A Machine Learning Based Search for Supersymmetry in All Hadronic Decays of the sTop Particle

by

Antony H. Adair

A Thesis Submitted in Partial Fulfillment of the Requirements for the Degree Doctor of Philosophy

Approved, Thesis Committee:

[Signatures]

B. Paul Padley, Chair
Professor of Physics and Astronomy

Karl Ecklund
Associate Professor of Physics and Astronomy

Moshe Vardi
Professor of Computer Science

Houston, Texas

May, 2016
ABSTRACT

A Machine Learning Based Search for Supersymmetry in All Hadronic Decays of the sTop Particle

by

Antony H. Adair

A search for signs of supersymmetry by means of all hadronic decays of the scalar top quark is presented. The data sample of proton-proton collisions used corresponds to an integrated luminosity of 19.6 fb$^{-1}$ collected at $\sqrt{s} = 8$ TeV with the CMS detector at the LHC. The investigation features machine learning based background suppression and prediction techniques, developed through an analogous 18.9 fb$^{-1}$ study. The data is found to be in agreement with the predicted backgrounds and no evidence of supersymmetry is observed. Exclusion limits are set, but found to be in general agreement with the previous study.
Acknowledgments

It is difficult to overstate my gratitude for my mentor, Dr. Paul Padley, with his enthusiasm, his inspiration, his patience, and his great effort to explain things clearly and simply. He has continuously provided encouragement, sound advice, good teaching, and lots of good ideas.

I would like to acknowledge the people of the USCMS and the CMS, for all their support and the wonderful learning opportunities provided to me as a member of the CMS collaboration. Being a member of the CMS collaboration has not only allowed me to work on ground breaking researching, but allowed me to experience new cultures and new lands. I am forever grateful to have been afforded such opportunities.

Lastly, I would like to express my utmost gratitude to Dr. Paul Padley, Dr. Karl Ecklund, and Dr. Moshe Vardi for taking the time out of their busy schedules to serve on my thesis committee.
Contents

Abstract ii
Acknowledgments iii
List of Illustrations viii
List of Tables xx

1 Introduction 1

2 The Physics 4
  2.1 The Standard Model of Particle Physics . . . . . . . . . . . . . . . . 4
    2.1.1 The Limitations . . . . . . . . . . . . . . . . . . . . . . . . . . 12
  2.2 Supersymmetry . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 14
  2.3 Supersymmetric Top Quark Search . . . . . . . . . . . . . . . . . . . 17
    2.3.1 The Signal . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 17
    2.3.2 The Background . . . . . . . . . . . . . . . . . . . . . . . . . 21
  2.4 Summary . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 25

3 The Experimental Apparatus 27
  3.1 The Large Hadron Collider . . . . . . . . . . . . . . . . . . . . . . . . 27
    3.1.1 Collider Physics . . . . . . . . . . . . . . . . . . . . . . . . . . 30
  3.2 The Compact Muon Solenoid . . . . . . . . . . . . . . . . . . . . . . . 32
    3.2.1 The Tracker . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 35
    3.2.2 The Electromagnetic Calorimeter . . . . . . . . . . . . . . . . . 38
    3.2.3 The Hadronic Calorimeter . . . . . . . . . . . . . . . . . . . . . 41
    3.2.4 The Magnet . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 44
4 The Object Identification and Reconstruction

4.1 Vertices ............................................ 50
4.2 Particle Flow ....................................... 53
4.3 Missing Energy .................................... 55
4.4 Isolation ............................................. 57
4.5 Jets .................................................. 58
  4.5.1 AK5 Jets ....................................... 61
  4.5.2 Prelude to Picky Jets .......................... 63
  4.5.3 Picky Jets ..................................... 66
  4.5.4 Picky Jet Corrections .......................... 68
4.6 B Flavor Tagging .................................... 73
  4.6.1 B-Tagging AK5 Jets ............................ 77
  4.6.2 Quark/Gluon Discrimination for AK5 Jets .... 79
  4.6.3 B-Tagging Picky Jets .......................... 80
  4.6.4 Quark/Gluon Shape Corrections ............... 81
4.7 Top Reconstruction ................................. 83
  4.7.1 Jet Seed ....................................... 83
  4.7.2 Top Candidate ................................. 84
  4.7.3 Top Pairs ..................................... 88
4.8 Lepton Control Regions ............................ 90
  4.8.1 Electron Identification ........................ 90
  4.8.2 Muon Identification ........................... 91
4.9 Lepton Veto BDTs ................................... 93
  4.9.1 Electron Veto ................................ 93
  4.9.2 Muon Veto ................................... 95
5 The Signal
5.1 Event Selection ................................................. 106
  5.1.1 BDT Preselection ........................................ 106
  5.1.2 BDT Input Variables .................................. 107
  5.1.3 BDT Training .............................................. 110
5.2 Search Region Selection .................................... 111

6 Top, W, and Z(→ν¯ν) + jets backgrounds .......... 117
  6.1 Leptonic Control Regions ................................. 118
  6.2 SFS for Kinematic Corrections ......................... 120
    6.2.1 Effect of MC-Reweighting on Control Distributions . . 121
  6.3 Closure Test for Leptonic Control Regions .......... 126
    6.3.1 Closure Test in the Single Lepton Control Region .. 126
    6.3.2 Closure Test in the Z → ℓ+ℓ− Control Region .... 130
  6.4 ET Triggered Control Region ............................. 136
  6.5 Prediction in Search Regions ............................ 137
    6.5.1 Uncertainties ....................................... 137
    6.5.2 Impact .............................................. 138
    6.5.3 Top Background ..................................... 139
    6.5.4 EWK backgrounds .................................. 139

7 QCD Multijet background ................................. 148

8 ttZ background ............................................. 151
  8.1 Yields and MC Generation Uncertainty .......... 152
  8.2 Cross Section Uncertainty .............................. 152
  8.3 Kinematic Uncertainty ................................. 156
8.4 Prediction in the Search Regions .................................................. 156

9 The Results And Interpretation ....................................................... 159
  9.1 Data Yields .............................................................................. 159
  9.2 Signal Efficiencies ................................................................... 165
  9.3 Signal Uncertainties .................................................................. 167
  9.4 Upper Limits on Data .................................................................. 177
  9.5 Best Regions and Quantities ......................................................... 196
  9.6 Limit calculation ........................................................................ 213

A Methods ......................................................................................... 230
  A.1 Tag and Probe ......................................................................... 230
  A.2 MC Reweighting ....................................................................... 231
    A.2.1 Process Kinematics ............................................................ 232
    A.2.2 Reconstruction Effects ...................................................... 236
  A.3 Machine Learning ...................................................................... 238
    A.3.1 Decision Tree Basics .......................................................... 240
    A.3.2 Boosted Decision Trees ...................................................... 244

B Samples ......................................................................................... 246
  B.1 Preselection .............................................................................. 246
  B.2 The Data Samples ...................................................................... 246
  B.3 The Simulated Samples ............................................................... 248
    B.3.1 The Simulated Signal Samples ............................................ 250
    B.3.2 The Simulated Background Related Samples ...................... 251
  B.4 Trigger Efficiency ...................................................................... 253

Bibliography ..................................................................................... 257
Illustrations

2.1 The Standard model of particle physics: 12 fermions and 5 bosons.
Brown loops indicate which bosons (red) couple to which fermions
(purple and green) [1]. .................................................. 5

2.2 One-loop fermion contribution to the Higgs boson mass ............. 13

2.3 One-loop SUSY scalar “s” contribution to the Higgs boson mass ... 17

2.4 T2tt Signal ............................................................. 19

2.5 All Hadronic Decay of T2tt Signal .................................. 20

2.6 T2bW Signal ............................................................ 21

2.7 All Hadronic Decay of T2bW Signal .................................. 22

2.8 Theory cross sections for selected SUSY processes at $\sqrt{s} = 8$ TeV.
Green line is the expected stop pair production cross section [2–4] ... 23

2.9 Electroweak backgrounds. Left: A leptonically decaying W boson
($W^\pm \rightarrow l^\pm \nu_l$). Right: Invisible decays of Z boson ($Z \rightarrow \nu\bar{\nu}$) .... 24

2.10 $t\bar{t}Z (Z \rightarrow \nu\bar{\nu})$ Decay ....................................... 25

3.1 Simplified LHC accelerator layout for proton-proton and ion-ion
collisions ................................................................. 28

3.2 Overview of the CMS detector [5] ....................................... 34

3.3 Segment (radially) of the CMS subdetectors and how particles
interact with them [6]. .................................................. 35
3.4 The tracker layout (quarter view along z). The Pixel Tracker is shown in green. Single-sided Silicon Strip Tracker modules are in red. Double-sided Silicon Strip Tracker modules are in blue. 36
3.5 The ECAL layout (quarter view along z) [7] 39
3.6 The HCAL layout (quarter view along z). HB = HCAL Barrel, HE = HCAL Endcap, and HF = HCAL Forward. [7] 42
3.7 The Muon Detector layout (quarter view along z) [7] 46
3.8 The muon momentum resolution for the CMS [8] 48

4.1 A proton-proton collision and resulting collimated spray of particles, a “jet” [9] 59
4.2 Jet features: Collinear-Safety and Infrared-Safety (non-Infrared safety shown) [10] 60
4.3 One “fat” top jet (red dots) splitting and one “thin” top jet (blue dots) merging. The Grey areas are “jets” produced by the Anti-kT 0.5 jet clustering algorithm [11] 65
4.4 Picky Jets by profile. Top Left: sub jet s1 (higher $p_T$) and s2 (lower $p_T$). Top Right: 1D $p_T$ profile of top left. Bottom right: Gaussian kernal of top right indicating the jet should be split. Bottom left: An example of a jet scenario that shouldn’t split. [11] 100
4.5 The efficiency of jet reconstruction algorithms (before PU corrections). The left is the more collimated top quark in SUSY stop pair production. The right is the less collimated top quark in SUSY stop pair production. The SUSY stop pair production uses a stop mass of 600 GeV and LSP mass of 50 GeV. [12] 101
4.6 Cumulative distribution function (cdf) mapping method for correcting jet shapes. [13] 101
4.7 The Top tagging efficiencies for constructing at least one top (left) and constructing both tops (right) for CORRAL, CTT, and HTT. [12] 102

4.8 Electron Veto. Left: BDT discriminator for signal (leptons from $W$ decays in $t\bar{t}b$, either coming directly from the $W$ or via an intermediate $\tau$ decay) and for background (true leptons coming from hadron decays or charged hadrons mis-identified as leptons). Right: ROC curve for BDT compared with reference working points of other veto algorithms available in CMS. The working point of the BDT has been chosen to be 98% efficient on signal and 5% on background, corresponding to a discriminator cut of -0.5. [13] 103

4.9 Muon Veto [13]. Left: BDT discriminator for signal (leptons from $W$ decays in $t\bar{t}b$, either coming directly from the $W$ or via an intermediate $\tau$ decay) and for background (true leptons coming from hadron decays or charged hadrons mis-identified as leptons). Right: ROC curve for BDT compared with reference working points of other veto algorithms available in CMS. The working point of the BDT has been chosen to be 98% efficient on signal and 5% on background, corresponding to a discriminator cut of -0.5. [13] 104

4.10 Tau Veto. Top Left: The $m_T$ distribution constructed for using all charged tracks in signal and for hadronic $\tau$'s in $t\bar{t}$ events. The veto is set for $m_T < 68 GeV$. Top Right: The maximum $\tau$ veto discriminator value for all $\tau$ candidates in an event. Bottom: The event level $t\bar{t}$ veto efficiency as a function of signal efficiency for stop and LSP masses of 620 GeV and 40 GeV, respectively. [13] 105

6.1 Simulated over data yields as a function of jet multiplicity ($p_T \geq 30 GeV, |\eta| \leq 2.4$) for various event selections as stated in the titles. [13] 123
6.2 Simulated over data yields as a function of jet multiplicity
\((p_T \geq 30\,\text{GeV}, |\eta| \leq 2.4)\) for various event selections as stated in the titles. [13] ................................................................. 124

6.3 Simulated over data yields as a function of jet multiplicity
\((p_T \geq 30\,\text{GeV}, |\eta| \leq 2.4)\) for various event selections as stated in the titles. [13] ................................................................. 125

6.4 Closure in the single lepton control region, where the lepton has been removed from all quantities (jets and \(E_T\)). Observed over predicted yields as a function of the BDT discriminator values for various T2bW search regions. [13] ................................................................. 127

6.5 Closure in the single lepton control region, where the lepton has been removed from all quantities (jets and \(E_T\)). Observed over predicted yields as a function of the BDT discriminator values for various T2tt search regions. [13] ................................................................. 128

6.6 Worst variable closure in the single lepton control region, where the lepton has been removed from all quantities (jets and \(E_T\)). Left: T2bW variables. Right: T2tt variables. [13] ................................................................. 129

6.7 Closure in the \(Z \rightarrow \ell^+\ell^-\) control region, where the lepton has been removed from all quantities (jets and \(E_T\)). Observed over predicted yields as a function of the BDT discriminator values for various T2bW search regions. [13] ................................................................. 131

6.8 Closure in the \(Z \rightarrow \ell^+\ell^-\) control region, where the lepton has been removed from all quantities (jets and \(E_T\)). Observed over predicted yields as a function of the BDT discriminator values for various T2tt search regions. [13] ................................................................. 132
6.9 Closure in the $Z \rightarrow \ell^{+}\ell^{-}$ control region, after a relaxed preselection, where the lepton has been removed from all quantities (jets and $E_{T}$).

Observed over predicted yields as a function of the BDT discriminator values for various T2bW search regions. [13] . . . . . . . 133

6.10 Closure in the $Z \rightarrow \ell^{+}\ell^{-}$ control region, after a relaxed preselection, where the lepton has been removed from all quantities (jets and $E_{T}$).

Observed over predicted yields as a function of the BDT discriminator values for various T2tt search regions. [13] . . . . . . . 134

6.11 Observed and predicted yields as a function of $H_{T}^{\text{along}}/H_{T}^{\text{away}}$ for events in the $Z \rightarrow \ell^{+}\ell^{-}$ control region and various T2bW search regions. 135

6.12 Worst variable closure in the $Z \rightarrow \ell^{+}\ell^{-}$ control region, where the lepton has been removed from all quantities (jets and $E_{T}$). Left: T2bW variables. Right: T2tt variables. [13] . . . . . . . . . . . . . . . . . . . 136

8.1 Yields in $t\bar{t}Z$ related control regions. Left: Simulated and data yields in the $t\bar{t}Z$ control region as a function of the number of picky jets with $p_{T} > 20 GeV$ and $|\eta| < 2.4$. Right: Simulated and data yields in the di-boson control region as a function of the number of picky jets with $p_{T} > 20 GeV$ and $|\eta| < 2.4$. [13] . . . . . . . . . . . . . . . . . . . . . . . . . . 154

9.1 T2bW yields with all lepton vetos applied. The bottom panes give the ratio of observed over predicted yields. The error on the data line is the error on the ratio of the data statistical uncertainty and simulation statistical uncertainty. The blue error band is the relative systematic error on the prediction yield. [13] . . . . . . . . . . . . . . . . . . . 161
9.2 T2tt yields with all lepton vetos applied. The bottom panes give the ratio of observed over predicted yields. The error on the data line is the error on the ratio of the data statistical uncertainty and simulation statistical uncertainty. The blue error band is the relative systematic error on the prediction yield. [13]...

9.3 T2bW yields with no lepton vetos applied. The bottom panes give the ratio of observed over predicted yields. The error on the data line is the error on the ratio of the data statistical uncertainty and simulation statistical uncertainty. The blue error band is the relative systematic error on the prediction yield. [13]...

9.4 T2tt yields with no lepton vetos applied. The bottom panes give the ratio of observed over predicted yields. The error on the data line is the error on the ratio of the data statistical uncertainty and simulation statistical uncertainty. The blue error band is the relative systematic error on the prediction yield. [13]...

