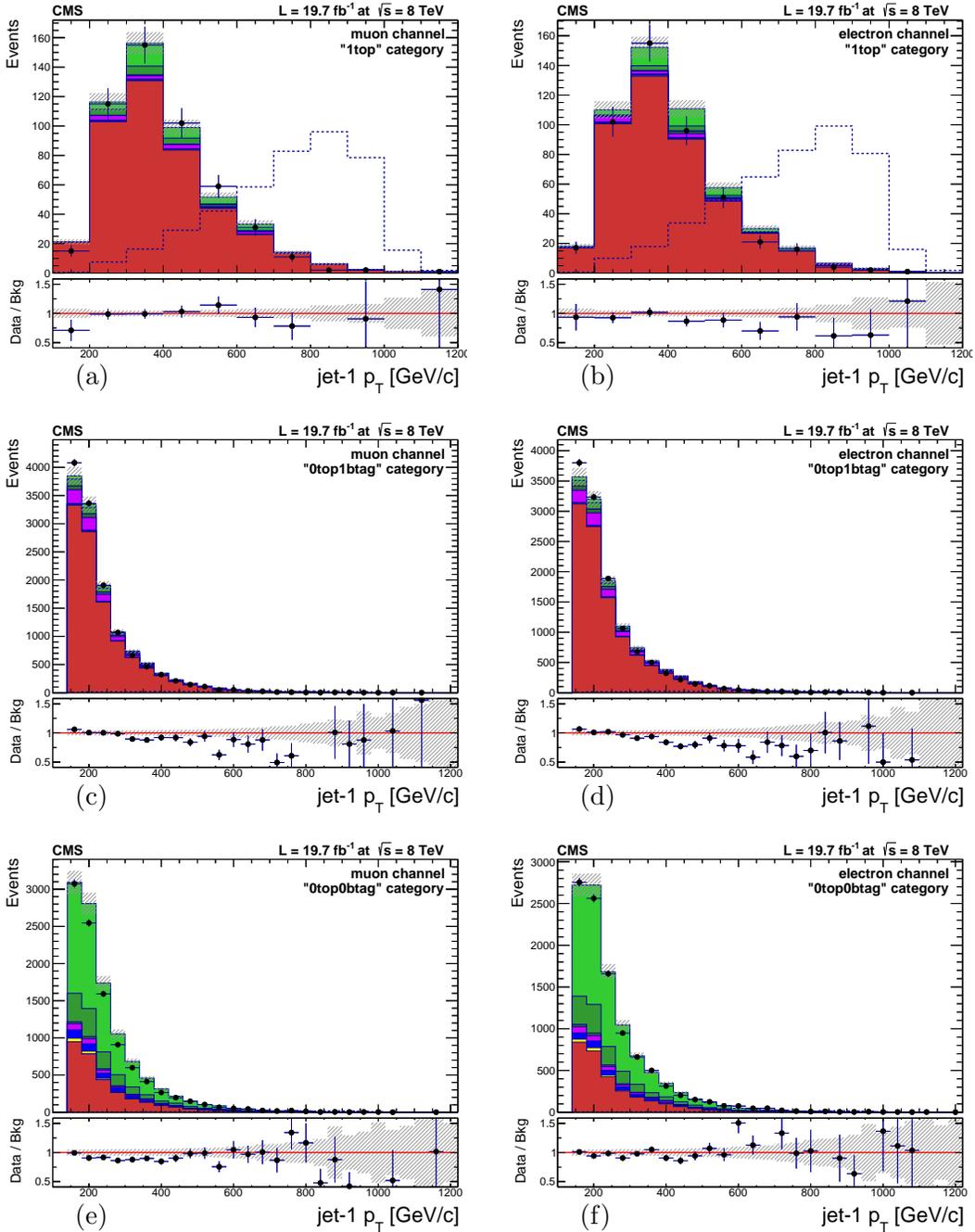


Figure B.7: Data/MC comparison for the leading AK5-jet p_T after all analysis cuts are applied. Events are split into the following categories: muon “1top” (a), electron “1top” (b), muon “0top1btag” (c), electron “0top1btag” (d), muon “0top0btag” (e), electron “0top0btag” (f). Each background process is corrected by scale factors from the maximum-likelihood fit to data as described in Chap. 8. The signal MC samples are normalized to a cross section of 1 pb. The shaded band shows the error associated with the background expectation and includes both the MC statistical uncertainty and the post-fit systematic uncertainty described in Chap. 8.

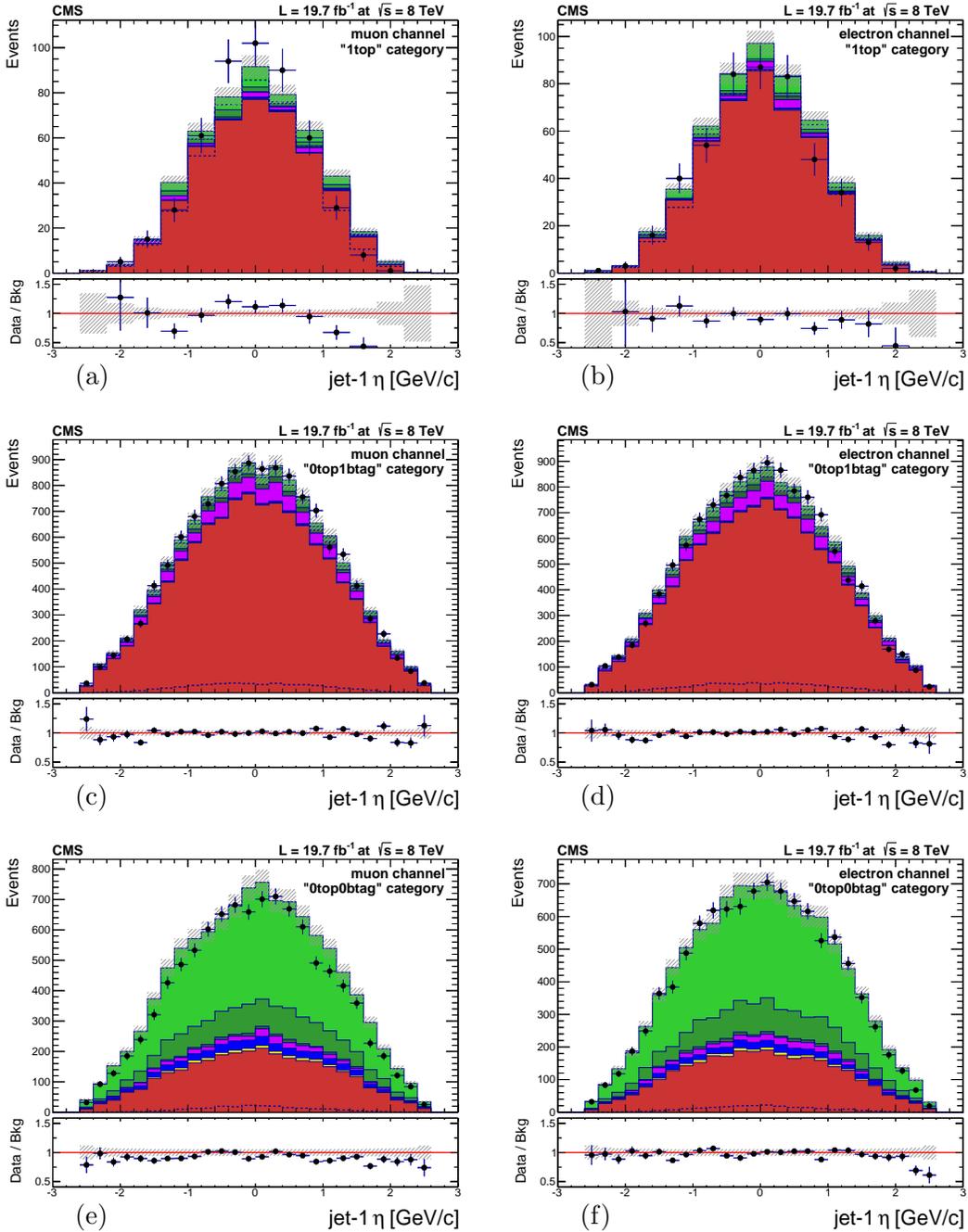


Figure B.8: Data/MC comparison for the leading AK5-jet η after all analysis cuts are applied. Events are split into the following categories: muon “1top” (a), electron “1top” (b), muon “0top1btag” (c), electron “0top1btag” (d), muon “0top0btag” (e), electron “0top0btag” (f). Each background process is corrected by scale factors from the maximum-likelihood fit to data as described in Chap. 8. The signal MC samples are normalized to a cross section of 1 pb. The shaded band shows the error associated with the background expectation and includes both the MC statistical uncertainty and the post-fit systematic uncertainty described in Chap. 8.

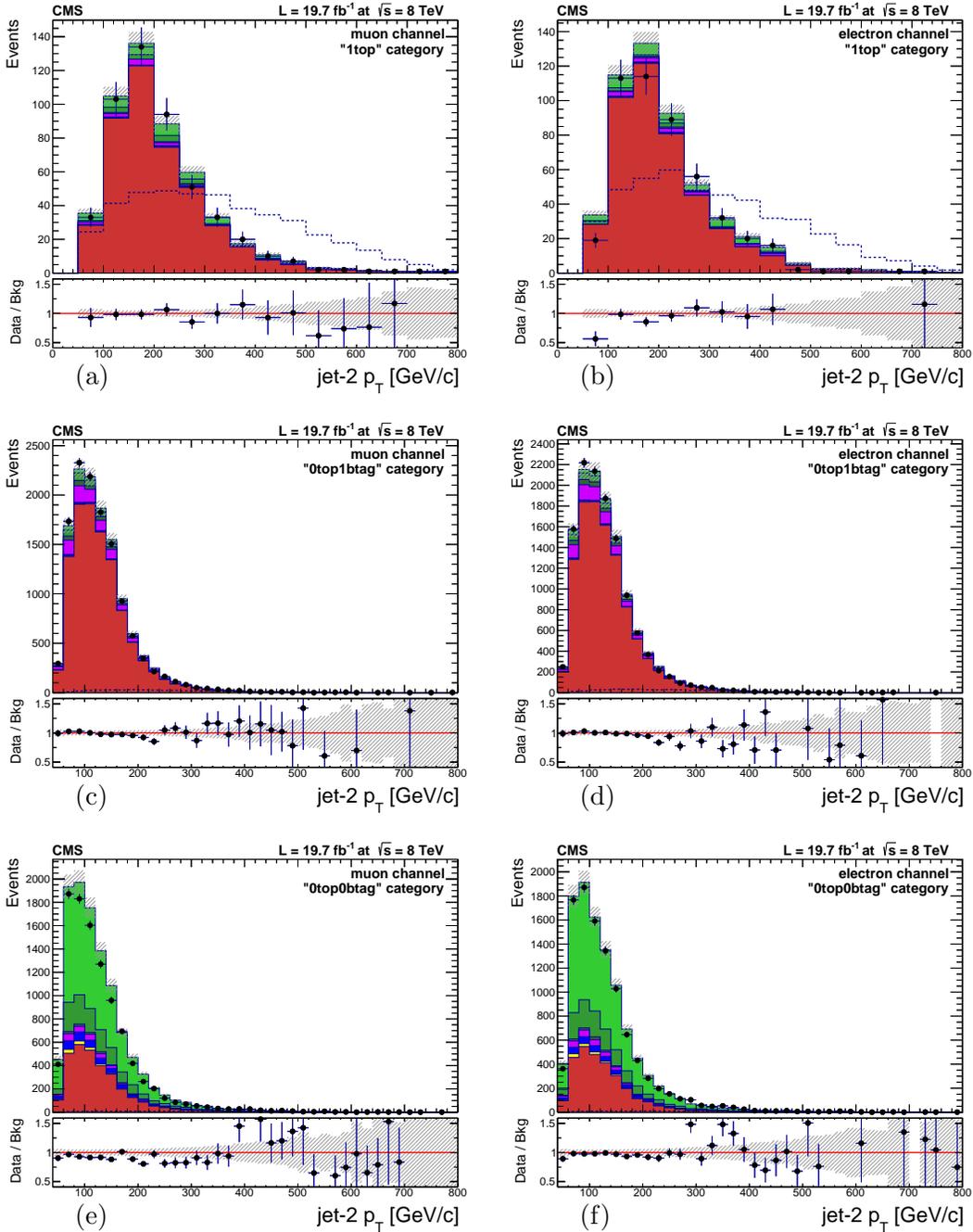


Figure B.9: Data/MC comparison for the 2nd-leading AK5-jet p_T after all analysis cuts are applied. Events are split into the following categories: muon “1top” (a), electron “1top” (b), muon “0top1btag” (c), electron “0top1btag” (d), muon “0top0btag” (e), electron “0top0btag” (f). Each background process is corrected by scale factors from the maximum-likelihood fit to data as described in Chap. 8. The signal MC samples are normalized to a cross section of 1 pb. The shaded band shows the error associated with the background expectation and includes both the MC statistical uncertainty and the post-fit systematic uncertainty described in Chap. 8.

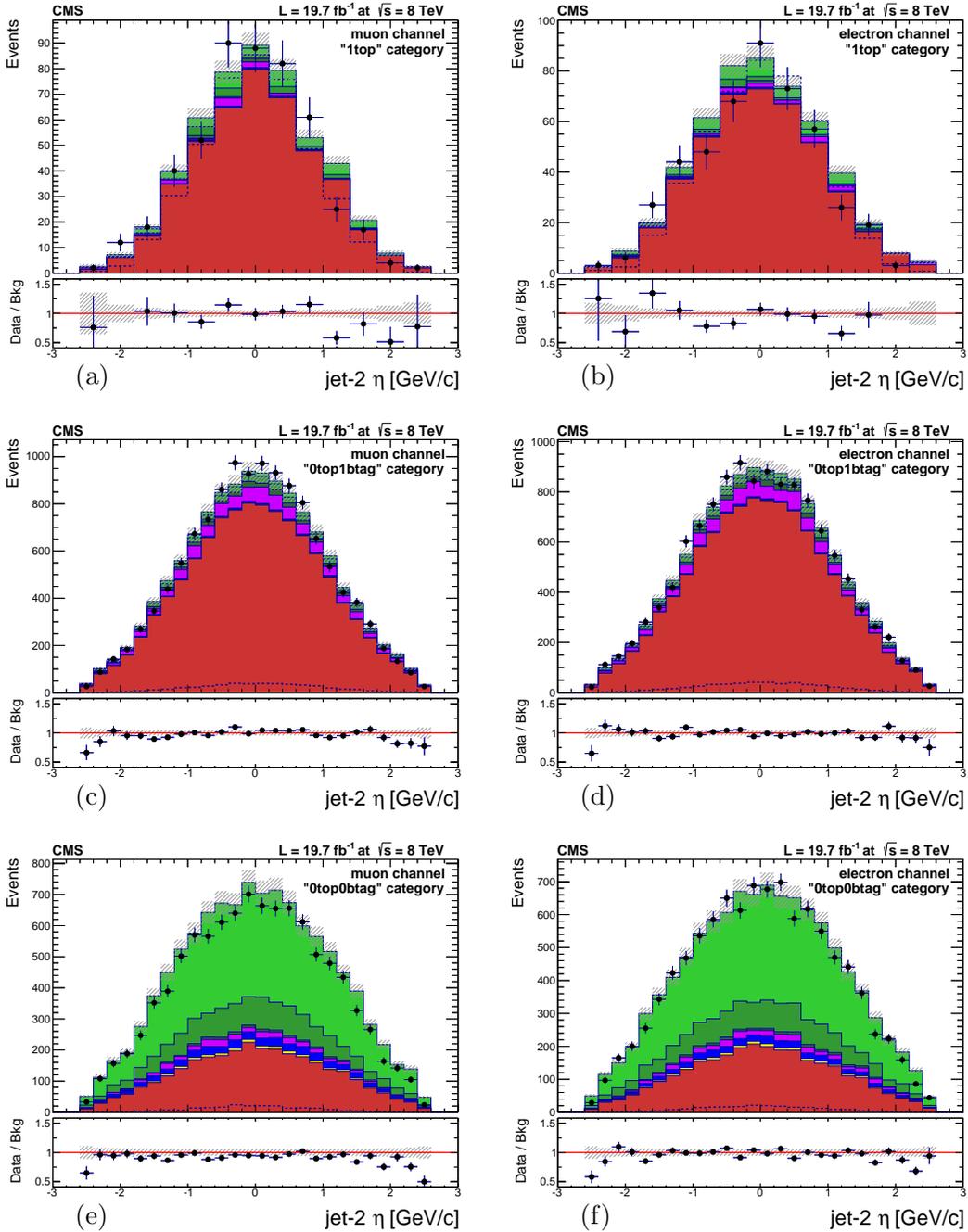


Figure B.10: Data/MC comparison for the 2nd-leading AK5-jet η after all analysis cuts are applied. Events are split into the following categories: muon “1top” (a), electron “1top” (b), muon “0top1btag” (c), electron “0top1btag” (d), muon “0top0btag” (e), electron “0top0btag” (f). Each background process is corrected by scale factors from the maximum-likelihood fit to data as described in Chap. 8. The signal MC samples are normalized to a cross section of 1 pb. The shaded band shows the error associated with the background expectation and includes both the MC statistical uncertainty and the post-fit systematic uncertainty described in Chap. 8.

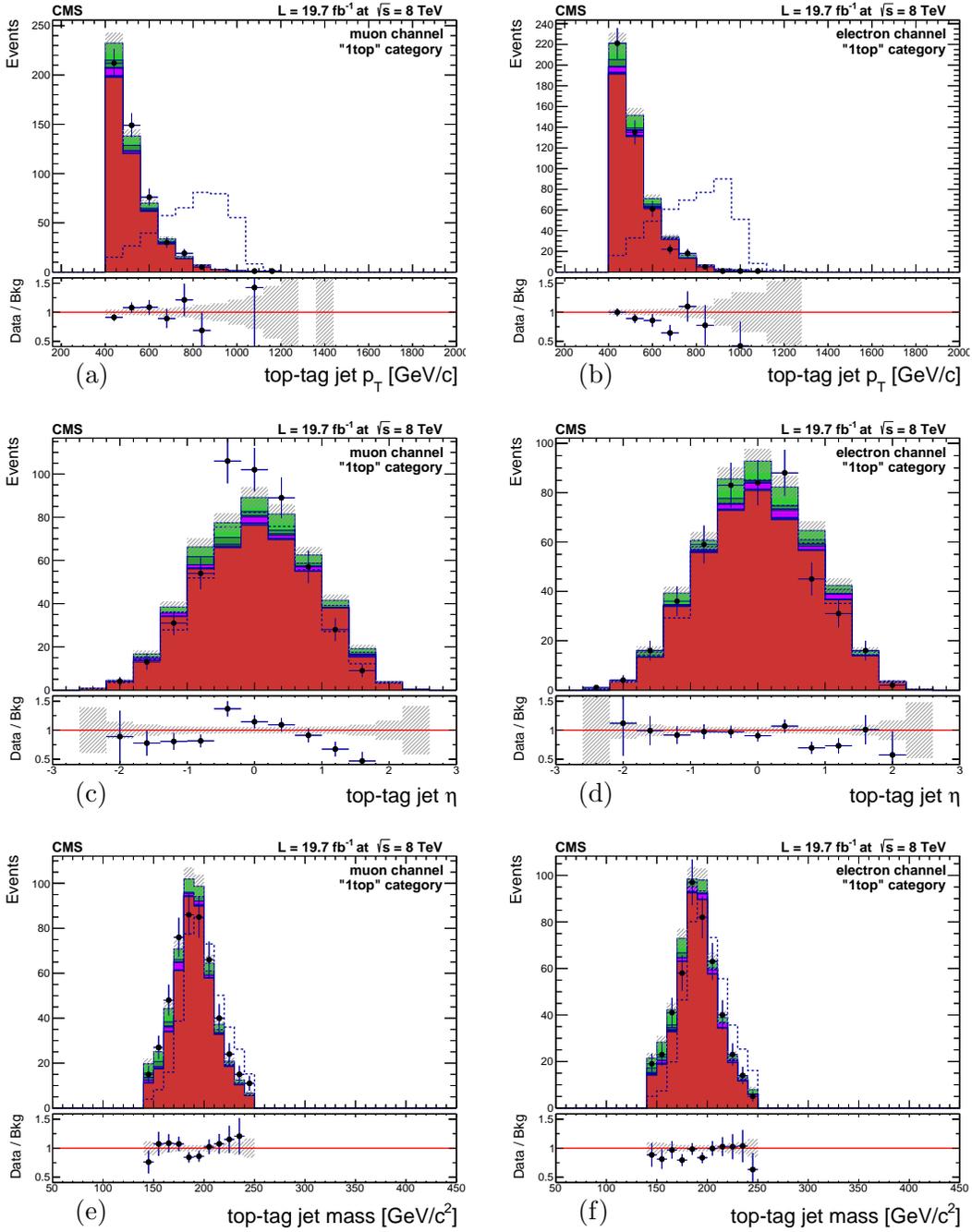


Figure B.11: Data/MC comparison for kinematic variables of the “top-tagged” CA8-jet in the “1top” event category: top-tag jet p_T (muon channel) (a), top-tag jet p_T (electron channel) (b), top-tag jet η (muon channel) (c), top-tag jet η (electron channel) (d), top-tag jet mass (muon channel) (e), top-tag jet mass (electron channel) (f). Each background process is corrected by scale factors from the maximum-likelihood fit to data as described in Chap. 8. The signal MC samples are normalized to a cross section of 1 pb. The shaded band shows the error associated with the background expectation and includes both the MC statistical uncertainty and the post-fit systematic uncertainty described in Chap. 8.

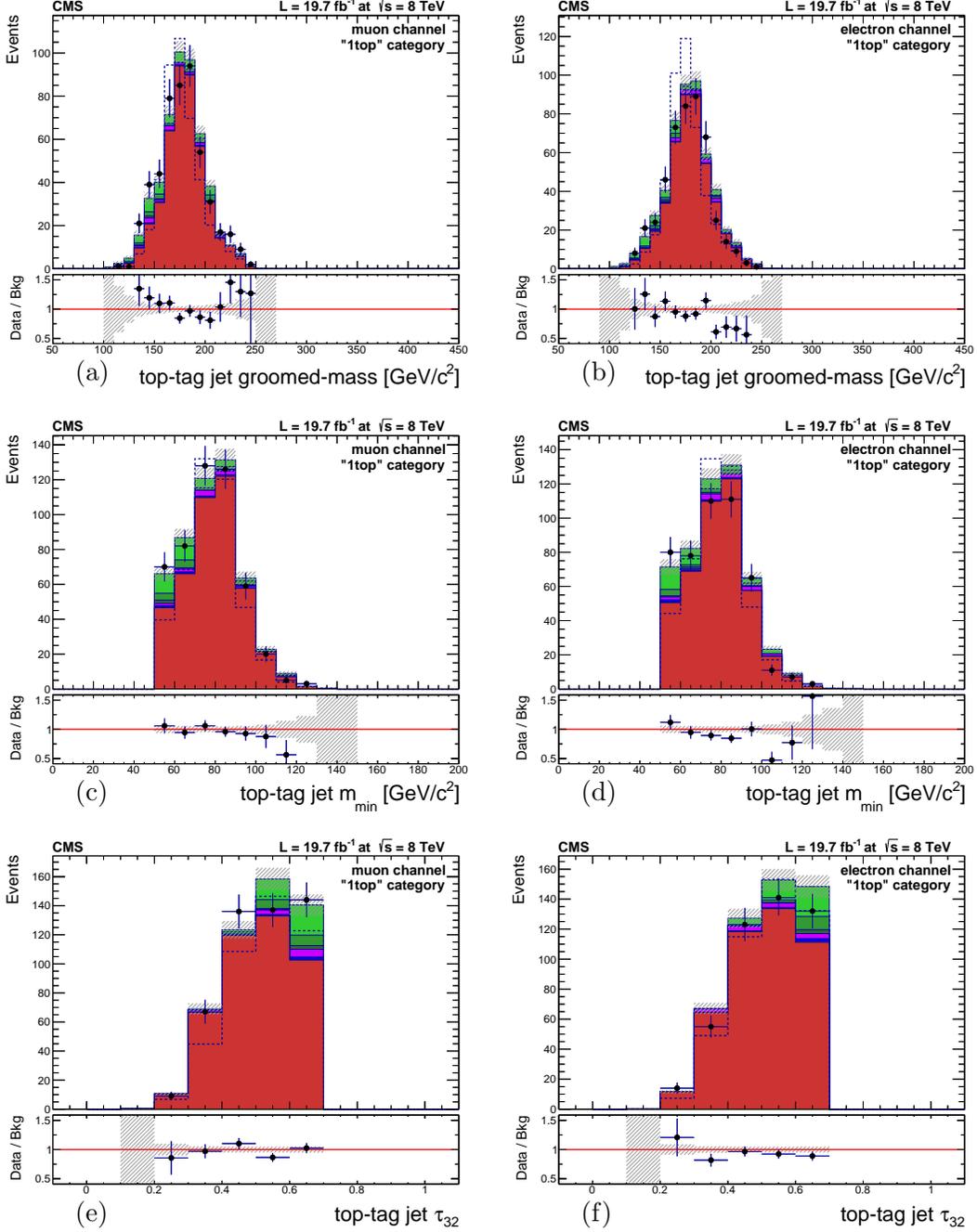


Figure B.12: Data/MC comparison for kinematic variables of the “top-tagged” CA8-jet in the “1top” event category: top-tag jet groomed mass (muon channel) (a), top-tag jet groomed mass (electron channel) (b), top-tag jet minimum subjet-pairwise mass (muon channel) (c), top-tag jet minimum subjet-pairwise mass (electron channel) (d), top-tag jet τ_{32} variable (muon channel) (e), top-tag jet τ_{32} variable (electron channel) (f). Each background process is corrected by scale factors from the maximum-likelihood fit to data as described in Chap. 8. The signal MC samples are normalized to a cross section of 1 pb. The shaded band shows the error associated with the background expectation and includes both the MC statistical uncertainty and the post-fit systematic uncertainty described in Chap. 8.

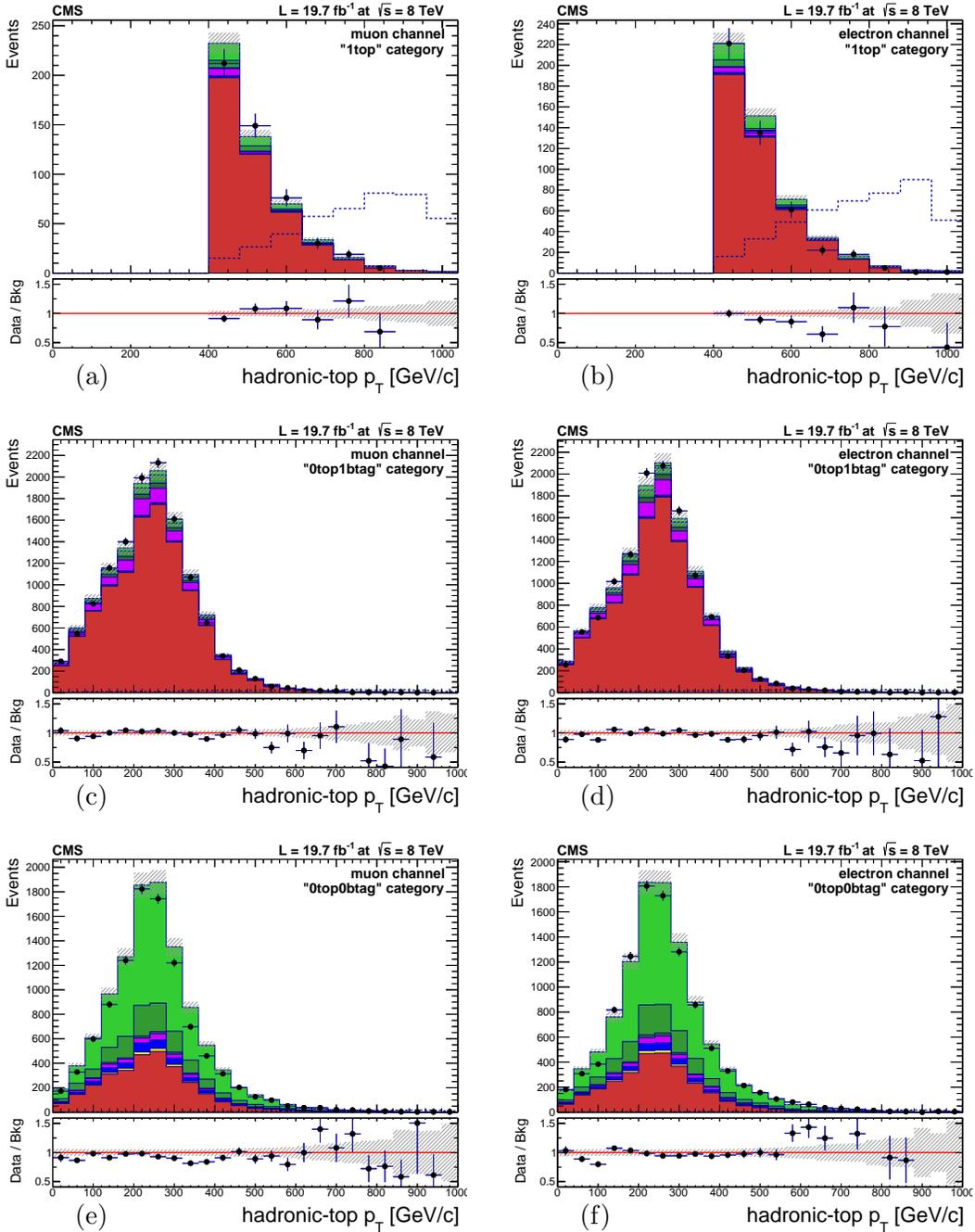


Figure B.13: Data/MC comparison for the hadronic-top p_T after all analysis cuts are applied. Events are split into the following categories: muon “1top” (a), electron “1top” (b), muon “0top1btag” (c), electron “0top1btag” (d), muon “0top0btag” (e), electron “0top0btag” (f). Each background process is corrected by scale factors from the maximum-likelihood fit to data as described in Chap. 8. The signal MC samples are normalized to a cross section of 1 pb. The shaded band shows the error associated with the background expectation and includes both the MC statistical uncertainty and the post-fit systematic uncertainty described in Chap. 8.