9.5 Simulated $t\bar{t} \rightarrow t\bar{t}\tilde{\chi}_1^0\tilde{\chi}_1^0$ Signal Efficiencies. Top Left: LM. Top Right: MM. Bottom Left: HM. Bottom Right: VHM.

9.6 Simulated $t\bar{t} \rightarrow \tilde{\chi}_1^+\tilde{\chi}_1^-b\bar{b} (x = 0.25)$ Signal Efficiencies. Top Left: LX. Top Right: LM. Middle Left: MXHM. Middle Right: VHM. Bottom: HXHM.

9.7 Simulated $t\bar{t} \rightarrow \tilde{\chi}_1^+\tilde{\chi}_1^-b\bar{b} (x = 0.50)$ Signal Efficiencies. Top Left: LX. Top Right: LM. Middle Left: MXHM. Middle Right: VHM. Bottom: HXHM.

9.8 Simulated $t\bar{t} \rightarrow \tilde{\chi}_1^+\tilde{\chi}_1^-b\bar{b} (x = 0.75)$ Signal Efficiencies. Top Left: LX. Top Right: LM. Middle Left: MXHM. Middle Right: VHM. Bottom: HXHM.

9.9 Simulated $t\bar{t} \rightarrow t\bar{t}\tilde{\chi}_1^0\tilde{\chi}_1^0$ Signal Uncertainties on Efficiency. Top Left: LM. Top Right: MM. Bottom Left: HM. Bottom Right: VHM.
9.10 Simulated $\tilde{t}\tilde{t} \to \tilde{\chi}^+ \tilde{\chi}^- b\bar{b}$ ($x = 0.25$) Signal Uncertainties on Efficiency.  
Top Left: LX. Top Right: LM. Middle Left: MXHM. Middle Right: VHM. Bottom: HXHM ................................. 174

9.11 Simulated $\tilde{t}\tilde{t} \to \tilde{\chi}^+ \tilde{\chi}^- b\bar{b}$ ($x = 0.50$) Signal Uncertainties on Efficiency.  
Top Left: LX. Top Right: LM. Middle Left: MXHM. Middle Right: VHM. Bottom: HXHM ................................. 175

9.12 Simulated $\tilde{t}\tilde{t} \to \tilde{\chi}^+ \tilde{\chi}^- b\bar{b}$ ($x = 0.75$) Signal Uncertainties on Efficiency.  
Top Left: LX. Top Right: LM. Middle Left: MXHM. Middle Right: VHM. Bottom: HXHM ................................. 176

9.13 95% CL$_s$ limit on $N_{\text{obs}}$ for $\tilde{t}\tilde{t} \to t\tilde{t}\chi^{0}\chi^{0}$. Top Left: LM. Top Right: MM. Bottom Left: HM. Bottom Right: VHM. ................................. 180

9.14 95% CL$_s$ limit on $N_{\text{obs}}$ for $\tilde{t}\tilde{t} \to \tilde{\chi}^+ \tilde{\chi}^- b\bar{b}$ ($x = 0.25$). Top Left: LX. Top Right: LM. Middle Left: MXHM. Middle Right: VHM. Bottom: HXHM ................................. 181

9.15 95% CL$_s$ limit on $N_{\text{obs}}$ for $\tilde{t}\tilde{t} \to \tilde{\chi}^+ \tilde{\chi}^- b\bar{b}$ ($x = 0.50$). Top Left: LX. Top Right: LM. Middle Left: MXHM. Middle Right: VHM. Bottom: HXHM ................................. 182

9.16 95% CL$_s$ limit on $N_{\text{obs}}$ for $\tilde{t}\tilde{t} \to \tilde{\chi}^+ \tilde{\chi}^- b\bar{b}$ ($x = 0.75$). Top Left: LX. Top Right: LM. Middle Left: MXHM. Middle Right: VHM. Bottom: HXHM ................................. 183

9.17 95% CL$_s$ limit on $N_{\text{exp}}$ for $\tilde{t}\tilde{t} \to t\tilde{t}\chi^{0}\chi^{0}$. Top Left: LM. Top Right: MM. Bottom Left: HM. Bottom Right: VHM. ................................. 184

9.18 95% CL$_s$ limit on $N_{\text{exp}}$ for $\tilde{t}\tilde{t} \to \tilde{\chi}^+ \tilde{\chi}^- b\bar{b}$ ($x = 0.25$). Top Left: LX. Top Right: LM. Middle Left: MXHM. Middle Right: VHM. Bottom: HXHM ................................. 185

9.19 95% CL$_s$ limit on $N_{\text{exp}}$ for $\tilde{t}\tilde{t} \to \tilde{\chi}^+ \tilde{\chi}^- b\bar{b}$ ($x = 0.50$). Top Left: LX. Top Right: LM. Middle Left: MXHM. Middle Right: VHM. Bottom: HXHM ................................. 186
9.20 95% CLs limit on $N_{\text{exp}}$ for $\bar{t}t \rightarrow \tilde{\chi}^+ \tilde{\chi}^- b \bar{b}$ ($x = 0.75$). Top Left: LX.  
Top Right: LM. Middle Left: MXHM. Middle Right: VHM. Bottom:  
HXHM . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 187

9.21 95% CLs limit on $N_{\text{exp-std.}}$ for $\bar{t}t \rightarrow t\bar{t}\tilde{\chi}_1^0 \tilde{\chi}_1^0$. Top Left: LM. Top Right:  
MM. Bottom Left: HM. Bottom Right: VHM. . . . . . . . . . . . . . . . . 188

9.22 95% CLs limit on $N_{\text{exp-std.}}$ for $\bar{t}t \rightarrow t\bar{t}\tilde{\chi}^+ \tilde{\chi}^- b \bar{b}$ ($x = 0.25$). Top Left: LX.  
Top Right: LM. Middle Left: MXHM. Middle Right: VHM. Bottom:  
HXHM . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 189

9.23 95% CLs limit on $N_{\text{exp-std.}}$ for $\bar{t}t \rightarrow \tilde{\chi}^+ \tilde{\chi}^- b \bar{b}$ ($x = 0.50$). Top Left: LX.  
Top Right: LM. Middle Left: MXHM. Middle Right: VHM. Bottom:  
HXHM . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 190

9.24 95% CLs limit on $N_{\text{exp-std.}}$ for $\bar{t}t \rightarrow \tilde{\chi}^+ \tilde{\chi}^- b \bar{b}$ ($x = 0.75$). Top Left: LX.  
Top Right: LM. Middle Left: MXHM. Middle Right: VHM. Bottom:  
HXHM . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 191

9.25 95% CLs limit on $N_{\text{exp+std.}}$ for $\bar{t}t \rightarrow t\bar{t}\tilde{\chi}_1^0 \tilde{\chi}_1^0$. Top Left: LM. Top Right:  
MM. Bottom Left: HM. Bottom Right: VHM. . . . . . . . . . . . . . . . . 192

9.26 95% CLs limit on $N_{\text{exp+std.}}$ for $\bar{t}t \rightarrow \tilde{\chi}^+ \tilde{\chi}^- b \bar{b}$ ($x = 0.25$). Top Left:  
LX. Top Right: LM. Middle Left: MXHM. Middle Right: VHM.  
Bottom: HXHM . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 193

9.27 95% CLs limit on $N_{\text{exp+std.}}$ for $\bar{t}t \rightarrow \tilde{\chi}^+ \tilde{\chi}^- b \bar{b}$ ($x = 0.50$). Top Left:  
LX. Top Right: LM. Middle Left: MXHM. Middle Right: VHM.  
Bottom: HXHM . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 194

9.28 95% CLs limit on $N_{\text{exp+std.}}$ for $\bar{t}t \rightarrow \tilde{\chi}^+ \tilde{\chi}^- b \bar{b}$ ($x = 0.75$). Top Left:  
LX. Top Right: LM. Middle Left: MXHM. Middle Right: VHM.  
Bottom: HXHM . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 195

9.29 95% CLs limit on $\sigma_{\text{obs}}$ for $\bar{t}t \rightarrow t\bar{t}\tilde{\chi}_1^0 \tilde{\chi}_1^0$. Top Left: LM. Top Right:  
MM. Bottom Left: HM. Bottom Right: VHM. . . . . . . . . . . . . . . . . 197
9.30 95% CL_{\alpha} limit on \sigma_{\text{obs}} for \bar{t}t \rightarrow \tilde{\chi}^+ \tilde{\chi}^- b \bar{b} (x = 0.25). Top Left: LX.
Top Right: LM. Middle Left: MXHM. Middle Right: VHM. Bottom: HXHM

9.31 95% CL_{\alpha} limit on \sigma_{\text{obs}} for \bar{t}t \rightarrow \tilde{\chi}^+ \tilde{\chi}^- b \bar{b} (x = 0.50). Top Left: LX.
Top Right: LM. Middle Left: MXHM. Middle Right: VHM. Bottom: HXHM

9.32 95% CL_{\alpha} limit on \sigma_{\text{obs}} for \bar{t}t \rightarrow \tilde{\chi}^+ \tilde{\chi}^- b \bar{b} (x = 0.75). Top Left: LX.
Top Right: LM. Middle Left: MXHM. Middle Right: VHM. Bottom: HXHM

9.33 95% CL_{\alpha} limit on \sigma_{\text{exp}} for \bar{t}t \rightarrow t\tilde{t}^0\tilde{t}^0. Top Left: LM. Top Right: MM. Bottom Left: HM. Bottom Right: VHM

9.34 95% CL_{\alpha} limit on \sigma_{\text{exp}} for \bar{t}t \rightarrow \tilde{\chi}^+ \tilde{\chi}^- b \bar{b} (x = 0.25). Top Left: LX.
Top Right: LM. Middle Left: MXHM. Middle Right: VHM. Bottom: HXHM

9.35 95% CL_{\alpha} limit on \sigma_{\text{exp}} for \bar{t}t \rightarrow \tilde{\chi}^+ \tilde{\chi}^- b \bar{b} (x = 0.50). Top Left: LX.
Top Right: LM. Middle Left: MXHM. Middle Right: VHM. Bottom: HXHM

9.36 95% CL_{\alpha} limit on \sigma_{\text{exp}} for \bar{t}t \rightarrow \tilde{\chi}^+ \tilde{\chi}^- b \bar{b} (x = 0.75). Top Left: LX.
Top Right: LM. Middle Left: MXHM. Middle Right: VHM. Bottom: HXHM

9.37 95% CL_{\alpha} limit on \sigma_{\text{exp-std.}} for \bar{t}t \rightarrow \tilde{t}t\tilde{t}^0\tilde{t}^0. Top Left: LM. Top Right: MM. Bottom Left: HM. Bottom Right: VHM

9.38 95% CL_{\alpha} limit on \sigma_{\text{exp-std.}} for \bar{t}t \rightarrow \tilde{\chi}^+ \tilde{\chi}^- b \bar{b} (x = 0.25). Top Left: LX.
Top Right: LM. Middle Left: MXHM. Middle Right: VHM. Bottom: HXHM

9.39 95% CL_{\alpha} limit on \sigma_{\text{exp-std.}} for \bar{t}t \rightarrow \tilde{\chi}^+ \tilde{\chi}^- b \bar{b} (x = 0.50). Top Left: LX.
Top Right: LM. Middle Left: MXHM. Middle Right: VHM. Bottom: HXHM
<table>
<thead>
<tr>
<th>Page</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>9.40</td>
<td>95% CL limit on $\sigma_{\text{exp-std.}}$ for $\tilde{t}\tilde{t} \rightarrow \tilde{\chi}^+\tilde{\chi}^- b\bar{b}$ ($x = 0.75$). Top Left: LX. Top Right: LM. Middle Left: MXHM. Middle Right: VHM. Bottom: HXHM.</td>
</tr>
<tr>
<td>9.41</td>
<td>95% CL limit on $\sigma_{\text{exp+std.}}$ for $\tilde{t}\tilde{t} \rightarrow t\tilde{t}\tilde{\chi}_1^0\tilde{\chi}_1^0$. Top Left: LM. Top Right: MM. Bottom Left: HM. Bottom Right: VHM.</td>
</tr>
<tr>
<td>9.42</td>
<td>95% CL limit on $\sigma_{\text{exp+std.}}$ for $\tilde{t}\tilde{t} \rightarrow \tilde{\chi}^+\tilde{\chi}^- b\bar{b}$ ($x = 0.25$). Top Left: LX. Top Right: LM. Middle Left: MXHM. Middle Right: VHM. Bottom: HXHM.</td>
</tr>
<tr>
<td>9.43</td>
<td>95% CL limit on $\sigma_{\text{exp+std.}}$ for $\tilde{t}\tilde{t} \rightarrow \tilde{\chi}^+\tilde{\chi}^- b\bar{b}$ ($x = 0.50$). Top Left: LX. Top Right: LM. Middle Left: MXHM. Middle Right: VHM. Bottom: HXHM.</td>
</tr>
<tr>
<td>9.44</td>
<td>95% CL limit on $\sigma_{\text{exp+std.}}$ for $\tilde{t}\tilde{t} \rightarrow \tilde{\chi}^+\tilde{\chi}^- b\bar{b}$ ($x = 0.75$). Top Left: LX. Top Right: LM. Middle Left: MXHM. Middle Right: VHM. Bottom: HXHM.</td>
</tr>
<tr>
<td>9.45</td>
<td>Best Search Regions. Top Left: $\tilde{t}\tilde{t} \rightarrow t\tilde{t}\tilde{\chi}_1^0\tilde{\chi}_1^0$. Top Right: $\tilde{t}\tilde{t} \rightarrow \tilde{\chi}^+\tilde{\chi}^- b\bar{b}$ ($x = 0.25$). Bottom Left: $\tilde{t}\tilde{t} \rightarrow \tilde{\chi}^+\tilde{\chi}^- b\bar{b}$ ($x = 0.50$). Bottom Right: $\tilde{t}\tilde{t} \rightarrow \tilde{\chi}^+\tilde{\chi}^- b\bar{b}$ ($x = 0.75$). For $\tilde{t}\tilde{t} \rightarrow t\tilde{t}\tilde{\chi}_1^0\tilde{\chi}_1^0$: 1-LM, 2-MM, 3-HM, and 4-VHM. For $\tilde{t}\tilde{t} \rightarrow \tilde{\chi}^+\tilde{\chi}^- b\bar{b}$: 1-LX, 2-LM, 3-MXHM, 4-VHM, and 5-HXHM.</td>
</tr>
<tr>
<td>9.46</td>
<td>Best Signal Efficiencies. Top Left: $\tilde{t}\tilde{t} \rightarrow t\tilde{t}\tilde{\chi}_1^0\tilde{\chi}_1^0$. Top Right: $\tilde{t}\tilde{t} \rightarrow \tilde{\chi}^+\tilde{\chi}^- b\bar{b}$ ($x = 0.25$). Bottom Left: $\tilde{t}\tilde{t} \rightarrow \tilde{\chi}^+\tilde{\chi}^- b\bar{b}$ ($x = 0.50$). Bottom Right: $\tilde{t}\tilde{t} \rightarrow \tilde{\chi}^+\tilde{\chi}^- b\bar{b}$ ($x = 0.75$).</td>
</tr>
<tr>
<td>9.47</td>
<td>Best Signal Uncertainties. Top Left: $\tilde{t}\tilde{t} \rightarrow t\tilde{t}\tilde{\chi}_1^0\tilde{\chi}_1^0$. Top Right: $\tilde{t}\tilde{t} \rightarrow \tilde{\chi}^+\tilde{\chi}^- b\bar{b}$ ($x = 0.25$). Bottom Left: $\tilde{t}\tilde{t} \rightarrow \tilde{\chi}^+\tilde{\chi}^- b\bar{b}$ ($x = 0.50$). Bottom Right: $\tilde{t}\tilde{t} \rightarrow \tilde{\chi}^+\tilde{\chi}^- b\bar{b}$ ($x = 0.75$).</td>
</tr>
<tr>
<td>9.48</td>
<td>Best 95% CL limit on $\sigma_{\text{obs.}}$. Top Left: $\tilde{t}\tilde{t} \rightarrow t\tilde{t}\tilde{\chi}_1^0\tilde{\chi}_1^0$. Top Right: $\tilde{t}\tilde{t} \rightarrow \tilde{\chi}^+\tilde{\chi}^- b\bar{b}$ ($x = 0.25$). Bottom Left: $\tilde{t}\tilde{t} \rightarrow \tilde{\chi}^+\tilde{\chi}^- b\bar{b}$ ($x = 0.50$). Bottom Right: $\tilde{t}\tilde{t} \rightarrow \tilde{\chi}^+\tilde{\chi}^- b\bar{b}$ ($x = 0.75$).</td>
</tr>
</tbody>
</table>
9.49 Best 95% CLs limit on $\sigma_{\text{exp}}$. Top Left: $\tilde{t}\tilde{t} \rightarrow t\tilde{t}\chi_1^0\chi_1^0$. Top Right: $\tilde{t}\tilde{t} \rightarrow \tilde{t}\tilde{t}\chi_1^0\chi_1^0$ (x = 0.25). Bottom Left: $\tilde{t}\tilde{t} \rightarrow \tilde{t}\tilde{t}\chi_1^0\chi_1^0$ (x = 0.50). Bottom Right: $\tilde{t}\tilde{t} \rightarrow \tilde{t}\tilde{t}\chi_1^0\chi_1^0$ (x = 0.75). . . . . . . . . . . . . . . . . 218

9.50 Best 95% CLs limit on $\sigma_{\text{exp-std}}$. Top Left: $\tilde{t}\tilde{t} \rightarrow \tilde{t}\tilde{t}\chi_1^0\chi_1^0$. Top Right: $\tilde{t}\tilde{t} \rightarrow \tilde{t}\tilde{t}\chi_1^0\chi_1^0$ (x = 0.25). Bottom Left: $\tilde{t}\tilde{t} \rightarrow \tilde{t}\tilde{t}\chi_1^0\chi_1^0$ (x = 0.50). Bottom Right: $\tilde{t}\tilde{t} \rightarrow \tilde{t}\tilde{t}\chi_1^0\chi_1^0$ (x = 0.75). . . . . . . . . . . . . . . . . 219

9.51 Best 95% CLs limit on $\sigma_{\text{exp-std}}$. Top Left: $\tilde{t}\tilde{t} \rightarrow \tilde{t}\tilde{t}\chi_1^0\chi_1^0$. Top Right: $\tilde{t}\tilde{t} \rightarrow \tilde{t}\tilde{t}\chi_1^0\chi_1^0$ (x = 0.25). Bottom Left: $\tilde{t}\tilde{t} \rightarrow \tilde{t}\tilde{t}\chi_1^0\chi_1^0$ (x = 0.50). Bottom Right: $\tilde{t}\tilde{t} \rightarrow \tilde{t}\tilde{t}\chi_1^0\chi_1^0$ (x = 0.75). . . . . . . . . . . . . . . . . 220

9.52 Best 95% CLs limit on $\sigma_{\text{obs}}/\sigma_{\text{SUSY}}$. Top Left: $\tilde{t}\tilde{t} \rightarrow \tilde{t}\tilde{t}\chi_1^0\chi_1^0$. Top Right: $\tilde{t}\tilde{t} \rightarrow \tilde{t}\tilde{t}\chi_1^0\chi_1^0$ (x = 0.25). Bottom Left: $\tilde{t}\tilde{t} \rightarrow \tilde{t}\tilde{t}\chi_1^0\chi_1^0$ (x = 0.50). Bottom Right: $\tilde{t}\tilde{t} \rightarrow \tilde{t}\tilde{t}\chi_1^0\chi_1^0$ (x = 0.75). . . . . . . . . . . . . . . . . 221

9.53 Best 95% CLs limit on $\sigma_{\text{obs-std}}/\sigma_{\text{SUSY}}$. Top Left: $\tilde{t}\tilde{t} \rightarrow \tilde{t}\tilde{t}\chi_1^0\chi_1^0$. Top Right: $\tilde{t}\tilde{t} \rightarrow \tilde{t}\tilde{t}\chi_1^0\chi_1^0$ (x = 0.25). Bottom Left: $\tilde{t}\tilde{t} \rightarrow \tilde{t}\tilde{t}\chi_1^0\chi_1^0$ (x = 0.50). Bottom Right: $\tilde{t}\tilde{t} \rightarrow \tilde{t}\tilde{t}\chi_1^0\chi_1^0$ (x = 0.75). . . . . . . . . . . . . . . . . 222

9.54 Best 95% CLs limit on $\sigma_{\text{obs-std}}/\sigma_{\text{SUSY}}$. Top Left: $\tilde{t}\tilde{t} \rightarrow \tilde{t}\tilde{t}\chi_1^0\chi_1^0$. Top Right: $\tilde{t}\tilde{t} \rightarrow \tilde{t}\tilde{t}\chi_1^0\chi_1^0$ (x = 0.25). Bottom Left: $\tilde{t}\tilde{t} \rightarrow \tilde{t}\tilde{t}\chi_1^0\chi_1^0$ (x = 0.50). Bottom Right: $\tilde{t}\tilde{t} \rightarrow \tilde{t}\tilde{t}\chi_1^0\chi_1^0$ (x = 0.75). . . . . . . . . . . . . . . . . 223

9.55 Best 95% CLs limit on $\sigma_{\text{exp}}/\sigma_{\text{SUSY}}$. Top Left: $\tilde{t}\tilde{t} \rightarrow \tilde{t}\tilde{t}\chi_1^0\chi_1^0$. Top Right: $\tilde{t}\tilde{t} \rightarrow \tilde{t}\tilde{t}\chi_1^0\chi_1^0$ (x = 0.25). Bottom Left: $\tilde{t}\tilde{t} \rightarrow \tilde{t}\tilde{t}\chi_1^0\chi_1^0$ (x = 0.50). Bottom Right: $\tilde{t}\tilde{t} \rightarrow \tilde{t}\tilde{t}\chi_1^0\chi_1^0$ (x = 0.75). . . . . . . . . . . . . . . . . 224

9.56 Best 95% CLs limit on $\sigma_{\text{exp-std}}/\sigma_{\text{SUSY}}$. Top Left: $\tilde{t}\tilde{t} \rightarrow \tilde{t}\tilde{t}\chi_1^0\chi_1^0$. Top Right: $\tilde{t}\tilde{t} \rightarrow \tilde{t}\tilde{t}\chi_1^0\chi_1^0$ (x = 0.25). Bottom Left: $\tilde{t}\tilde{t} \rightarrow \tilde{t}\tilde{t}\chi_1^0\chi_1^0$ (x = 0.50). Bottom Right: $\tilde{t}\tilde{t} \rightarrow \tilde{t}\tilde{t}\chi_1^0\chi_1^0$ (x = 0.75). . . . . . . . . . . . . . . . . 225

9.57 Best 95% CLs limit on $\sigma_{\text{exp-std}}/\sigma_{\text{SUSY}}$. Top Left: $\tilde{t}\tilde{t} \rightarrow \tilde{t}\tilde{t}\chi_1^0\chi_1^0$. Top Right: $\tilde{t}\tilde{t} \rightarrow \tilde{t}\tilde{t}\chi_1^0\chi_1^0$ (x = 0.25). Bottom Left: $\tilde{t}\tilde{t} \rightarrow \tilde{t}\tilde{t}\chi_1^0\chi_1^0$ (x = 0.50). Bottom Right: $\tilde{t}\tilde{t} \rightarrow \tilde{t}\tilde{t}\chi_1^0\chi_1^0$ (x = 0.75). . . . . . . . . . . . . . . . . 226
9.58 For the 19.6 fb$^{-1}$ study, the 95% CL$_{s}$ limit on the $\tilde{t}\tilde{t}$ production cross section and expected exclusion areas in the $m_{\tilde{t}} - m_{\chi^0}$ plane. Top

Left: $\tilde{t}\tilde{t} \rightarrow t\tilde{t}\chi_{1L}^0\tilde{\chi}_{1U}^0$. Top Right: $\tilde{t}\tilde{t} \rightarrow \tilde{t}
\tilde{t} \rightarrow \tilde{\chi}^+\tilde{\chi}^- b\bar{b} (x = 0.25)$. Bottom Left: $\tilde{t}\tilde{t} \rightarrow \tilde{\chi}^+\tilde{\chi}^- b\bar{b} (x = 0.50)$. Bottom Right: $\tilde{t}\tilde{t} \rightarrow \tilde{\chi}^+\tilde{\chi}^- b\bar{b} (x = 0.75)$. ...

9.59 For the 18.9 fb$^{-1}$ study, the 95% CL$_{s}$ limit on the $\tilde{t}\tilde{t}$ production cross section and expected exclusion areas in the $m_{\tilde{t}} - m_{\chi^0}$ plane. Top

Left: $\tilde{t}\tilde{t} \rightarrow t\tilde{t}\chi_{1L}^0\tilde{\chi}_{1U}^0$. Top Right: $\tilde{t}\tilde{t} \rightarrow \tilde{t}
\tilde{t} \rightarrow \tilde{\chi}^+\tilde{\chi}^- b\bar{b} (x = 0.25)$. Bottom Left: $\tilde{t}\tilde{t} \rightarrow \tilde{\chi}^+\tilde{\chi}^- b\bar{b} (x = 0.50)$. Bottom Right: $\tilde{t}\tilde{t} \rightarrow \tilde{\chi}^+\tilde{\chi}^- b\bar{b} (x = 0.75)$. Here LSP on the y-axis corresponds to $\tilde{\chi}_{1L}^0$. [13]. ...

A.1 Simple Decision Tree ........................................ 241
A.2 Simple Decision Tree - Predictor Space .................. 243

B.1 Search trigger efficiency as a function of offline $E_{T}$. [13]. .......................... 255
B.2 The Trigger turn-on curves as a function of BDT (MVA)

discriminator values. Top is T2bW. Bottom is T2tt. [13]. .......................... 256
### Tables

2.1 Fundamental Interactions ................................................. 4

4.1 Summary of picky jet $p_T$ corrections. ............................. 72

5.1 List of input signal points used for training the search region boosted
decision trees ................................................................. 114

5.2 Search Region Definitions ................................................ 115

5.3 T2bW variable rank ....................................................... 115

5.4 T2tt variable rank ......................................................... 116

6.1 Scale factor sets ............................................................ 140

6.2 Corrections applied to the simulation in the lepton closure tests. . . . 141

6.3 Systematic uncertainty and correction due to closure of the MC
reweighting prediction, for the $t\bar{t}$ and W processes. ................. 142

6.4 For T2bW search regions, the worst variable closure in the $Z \to \ell^+\ell^-$
control region, where the lepton has been removed from all quantities
(jets and $E_T$). ................................................................. 143

6.5 For T2tt search regions, the worst variable closure in the $Z \to \ell^+\ell^-$
control region, where the lepton has been removed from all quantities
(jets and $E_T$). ................................................................. 144
6.6 Impact of the various corrections on the total predicted yields of top and EWK processes. Picky jet CSV discriminator shape corrections have no impact on T2bW search regions, which only utilizes anti-$k_T$ jets.

6.7 Impact of various kinematics reweighting scale factor sets on the total predicted yields of top and EWK processes.

6.8 MC predicted single top yields and uncertainties (in events).

6.9 MC reweighting predicted $t\bar{t}$ and $W$ yields and uncertainties (in events).

6.10 MC reweighting predicted $Z \rightarrow \nu\bar{\nu}$ yields and uncertainties (in events).

7.1 Results of the QCD prediction.

8.1 Simulated yields for $t\bar{t}Z$ for both the Madgraph and MCatNLO samples.

8.2 Yields for data and simulation in both the $t\bar{t}Z$ and di-boson control regions. The uncertainty on $SF_{t\bar{t}Z}$ is taken as the cross section systematic uncertainty.

8.3 Comparison of data-simulation kinematic uncertainties. The first column shows the change of $t\bar{t}Z$ yields after applying the $t\bar{t}$ kinematic SFs. The second column shows the kinematic closures derived in Section 6.3 for $t\bar{t} + W$. The third column show the kinematic closures derived in Section 6.3 for $Z$. Finally, the last column is the kinematics uncertainty that we apply to the $t\bar{t}Z$ sample.

8.4 $t\bar{t}Z$ prediction and uncertainties.

9.1 T2tt Search Region Yields for 18.9 fb$^{-1}$ study.

9.2 T2bW Search Region Yields for 18.9 fb$^{-1}$ study.

9.3 T2tt Search Region Yields for 19.6 fb$^{-1}$ study.

9.4 T2bW Search Region Yields for 19.6 fb$^{-1}$ study.

9.5 Summary of T2tt Search Region Yields.
9.6 Summary of Search Region Yields ............................................ 167

B.1 CMS: Run 1 - Dataset .............................................................. 247
B.2 CMS: Run 1 - Parked Dataset ................................................... 247
B.3 Single Electron Dataset ............................................................ 247
B.4 Single Muon Dataset ............................................................... 248
B.5 $H_T$ Dataset ......................................................................... 248
B.6 Triple Lepton CR Dataset .......................................................... 249
B.7 Luminosity and trigger used to collect the triple lepton control regions 250
B.8 T2tt simulated dataset ............................................................. 251
B.9 T2bW simulated dataset ............................................................ 251
B.10 Z simulated dataset ( $Z \rightarrow \nu\bar{\nu}$, $Z \rightarrow l^+l^-$ ) ..................... 252
B.11 W simulated dataset .............................................................. 252
B.12 $t\bar{t}W$ simulated dataset ....................................................... 253
B.13 $t$ simulated dataset ............................................................... 253
B.14 $t\bar{t}$ simulated dataset ......................................................... 254
B.15 Multijet QCD simulated dataset .................................................. 254
B.16 $t\bar{t}Z$ simulated dataset ........................................................... 254
B.17 Diboson simulated dataset ........................................................ 254
Chapter 1

Introduction

One of the crowning achievements of modern science is the ability to explain matter in terms of fundamental particles and interactions. The current gold standard for this is the Standard Model of particle physics (SM) [14] (Section 2.1). However, the SM has many faults: lacking a theory of quantum gravity, inability to explain dark energy [15], and dark matter [16]. Supersymmetry (SUSY) (Section 2.2) is an extension to the SM which remedies some of these concerns. This analysis is a search for a particular SUSY signal: stop-antistop pair production by means of all hadronic final states (Section 2.3). Furthermore, this analysis is conducted using 19.6 \text{ fb}^{-1} (Section 3.1.1) of \sqrt{s} = 8 \text{ GeV} proton-proton collision collected at the Compact Muon Solenoid (Section 3).

The primary motivation for this analysis is the desire to differentiate between SUSY stop signals from SM top signals. Naturally, in terms of the “big picture”, identification of stops would provide evidence to support the validity of SUSY or physics Beyond the standard model (BSM). What sets this analysis apart from many is its extensive use of boosted decision trees (BDT) (Section A.3.2) for signal as well as background identification. Furthermore, this analysis is complementary to “A search for direct production of supersymmetric top squark pairs decaying to all-hadronic
final states in pp collisions at $\sqrt{s} = 8$ TeV” [13,17].

My main contribution to that analysis was a method for electroweak background identification, “embedding” (Section 9 in [13]). Additionally, I measured the muon (electron) trigger, identification, and reconstruction efficiencies used for that method (Appendix C in [13]). In particular, the study was helpful in validating the results of the MC reweighing method that is used (Section 6) and restricting the input variables used in the lepton veto BDTs (Section 4.9). In addition, I performed a complete reanalysis of the data that included data that was not part of the original version of the analysis. My version of the analysis uses a similar methodology as [13], but incorporates an additional 657 pb$^{-1}$ of data, in hopes of achieving a more refined result. Additionally, in lieu of re-deriving all background based predictions, this analysis simply adapts them for our increased 657 pb$^{-1}$ of data. Unless otherwise specified the correction factors and background yields in sections 6, 7, and 8 correspond to those derived for a luminosity of 18.9 pb$^{-1}$. The predicted yields are then normalized to our 19.6 pb$^{-1}$ luminosity in Section 9. This gave me the opportunity to verify the original analysis and led to a slight increase in the sensitivity for new physics.