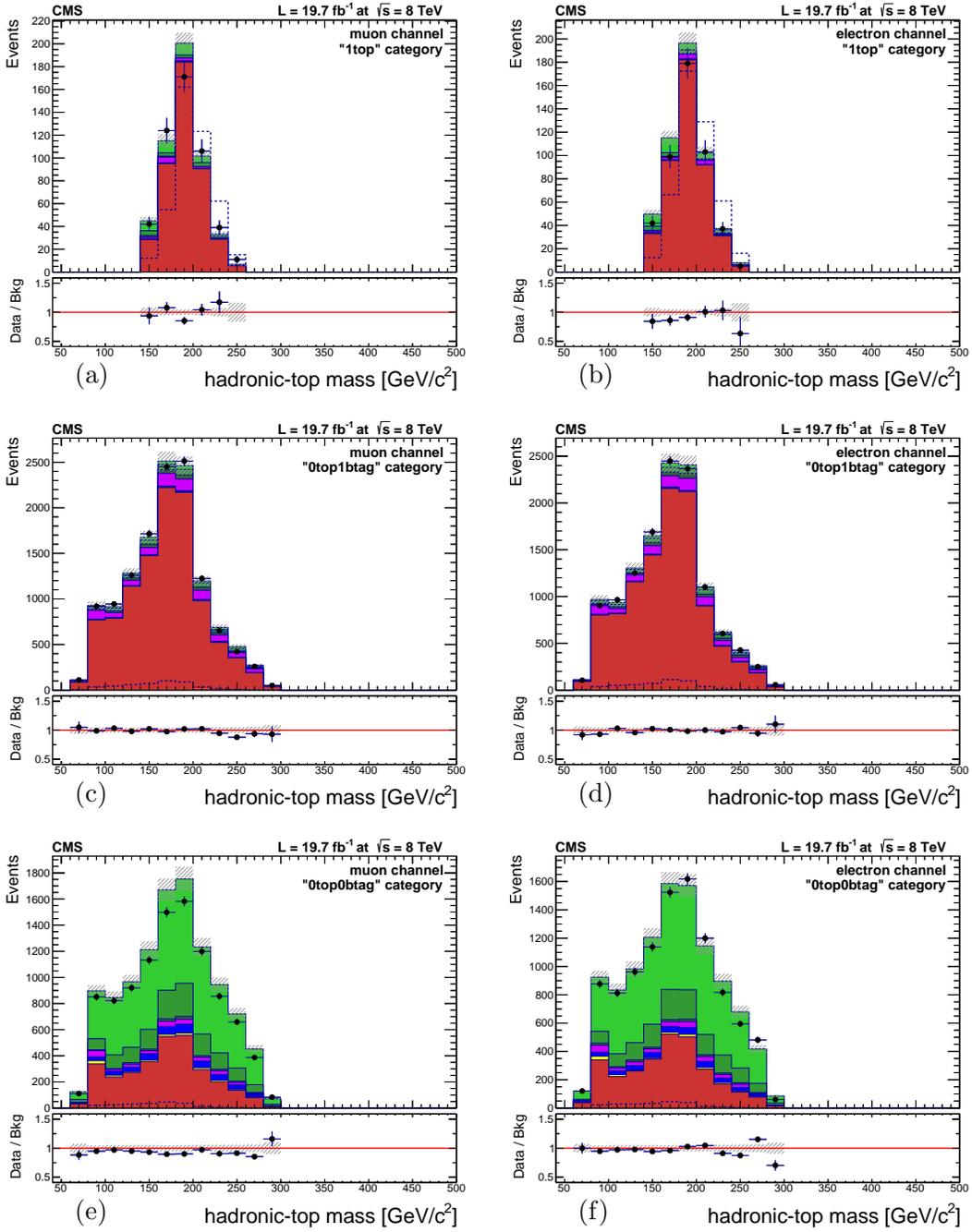


Figure B.14: Data/MC comparison for the hadronic-top mass after all analysis cuts are applied. Events are split into the following categories: muon “1top” (a), electron “1top” (b), muon “0top1btag” (c), electron “0top1btag” (d), muon “0top0btag” (e), electron “0top0btag” (f). Each background process is corrected by scale factors from the maximum-likelihood fit to data as described in Chap. 8. The signal MC samples are normalized to a cross section of 1 pb. The shaded band shows the error associated with the background expectation and includes both the MC statistical uncertainty and the post-fit systematic uncertainty described in Chap. 8.

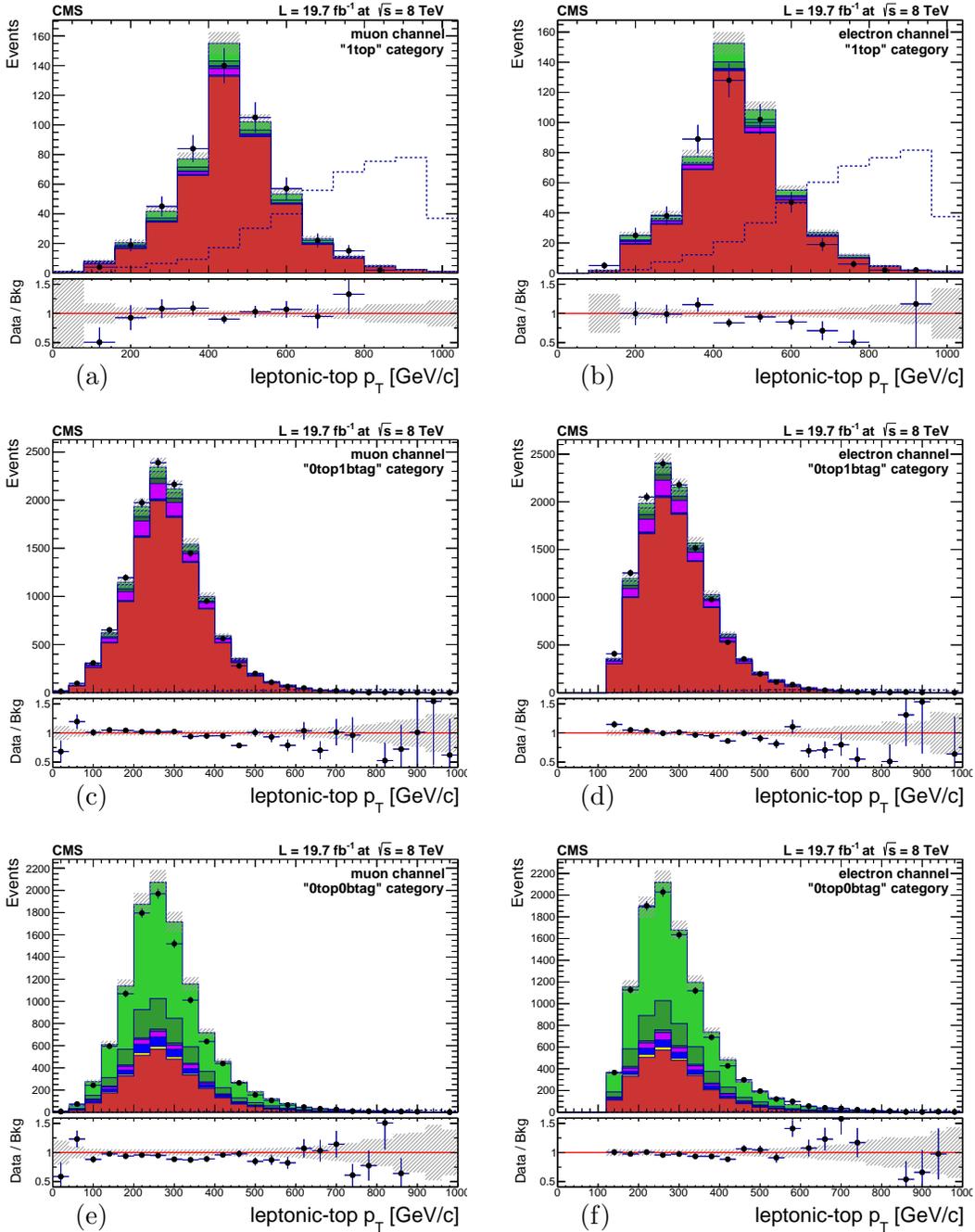


Figure B.15: Data/MC comparison for the leptonic-top p_T after all analysis cuts are applied. Events are split into the following categories: muon "1top" (a), electron "1top" (b), muon "0top1btag" (c), electron "0top1btag" (d), muon "0top0btag" (e), electron "0top0btag" (f). Each background process is corrected by scale factors from the maximum-likelihood fit to data as described in Chap. 8. The signal MC samples are normalized to a cross section of 1 pb. The shaded band shows the error associated with the background expectation and includes both the MC statistical uncertainty and the post-fit systematic uncertainty described in Chap. 8.

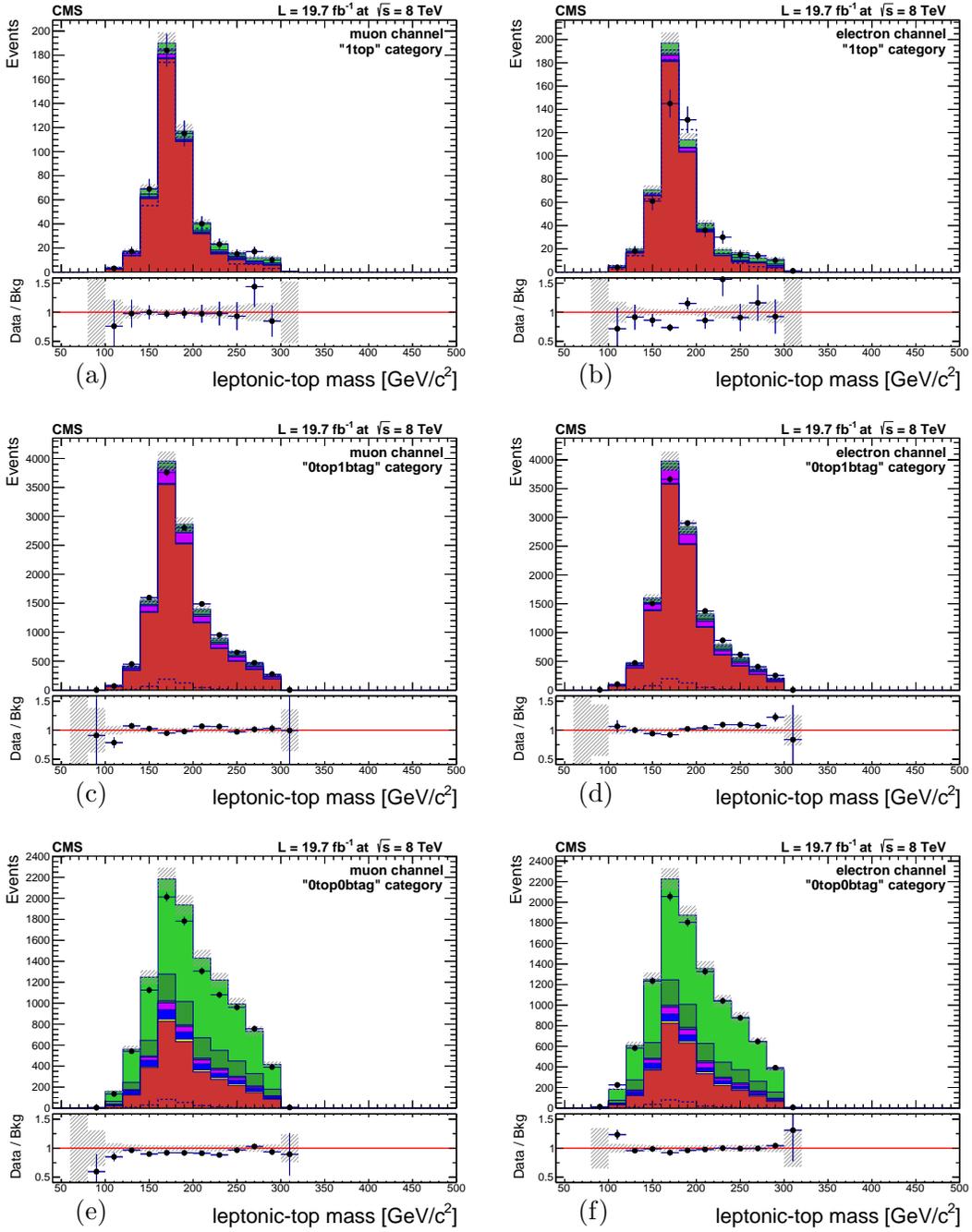


Figure B.16: Data/MC comparison for the leptonic-top mass after all analysis cuts are applied. Events are split into the following categories: muon “1top” (a), electron “1top” (b), muon “0top1btag” (c), electron “0top1btag” (d), muon “0top0btag” (e), electron “0top0btag” (f). Each background process is corrected by scale factors from the maximum-likelihood fit to data as described in Chap. 8. The signal MC samples are normalized to a cross section of 1 pb. The shaded band shows the error associated with the background expectation and includes both the MC statistical uncertainty and the post-fit systematic uncertainty described in Chap. 8.

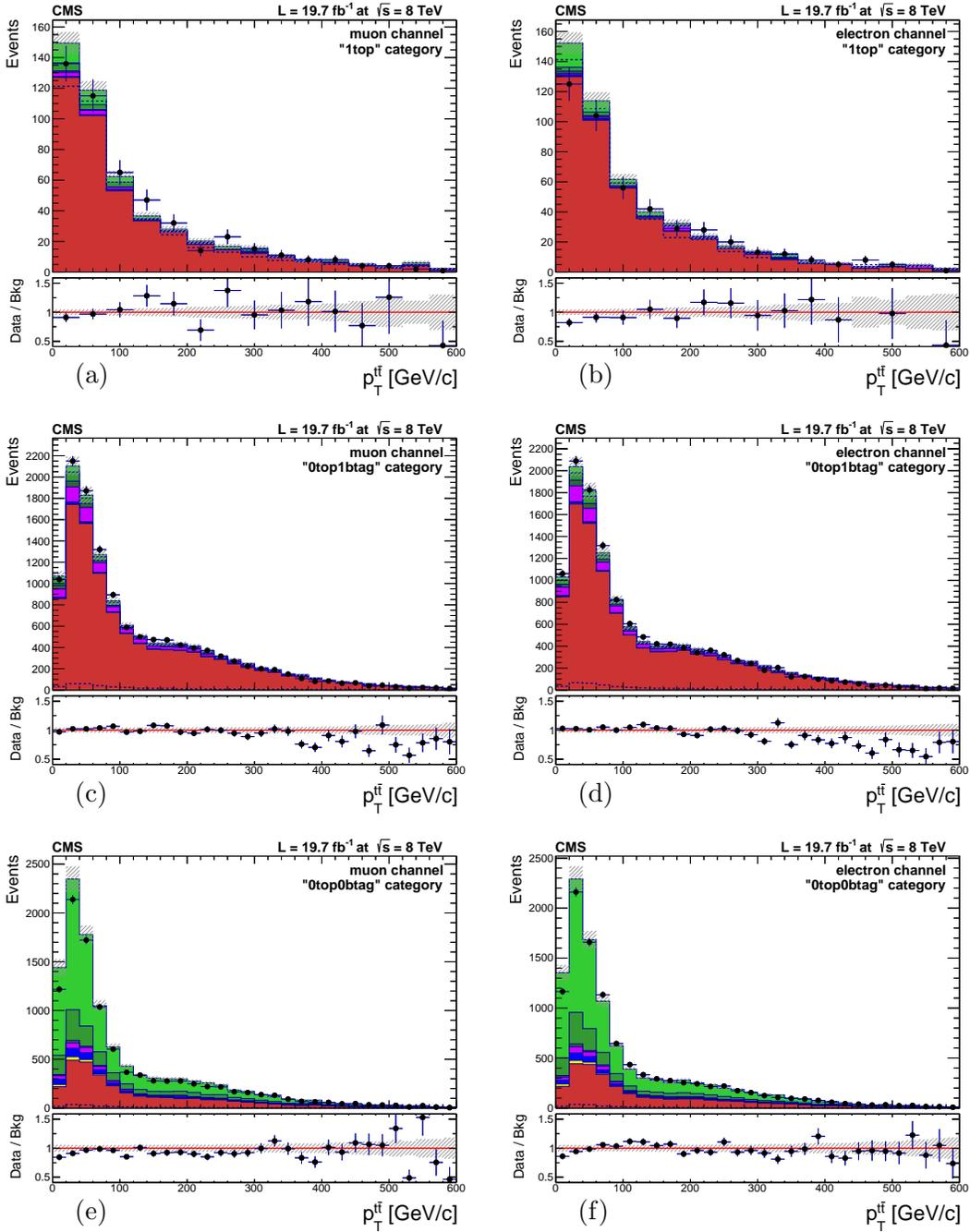


Figure B.17: Data/MC comparison for the p_T of the $t\bar{t}$ pair after all analysis cuts are applied. Events are split into the following categories: muon “1top” (a), electron “1top” (b), muon “0top1btag” (c), electron “0top1btag” (d), muon “0top0btag” (e), electron “0top0btag” (f). Each background process is corrected by scale factors from the maximum-likelihood fit to data as described in Chap. 8. The signal MC samples are normalized to a cross section of 1 pb. The shaded band shows the error associated with the background expectation and includes both the MC statistical uncertainty and the post-fit systematic uncertainty described in Chap. 8.

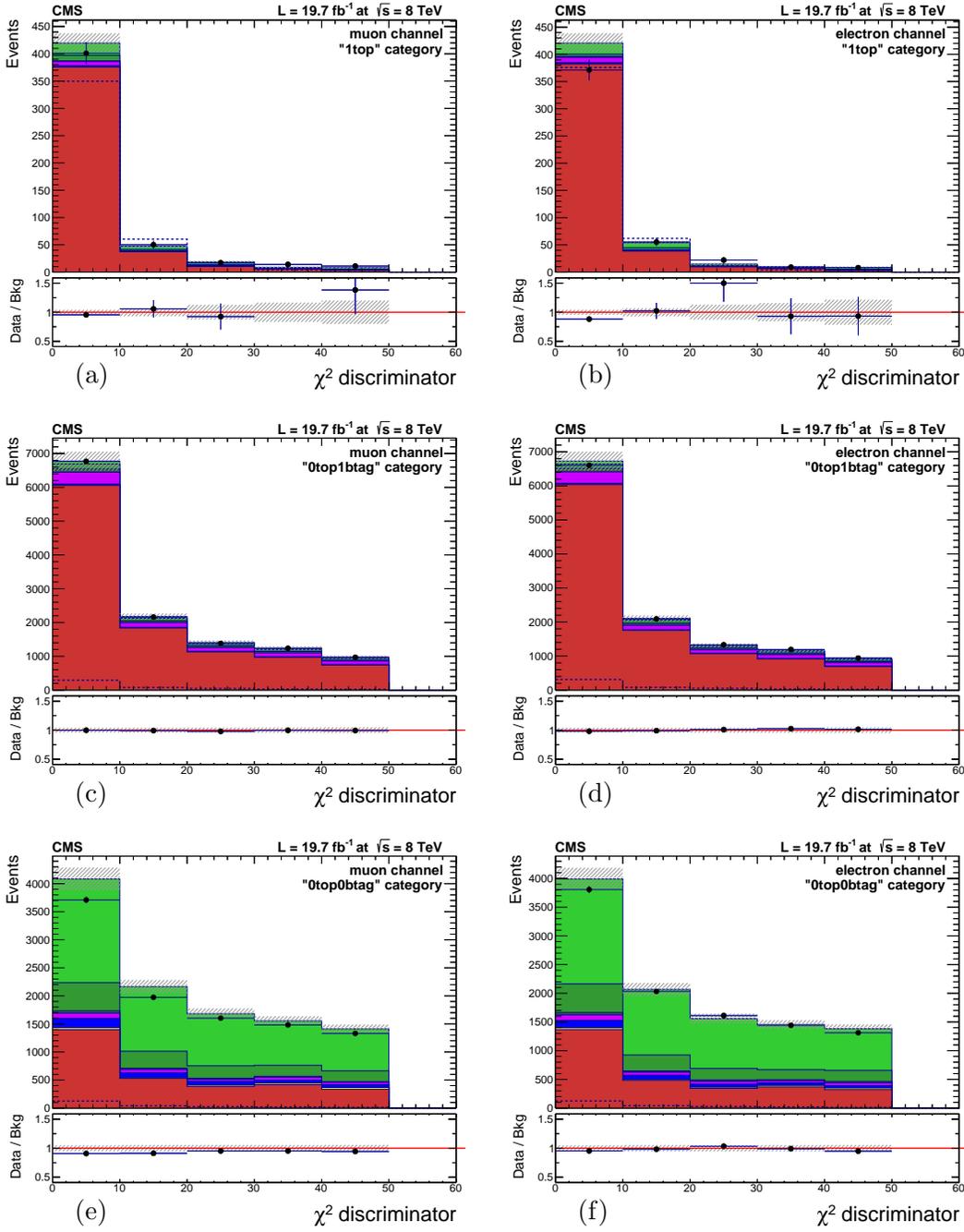


Figure B.18: Data/MC comparison for the χ^2 of the reconstructed $t\bar{t}$ pair after all analysis cuts are applied. Events are split into the following categories: muon “1top” (a), electron “1top” (b), muon “0top1btag” (c), electron “0top1btag” (d), muon “0top0btag” (e), electron “0top0btag” (f). Each background process is corrected by scale factors from the maximum-likelihood fit to data as described in Chap. 8. The signal MC samples are normalized to a cross section of 1 pb. The shaded band shows the error associated with the background expectation and includes both the MC statistical uncertainty and the post-fit systematic uncertainty described in Chap. 8.

C. SYSTEMATIC UNCERTAINTY TEMPLATES

This appendix contains the templates used to model the $\pm 1\sigma$ variation of several systematic uncertainties described in Sec. 8.1.2 in the muon channel. Similar templates are obtained for the electron channel but are omitted for conciseness. The effect of the systematic uncertainties is shown for the sum of SM MC backgrounds and a narrow Z' signal with $M_{Z'} = 2$ TeV.

Figures C.1–C.10 show the $\pm 1\sigma$ effect of each systematic on the $t\bar{t}$ invariant mass distribution, separately in the three event categories used to set the final exclusion limits.

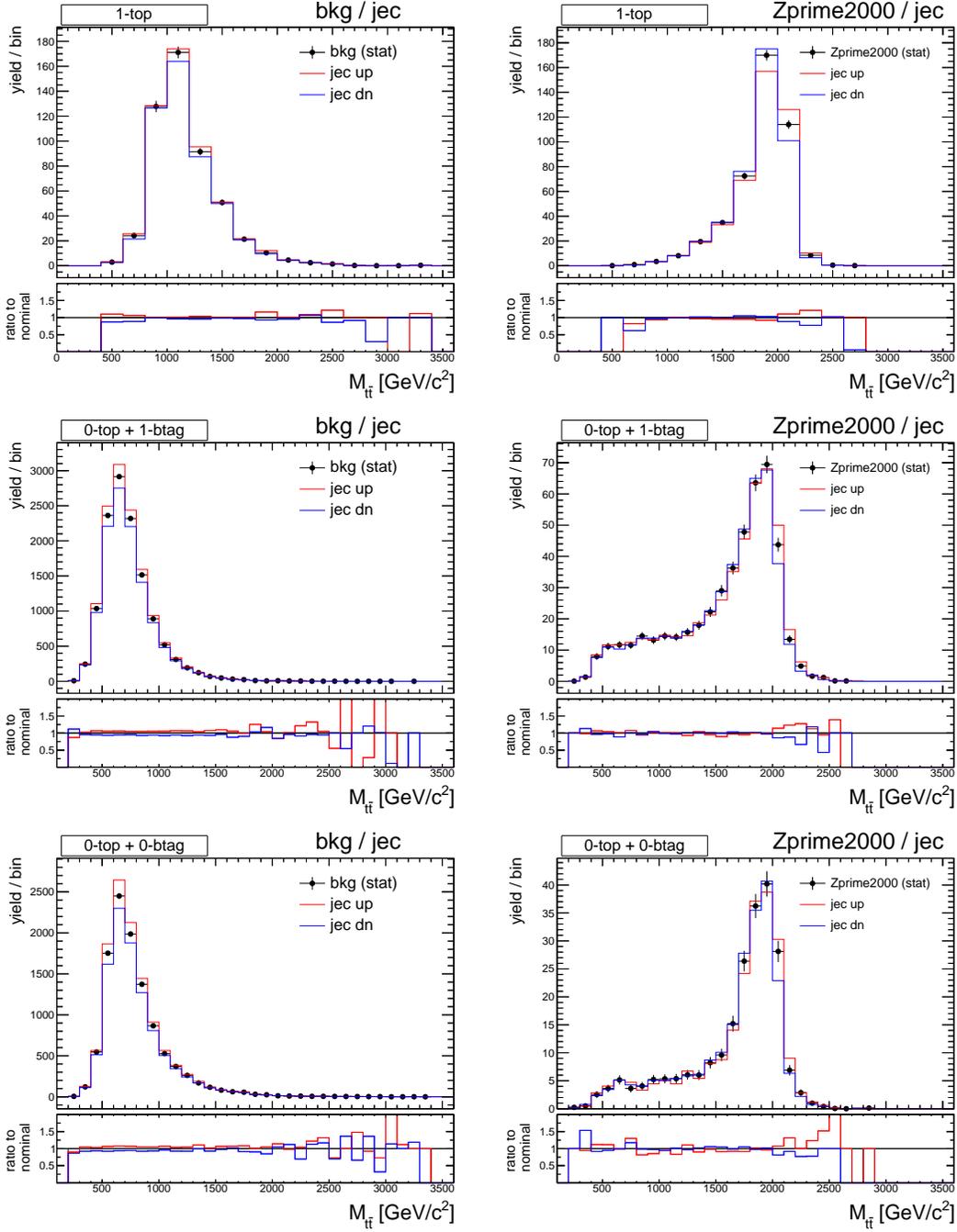


Figure C.1: $\pm 1\sigma$ systematic variation on the $M_{t\bar{t}}$ distribution due to the Jet Energy Scale: (left) total background, (right) narrow Z' signal with $M_{Z'} = 2$ TeV. Plots are shown for the three final categories in the muon channel: (from top to bottom) “1top”, “0top1btag”, and “0top0btag”. The nominal distribution (black line) is plotted with its statistical uncertainty, the $\pm 1\sigma$ templates account for differences in both shape and normalization with respect to the nominal sample.