This thesis is outlined as follows:

- Chapter 2 is a review of the relevant physics and related motivations: The Standard Model of Particle Physics, Supersymmetry, and the Supersymmetric Top Quark search itself.

- Chapter 3 describes the experimental set-up used to conduct this analysis: The
European Center for Nuclear Research, the Large Hadron Collider, and the Compact Muon Solenoid Detector.

- Chapter 4 describes reconstitution: how the objects used in our analysis are constructed and identified from proton-proton collision at the CMS.

- Chapter 5 describes how we identify the Supersymmetric Top Quark signals, as well as the optimal regions we use to search for it.

- Chapters 6, 7, and 8, describe how we identify the backgrounds to our signal: W misidentification and Z(→ ν¯¯ν) + jets, QCD Multijet background, and t¯tZ.

- Chapter 9 are the results: Combines our information about the signal prediction, background predictions, and data yields, to make predictions on the existence of SUSY.
Chapter 2

The Physics

2.1 The Standard Model of Particle Physics

There are four known fundamental interaction types in the universe are gravity, the weak nuclear interaction, the electromagnetic interaction, and the strong nuclear interaction. Figure 2.1 summarizes the relative strength, range, and name of the mediating particles (except for gravity) for these forces. We currently have no way to detect gravity at a quantum level. The Standard Model of particle physics (SM) [14] is a theory of concerning the weak nuclear interaction, the electromagnetic interaction, and the strong nuclear interaction. The SM classifies all known subatomic particles into various schemes by property and interaction types (Figure 2.1) [18]. Currently, the SM is the closest theory we have to a grand unified theory of everything (GUT). The SM is unprecedented in terms of explanatory and predictive power.

<table>
<thead>
<tr>
<th>Name</th>
<th>Strength</th>
<th>Range</th>
<th>Mediating Particles</th>
</tr>
</thead>
<tbody>
<tr>
<td>Strong</td>
<td>1</td>
<td>$10^{-15}$ m (diameter of a medium size nucleus)</td>
<td>Gluon</td>
</tr>
<tr>
<td>Electromagnetic</td>
<td>$\frac{1}{137}$</td>
<td>Infinite</td>
<td>Photon</td>
</tr>
<tr>
<td>Weak</td>
<td>$10^{-6}$</td>
<td>$10^{-18}$ m (0.1% of the diameter of a proton)</td>
<td>W and Z bosons</td>
</tr>
<tr>
<td>Gravity</td>
<td>$6 \times 10^{-39}$</td>
<td>Infinite</td>
<td>Unknown?</td>
</tr>
</tbody>
</table>

Table 2.1: Fundamental Interactions
Figure 2.1: The Standard model of particle physics: 12 fermions and 5 bosons. Brown loops indicate which bosons (red) couple to which fermions (purple and green) [1].

The Standard Model is a Quantum field theory (QFT) [19]. A QFT is a theoretical framework for constructing models of quantum particles and their interactions. A QFT brings together special relativity (SR), quantum mechanics (QM), and field theory (FT). Gravity is not incorporated into the SM because we do not have a quantum description of gravity. In QFT the fundamental objects of nature are quantum fields. Additionally, particles are simply field quanta or excited states of these quantum fields. Interactions between particles are described by interactions of these fields.
Quantum Electrodynamics (QED) [20] was the first Quantum Field theory to be developed and explains electromagnetism to great precision. QED served as a template for all later QFTs. Quantum Chromodynamics (QCD) [21] is the QFT concerned with the strong nuclear interactions. The weak nuclear interaction is better understood in terms of electroweak theory (EWK) [22]. EWK is the quantum field theory that unites the electromagnetic interaction and the weak nuclear interaction. The Standard Model of particle physics (SM) is the culmination of QCD and EWK.

In general, the Lagrangian or Hamiltonian formalism, in conjunction with the least action principle, can be used to determine equations of motions and states of the system [23]. In QFT the Lagrangian and Hamiltonian (functions of generalized coordinates) are replaced with Lagrangian density and Hamiltonian density (a function of fields, their derivatives, and space-time coordinates). However, exact solutions for such systems usually don’t exist. Hence, perturbation theory is often used. Perturbation theory is a set of approximation schemes for describing a complicated system in terms of a simpler one. The simple system is “perturbed” by additional terms representing a weak disturbance. Solutions of the perturbed system are expressed as corrections to the solutions of the simple system.

In QFT, we are often concerned with transitions between some initial quantum state to some final quantum state, by means of perturbation theory [23]. The transition probability amplitude between states can be used to calculate observables such as scattering cross sections and decay widths. The transition probability amplitude
between states can be represented by means of a scattering matrix (S-matrix) between the initial and final states. S-matrix elements can be perturbatively calculated through knowledge of the Lagrangian density or Hamiltonian density. The “Feynman diagrams” provide a pictorial representations for the perturbative contributions to the transition probability amplitude [24].

Symmetries are very important in physics, because it can be related to conversed quantities. Noether’s theorem proves that for every continuous symmetry in a system, there are corresponding quantities whose values are conserved [25]. In its most general terms; a symmetry means that some particular mathematical object is invariant under some particular transformation. There are two types of symmetry breaking: Explicit and Spontaneous [26,27]. Explicit symmetry breaking occurs when the defining equations are thought to be symmetric in nature, but are not (solutions are not symmetrical). This often happens from introduction of small symmetry-breaking terms. An example of explicit symmetry breaking is the spectral line splitting due to magnetic perturbation. This is called the Zeeman effect [28]. Spontaneous symmetry breaking occurs when the defining equations are symmetrical (solutions are symmetrical), yet observation only reflects one particular solution. In this sense, the systems is fundamentally symmetric, yet splitting occurred somewhere. This results in observation of a state that does not indicate the “hidden” symmetry.

The SM is what we call a “gauge” QFT [24]. In its most general sense the term “gauge” refers to a redundant degree of freedom in the Lagrangian. A global gauge
transform is one in which an identical phase transformation is performed at every point in spacetime. A physical model is said to be globally gauge invariant (it has a global gauge symmetry), if its Lagrangian remains unchanging under a global gauge transform. A local gauge transform is one in which an local phase transformation is performed at every point in spacetime. This means every point in spacetime is undergoing a different transformation. A physics model is said to be locally gauge invariant, it has a local gauge symmetry, if its Lagrangian remains unchanged under continuous local gauge transformation.

Local Gauge symmetry is at the heart of the SM. The interactions between particles are entirely determined by the local gauge symmetries of the model. In order to achieve local gauge invariance for the “matter” fields in the SM, carefully arranged gauge fields need to exist in the Lagrangian. Gauge fields correspond to the “force carriers”. Furthermore, these gauge transformations and gauge fields are analogous to the mathematics of objects called “lie groups” [29]. A lie group is a mathematical “group” that is also a differentiable manifold. Lie groups themselves being associated with a group generator. The group generator is the generation set of the group. Furthermore, in a quantized gauge theory the gauge bosons (spin 1 particles) are quanta of the gauge fields. Consequently, the number of gauge bosons of the gauge field correspond to the number of generators of the lie group. The corresponding lie groups are often called the “gauge groups” or the “symmetry groups” of the corresponding Lagrangian/model. In QED the local gauge symmetry corresponds to the
gauge group U(1): one gauge boson, the photon (\(\gamma\)). The photon mediates the electromagnetic interaction: electric charge. In QCD local gauge symmetry corresponds to the gauge group SU(3): eight generators, eight gluons (\(g\)) (eight color states). The gluon mediates the strong nuclear interaction: color charge.

The weak nuclear interaction violates parity symmetry. Parity symmetry is a transformation that changes a left-handed coordinate systems into a right-handed one and visa versa. Indeed, experiment has shown that neutrinos always have their intrinsic spin pointed in the direction opposite their velocity [30]. The gauge group of EWK is SU(2)\(_L\) \(\otimes\) U(1)\(_{Y_w}\). The \(\otimes\) symbol denotes the gauge group it’s the product of both lie groups. Subscript L denotes that the gauge fields only couple to left-handed fermions, in accordance with the weak interactions parity symmetry violation. The subscript \(Y_w\) denotes “weak hypercharge” [24] is the relevant quantity being mediated. SU(2)\(_L\) actually has three generators and subsequently three gauge fields, \(W_1, W_2, W_3\). U(1)\(_{Y_w}\) has only one generator and one gauge field, \(B\). However, these gauge fields and generators don’t correspond to what we actually observe, \(\gamma\) field (\(A\)), Z field (\(Z\)), and W field (\(W^{\pm}\)). Instead, EWK describes these as superpositions of the former:

\[
Z = \cos\theta_w W_3 - \sin\theta_w B, \quad (2.1)
\]

\[
\gamma = \sin\theta_w W_3 + \cos\theta_w B, \quad (2.2)
\]
\[ W^\pm = \frac{1}{\sqrt{2}}(W_1 \mp iW_2), \]  
\hspace{1cm} (2.3)

where \( \theta_w \) is “the weak mixing angle” (empirically determined). This mixing arises via the “Higgs mechanism” [24]. This corresponds to the mass generation for the gauge bosons through electroweak spontaneous symmetry breaking. The Higgs Boson is a quanta of the Higgs field (\( \varphi \)). The W and Z boson arise as a results of the Higgs mechanism being applied to the mass-less fields of U(1) \( \otimes \) SU(2).

In total the SM has four types of fundamental quantum fields [24]. These are the fermion field (\( \psi \)), the four electroweak boson fields (\( W_1, W_2, W_3, B \)), the gluon field (\( G_a \)), and the higgs field (\( \varphi \)). As their names suggest they produce quanta of fermions, electroweak bosons, gluons, and the Higgs boson respectively. The gauge group of the standard model is SU(3) \( \otimes \) SU(2) \( \otimes \) U(1). SU(3) acts on \( G_a \); SU(2) acts on \( W_1, W_2, W_3 \), and \( \varphi \); and U(1) acts on \( B \) and \( \varphi \).

The fermions (1/2 spin particles) are referred to as “matter” particles [24]. Matter is further classified into “quarks” (\( q \)) and “leptons” (\( l \)). Quarks carry electric charge, color charge, and flavor. Accordingly, quarks can interact via the electromagnetic force, are the only known type of matter particle that interacts via the strong nuclear interaction, and can interact via the weak nuclear interaction. Quarks are color confined. This means any particle with color charge can never be in isolation. Subsequently, quarks commonly combine to form composite particles called hadrons. Hadrons come in two common types, mesons (quarks - antiquark (see below)) and
baryons (three quarks). However, some exotic baryon combinations are theoretically possible and the pentaquark (five quark state) has already been observed [31]. Some hadrons can exist as nucleons or analogs to nucleons. Quarks and gluons are often referred to as partons. Leptons carry electric charge, no color charge, and flavor. Leptons are referred to as either “charged-type” (-1 electric charge) and “neutral-type” (0 electric charge). Charged leptons interact via the weak force and the electromagnetic force. Charged leptons can form shells around nuclei, analogous to the way a electron forms shells around nuclei. Neutrinos do not interact electromagnetically. Neutrinos (neutral-type) only act via the weak nuclear force. Since the weak nuclear interaction is extremely short ranged, neutrinos typically pass through the universe undetected and unimpeded.

Additionally, for every fundamental particle there is an “antiparticle” [24]. Particles and antiparticles carry the same mass and same spin, yet have opposite electrical charge, opposite color charge, and opposite lepton flavor (when applicable). An antiparticle is denoted by adding the prefix “anti” to the particle name and by adding a bar over the particle symbol, ex. The antielectron (τ), the antiup (π), etc. The Higgs boson is its own antiparticle. Every fermion has an unique antiparticle. All gauge bosons, except the W, are their own antiparticle. The W boson actually comes in a +1 electrical charge (W^{+1}) and -1 electrical charge (W^{-1}) version. These two types of W bosons form a particle and antiparticle. The W boson is often written as W^{±}.
2.1.1 The Limitations

For all its success the SM is far from a theory of everything. In addition to lacking a theory of quantum gravity, the SM fails at explaining dark energy [15] and dark matter [16]. Dark energy is associated with the cosmological evidence for the accelerating expansion of the universe. Cosmological observation concludes that a large amount of mass in the universe is unobservable. The matter corresponding to this unobservable mass is referred to as dark matter.

In addition to these yet to be explained phenomena, the SM leads to a few conclusions that contradict observation. Perhaps the most pressing of these issues is referred to as the hierarchy problem [32–34]. This encompasses, the inability to explain the large gap between the plank scale (massplank = \(\sim 10^{18}\text{GeV}\)) and the electroweak symmetry breaking scale (\(\sim 10^2\text{GeV}\)). The plank mass is maximum allowed mass for point particles. At greater than the plank mass a point particle could spontaneously form a black hole. The hierarchy problem is best demonstrated when trying to calculate the mass of the Higgs boson to all orders of perturbation [35]:

\[ m_H^2 = m_{H_0}^2 + \Delta m_H^2, \]  

(2.4)

where \(m_H\) is the real physical mass, \(m_{H_0}\) is the bare mass, and \(\Delta m_H\) is the quantum loop correction to the mass.

Most next-to-leading order (NLO) contributions to the Higgs boson mass cancel each other out. However, the one-loop fermion contribution does not cancel out
(Equation 2.5 and Figure 2.2).

\[
(\Delta m_H^2)_f = N_f \frac{|\lambda_f|}{8\pi} \left[ \Lambda^2 + 6 m_f^2 \log \left( \frac{\Lambda}{m_f} - 2m_f^2 \right) \right] + O\left( \frac{1}{\Lambda} \right) \tag{2.5}
\]

Here \(\Lambda\) is the momentum scale, \(m_f\) is the mass of the fermion, \(N_f\) is fermionic degrees of freedom, and \(\lambda_f\) is the yukawa coupling.

Obviously \(\Delta m_H^2\) quadratically diverges as a function of the momentum scale. We therefore impose and cut off point. If the SM is the only physics until the plank scale, then the ultraviolet cut-off = plank mass. However, \(m_H\) still needs to correspond to the actual observable Higgs mass (\(\approx 125\) GeV [18]). Hence, a “fine tuning” of over 30 magnitudes would need to occur to get \(m_H\) at the right Higgs mass value. Fine tuning means that the parameters would have to be adjusted very precisely by hand. In lieu of fine tuning, a physical model that introduces a tuning is preferred.
2.2 Supersymmetry

Supersymmetry (SUSY) [36–40] is a model of particle physics beyond the SM (BSM). BSM refers to a theory that attempts to explain phenomena the SM is unable to explain, e.g., the hierarchy problem. The symmetry of SUSY introduces an operation that transforms bosonic states to fermionic states (and visa versa).

SUSY is a very general term and can mean any one of a near infinite set of models where SM particles have SUSY “superpartners” (sparticles). This can even encompass models in which a partner has multiple SUSY superpartners. The key characteristic between a particle and it’s sparticles, is that their spin differ by half an integer. SUSY is fundamentally a theory about an underlying symmetry between fermions and bosons. If SUSY were “perfect” then all internal quantum numbers (besides spin) of particles and sparticles would remain invariant under the transform. We would observe sparticles with the same mass as their corresponding particle. However, we do not observe this. Hence, if SUSY exists the symmetry is said to be “broken” along mass.