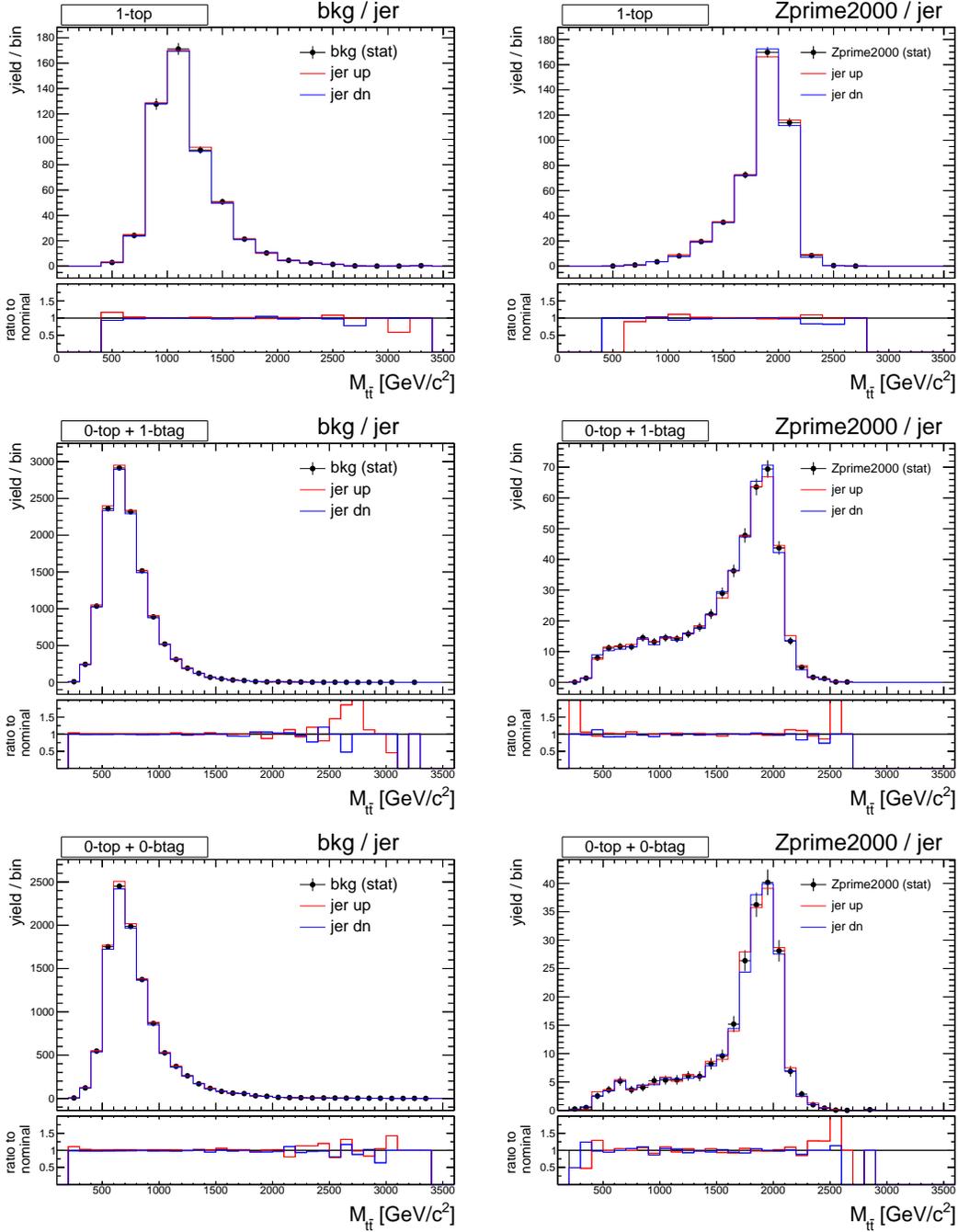


Figure C.2: $\pm 1\sigma$ systematic variation on the $M_{t\bar{t}}$ distribution due to the Jet Energy Resolution: (left) total background, (right) narrow Z' signal with $M_{Z'} = 2$ TeV. Plots are shown for the three final categories in the muon channel: (from top to bottom) “1top”, “0top1btag”, and “0top0btag”. The nominal distribution (black line) is plotted with its statistical uncertainty, the $\pm 1\sigma$ templates account for differences in both shape and normalization with respect to the nominal sample.

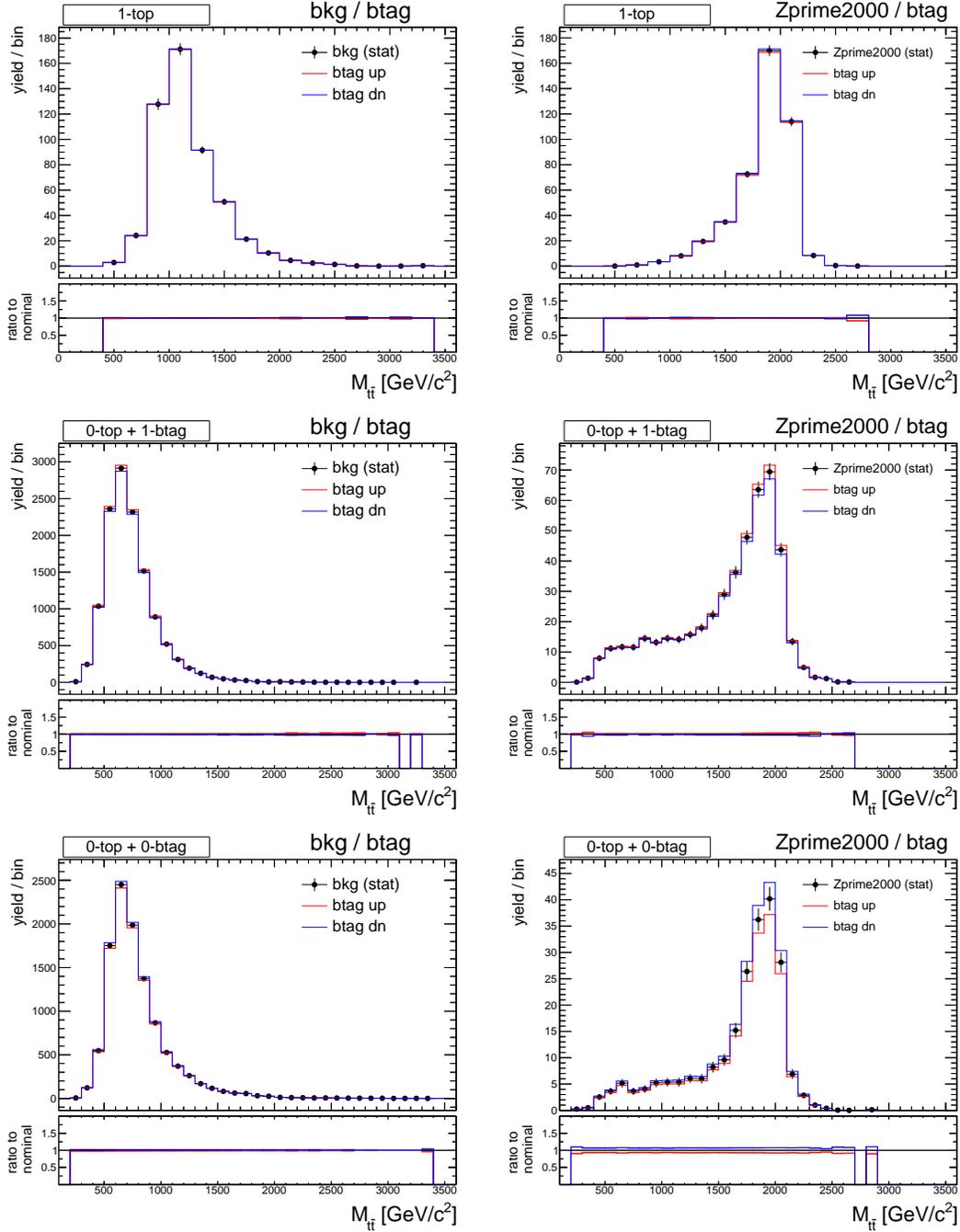


Figure C.3: $\pm 1\sigma$ systematic variation on the $M_{t\bar{t}}$ distribution due to the b -tagging SF: (left) total background, (right) narrow Z' signal with $M_{Z'} = 2$ TeV. Plots are shown for the three final categories in the muon channel: (from top to bottom) “1top”, “0top1btag”, and “0top0btag”. The nominal distribution (black line) is plotted with its statistical uncertainty, the $\pm 1\sigma$ templates account for differences in both shape and normalization with respect to the nominal sample.

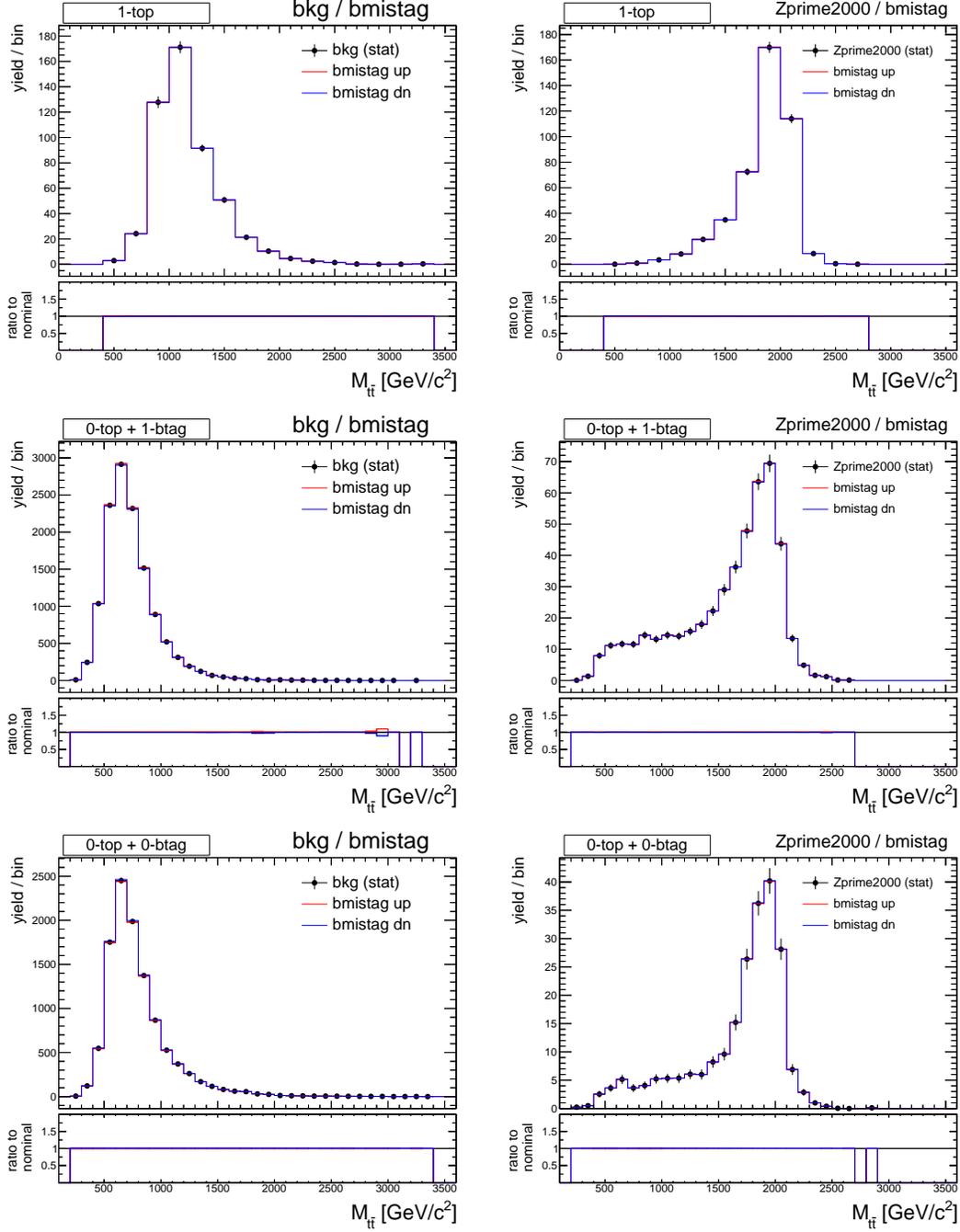


Figure C.4: $\pm 1\sigma$ systematic variation on the $M_{t\bar{t}}$ distribution due to the b -mistag SF: (left) total background, (right) narrow Z' signal with $M_{Z'} = 2$ TeV. Plots are shown for the three final categories in the muon channel: (from top to bottom) “1top”, “0top1btag”, and “0top0btag”. The nominal distribution (black line) is plotted with its statistical uncertainty, the $\pm 1\sigma$ templates account for differences in both shape and normalization with respect to the nominal sample.

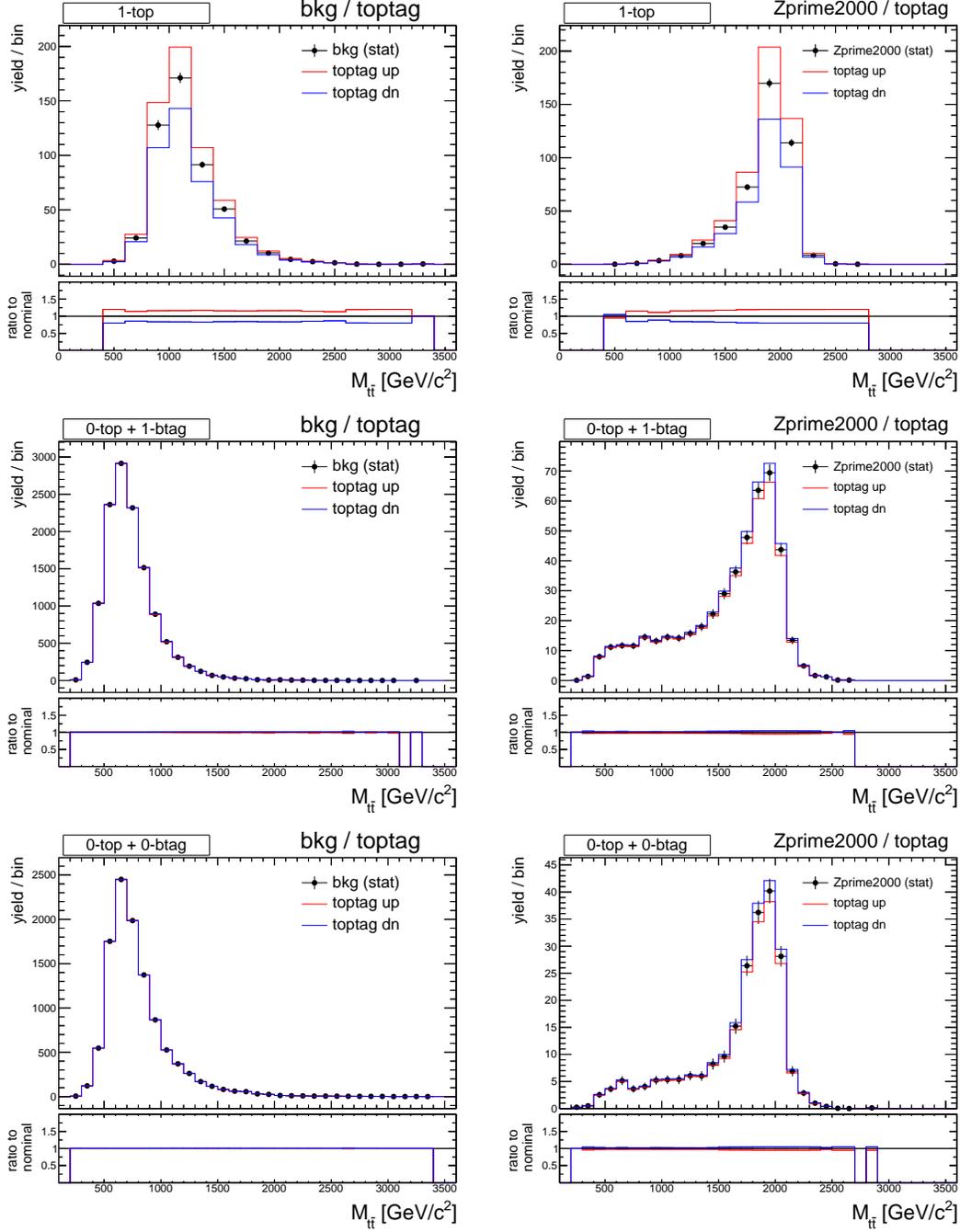


Figure C.5: Effect on the $M_{t\bar{t}}$ distribution due to a 20% systematic variation of the top-tagging SF: (left) total background, (right) narrow Z' signal with $M_{Z'} = 2$ TeV. Plots are shown for the three final categories in the muon channel: (from top to bottom) “1top”, “0top1btag”, and “0top0btag”. The nominal distribution (black line) is plotted with its statistical uncertainty, the $\pm 1\sigma$ templates account for differences in both shape and normalization with respect to the nominal sample.

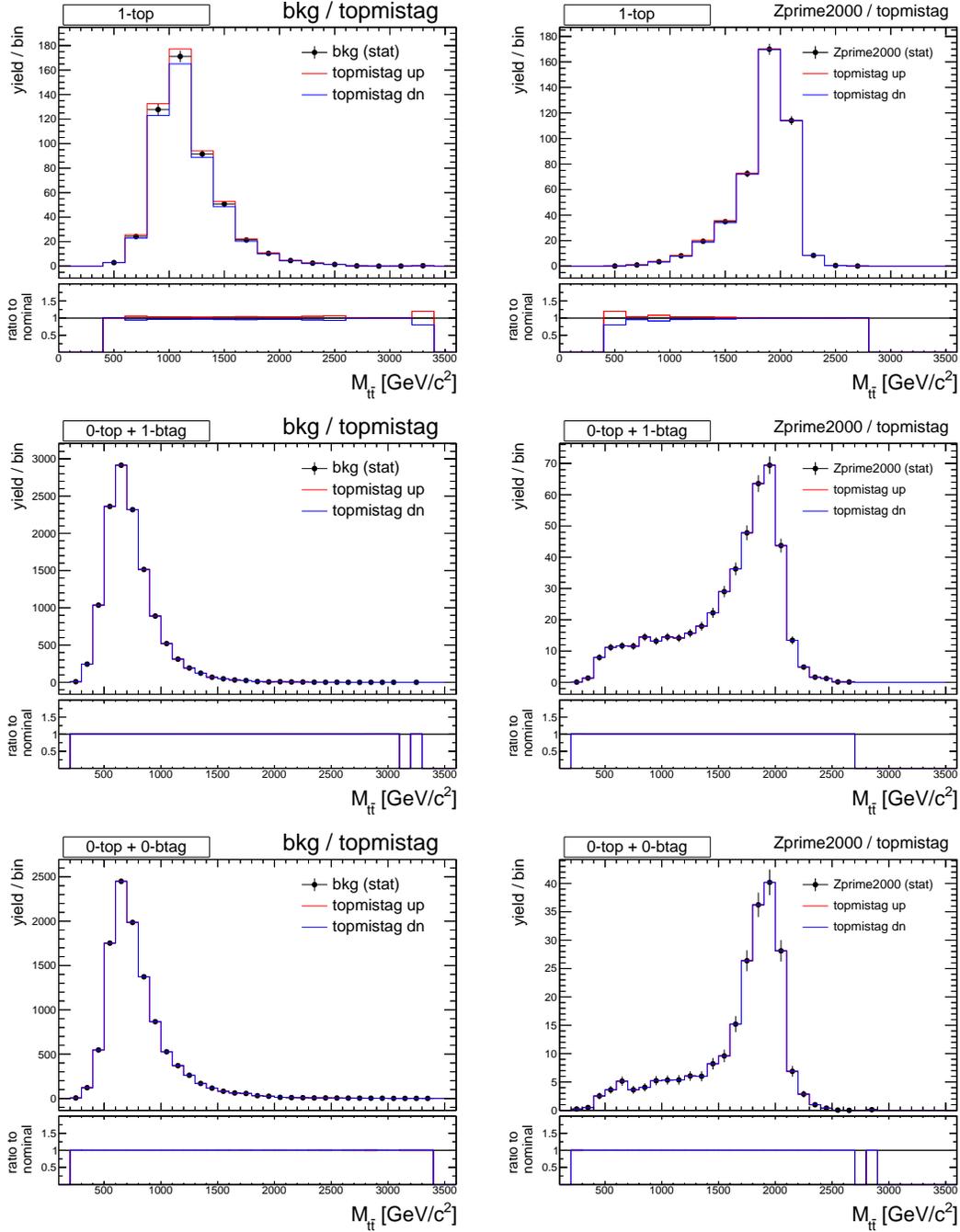


Figure C.6: $\pm 1\sigma$ systematic variation on the $M_{t\bar{t}}$ distribution due to the top-mistag SF: (left) total background, (right) narrow Z' signal with $M_{Z'} = 2$ TeV. Plots are shown for the three final categories in the muon channel: (from top to bottom) “1top”, “0top1btag”, and “0top0btag”. The nominal distribution (black line) is plotted with its statistical uncertainty, the $\pm 1\sigma$ templates account for differences in both shape and normalization with respect to the nominal sample.

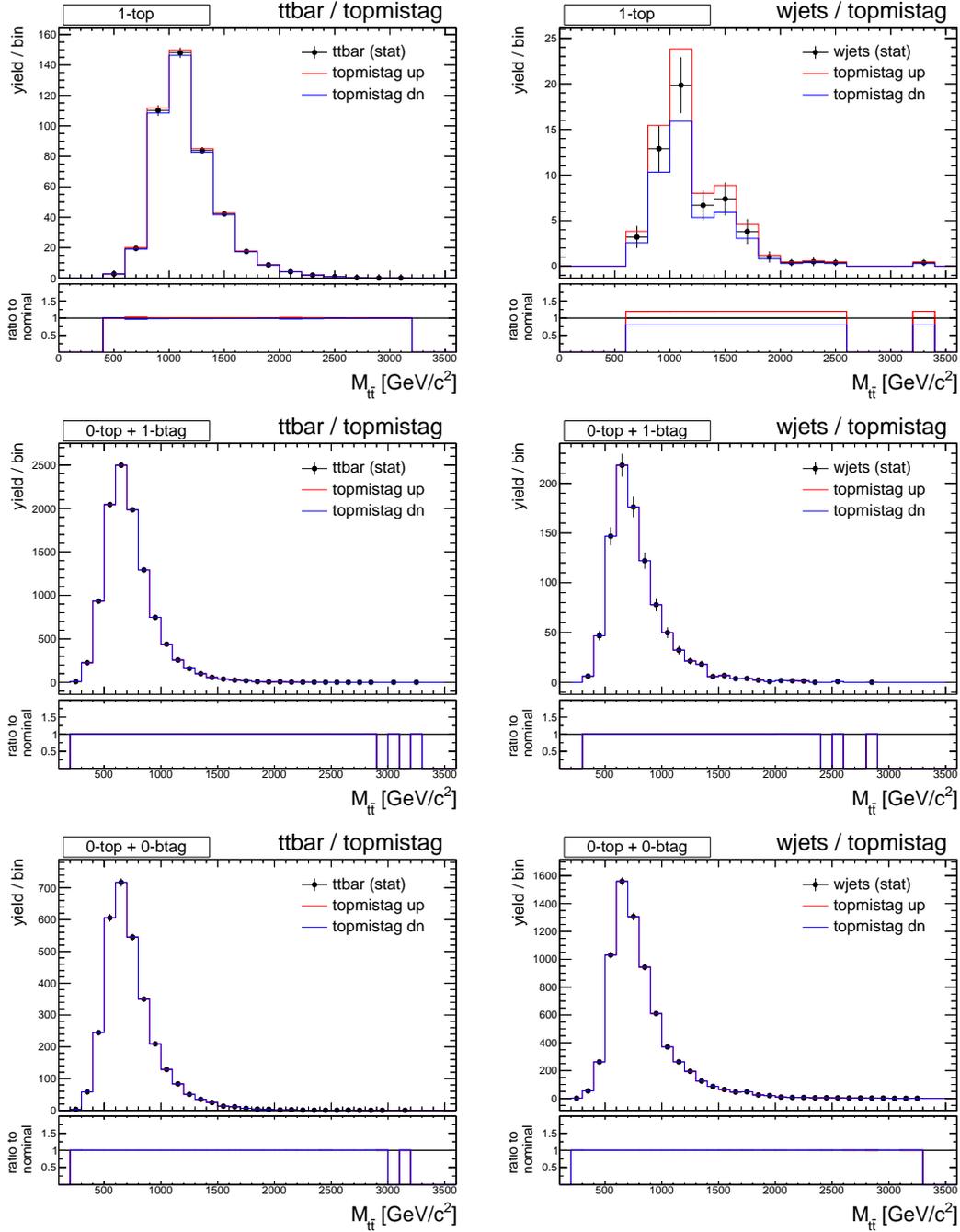


Figure C.7: $\pm 1\sigma$ systematic variation on the $M_{t\bar{t}}$ distribution due to the top-mistag SF: (left) SM $t\bar{t}$, (right) SM $W + \text{jets}$. Plots are shown for the three final categories in the muon channel: (from top to bottom) “1top”, “0top1btag”, and “0top0btag”. The nominal distribution (black line) is plotted with its statistical uncertainty, the $\pm 1\sigma$ templates account for differences in both shape and normalization with respect to the nominal sample.

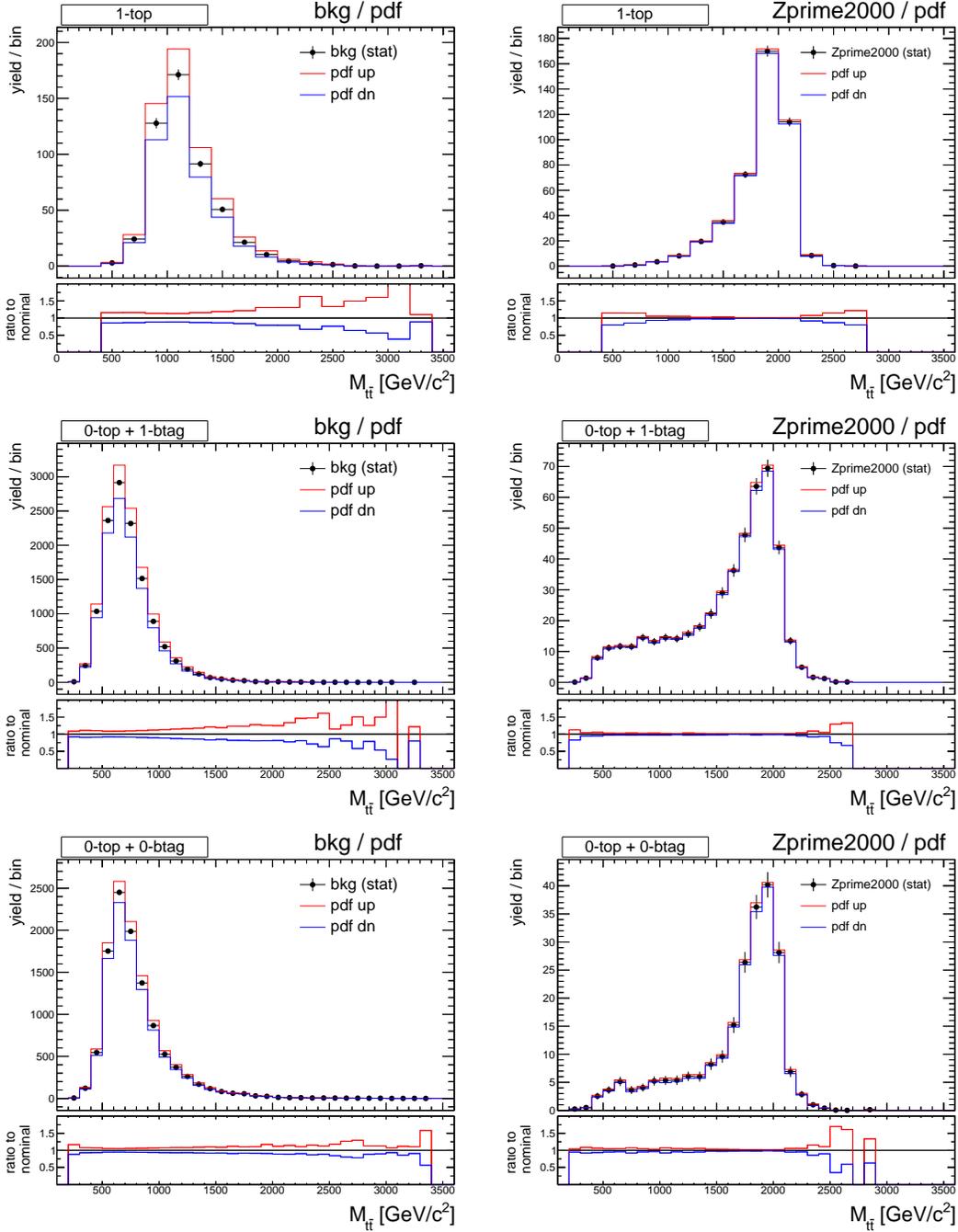


Figure C.8: $\pm 1\sigma$ systematic variation on the $M_{t\bar{t}}$ distribution due to the Parton Distribution Functions: (left) total background, (right) narrow Z' signal with $M_{Z'} = 2$ TeV. Plots are shown for the three final categories in the muon channel: (from top to bottom) "1top", "0top1btag", and "0top0btag". The nominal distribution (black line) is plotted with its statistical uncertainty, the $\pm 1\sigma$ templates account for differences in both shape and normalization with respect to the nominal sample.