It somewhat vague to talk about “SUSY” without referring to some particular type of SUSY. This analysis exclusively uses the minimal supersymmetric standard model (MSSM) [41]. The MSSM is the minimal SUSY extension to the SM. It is the most researched type of SUSY. A sparticle is labeled by adding a tilde above the name of its corresponding partner. The sparticles of fermions are named by prefixing “s” to the relevant term. For example, sfermion, slepton (\(\tilde{l}\)), squark (\(\tilde{q}\)), stop squark
(\tilde{t})$, etc. While the sparticles of bosons are named by modifying the end of the term and suffixing it with “ino”. Hence, bosino, photino ($\tilde{\gamma}$), Zino ($\tilde{Z}$), Wino ($\tilde{W}$), gluino ($\tilde{g}$), Higgsino ($\tilde{H}^0$), etc. Although there is some leeway in these naming conventions (ex. sboson). Hence, each SM fermion has a SUSY sparticle sfermion (a boson) with spin 0 and each SM boson has a SUSY sparticle bosino (a fermion) with spin 1/2. The neutralinos ($\tilde{\chi}_0^1$, $\tilde{\chi}_0^2$, $\tilde{\chi}_0^3$, $\tilde{\chi}_0^4$) are a set of four electrically neutral fermions. The lightest neutralino ($\tilde{\chi}_0^1$) is stable. The neutralinos states are mixtures of the bino, neutral wino, and neutral higgsino. They are majorana fermions, a fermion that is its own antiparticle. The charginos ($\tilde{\chi}_1^\pm$, $\tilde{\chi}_2^\pm$) are the mass eigenstates of any new electrically charged spin 1/2 fermion predicted by SUSY. The charginos are mixtures of charged wino and charged higgsino. Charginos decay as $\tilde{\chi}_2^\pm \rightarrow \tilde{\chi}_1^\pm Z$, $\tilde{\chi}_2^\pm \rightarrow \tilde{\chi}_1^0 W^\pm$, and $\tilde{\chi}_1^\pm \rightarrow \tilde{\chi}_1^0 W^\pm$. The chargino and neutralino subscript denotes the order of increasing mass.

In the MSSM, the baryon number (number of baryons minus the number of antibaryons) and lepton number (number of leptons minus the number of antileptons) are not conserved by some of the coupling terms. However, since baryon number and lepton number conservation has been extensively experimentally confirmed, this causes a contradiction. In order to match experiment, we need eliminate these terms. 

R-party is a Z2 MSSM [42] symmetry concerning the leptons and bosons designed to do just that. Conservation of R-party forbids these couplings from happening. 

R-parity is defined to be:
\[ P_R = (-1)^{3B+L+2s}, \tag{2.6} \]

where \( s \) is spin, \( B \) is baryon number, and \( L \) is lepton number. SM particles have R-parity of +1, while SUSY particles have R-parity of -1. Conservation of R-parity implies that SUSY particles must be pair produced and have lowest order entry into the SM through loop diagrams. Consequently, this implies that the direct production rates of SUSY particles are very small and the indirect effects on the SM are very small. This would explain why they have not been detected yet.

The Lightest supersymmetric particle (LSP) is the lightest superpartner in SUSY. Furthermore, conservation of R-parity implies the existence of a stable and likely electrically neutral LSP \[43, 44\]. This particular type of LSP is expected to appear at the end of SUSY decay cascades and cannot be detected directly. Instead, its presence is inferred from missing energy. This type of electrically neutral weakly interacting LSP is a very attractive candidate for Dark Matter \[45, 46\].

The MSSM arose as an attempt to stabilize the weak scale and solve the hierarchy problem \[47\] (Section 2.1.1). When trying to calculate the Higgs mass, we get a one-loop SUSY scalar contribution (Picture 2.3) supersymmetric to the original SM one-loop contribution (Picture 2.2). Hence, the Higgs mass correction term becomes \[35\]:

\[ (\Delta m^2_{H_f}) = N_f \frac{\lambda_f^2}{4\pi} \left[ (m_f^2 - m_s^2) \log \left( \frac{\Lambda}{m_s} + 3m_f^2 \log \frac{m_s}{m_f} \right) + m_f^2 \right], \tag{2.7} \]

where \( m_s \) is the mass of SUSY scalar and the rest of variables given below Equation
This equation no longer contains quadratic divergences, only a small logarithmic divergence. The logarithmic divergence requires fine tuning as well. However, the level of tuning can adjusted by requiring limits on the masses of the sfermions. If SUSY were exact, this divergence would vanish and the Higgs mass would becomes independent of the energy scale. When SUSY is broken, as we know it would be, the hierarchy problem returns when the differences in masses between fermion and sfermion are too large. To date, no sparticles have been observed. Hence, sparticles are generally expected to be more massive than their SM counterparts.

### 2.3 Supersymmetric Top Quark Search

#### 2.3.1 The Signal

The stop (\(\tilde{t}\)) and the sbottom (\(\tilde{b}\)) squarks are expected to be among the lightest sparticles [48–51]. Hence, the stop and sbottom are potentially the most accessible to the LHC. Additionally, it follows from conservation of R-parity [43,44] that the SUSY
particles are produced in pairs, the LSP has to be in every SUSY decay, and the LSP is necessarily stable. If the lightest neutralino ($\tilde{\chi}_1^0$) is indeed the stable LSP, it is a leading candidate for dark matter [45]. With these factors taken into consideration, the LHC has taken particular interest to look for evidence of $t\bar{t}$ production, through means of $\tilde{t}$ and $\tilde{\bar{t}}$ decay chains ending is SM particles and LSPs.

If MSSM is a solution to the hierarchy problem, then the stop squark mass scale must be similar to that of the top quark. Furthermore, the minimal difference between a stop decaying via a top ($\tilde{t} \rightarrow t + \text{LSP} \rightarrow \text{top decay products} + \text{LSP}$) and a top decay ($t \rightarrow \text{top decay products}$), is just the LSP. Hence, identification of the LSP could be used to differentiate between a top decay and a stop decaying via a top. This analysis is motivated by the notion that these stop via top decays may be going misidentified because of inadequate identification of the LSP. Using the “simplified model spectra” (SMS) [52–54] this particular stop via top decay model is referred to as “T2tt” (Figure 2.4). The lightest neutralino ($\tilde{\chi}_1^0$) corresponding to the LSP.

Top quarks decay before they get a chance to hadronize. Top quarks can only decay to a W boson and a down-type quark. Furthermore, the dominant decay is a W boson and b quark (91%) [18]. This makes the simultaneous identification of a W boson and b quark a good way to identify tops. Hence, we further narrow our search signal to $\tilde{t} \rightarrow t + \tilde{\chi}_1^0 \rightarrow bW + \tilde{\chi}_1^0$. Additionally, the W can decay into a lepton and neutrino (32%) [18], or to an up-type quark (not top) and a down-type quark (68%)
Neutrinos and LSPs both have the same signature, “Missing transverse energy” ($E_T$), in proton-proton collisions. Hence, we use all hadronic $W$ decays, as they provide the largest branching ratio and eliminate the possibility of confusing a neutrino with the LSP. The decay chain is therefore $\tilde{t} \to t + \tilde{\chi}_1^0 \to bW + \tilde{\chi}_1^0 \to bq + \tilde{\chi}_1^0$. The corresponding all hadronic final state $T2tt$ decay (Figure 2.5) is one of the signals we look for in this analysis. Since LSPs are identified through $E_T$, the total signature corresponding to this signal is $b\bar{b}qq + E_T$.

However, the all hadronic final state of the SMS “$T2bW$” produces the same...
signature as the all hadronic final state of T2tt. The corresponding T2bW decay chain is $\tilde{t} \rightarrow b + \tilde{\chi}_1^+ \rightarrow b + W^+\tilde{\chi}_1^0$. T2bW is shown in Figure 2.6. The all hadronic final state T2bW decay chain is $\tilde{t} \rightarrow b + \tilde{\chi}_1^+ \rightarrow b + W^+\tilde{\chi}_1^0 \rightarrow bq + \tilde{\chi}_1^0$. The all hadronic final state of T2bW is shown in Figure 2.7.

The all hadronic final states of the SMS T2tt and T2bW models are the two signals we seek in this analysis. The signature for these signals is $b\bar{b}qq + \not\! E_T$. Chapter 5 further describes how we identify these signals.
The background is any process that can be misidentified as the signal. This happens when a non-signal process produces the same signature, $b\bar{b}qq\bar{q}q + \not{E}_T$, as the signal. The relative strength of the signal or background can be expressed as a “cross section”, discussed in Section 3.1.1. The expected stop pair ($\tilde{t}\tilde{t}$) production cross section (Figure 2.8) ($10^{-4}$ pb to 1 pb) is much smaller than the top pair ($t\bar{t}$) production cross section ($\sim 250$ pb [55]). Hence, even a few top pairs misidentified as stop pairs can produce a huge background for this analysis. Therefore, the job of background
identification and suppression is of utmost importance to this analysis. In practice, the backgrounds for this analysis fit into four categories, W misidentification, $Z(\to \nu \bar{\nu})$ + jets, QCD multijets, and $t\bar{t}Z (Z \to \nu \bar{\nu})$. A “Jet” in this context corresponds to identification of a quark. Jets are further described in Section 4.5.

W misidentification is when a leptonically decaying W (Figure 2.9 Left) is misidentified as a hadronically decaying W. This happens when the charged-lepton is not properly identified, yet the neutral lepton is misidentified and its energy is mistakenly attributed to the LSP. This can happen from a W or a top quark with a W in its
Figure 2.8: Theory cross sections for selected SUSY processes at $\sqrt{s} = 8$ TeV. Green line is the expected stop pair production cross section [2–4].

decay chain. With addition of the needed jets this scenario might misidentified as the signal. This misidentification can occur in leptonically decaying $W +$ jets, single top $(t) +$ jets, and semileptonic and dileptonic decays of $t\bar{t} +$ jets. All these backgrounds is further discussed in Section 6.

The invisible decays of $Z$ ($Z \rightarrow \nu\bar{\nu}$) (Figure 2.9 Right) with the right combination of jets ($b\bar{b}qq\bar{q}q\bar{q}$) can also be misidentified as the signal. Here the neutrinos once again give the false impression of the LSP. This background is further discussed in Section 6.
A QCD multijet event is an event with more than 3 QCD jets. QCD multijets happen frequently and contribute a lot of $\mathcal{E}_T$, b-jets, quark-jets to an event. The right configuration of QCD multijets can indeed lead to $b\bar{b}qqq + \mathcal{E}_T$. This background is further discussed in Section 7.

The process $t\bar{t}Z (Z \rightarrow \nu \bar{\nu})$ (Figure 2.10) is similar to the $Z \rightarrow \nu \bar{\nu} +$ jets scenarios. Here the neutrinos yet again give the false impression of the LSP. If the tops decays into $b\bar{b}qqq + \mathcal{E}_T$ (and possibly LSP), then this scenario is misidentified as our signal. The process $t\bar{t}Z$ has a small production cross section ($\sim 242$ fb [56]). Furthermore, the branching ratio of $Z \rightarrow \nu \bar{\nu}$ is $\sim 20\%$ [18]. This brings the expected cross section for the process $t\bar{t}Z (Z \rightarrow \nu \bar{\nu})$ to $\sim 48$ fb. This background is significantly larger that the expected stop pair production cross section (Figure 2.8) ($10^{-4}$ pb to 1 pb). This background is further discussed in Section 8.
Figure 2.10: $t \bar{t} Z (Z \to \nu \bar{\nu})$ Decay

2.4 Summary

In this analysis we are looking for SUSY in the T2tt and T2bW channels. Both of these decay modes lead to the $b\bar{b}q\bar{q}q\bar{q} + \not{E}_T$ final state. Additionally, there are a number of known backgrounds that could be mistaken for this signal (Sections 6-8). The goal of this analysis is to classify possible T2tt and T2bW events (Section 5) in 19.6 fb$^{-1}$ collected at $\sqrt{s} = 8$ TeV with the CMS detector at the LHC, estimate how much of this is expected background, and look for any significant excess (signs of signal). This is often described as the “cut and count” method. This method takes
the total number of events passing a cut, subtracts the expected backgrounds, and observes the excess as an indication of signal.
Chapter 3

The Experimental Apparatus

This analysis was conducted using Run 1 proton-proton collision data from the Compact Muon Solenoid (CMS) [7] particle physics detector at the Large Hadron Collider (LHC) [57]. The Large Hadron Collider is a particle accelerator located at the European Center of Nuclear Research (CERN). This section is intended to be a brief outline on how events are measured at the CMS at the detector level.

3.1 The Large Hadron Collider

The LHC is composed of a main collider ring, physics detectors situated on that ring, and the associated infrastructure. The collider ring is essentially a ring torus (27 km circumference, 3.8 m wide inner tunnel) buried underground (between 50 and 175 meters). The collider ring contains two parallel beam pipes, each for one “beam” traveling in opposite directions around the ring. The main ring is designed to accelerate these beams, as well as provide beam collisions at four main points on the ring. Located at these four points are the major physics detectors: CMS, ATLAS, ALICE, and LHCb.

The LHC is designed to operate in proton-proton collision, heavy-ion collision, and proton-ion collision mode. Beams are not created in the main ring, but rather in other
Figure 3.1: Simplified LHC accelerator layout for proton-proton and ion-ion collisions specially designed particle accelerators. The proton source is a bottle of hydrogen gas at one end of LINAC 2 [58]. An electric field is used to strip electrons from the gas, leaving only protons. The protons have an energy of 50 MeV by the time they reach the other end of the LINAC 2 [59]. After LINAC 2 the beams are run through a series of intermediate accelerators, before finally being dumped into the main LHC ring. The sequence of proton beam acceleration is as follows: The Proton Synchrotron Booster (PSB) [59,60] accelerates them to 1.4 GeV, The Proton Synchrotron (PS) [61] accelerates them to 26 GeV, The Super Proton Synchrotron (SPS) [61] accelerates
them to 450 GeV, and the main LHC ring [57] further accelerates them to 6.5 GeV (4 TeV for this analysis). A simplified picture of the CERN accelerator complex is given in 3.1.

Once the beams are up to the desired energy and configuration, they are collided. The collisions are recorded by the particle detectors. This thesis was done using the Compact Muon Solenoid particle detector. At its current energy, 6.5 TeV per proton and 13 TeV center-of-mass energy (\(\sqrt{s}\)), the protons in the LHC proton-proton collisions have a Lorentz factor of about 6,930 and travel at about 0.999999990 c. However, this analysis was done with using 2012 data, which has 4 TeV per proton and \(\sqrt{s} = 8\) TeV. This gives us a Lorentz factor of about 4,263 around 0.999999972 c. It is also very important to note that the LHC beams are not “continuous”. The LHC proton beams are actually protons bunched together in “discrete” blocks moving in the beam pipes. The distance between bunches is 25 ns (\(\sim 7\) m distance for protons moving near the speed of light). Each LHC proton beam can currently accommodate 2,808 bunches with roughly 115 billion protons per bunch. Interactions between the bunches from each proton beam take place at discrete intervals, known as the “bunch crossing” (BX) [18]. This occurs at the same rate as the bunch spacing, every 25 nanoseconds. Using the number of bunches \((N_b)\) and the revolution frequency of the protons \((f_R)\), the average crossing rate \(CR_{ave}\) can be estimated as follows:

\[
CR_{ave} = N_b \times f_R, \tag{3.1}
\]
where
\[ f_R = \frac{c}{27 \times 10^3 m}. \] (3.2)

Here \( 27 \times 10^3 m \) is the circumference of the LHC. Combining this information with the number of events per crossing allows \( (N_x) \) the calculation of the number of events per second \( (N_s) \) produced by the proton-proton collision is

\[ N_s = N_x \times CR_{ave}. \] (3.3)

When protons from each beam collide, it is referred to as an “event”. However, since we have a large number of protons in each bunch, there is a non-negligible probably of getting several simultaneous events. This is know as “pileup”. Special care must be taken to combat pileup effects in any analysis and described further in Section 4.

### 3.1.1 Collider Physics

In this section we go through the definitions of luminosity, cross section \( (\sigma) \), and pseudorapidity \( (\eta) \) [18]. These general collider physics concepts are crucial to this analysis.

Instantaneous luminosity \( (L) \) is proportional to the rate of proton-proton collisions per area per time at the interaction point. The luminosity from the LHC can be estimated through the following:
Here

- \( f \) is the frequency of the protons (speed/circumference).
- \( E \) is the beam energy.
- \( \varepsilon_n \) is the beam emittance. This is related to getting the LHC beams from the injector to the final ring.
- \( N_b \) is the number of bunches.
- \( N_p \) is the number of protons per bunch.
- \( \beta^* \) is the strength factor. This is related to focusing the beam.