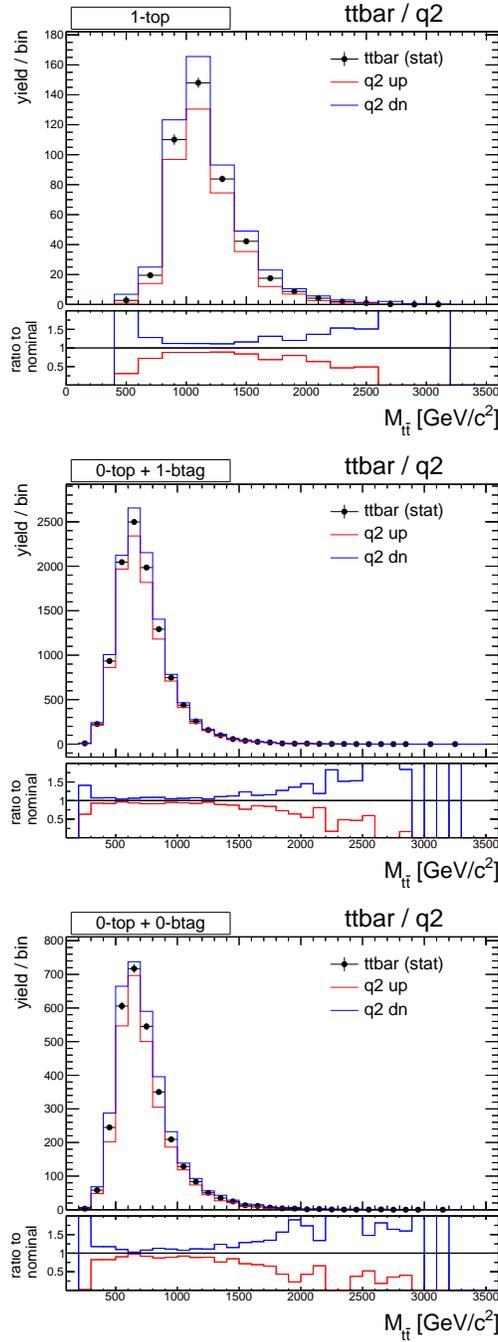


Figure C.9: $\pm 1\sigma$ systematic variation on the $M_{t\bar{t}}$ distribution for the $t\bar{t}$ background, due to the choice of the renormalization scale. Plots are shown for the three final categories in the muon channel: (from top to bottom) “1top”, “0top1btag”, and “0top0btag”. The nominal distribution (black line) is plotted with its statistical uncertainty, the $\pm 1\sigma$ templates account for differences in both shape and normalization with respect to the nominal sample.

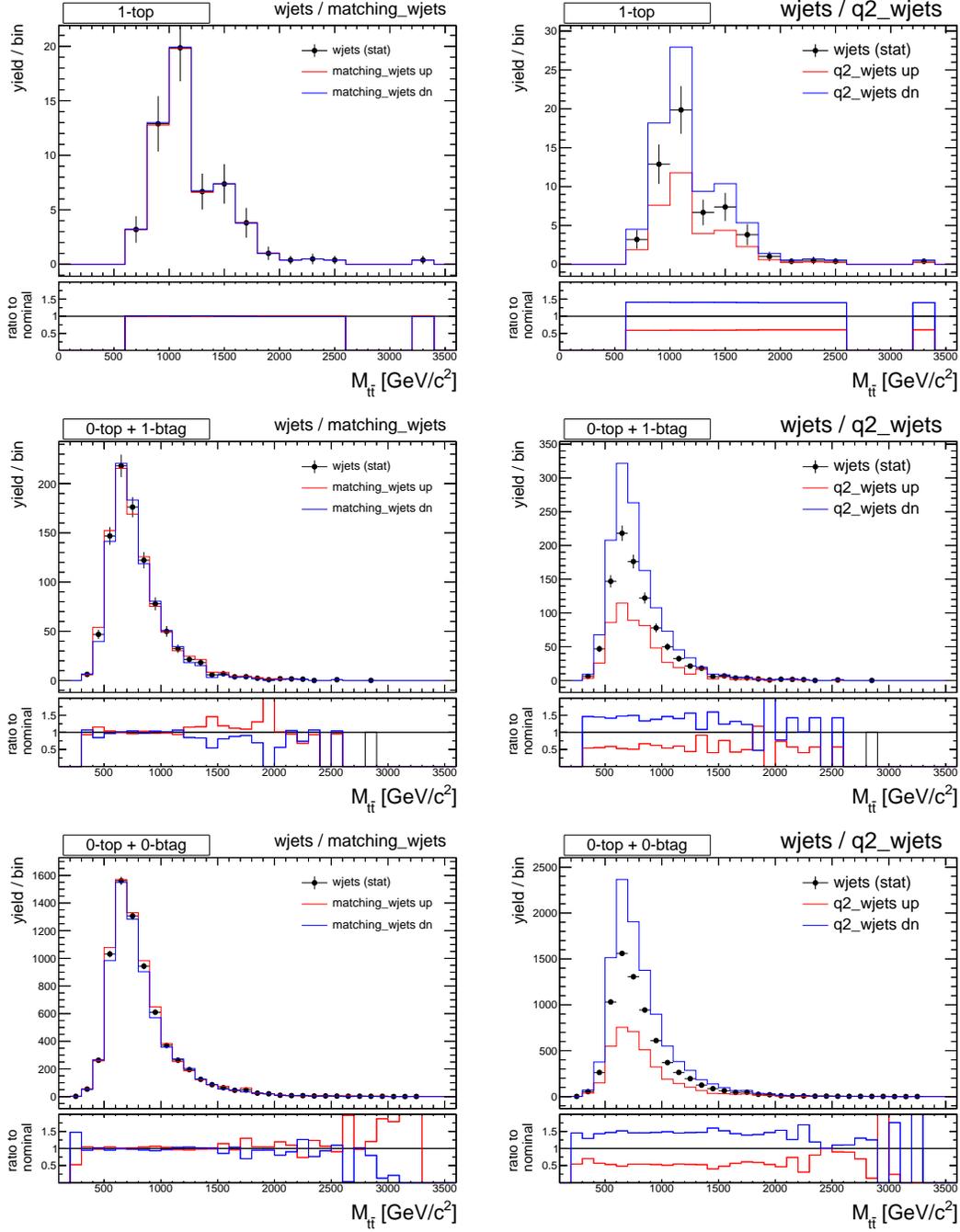


Figure C.10: $\pm 1\sigma$ systematic variation on the $M_{t\bar{t}}$ distribution for the $W + \text{jets}$ background, due to the choice of (left) the matching scale and (right) the renormalization scale. Plots are shown for the three final categories in the muon channel: (from top to bottom) “1top”, “0top1btag”, and “0top0btag”. The nominal distribution (black line) is plotted with its statistical uncertainty, the $\pm 1\sigma$ templates account for differences in both shape and normalization with respect to the nominal sample.

CITED LITERATURE

- [1] CMS Collaboration. Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC. *Phys. Lett. B*, 716(arXiv:1207.7235. CMS-HIG-12-028. CERN-PH-EP-2012-220):30–61. 59 p, Jul 2012.
- [2] ATLAS Collaboration. Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC. *Phys. Lett. B*, 716(arXiv:1207.7214. CERN-PH-EP-2012-218):1–29. 39 p, Aug 2012.
- [3] DONUT Collaboration. Observation of tau neutrino interactions. *Physics Letters B*, 504(3):218 – 224, 2001.
- [4] CDF Collaboration. Observation of Top Quark Production in $\bar{p}p$ Collisions with the Collider Detector at Fermilab. *Phys. Rev. Lett.*, 74:2626–2631, Apr 1995.
- [5] DØ Collaboration. Observation of the top quark. *Phys. Rev. Lett.*, 74:2632–2637, Apr 1995.
- [6] Eric Drexler. Elementary particle interactions in the Standard Model. Wikimedia Commons, Apr 2014.
http://en.wikipedia.org/wiki/File:Elementary_particle_interactions_in_the_Standard_Model.png.
- [7] Noether, E. Invariant Variation Problems. *Transport Theory and Statistical Mechanics*, 1:183–207, 1971.
- [8] Latham Boyle. Standard Model Of Particle Physics, Most Complete Diagram. Wikimedia Commons, Feb 2014.
http://en.wikipedia.org/wiki/File:Standard_Model_Of_Particle_Physics_Most_Complete_Diagram.jpg.
- [9] P.H. Frampton and S.L. Glashow. Chiral color: An alternative to the standard model. *Phys. Lett. B*, 190:157, 1987.
- [10] D. Choudhury et al. Top production at the Tevatron/LHC and nonstandard, strongly interacting spin one particles. *Phys. Lett. B*, 657:69, 2007.
- [11] C.T. Hill. Topcolor: Top quark condensation in a gauge extension of the standard model. *Phys. Lett. B*, 266:419, 1991.
- [12] C.T. Hill and S.J. Parke. Top production: Sensitivity to new physics. *Phys. Rev. D*, 49:4454, 1994.

- [13] C.T. Hill. Topcolor Assisted Technicolor. *Phys. Lett. B*, 345:483, 1995.
- [14] R. M. Harris and S. Jain. Cross sections for leptophobic topcolor Z' decaying to top-antitop. *Eur. Phys. J. C*, 72:2072, 2012.
- [15] J.L. Rosner. Prominent decay modes of a leptophobic Z' . *Phys. Lett. B*, 387:113, 1996.
- [16] K.R. Lynch et al. Finding Z-prime bosons coupled preferentially to the third family at LEP and the Tevatron. *Phys. Rev. D*, 63:035006, 2001.
- [17] M.S. Carena et al. Z-prime gauge bosons at the Tevatron. *Phys. Rev. D*, 70:093009, 2004.
- [18] D. Dicus, A. Stange, and S. Willenbrock. Higgs decay to top quarks at hadron colliders. *Phys. Lett. B*, 333:126, 1994.
- [19] L. Randall and R. Sundrum. Large Mass Hierarchy from a Small Extra Dimension. *Phys. Rev. Lett.*, 83:3370, 1999.
- [20] K. Agashe et al. LHC Signals from Warped Extra Dimensions. *Phys. Rev. D*, 77:015003, 2008.
- [21] H. Davoudiasl, J.L. Hewett, and T.G. Rizzo. Phenomenology of the Randall-Sundrum Gauge Hierarchy Model. *Phys. Rev. Lett.*, 84:2080, 2000.
- [22] CDF Collaboration. Limits on the production of narrow $t\bar{t}$ resonances in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV. *Phys. Rev. D*, 77:051102, 2008.
- [23] CDF Collaboration. Search for resonant $t\bar{t}$ production in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV. *Phys. Rev. Lett.*, 100:231801, 2008.
- [24] CDF Collaboration. A Search for resonant production of $t\bar{t}$ pairs in 4.8fb^{-1} of integrated luminosity of $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV. *Phys. Rev. D*, 84:072004, 2011.
- [25] DØ collaboration. Search for a Narrow $t\bar{t}$ Resonance in $p\bar{p}$ Collisions at $\sqrt{s} = 1.96$ TeV. *Phys. Rev. D*, 85:051101, 2012.
- [26] CDF Collaboration. Search for resonant production of $t\bar{t}$ decaying to jets in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV. *Phys. Rev. D*, 84:072003, 2011.
- [27] DØ collaboration. Search for $t\bar{t}$ resonances in the lepton plus jets final state in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV. *Phys. Lett. B*, 668:98, 2008.
- [28] CMS Collaboration. Search for anomalous $t\bar{t}$ production in the highly-boosted all-hadronic final state. *JHEP*, 09:029, 2012.
- [29] ATLAS Collaboration. A search for $t\bar{t}$ resonances in lepton+jets events with highly boosted top quarks collected in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector. *JHEP*, 09:041, 2012.

- [30] ATLAS Collaboration. Search for $t\bar{t}$ resonances in the lepton plus jets final state with ATLAS using 4.7 fb^{-1} of pp collisions at $\sqrt{s} = 7 \text{ TeV}$. *Phys. Rev. D.*, 88:012004, 2013.
- [31] CMS Collaboration. Search for resonant $t\bar{t}$ production in lepton+jets events in pp collisions at $\sqrt{s} = 7 \text{ TeV}$. *JHEP*, 12(CMS-TOP-12-017, CERN-PH-EP-2012-264):015, 2012.
- [32] CMS Collaboration. Search for Z' resonances decaying to $t\bar{t}$ in dilepton+jets final states in pp collisions at $\sqrt{s} = 7 \text{ TeV}$. *Phys. Rev. D*, 87:072002, 2013.
- [33] CMS Collaboration. Searches for new physics using the $t\bar{t}$ invariant mass distribution in pp collisions at $\sqrt{s}=8 \text{ TeV}$. *Phys. Rev. Lett.*, 111(21):211804, 2013.
- [34] Liu, Chun. Top-color-assisted supersymmetry. *Phys. Rev. D*, 61:115001, May 2000.
- [35] G Cvetič. Top quark condensation: a review. *Rev. Mod. Phys.*, 71(hep-ph/9702381. DO-TH-97-03):513–574. 165 p, Feb 1997.
- [36] Haffner, Julie. The CERN accelerator complex. Complexe des accélérateurs du CERN. , (OPEN-PHO-ACCEL-2013-056), Oct 2013.
- [37] AC Team. Diagram of an LHC dipole magnet. Schéma d'un aimant dipôle du LHC. (CERN-DI-9906025), Jun 1999.
- [38] CMS Collaboration. The CMS experiment at the CERN LHC. *J. Instrum.*, 3:S08004. 361 p, 2008.
- [39] CERN AC. Layout of CMS. , (CERN-DI-9803027), Mar 1998.
- [40] D. Barney. Interactive Slice of the CMS detector. , (CMS-doc-4172-v2), 2004.
- [41] CMS Collaboration. Operation and performance of the CMS tracker. (arXiv:1402.0675), Feb 2014. Comments: accepted for publication in J. of Instrum.
- [42] CMS Collaboration. Description and performance of track and primary-vertex reconstruction with the CMS tracker. *J. Instrum.*, 9(arXiv:1405.6569. CERN-PH-EP-2014-070. CMS-TRK-11-001):P10009. 80 p, May 2014.
- [43] CMS Collaboration. Energy calibration and resolution of the CMS electromagnetic calorimeter in pp collisions at $\sqrt{s} = 7 \text{ TeV}$. *J. Instrum.*, 8(CMS-EGM-11-001, CERN-PH-EP-2013-097):P09009, 2013.
- [44] CMS Collaboration. Performance of CMS muon reconstruction in pp collision events at $\sqrt{s} = 7 \text{ TeV}$. *J. Instrum.*, 7(CMS-MUO-10-004, CERN-PH-EP-2012-173):P10002, 2012.
- [45] CMS Collaboration. Performance of CMS muon reconstruction in pp collision events at $\sqrt{s} = 7 \text{ TeV}$. *J. Instrum.*, 7(arXiv:1206.4071. CMS-MUO-10-004. CERN-PH-EP-2012-173):P10002. 81 p, Jun 2012. Comments: Submitted to the J. of Instrum.

- [46] CMS Collaboration. Particle-Flow Event Reconstruction in CMS and Performance for Jets, Taus, and MET. *CMS Physics Analysis Summary*, CMS-PAS-PFT-09-001(CMS-PAS-PFT-09-001), 2009.
- [47] CMS Collaboration. Tracking and Primary Vertex Results in First 7 TeV Collisions. Technical Report CMS-PAS-TRK-10-005, CERN, 2010. Geneva, 2010.
- [48] Chabanat, E. and Estre, N. Deterministic Annealing for Vertex Finding at CMS. (DOI10.5170/CERN-2005-002.287), 2005.
- [49] Früwirth, R and Waltenberger, Wolfgang and Vanlaer, Pascal. Adaptive Vertex Fitting. Technical Report CMS-NOTE-2007-008, CERN, Geneva, Mar 2007.
- [50] CMS Collaboration. Performance of electron reconstruction and selection with the CMS detector in proton-proton collisions at $\sqrt{s} = 8$ TeV. (CERN-PH-EP-2015-004. CMS-EGM-13-001-003. arXiv:1502.02701), Feb 2015. Comments: Submitted to J. Instrum.
- [51] Cacciari, Matteo and Salam, Gavin P. and Soyez, Gregory. The anti- k_t jet clustering algorithm. *JHEP*, 04:063, 2008.
- [52] Dokshitzer, Yuri L. and Leder, G. D. and Moretti, S. and Webber, B. R. Better jet clustering algorithms. *JHEP*, 08:001, 1997.
- [53] CMS Collaboration. Identification of b-quark jets with the CMS experiment. *J. Instrum.*, 8(arXiv:1211.4462. CMS-BTV-12-001. CERN-PH-EP-2012-262):P04013. 67 p, Nov 2012.
- [54] Kaplan, David E. and Rehermann, Keith and Schwartz, Matthew D. and Tweedie, Brock. Top Tagging: A Method for Identifying Boosted Hadronically Decaying Top Quarks. *Phys.Rev.Lett.*, 101:142001, 2008.
- [55] CMS Collaboration. A Cambridge-Aachen (C-A) based Jet Algorithm for boosted top-jet tagging. *CMS Physics Analysis Summary*, (CMS-PAS-JME-09-001), 2009.
- [56] Ellis, Stephen D. and Vermilion, Christopher K. and Walsh, Jonathan R. Techniques for improved heavy particle searches with jet substructure. *Phys.Rev.*, D80:051501, 2009.
- [57] Ellis, Stephen D. and Vermilion, Christopher K. and Walsh, Jonathan R. Recombination Algorithms and Jet Substructure: Pruning as a Tool for Heavy Particle Searches. *Phys.Rev.*, D81:094023, 2010.
- [58] Thaler, Jesse and Van Tilburg, Ken. Identifying Boosted Objects with N-subjettiness. *JHEP*, 1103(MIT-CTP-4191):015, 2011.
- [59] S. Agostinelli and others. Geant4 — a simulation toolkit. *Nucl. Instrum. Meth. A*, 506:250, 2003.

- [60] Jerrum, Mark and Sinclair, Alistair. The Markov chain Monte Carlo method: an approach to approximate counting and integration. *Approximation algorithms for NP-hard problems*, pages 482–520, 1996.
- [61] Frixione, Stefano and Nason, Paolo and Ridolfi, Giovanni. A Positive-Weight Next-to-Leading-Order Monte Carlo for Heavy Flavour Hadroproduction. *JHEP*, 09:126, 2007.
- [62] Nason, Paolo. A new method for combining NLO QCD with shower Monte Carlo algorithms. *JHEP*, 11:040, 2004.
- [63] Frixione, Stefano and Nason, Paolo and Oleari, Carlo. Matching NLO QCD computations with Parton Shower simulations: the POWHEG method. *JHEP*, 11:070, 2007.
- [64] Alioli, Simone and Nason, Paolo and Oleari, Carlo and Re, Emanuele. A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX. *JHEP*, 06:043, 2010.
- [65] Sjöstrand T. and Mrenna S. and Skands P. PYTHIA 6.4 Physics and Manual . *JHEP*, 0605:026, 2006.
- [66] Alwall, J. and Herquet, M. and Maltoni, F. and Mattelaer, O. and Stelzer, T. MadGraph 5 : Going Beyond. *JHEP*, 06:128, 2011.
- [67] Sjöstrand T. and Lönnblad L. and Mrenna S. PYTHIA 6.2 Physics and Manual . (arXiv:hep-ph/0108264), 2001.
- [68] Alwall, Johan and Demin, Pavel and de Visscher, Simon and Frederix, Rikkert and Herquet, Michel and others. MadGraph/MadEvent v4: The New Web Generation. *JHEP*, 0709(SLAC-PUB-12603, CP3-07-17):028, 2007.
- [69] Mangano, Michelangelo L. and Moretti, Mauro and Piccinini, Fulvio and Treccani, Michele. Matching matrix elements and shower evolution for top-quark production in hadronic collisions. *JHEP*, 0701(CERN-PH-TH-2006-232, FNT-T-2006-09):013, 2007.
- [70] Sjöstrand T. and Mrenna S. and Skands P. A Brief Introduction to PYTHIA 8.1. *Comput. Phys. Commun.*, 178, 2008.
- [71] Nadolsky, Pavel M. and Lai, Hung-Liang and Cao, Qing-Hong and Huston, Joey and Pumplin, Jon and others. Implications of CTEQ global analysis for collider observables. *Phys.Rev.*, D78(MSUHEP-080125, UCRHEP-T447):013004, 2008.
- [72] Gao, Jun and Li, Chong Sheng and Li, Bo Hua and Yuan, C.-P. and Zhu, Hua Xing. Next-to-leading order QCD corrections to the heavy resonance production and decay into top quark pair at the LHC. *Phys.Rev.*, D82:014020, 2010.
- [73] Olive, K.A. and others. Review of Particle Physics. *Chin.Phys.*, C38:090001, 2014.

- [74] J. Beringer et al. 2012 Review of Particle Physics - Gauge and Higgs Bosons. *Phys. Rev. D*, 86, 2012.
- [75] CMS Collaboration. Search for high-mass resonances decaying to $t\bar{t}$ in the electron+jets channel. *CMS Physics Analysis Summary*, CMS-PAS-EXO-11-092, 2011.
- [76] B. Abbott et al. Measurement of W and Z boson production cross sections in $p\bar{p}$ collisions at $\sqrt{s} = 1.8$ TeV. *Phys. Rev. D*, 60:052003, Aug 1999.
- [77] CMS Collaboration. The performance of the CMS muon detector in proton-proton collisions at $\sqrt{s} = 7$ TeV at the LHC. *J. Instrum.*, 8(arXiv:1306.6905. CMS-MUO-11-001. CERN-PH-EP-2013-072):P11002, Jun 2013. Comments: Submitted to J. Instrum.
- [78] CMS Collaboration. Performance of b tagging at $\sqrt{s}=8$ TeV in multijet, $t\bar{t}$ and boosted topology events. (CMS-PAS-BTV-13-001), 2013.
- [79] CMS Collaboration. Boosted Top Jet Tagging at CMS. *CMS Physics Analysis Summary*, CMS-PAS-JME-13-007, 2014.
- [80] Ott, J. theta-auto testing documentation.
<http://www-ekp.physik.uni-karlsruhe.de/~ott/theta/theta-auto/>.
- [81] John Conway. Nuisance Parameters in Likelihoods for Searches. Presentation at PhyStat 2011, 2011.
- [82] Barlow, Roger J. and Beeston, Christine. Fitting using finite Monte Carlo samples. *Comput.Phys.Commun.*, 77(MAN-HEP-93-1):219–228, 1993.
- [83] CMS Collaboration. Measurement of differential top-quark pair production cross sections in the lepton+jets channel in pp collisions at 8 TeV. *CMS Physics Analysis Summary*, CMS-PAS-TOP-12-027(CMS-PAS-TOP-12-027), 2013.
- [84] CMS Collaboration. Differential cross section measurements for the production of a W boson in association with jets in proton-proton collisions at $\sqrt{s} = 7$ TeV. (CMS-SMP-12-023, CERN-PH-EP-2014-134), 2014.
- [85] CMS Collaboration. Measurement of the production cross section for a W boson and two b jets in pp collisions at $\sqrt{s} = 7$ TeV. (CMS-SMP-12-026, CERN-PH-EP-2013-223), 2013.
- [86] CMS Collaboration. Measurement of the differential production cross section of Z bosons in association with jets in pp collisions at $\sqrt{s} = 8$ TeV. *CMS Physics Analysis Summary*, CMS-PAS-SMP-13-007(CMS-PAS-SMP-13-007), 2014.
- [87] CMS Collaboration. Observation of the associated production of a single top quark and a W boson in pp collisions at $\sqrt{s} = 8$ TeV. *Phys. Rev. Lett.*, 112(CMS-TOP-12-040, CERN-PH-EP-2013-237):231802, 2014.

- [88] CMS Collaboration. Measurement of $W+W^-$ and ZZ production cross sections in pp collisions at $\sqrt{s} = 8$ TeV. *Phys. Lett.*, B721(CMS-SMP-12-024, CERN-PH-EP-2012-376):190–211, 2013.
- [89] CMS Collaboration. Measurement of WZ and ZZ production in pp collisions at $\sqrt{s} = 8$ TeV in final states with b -tagged jets. (CMS-SMP-13-011, CERN-PH-EP-2014-022), 2014.
- [90] CMS Collaboration. CMS Luminosity Based on Pixel Cluster Counting - Summer 2013 Update. *CMS Physics Analysis Summary*, CMS-PAS-LUM-13-001(CMS-PAS-LUM-13-001), 2013.
- [91] CMS Collaboration. Search for $t\bar{t}$ resonances in semileptonic final states in pp collisions at $\sqrt{s} = 8$ TeV. *CMS Physics Analysis Summary*, CMS-PAS-B2G-12-006(CMS-PAS-B2G-12-006), 2012.

VITA

Name Paul Jonathan Turner

Education Ph.D., Physics
University of Illinois at Chicago, Chicago, Illinois
2015

B.S., Physics
Boise State University, Boise, Idaho
2008

Workshops CMS Data Analysis School (CMSDAS)
Fermilab LPC, Jan. 12-16, 2015
Coordinator

Jet Substructure Hands-on Advanced Tutorial Session
Fermilab LPC, Jun. 18-19, 2014

Statistics Hands-on Advanced Tutorial Session
Fermilab LPC, Jun. 2-4, 2014

Teaching Teaching Assistant
University of Illinois at Chicago, Chicago, Illinois
2009 – 2010

Experience Research Assistant (CMS Experiment)
University of Illinois at Chicago, Chicago, Illinois
2010 – 2015

NSF PIRE Program, Summer Research Abroad
ETH Zürich, 2013

NSF PIRE Program, Summer Research Abroad
Paul Scherrer Institute, 2012

NSF PIRE Program, Summer Research Abroad
Paul Scherrer Institute, 2011

- Conference Proceedings** P. Turner, “Search for $t\bar{t}$ resonances in semileptonic final states in pp collisions at $\sqrt{s} = 8$ TeV”
The meeting of the American Physical Society, Division of Particles and Fields 2013
- P. Turner, “X-Ray VCal Calibration Methods and Simulations”
PIRE Research Report 2013
- P. Turner, “Design and Implementation of a High Resolution Telescope for CMS Pixel Detector Characterization”
PIRE Research Report 2012
- P. Turner, “Design Of A High Resolution Pixel Telescope”
PIRE Research Report 2011
- Publications** CMS Collaboration, “A search for resonant $t\bar{t}$ production in pp collisions at $\sqrt{s} = 8$ TeV”
CMS-PAS-B2G-13-008, (In Prep., Target Journal is Phys. Rev. D)
- CMS Collaboration, “Searches for new physics using the $t\bar{t}$ invariant mass distributions in pp collisions at $\sqrt{s}=8$ TeV”
CMS-B2G-13-001-003, <http://cds.cern.ch/record/1599045>
Phys. Rev. Lett., 111(21):211804, 2013.
- CMS Collaboration, “Search for $t\bar{t}$ resonances in semileptonic final state”
CMS-PAS-B2G-12-006, <http://cds.cern.ch/record/1543467>
- M. Adams, V. Bazterra, C. Gerber, J. Haller, S. Khalatyan, R. Kogler, M. Meyer, T. Peiffer, D. Sandoval, D. Strom, P. Turner, “Search for $t\bar{t}$ resonances in semileptonic final states”
CMS-AN-12-277, (Internal CMS Analysis Note)
- M. Adams, D. Berry, C. Gerber, D. Gonzalez, J. Haller, R. Kogler, V. Kripas, M. Missiroli, T. Peiffer, D. Sandoval, P. Turner, “Search for $t\bar{t}$ resonances in boosted semileptonic final states”
CMS-AN-14-035, (Internal CMS Analysis Note)