The Integrated Luminosity (\( L_{\text{int}} \)) is simply the integral of luminosity with respect to time:

\[
L_{\text{int}} = \int Ldt.
\]  \hfill (3.5)

The cross section is a measure associated with the probability of an interaction or particle. The cross section is an effective area that represents the intrinsic likelihood of said interactions or particles. The cross section is denoted by \( \sigma \) and measured in units of area (meters or barn). A barn is defined as \( 10^{-28} \text{m}^2 \). Cross-section is detector independent and can be calculated strictly (in theory) from quantum field theory, scattering theory, and the standard model. The relationship between luminosity and cross section is as follows:

\[
L = \frac{1}{\sigma} \frac{dN}{dt},
\]  \hfill (3.6)
where $N$ is the number of events in a certain time $t$ and $\sigma$ is the interaction cross section.

Another important quantity of interest in particle physics is “rapidity” of a particle. The rapidity of a particle in a particle collider is defined as:

$$y_{\text{rapidity}} = \frac{1}{2} \ln \left( \frac{E + p_z}{E - p_z} \right),$$

(3.7)

where $E$ is the energy of the particle and $p_z$ is the momentum of particle in $z$ (beam axis). Differences between the rapidity of particles are Lorentz Invariant under boosts along the $z$-axis (beam axis), making it a quantity of interest. However, it is often hard to measure rapidity as it is often hard to get the total energy or momentum of a particle. Hence, we define an approximate rapidity or “pseudorapidity” of a particle as follows:

$$\eta = -\ln \left[ \tan \left( \frac{\theta}{2} \right) \right],$$

(3.8)

where $\theta$ is the polar angle measured from the $z$-axis (beam axis). Pseudorapidity is approximately equal to rapidity for highly relativistic particles.

### 3.2 The Compact Muon Solenoid

The Compact Muon Solenoid detector (CMS) [7] is one of two general purposes detectors at the LHC, the other being the ATLAS detector [62]. The CMS is located in an underground cavern in Cessy, France. The complex housing the CMS is referred to as “point 5”, the fifth major complex going clock-wise around the LHC ring (as
viewed from space). It is the complex furthest from where the beams originate, the SPS.

CMS gets its name from its design features. The central feature of the CMS is a large powerful superconducting solenoid electromagnet, hence “solenoid”. CMS was designed so most of the detector would be “compact” within this magnet, while another feature of the CMS is its extensive “muon” detection system.

CMS is roughly cylindrical (15 m diameter). It has its cylinder height (21.6 m) situated parallel to the ground. The CMSs cylindrical axis of rotation is along the LHC beam axis. Particle collisions occur at the center of the CMS cylinder. A visualization of the CMS is given in Figure 3.2.

The CMS defines a relative Cartesian coordinate system \((x,y,z)\):

- **Origin**: Center of the CMS cylinder
- **\(x\)**: Pointing radially inwards towards the center of the LHC ring
- **\(y\)**: Pointing vertically upward into space
- **\(z\)**: Along the LHC ring pointing counter-clockwise (as observed from space)

It is often more convenient to use coordinates more suited for particle physics calculations \((r, \eta, \phi)\):

- **\(R\)**: Radius from the center of the CMSs cylindrical axis of rotation
- **\(\eta\)**: “Pseudorapidity” (Section 3.1.1) is preferred over the polar angle
- **\(\phi\)**: Azimuthal angle measured from the \(x\)-axis in the \(x-y\) plane

Another term of note is “transverse”. Transverse is used to indicate a measurement or component perpendicular to the beam-line (the \(x-y\) plane). While “longitudinal”
Figure 3.2: Overview of the CMS detector [5]

is used to indicate the $r$-$z$ plane. Furthermore, an “impact parameter” in the distance of closest approach relative to some point in space.

CMS is subdivided into a series of five roughly hollow cylindrical like layers, the innermost layer being the one with the smallest radius, the outer most layer having the largest radius, every layer being encompassed by the next one. The term “Endcap” is used to describe components located on the circular endplane of the detector. The endcaps are designed to measure particles traveling primarily through the $z$-axis, that is through the ends of the cylinders. The term “Barrel” is used to describe com-
ponents on the curved surface of the detector cylinder. The barrel detector measures particles as they travel radially. The following subsections briefly go over the various subdetector elements and how particles are measured within the CMS. A general picture of how particles interact with the CMS layers is given in Figure 3.3.

![Figure 3.3](image)

Figure 3.3: Segment (radially) of the CMS subdetectors and how particles interact with them [6].

### 3.2.1 The Tracker

The innermost layer of the CMS detector is the Tracker [63]. The tracker helps to extract trajectories of charged particles. The trajectories can be used to calculate momentum. The tracker has a volume given by a cylinder of length 5.4 m and radius
1.1 m. It measures charged particles within $|\eta| \leq 2.5$. A diagram of the tracker is shown in Figure 3.4.

![Diagram of the tracker](image)

Figure 3.4: The tracker layout (quarter view along z). The Pixel Tracker is shown in green. Single-sided Silicon Strip Tracker modules are in Red. Double-sided Silicon Strip Tracker modules are in blue.

The tracker is based on silicon technology, which is designed to endure violent radiation. The tracker is made of "silicon pixels" and "silicon strips". The silicon pixels are at the core of the detector (closest to the interaction point), so they receive the highest intensity of particles. There are about 66 million silicon pixels. Each silicon pixel is about $100 \times 150 \ \mu m^2$ large. The silicon pixels have a resolution of $10 \mu m$ in $r \times \phi$ and $20 \ \mu m$ in $z$. The silicon pixels are surrounded by the silicon strips. There are 9.6 million silicon strips. The pixels and strips are designed to be extremely light weight, so they disturb the particle as little as possible. The tracker is designed to allow all particles to pass to the next layer (if they kinematically can).
If a particle passes through the tracker, a tiny charge is introduced and recorded. These correspond to detector “hits”. Offline software combines these hits using a fitting procedure to estimate a “local inner track”. Local charge tracks can later be extrapolated upon by other layers to form “global tracks” using the same fitting procedures. Many track fitting procedures exist. The baseline procedure for track reconstruction is the “Kalman filter” (KF) [64]. The KF is a linear least-square estimator. It is optimal for scenarios where all probability densities encountered during track reconstruction are Gaussian. A non-linear generalization of the KF is Gaussian-sum filter (GSF) [65].

The Lorentz force law implies that trajectories of charged particles are bent by a magnetic field. Equating the Lorentz force to centripetal force, allows us to calculate the traverse momentum of a charged particle in a magnetic field:

\[ p_T [\text{GeV/c}] = 0.2998 \times R [\text{m}] \times B [\text{T}], \]  

where \( R \) is the radius of curvature and \( B \) is the magnetic field. The main contributions of \( p_T \) resolution come from local resolution of the individual position measurements and multiple scattering. The general relationship between the transverse momentum \( (p_T) \) and its resolution \( (\sigma_{p_T}) \) [66,67] is given by:

\[ \left( \frac{\sigma_{p_T}}{p_T} \right)^2 = \left( \frac{b}{\sin[\theta]} \right)^2 + (a p_T)^2, \]  

where \( a \) is a geometric constant, \( b \) is a material constant, and \( \theta \) is the polar angle between \( z \) axis and track. The first term corresponds to the local track resolution...
and the second term corresponds to multiple scattering. The track transverse impact parameter \((d_0)\) is the impact parameter of a track with respect to the beam line or primary vertex (Section 4.1). The track longitudinal impact parameter \((z_0)\) is defined as the point of intercept between the line joining the first two hits and the z axis. The general resolution of these impact parameters \((\sigma_{d_0} \text{ or } \sigma_{z_0})\) \cite{67} is given by the form:

\[
\sigma^2 = a^2 + \left(\frac{b}{p_T \sqrt{\sin(\theta)}}\right)^2.
\] (3.11)

As with the momentum resolution, the first term corresponds to the local track resolution and the second term corresponds to multiple scattering. Studies have shown that isolated muons of \(p_T = 100\) GeV emitted with \(|\eta| < 1.4\) have track resolutions of 2.8\% in \(p_T\), 10\(\mu m\) in \(d_0\), and 30\(\mu m\) in \(z_0\) \cite{68} using the CMS tracker.

### 3.2.2 The Electromagnetic Calorimeter

The second layer of the CMS detector is the Electromagnetic Calorimeter (ECAL) \cite{69}. A calorimeter (calo) in this sense is a device used to measure the energy of a particle. The ECAL is designed for measuring the energy of high-energy electrons and photons through full absorption, but can measure the energy of hadrons passing through it as well. Particles that interact through the electromagnetic or strong force, will deposit energy in the ECAL. The ECAL measures charged particles out to \(|\eta| \leq 3\). The ECAL barrel (EB) extends to \(|\eta| < 1.479\) and the ECAL endcaps (EE) covers \(1.479 < |\eta| < 3.0\). A ECAL preshower detector (ES) is located in front of the ECAL endcap from \(1.653 < |\eta| < 2.6\). A diagram of the ECAL is shown in 3.5.
The CMS ECAL is made of lead tungstate crystals (PbWO$_4$). These crystals are optically clear, heavier than stainless steel, and extremely dense (8.28 g/cm$^3$). A high-energy charged particle passing through this material will produce an electromagnetic shower (EM shower). An EM shower is a cascade of particles produced as the result of an initial high-energy electromagnetic particle interacting with dense matter. Ideally, ECAL will stop high-energy photons, electrons, and the shower, from making it to the next layer. This is in part to the small “radiation length” ($X_0$). $X_0$ is the mean length in which a high energy electron has lost all but 1/e of its original energy to bremsstrahlung radiation. For instance, energy loss (Stopping power) for electrons as a function of distance traveled traveled (x) and $X_0$ is [18]:

$$-rac{dE}{dx} = \frac{E}{X_0}.$$ (3.12)
A smaller radiation length implies better electromagnetic shower position resolution as well as better separation between neighboring electromagnetic showers. ECAL crystals have a small $X_0$ of 0.89cm. The EB crystals have a length of 23cm ($25.8X_0$) and the EE crystals have a length of 22cm ($24.7X_0$). The ES is designed specifically to aid the EE in differentiating between individual high-energy photons and photon pairs.

Furthermore, the PbWO$_4$ is designed to scintillate when charged particles pass through it. Scintillation is the property of luminescence (photons) when excited by ionizing radiation. The PbWO$_4$ scintillation process is very fast, 80% of the photons are emitted when the LHC bunch spacing is 25 ns. For a charged particle stopped by the ECAL, the total luminescence being produced from all the interactions is proportional to the initial particle energy, thus allowing us to calculate that particles initial energy. This type of calorimetry, in which a material is both the “absorber” and the “active medium”, is referred to as homogeneous calorimetry. Photodetectors are glued onto the back of each crystal to detect the scintillation light (or prompt photons) and convert it to an electrical signal that is amplified and sent for analysis.

An “island cluster” is simply a cluster of ECAL crystals believed to have energy deposited from that particular particle. The island algorithm starts with a “seed crystal” with energy above some defined threshold. From the seed position the island algorithm scans adjacent crystals and clusters them into an island. A “supercluster” is a collection of non-overlapping island clusters in the ECAL [70]. It ideally corresponds
to one particle transversing the ECAL.

The general relationship between particle energy ($E$) and its resolution ($\sigma$) in the ECAL can be parameterized as:

$$\left( \frac{\sigma}{E} \right)^2 = \left( \frac{S}{\sqrt{E}} \right)^2 + \left( \frac{N}{E} \right)^2 + (C)^2,$$

(3.13)

where $S$ is a stochastic term, $N$ is a noise term, and $C$ is a constant term [8]. These variables are determined by fitting. Test-beam studies on electrons in EB [71] place $S = 2.8\%$, $N = 127\,\text{MeV}$, and $C = 0.30\%$. Additionally, studies [72] have shown that the energy resolution of photons from Higgs decays vary across the barrel from 1.1\% to 2.6\% and from 2.2\% to 5\% in the endcaps.

### 3.2.3 The Hadronic Calorimeter

The third layer of the CMS detector is the Hadronic Calorimeter (HCAL) [73]. As its name suggests, the HCAL is concerned with calculating the energy of hadrons that pass through it. The HCAL is usually used for calculating the energy of high-energy hadrons. Although, charged hadrons that make it past the ECal will deposit energy in the HCAL as well. The HCAL covers hadrons out to $|\eta| < 5.2$. The HCAL Barrel (HB) and HCAL Outer Barrel (HO) cover $0 < |\eta| < 1.3$. The HCAL Endcaps (HE) cover $1.3 < |\eta| < 3.0$. The HCAL Forward (HF) extends the HCAL range out to $|\eta| < 5.2$. A diagram of the HCAL is shown in Figure 3.6.

The HE and HB are made of layers of alternating dense absorber material and active media (plastic scintillator). The absorber material is brass ($8.83\,\text{g/cm}^3$). A
high-energy hadron passing through a absorber layer will produce a hadronic shower. A hadronic shower is a cascade of particles produced as the result of an initial high-energy hadron interacting with dense matter. Charged particles passing though the plastic scintillator will produce luminescence. The light is collected by “wavelength-shifting” (WLS) fibers embedded in the plastic scintillators. This type of calorimetry, in which separate materials are used for the “absorber” and the “active medium”, is referred to as sampling calorimetry. HO is a tail catcher for HB. It is specifically designed so highly energetic hadronic showers in the barrel region can be fully contained. HO is outside the solenoidal magnet and only consists of the plastic scintillator tiles. The magnet is essentially the active medium for the HO. The HF is a unique among
the HCal in that it uses a special “quartz fiber calorimetry”. The HF is made of steel (8.05 g/cm³) embedded with quartz optical fibers. Cherenkov radiation from the quartz fibers forms the basis of signal generation in the HF. The amount of light in a given HCAL region is summed up over many layers in depth. This is called a “HCAL tower”. A single HCAL tower covers a 5x5 ECAL crystal window. The resulting collection, matching ECAL and HCAL clusters, is called a “Calorimeter tower” (CaloTower).

The primary length of interest when selecting an absorber material in sampling hadronic calorimetry is “nuclear interaction length” (\(\lambda_I\)). The nuclear interaction length is the mean distance traveled by a hadronic particle before undergoing an inelastic nuclear interaction. For heavy (analogous to high atomic number) material, \(\lambda_I\) is typically great than \(X_0\). This implies hadronic showers typically start later than electromagnetic showers and are more diffuse. Brass has an \(\lambda_I\) of 16.42 cm and a \(X_0\) of 1.49 cm. For HB the absorber thickness goes from \(5.82\lambda_I\) \((\eta = 0)\) to \(10.6\lambda_I\) \((|\eta| = 1.3)\). The HO increases the total depth to \(11.8\lambda_I\). The depth for HE along with EE provide \(\sim 10\lambda_I\).

In the CMS calorimeter (ECAL + HCAL), each calorimeter cell produces a four-vector. This four vector has an energy equal to the measured energy in the cell and points from the vertex to the center of the cell. The total energy resolution of the CMS calorimetry system can be parameterized analogous to ECAL. Performance studies on a pion test beam (2 to 300 GeV) [74] give the total energy resolution for
the CMS calorimeter as:

\[
\left( \frac{\sigma_E}{E} \right)^2 = \left( \frac{115.3\%}{\sqrt{E}} \right)^2 + (5.5\%)^2.
\]  

(3.14)

In addition to energy, the CMS calorimeter is responsible for determining missing transverse energy in an event (MET) \[75\]. MET is the imbalance in the transverse momentum of all visible particles in the final state of collisions. MET ($\vec{E}_T$) is defined as follows:

\[
\vec{E}_T = - \sum_i (\vec{E}_T)^i,
\]

(3.15)

where $\vec{E}_T$ is the MET four vector and $(\vec{E}_T)^i$ are the four vectors of the calorimeter cells. Since momentum is conserved along each direction, MET must be carried away by something invisible (not interacting through the electromagnetic or strong nuclear force). MET provides way to imply the existence of to particles the CMS is not able to directly detect. These include the neutrino and the LSP. Similar to fashion, it is useful to define the missing transverse momentum $H_T = | \sum \vec{p}_T |$ and scalar sum of the transverse energy $H_T = \sum E_T$. HF is needed for forward jet physics and calculation of good $\vec{E}_T$ resolution.

3.2.4 The Magnet

The fourth layer of the CMS detector is the magnet \[76\]. The magnet is a cylinder 13m long and 6m in diameter. The magnet is a solenoidal electromagnet comprised of superconducting niobium-titanium coils. It is the largest superconducting magnet in the world. As a superconductor the magnet requires cryogenics and is kept at a
temperature of $\sim 4.65\text{K}$ (only $\sim 1.65\text{K}$ warmer than outer space). The Magnet has an electric inductance of $14\text{H}$ and has a magnetic field of $3.8\text{T}$ ($\sim 100,000$ times stronger than Earth's magnetic field). The nominal current required to keep the Magnet at $3.8\text{T}$ is $\sim 19,500\text{A}$, giving the total stored energy in the Magnetic as $2.66\text{ GJ}$ (half a ton of TNT).

The purpose of the magnet is to curve the paths of charged particles emerging from collisions. Charged particles bend because of the Lorentz force. This allows us to calculate the momentum of a charged particle based on its curve. The tracker, ECAL, and HCAL are contained within the magnet.

### 3.2.5 The Muon Detector and Return Yoke

The fifth, and last, layer of the CMS detector is the muon detector \cite{77}. The muon detector covers $|\eta| \leq 2.4$. The muon detector is comprised of The Drift Tubes (DT), The Cathode Strip Chambers (CSC), and The Resistance Plate Chambers (RPC). The DTs cover $|\eta| < 1.2$, the CSCs cover $|\text{eta}| < 2.4$, and the RPC cover up to $|\eta| = 1.6$. The various sections of the various muon detectors are intertwined with “return yoke” plates. The return yoke is made of common structural steel. The role of the return yoke is to increase the magnetic field homogeneity in the tracker volume. By returning the magnetic flux of the solenoid, the return yoke also reduces stray magnetic fields. Furthermore, the yoke plates act as an extra barrier to any non-heavy particles that somehow got through the calorimeters. This ensures that
only muons (or other heavy stable charged particles) are reaching the muon detector. A diagram of the Muon system is shown in 3.7.

![Figure 3.7: The Muon Detector layout (quarter view along z) [7]](image)

The DT are used in the Barrel, the CSC are used in the Endcaps, and the RPC are used in both the Barrel and Endcaps. The basis for all three types of CMS Muon detectors is “gaseous ion detection” [78]. A gaseous ion detector is a detector filled with special gas that ionizes on contact with incoming charged particles. The DTs utilize a gas mixture of 85%Ar + 15%CO₂, the CSCs use 40%Ar + 50%CO₂ +
10%CF₄, and the RPCs use 96.2%C₂H₂F₄ + 3.5%C₄H₁₀ + 0.3%SF₆. There are two types of gaseous ion detectors worth mentioning. An “ionization chamber” has no gas multiplication taking place (no secondary ionizations). A “proportional chamber” has discrete ion avalanches (secondary ionizations). The CMS muon detectors are all proportional chambers. The DT uses a single positively charged wire to measure the ionization. The CSC use a series of anode wires crossed with cathode strips. The RPC use two oppositely charged plates on opposite sides of a chamber.

Hits in that muon system are used to construct a “muon track”. The muons identified from only the muon detection system are called “stand alone muons”. The momentum of stand alone muons can be calculated by incorporating information from the magnetic field and/or using pre-arranged track-to-momentum look-up tables. For each stand alone muon track, a matching track from the tracker is sought. A “global muon” track is then calculated by refitting all the hits in the stand alone muon track and the tracker track with a Kalman filter technique [64]. Global muons are used for this analysis.

The momentum resolution as a function of $p_T$ for muons is shown in Figure 3.8 [8]. The red line gives the resolution using the full CMS (including tracker). The green line is the resolution using only the CMS muon detectors. The blue line is the resolution for only the tracker. Studies [79] show the muon reconstitution and identification efficiency to be > 95% for muons with $p_T > 2$ and $|\eta| < 2.4$. 

3.2.6 The Trigger

The LHC has a BX every 25 ns or 50 ns. There are more than 20 interactions per bunch crossing, which translates to $\sim 10^9$ interactions per second. The average event size of is $\sim 10$ MB. Obviously, this is too much data for the CMS to record. The CMS is only able to record $\sim 300$ BX per second. The goal of CMS “Trigger system” is to discriminate the events to pick out the ones of interest. The trigger system disregards events in two steps, the Level-1 (L1) Trigger [80] and the High-Level Trigger (HLT) [81].

The L1 is implemented in customized electronics housed in the detectors and in a service cavern 90 m away from CMS experimental cavern. The L1 decision
making process (including transit time) occurs within 3.2 $\mu$s using high-resolution event data temporarily stored in a buffer. The L1 uses the buffered event data from the calorimeters and muon system to form “trigger primitives”. These are used to identify “trigger objects” (electrons, photons, jet, and muon candidates) and “global quantities” (Energy, $E_T$, jet multiplicities, $H_T$). L1 uses these objects and quantities to select for events of interest.

After the L1 decides to keep an event, the data in the buffer is streamed to a $\sim 1000$ processor computer farm. This computing farm runs the HLT software. The HLT has access to all the data of the event and can apply sophisticated algorithms to reconstruct the objects of interest. The HLT as a processor farm is highly beneficial, since it allows CMS to take advantage of advances in computing technology and ever evolving HLT reconstruction algorithms.
Chapter 4

The Object Identification and Reconstruction

We don’t actually ever get direct knowledge of any particle or particle interaction. What we get is detector observables that must be translated back into physics objects. The final stage of this process provides what we call “reconstructed physics objects”. However, the process of reconstructing these physics objects from observables is far from trivial.

The final state reconstructed physics objects we seek as signal in this analysis are; missing transverse energy, two b quark jets, and four lighter quark jets. This chapter describes all reconstructed physics objects used in this analysis.

4.1 Vertices

In general, a vertex is just a point in space associated with a particle decay or particle interaction. Vertex reconstruction generally depends on the ability to find correlations among tracks. Vertex reconstruction involves two general steps, vertex finding and vertex fitting. Vertex finding is done by grouping tracks into vertex candidates based on some discriminating algorithm. The vertex-finding algorithms vary depending on the type of physics analysis being performed and the type of vertex being sought. Vertex fitting is done by determining the best estimate of the vertex
parameters for a given set of tracks and calculating the quality of the fit.

When two protons collide at the LHC, the interaction point is referred to as a “proton-proton interaction vertex”. The CMS can produce a number of these proton-proton interaction vertices for any event. These include vertices associated with the pileup collisions as well as the actual “signal vertex” or “primary vertex” for the event of interest. “Secondary vertices”, “tertiary vertices”, etc, are associated with later interactions and later decays.

The remainder of this section is a summary of how the primary vertex can be obtain by means of the CMS tracker. A more detailed explanation of the methodology is available here [68]. The primary vertex candidates are found in three steps. These steps are selecting tracks, clustering of the tracks that seem to correspond to the same vertex, and finding the position of the vertex by fitting its associated tracks. The impact parameter (IP) of a track, is the impact parameter with respect to the primary vertex. The x-y and z components of the impact parameter are given by $d_{xy}$ and $d_z$, respectively. The significance of a variable is generally defined as the ratio of that variable to its uncertainty. The significance of the impact parameter is called SIP 3D.

Track selection for primary vertices, requires the tracks to be consistent with the region of primary interaction and the following criteria:

- The maximum impact parameter significance of the track to the beam spot be less that 5.
- The number of pixel layer hits being greater than 1.
- The number of pixel+strip hits being greater than 5.
The normalized $\chi^2$ fit to the trajectory be less that 20.

Clustering of the selected tracks is based on the $z$ coordinate of their closest approach to the beam line. The current clustering algorithm standard is the deterministic annealing (DA) clustering algorithm [82]. The DA clustering algorithm finds a list of “incomplete vertex candidates” and their respective tracks. We refer to these as “incomplete”, because they don’t contain full spatial information about the vertices. Spatially they only contain the $z$.

The incomplete vertex candidates with at least two tracks are then fitted using the adaptive vertex fit algorithm [83]. The adaptive vertex fit algorithm is used to obtain vertex candidates. We define “vertex candidates” as those with a full spatial position estimate. The adaptive vertex algorithm involves three main steps. The first step estimates the vertex candidates based on their tracks and the incomplete vertex candidates. The second step measures the compatibility between vertex candidates and their associated tracks. The compatibility measure goes from 0 (low compatibility) to 1 (high compatibility). The third step removes tracks (for every vertex candidate) with compatibility less than 0.5. This process is iterated until no new vertex candidates are produced. We refer to vertex candidates produced by the adaptive vertex fit algorithms as “primary vertex candidates”.

The primary vertex candidate with the highest $\Sigma (p_T^{\text{track}})^2$ is selected as the primary vertex, where $p_T^{\text{track}}$ is the traverse momenta of the tracks. While this methodology is consistent with the one used by our b-tagging algorithms [84], there are many different
methods that can be used to find the primary vertex in a CMS event. The different methodologies corresponding to the different type of physics and reconstruction being done.

It is important to note that the we refer to both “incomplete vertex candidates” and “vertex candidates” as simply “vertices candidates”. Additionally, all these terms in the context of searching for primary vertices could simply be called “primary vertex candidates”. We use only impose the distinctions here for simplicity.

4.2 Particle Flow

This analysis uses the CMS Particle Flow (PF) algorithm as a starting point for particle identification and reconstruction [85]. Particles arising from PF are simply referred to as PF candidates. The challenge of PF is to correctly associate all the deposits (energy) in the detectors and tracks in the detector, with particles. The goal of PF is to provide a single list of reconstructed photons, charged hadrons, neutral hadrons, muons, and electrons. The are five main ingredients of the PF algorithm are calorimeter clustering, tracking, tracker extrapolation to the calorimeters, muon identification, and electron identification.

Calorimeter clustering is simply a set of PF clustering algorithms designed to disentangle the overlapping showers in the calorimeters. The overall goal being to form a one-to-one correspondence between a cluster and an shower. Tracks are extrapolated through the calorimeters to look for associated clusters. The specific algorithms for
this extrapolation and tracks in general are given by the tracking Physics Object Group (POG) [86]. In general, POGs are responsible for maintaining the collaboration wide specifications for any particular particle identification. The set of a track and cluster(s) constitute a charged hadron in the PF framework.

Muon identification is done following the recommendations of the Muon POG group [87]. PF Muons are global Muons. However, muons must be identified before charged hadrons. This ensures that the muon track isn’t incorrectly identified as a charged hadron. Electron identification follows the guidelines of the e/gamma POG group [88]. However, electrons require a special track reconstruction, because of the frequent emission of Bremsstrahlung photons. These are photons produced by the deceleration of a particle when deflected by an electromagnetic field. This special reconstruction [89] is designed to properly attach the photon clusters and avoid energy double counting.

When all the tracks have been accounted for, any remaining clusters can then be associated with photons (ECAL clusters) or neutral hadrons (HCAL clusters). When all the deposits of a particle have been associated, the nature of the particle can be assessed. The subdetector information is combined to determine optimally the particles four-momentum. If the energy from the clusters is in excess of that derived from the track momentum by more than one sigma, the excess is attributed to a neutral particle (photon or hadron) carrying the energy difference.

The PF candidates are used as seeds for higher level reconstruction algorithms.
These higher level algorithms include jets, b jets, taus, missing transverse energy ($E_T$), and to compute charged lepton/photon isolation. Jets algorithms are described in Section 4.5. b jets are generally identified by examining the contents of jets. b jet algorithms are described in Section 4.6. Taus are generally identified using information about their decay products [90]. $E_T$ is generally calculated by summing over the transverse energy vectors of the reconstructed particles and simply taking the negative (see section 3.2.3 and 4.3). Isolation is explained in section 4.4.

The performance of the PF algorithm has been extensively studied and more information can be found here [85, 91]. The PF algorithm is currently the default standard for physics analysis at the CMS. Although, it is often merely the starting point for a more in-depth participle identification and reconstruction methodology.

### 4.3 Missing Energy

The general subdetector level based $E_T$ is given by equation 3.15. However, this analysis uses a particle flow definition of $E_T$ with type-0 corrections [92]. $E_T^{\text{type}-0}$ is the type-0 corrected negative vectorial sum of the missing energy of all Particle Flow Particles in the event:

$$
E_T^{\text{type}-0} = E_T^{\text{raw}} + 0.703 \times p_T^{ch} \times (1 + erf(-0.030 \times (p_T^{ch})^{0.099})).
$$

(4.1)

Here

- $E_T^{\text{type}-0}$ is the type-0 corrected negative vectorial some of the missing energy of all Particle Flow Particles in the event.
- $\vec{E}_T^{\text{raw}}$ is the negative vectorial sum of the missing energy of all particle flow particles in the event:

$$\vec{E}_T^{\text{raw}} = - \sum_{\text{PF particles}} \vec{p}_T.$$  

- $\vec{p}_T^{\text{ch}}$ is the vectorial momentum sum of all charged particles associated to pileup vertices.

- erf is the error function.

This type-0 correction is designed to reduce the influence of pileup on the $E_T$. The first $\vec{p}_T^{\text{ch}}$ term accounts for removing the charged component of the pileup interactions, but it also degrades the $E_T$ resolution. The second $\vec{p}_T^{\text{ch}}$ term is dependent on the magnitude of $\vec{p}_T^{\text{ch}}$, counteracts the $E_T$ resolution degradation, and accounts for the neutral component of the pileup interactions.

This analysis uses “Monte Carlo” (MC) based background predictions to improve the statistical precision of merely data based background prediction methods. MC are a specific class of programs used to generate synthetic data. More information on MC and synthetic samples we used in this analysis are summarized in Appendix B.3. MC based background prediction is especially important as the search regions we are dealing with have very low yields, including background yields. The overall goal of MC based prediction is a simulated background sample that can be used to accurately predict the backgrounds we would expect in data. However, the MC modeling of $E_T$ is a well known problem at the CMS [93]. The three major inadequacies of MC $E_T$ modeling are:

1. The average $E_T$ response (defined as the ratio of measured over true magnitudes)
is lower in data than in simulation.

2. $\mathcal{E}_T$ resolution is worse in data than simulation.

3. There are various $\phi$-dependent effects in the detector that are not fully modeled by simulation. These effects induce a $\phi$ modulation of the inherently flat $\mathcal{E}_T$ distribution that is stronger in data than simulation.

The first two problems impact the $\mathcal{E}_T$-lepton correlations we model in simulation. However, we have found that correcting/equalizing for the third problem has very little impact on any observables of interest to this analysis. This analysis uses a custom $\mathcal{E}_T$ scale and resolution correction procedure that can be found in section 3.2 of [13]. The procedure was developed through studies of $Z \rightarrow l^+ l^-$ [13] event samples, which lead to corrections to the simulation modeling of $\mathcal{E}_T$. These custom corrections are applied to our top/ EWK background prediction. The QCD backgrounds require more detailed corrections. These are further touched upon in Section 7.

4.4 Isolation

“Isolation” is particular measure of how isolated a particular particle is. Isolation is usually expressed as all the momentum/energy in a cone around a particle:

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

(4.3)

Isolation helps ensure that particles are identified correctly and helps to distinguish isolated particles from particles that are part of a jet. As with $\mathcal{E}_T$, isolation can be
derived on a number of different levels. The general subdetector based combined isolation is defined as:

$$\sum p_T(TRK) + \sum E_T(ECAL) + \sum E_T(HCAL),$$  \hspace{1cm} (4.4)$$

where the sums are taken over $p_T$ derived from tracks and the energies are from deposits in the ECAL/HCAL. The combined PF isolation (PFIso) is defined as:

$$\sum E_T(\text{charged hadrons from PV}) + \sum E_T(\text{neutral hadrons}) + \sum E_T(\text{photons}),$$  \hspace{1cm} (4.5)$$

where the sums are now taken over the collections of PF candidates. The terms of this equation can be referred to at the “relative PFIso on charged”, “relative PFIso on neutral”, and the “relative PFIso on photons”, accordingly. Unless otherwise specified, the isolations used in this analysis have a cone of $\Delta R \leq 0.3$.

### 4.5 Jets

Due to color confinement partons hadronize. A parton ends up being fundamentally responsible for the production of a plethora of subsequent particles and decays. The cone formed by these resultant particles is referred to as a “jet” (Figure 4.1). Approximately, these jets should be in correspondence with the initial partons. Jet finding is essentially an attempt to reverse engineer the hadronization process of the partons.

Jets are one of the main reconstructed physics objects we look for at the LHC. Searching for jets is not a trivial process. Many different jet algorithms and method-
ologies have been developed [94–96]. There are a variety of desired characteristics and properties to consider when choosing jet algorithms including computation speed, collinear-safety, infrared-safety, and softness 4.2. Collinear-safety means that collinear splitting shouldn’t change jets, while infrared-safety means that soft emissions shouldn’t change jets. Softness refer to the jet algorithms adaptability to the successive branching nature of QCD radiation. Soft-resilient jets are not influenced by soft radiation, while soft-adapTable jets have irregularities in the boundaries provoked by soft radiation.

There are two type of mainstream jet clustering algorithms: cone-type algorithms and sequential clustering algorithms. Cone-type algorithms arise from a simple geometric motivation. In cone-type algorithms we seek a stable cone: the total 4-momentum of all encompassed particles is along the cone axis. The stable cone is the jet. In sequential clustering algorithms we define a special distance metric. We
sequentially cluster the particles based on the metric. The Sequential clustering algorithm uses the distance metrics $d_{ij}$ and $d_{iB}$. $d_{ij}$ is the distance between any two particles $i$ and $j$:

$$d_{ij} = \min(k_{T_i}^2, k_{T_j}^2) \frac{\Delta_{ij}^2}{R^2}. \tag{4.6}$$

$d_{iB}$ is the distance between any particle $i$ and the beam $B$ (the beam distance):

$$d_{iB} = k_{Ti}^{2p}. \tag{4.7}$$

Here $\Delta_{ij}^2 = (y_i - y_j)^2 - (\phi_i - \phi_j)^2$, $k_{Ti}$ is the transverse momentum of particle $i$, $y_i$ is the rapidity of particle $i$, and $\phi_i$ is the azimuth of particle $i$. The Sequential clustering algorithm proceeds as follows:

1. **Metric**: For all particles $i$, find the smallest distance metric
   - If $d_{ij}$ is the smallest, combine $i$ and $j$ (sum over 4-momentum) to form a new particle $k$.
   - If $d_{iB}$ is the smallest, remove particle $i$, call it a “jet”.

2. **Iterate**: Repeat until all particles are in a jet.
The geometric factor, “p”, governs the relative power of energy vs geometric scale:

- 0 = Cambridge/Aachen (C/A) algorithm [95]
- 2 = kt algorithm [96]
- -2 = Anti-Kt (AK) algorithm [94]

The radius parameter, “R”, scales $d_{ij}$ with respect to $d_{iB}$ such that any pair of final jets i and j are at least $\Delta_{ij}^2 = R^2$ apart.

In general, sequential clustering algorithms are collinear-safe and infrared-safe by construction. Sequential clustering algorithms are currently the gold standard for many CMS analyses. Furthermore, the AK algorithm is soft-resilient. It is the only mainstream CMS jet algorithm to incorporate collinear-safety, infrared-safety, and soft-resilience. Hence, the AK jets are extremely desirable for our analysis.

This analysis uses two different jet clustering algorithms: AK5 and Picky.

### 4.5.1 AK5 Jets

The AK5 jets are used for our T2BW search to cluster PF particles into jets. “AK” in AK5 simply refers to the anti-kt algorithm [94], while “5” in AK5 refers to a distance parameter $\eta-\phi$ of 0.5 used in that algorithm. This analysis uses the FASTJET package [97] to implement the AK5 algorithms.

The AK5 jet reconstruction algorithms is just the first step in a robust jet reconstruction methodology. Additional “jet energy scale” (JES) corrections must be accounted for to make the jet more representative of the physics present and the initial parton responsible. There are three standard sequential corrections used for all
jets, referred to as level 1 (L1), level 2 (L2), and level 3 (L3) corrections, respectively.

L1 corrects our jets for pileup energy. This occurs when the energy resulting from a different primary vertex candidate is misattributed to the primary vertex of interest. Hence, extraneous energy may end up falsely attributed to our jet. This energy needs to be removed from our jets on an event-by-event basis. This is done in a two fold process. First, the “Charged Hadron Subtraction” (CHS) method [98] removes charged hadrons associated to pileup vertices. Only the remaining PF candidates are then used for jet formation. CHS is performed in this analysis a priori jet construction. Second, the pileup energy being mistaken for neutral particles is calculated (after jet clustering) using L1FastJet [99] parameterization. This mistaken neutral particle pileup energy is then subtracted using the baseline areamedian method [100]. The general formula for this pileup correction is given by:

\[ p^\mu (PU \rightarrow 0) = p^\mu - \rho A^\mu, \]

(4.8)

where \( p^\mu \) is the momentum 4-vector, PU \( \rightarrow 0 \) indicates the corrected quantity, \( \rho \) is the globally calculated pileup transverse-momentum flow per unit area, and \( A^\mu \) is the individual jets area.

Once the jet 4-momentum has been corrected for pileup, the jet momentum is corrected for in two more levels [101]. Calibration factors derived from simulation are used. These calibration factors account for the pseudorapidity dependence on jet energy (L2) and momentum dependence on jet energy (L3). The general L2 and L3
corrections to the 4-momentum are given by:

\[ p_\mu^{\text{corr}} = C \cdot p_\mu^{\text{raw}}, \tag{4.9} \]

where the correction \( C \) is composed of an offset correct \( C_{\text{offset}} \), a MC calibration factor \( C_{\text{MC}} \), a residual calibration for the relative energy scale \( C_{\text{rel}} \), and a residual calibration for the absolute energy scale \( C_{\text{abs}} \). The offset correction is designed to remove energy associated with pileup and noise. The MC correction removes the bulk of the non-linearity in \( p_T \) and non-uniformity in \( \eta \). The residual correction accounts for the differences in simulation and data. The total correction is done sequentially and can be written as:

\[ C = C_{\text{offset}}(p_T^{\text{raw}}) \cdot C_{\text{MC}}(p_T^{'}) \cdot C_{\text{rel}}(\eta) \cdot C_{\text{abs}}(p_T^{''}). \tag{4.10} \]

Here \( p_T^{'} \) is taken after applying \( C_{\text{offset}} \) and \( p_T^{''} \) is taken after applying everything before.

\section*{4.5.2 Prelude to Picky Jets}

The fully hadronic decays of T2tt produce two top quarks. These two top quarks decay into two b quarks and four light quarks. Each of these six quarks will generally produce multiple final state hadrons. Hence, there is a substantial amount of activity in the detector when searching for these signals. This activity can lead to complications in jet reconstruction algorithms, including the AK5 algorithm. Perhaps, the two important issues that arise in such a cluttered topology are jet “merging” and jet “splitting”.
Jet merging occurs when particles originating from different initial partons (different jets) are incorrectly associated to one initial parton (one jet) (Figure 4.3). Jet merging is prevalent when the initial partons are collimated. This a major effect in this analysis, as the boosting of top quarks can provide the partons with this collimation. Jet reconstruction algorithms with fixed distance parameter clustering, like the AK5, can be particularly susceptible to jet merging. If the distance parameter is too large, jets will naturally be merged. If the distance parameter is too small, jet merging can be reduced, but you also reducing the correct association of soft partons to jets. “Soft” specifying low-momentum transfer.

Jet splitting occurs when particles originating from one initial parton (one jet) are incorrectly associated with various other initial partons (other jets) (Figure 4.3). Jet splitting is prevalent when hadronization of low-$p_T$ partons produces particles that spread over distances comparable or greater than the distance parameter.

Jet merging and jet splitting provide a large challenge for this analysis. Our T2tt signal can lead to scenarios where jets can easily be merged or split. Figure 4.3 highlights this problem, the minimum $\Delta r(\eta, \phi)$ separation between the top-quark jets is often smaller than the distance scale used in AK5 jet clustering. Hence, AK5 clustering is insufficient for T2tt jet reconstitution. Furthermore, the T2tt topology can produce asymmetric momentum among the tops, which can lead to merged jets and split jets even in the event. Unless the tops are highly boosted, the top jets can be of both low and high $p_T$. 
In order to address all these concerns, a special type of variable radii jet clustering algorithm, “Picky Jets”, was developed with T2tt in mind. These Picky jets are part of the CORRAL framework [12] and have been used in a related analysis [13]. Picky jets are designed to resolve all six quark jets in the fully hadronic T2tt decay across a large range of possible top quark boosts.
4.5.3 Picky Jets

The picky jet reconstitution algorithm is a substructure-like method. The algorithm starts by clustering all PF particles not compatible with a pileup vertex. These particles are clustered into “proto-jets” by means of the C/A jet reconstitution algorithm with a distance parameter of 1.0 as opposed to the standard value of 0.5. This is done in an attempt to ensure we don’t leave out particles. The picky jet algorithm then seeks to iteratively break up these jets into sub jet pairs. Jets are split up until there is no more evidence of substructure. Determining whether a jet should be split and the structure of the resultant jets is a non-trivial matter involving the subjettness measures, \( p_T \) projection, and Boosted Decision Tree (BDT) discrimination.

First, the jet is decomposed into two sub jets. These sub jets are then determined by minimization of the N-subjettiness \([102]\) measure \( \tau_2 \). The general N-subjettiness measure is defined as:

\[
\tau_N = \frac{1}{d_0} \sum_k p_{T,k} \min(\Delta R_{1,k}, \Delta R_{2,k}, \ldots, \Delta R_{N,k}).
\]

(4.11)

Here \( k \) runs over the particles in the jet, \( p_{T,k} \) is their transverse momenta, and \( \Delta R_{J,k} = \sqrt{(\Delta \eta)^2 + (\Delta \phi)^2} \) is the distance in the \( \eta - \phi \) plane between the candidate sub jet \( J \) and particle \( k \). The normalization factor is defined as:

\[
d_0 = \sum_k p_{T,k} \Delta R_0,
\]

(4.12)

where \( R_0 \) is the characteristic jet radius used in the original jet clustering algorithm. Minimization of \( \tau_2 \) in an attempt to quantify how 2-subjetty a particular jet is. \( \tau_1 \)
corresponds to what degree a jet can be thought of as composed of two sub jets. Additionally, N-subjettiness measure $\tau_1$ is calculated. $\tau_1$ corresponds to what degree a jet can be thought of as composed of no sub jets.

Second, an axis is set up in $\eta - \phi$ space along the line containing the most extreme (highest $p_T$) particle for both subjets. The axis points from the lowest extreme particle (s1) to the highest extreme particle (s2). The zero value of the axis is chosen halfway between these particles. The 2-D $p_T$ density is projected onto this 1-D axis. The behavior of the critical points of this 1-D $p_T$ density can be used as an indicator of whether a jet should be split up or not. This method is shown, along with the critical points "L", "C", and "R" in Figure 4.4.

Third, a BDT is used to determine whether the "split" should be made. More information on machine learning and BDT’s can be found in Appendix A.3. The BDT uses the following variables:

- Jet mass
- $\tau_1$ and $\tau_2$
- the distance between sub jets ($\Delta R$)
- The L/C/R locations
- $p_T$ densities at the L/C/R locations

The BDT has highly efficient working points, 97% efficiency at splitting up merged jets and 9% probability of incorrectly splitting up already pure jets [11]. Figure 4.5 shows the performance of this working point in terms of efficiency. The left plot gives the efficiency of resolving jets from the decay of a more collimated top quark (t1) in
SUSY stop pair production. The right plot gives the efficiency of resolving jets from the decay of a more collimated top quark (t2) in SUSY stop pair production. The horizontal axis corresponds to different jet combinations. The first two bins are not used. The clustering algorithms compared are:

- antiKT0.3: Standard anti-kT jets with size parameter 0.3.
- antiKT0.5: Standard anti-kT jets with size parameter 0.5.
- caMassDrop1: Mass drop tagger with suggested parameters m-cut = 0.67 and y-cut = 0.15 [103].
- caPickyDeclus1: Picky jet algorithm starting from C/A jets (size parameter 1.0) and with subjets defined by declustering. Declustering is a different method that was investigated [12].
- caPickyNSub1: Picky jet algorithm starting from C/A jets (size parameter 1.0) and with subjets defined by minimization of N-subjettiness. This is the method we use in this analysis.

It is evident that the picky jets produce a better efficiency when reconstructing all three jets (right most bins). A more detailed look at the performance of picky jets can be found in Appendix K of [13] and section 4.2 of [12].

4.5.4 Picky Jet Corrections

The picky jet pileup correction is somewhat analogous to the AK5 pileup correction, although the implementation is more complicated. Pileup is corrected for by subtracting an appropriate energy associated with pileup interactions on an event-by-event basis. A detailed description of the method can be found here [12]. Picky jet pileup is corrected using a modified jet trimming method [104]. Jet trimming is a
Jet grooming is a systematic removal of a subset of jets. The goal of jet grooming is to remove soft and wide-angle radiation from jets. Grooming in the presence of pileup is used to reduce the jet mass dependence on pileup activity. The general pileup subtracted 4-momentum from jet grooming is an extension to the baseline areamedian method [100], given by something similar to:

$$p^\mu(\text{PU} \to 0) = p^\mu - \rho A^\mu - \rho_m A^\mu_m.$$  \hspace{1cm} (4.13)

The terms with the “m” are analogous to their unscripted counterparts and taken with respect to mass as opposed to $p_T$. The many different types of jet grooming include “pruning” [106], “soft drop” [107], “modified mass drop tagger” [108], and the aforementioned “trimming” [105].

Jet trimming removes particles in a jet that fall below some dynamic $p_T$ threshold. Jet trimming re-clusters the constituents of the jet using the kT algorithm with a radius $r_{sub}$. Only accepting sub jets that have:

$$p_{T_{sub}} > p_{Tfrac} \lambda_{hard},$$  \hspace{1cm} (4.14)

where $p_{Tfrac}$ is a dimensionless cutoff parameter, and $\lambda_{hard}$ is a hard QCD scale. $\lambda_{hard}$ is chosen to equal the $p_T$ of the original jet.

We use a jet trimming procedure for picky jets, because other methods like jet pruning are more appropriate for jets with significant substructure [98]. The picky
jet pileup subtracted 4-momentum we use is given by:

\[
\vec{p}_{\text{jet}}(\text{PU} \to 0) = \vec{p}_{\text{jet}} - \rho \sum_i \hat{e}_i,
\]

\[ (4.15) \]

\[
E_{\text{jet}}(\text{PU} \to 0) = E_{\text{jet}} - \rho \sum_i \sqrt{e^2_{i,z} + (\rho m_\rho + 1)^2}.
\]

\[ (4.16) \]

Here \( p_{\text{jet}} = (\vec{p}_{\text{jet}}, E_{\text{jet}}) \) is the measured 4-momentum of the jet and \( \hat{e}_i \) is the unit vector locations of a set of randomly distributed, infinitesimally soft “ghost” particles clustered along the jet. \( e_{i,z} \) is the z component of \( \hat{e}_i \). The ghost particles are generated from MC integration techniques and arise when doing the area calculation [109,110]. The ghost particles have an average separation of 0.001 in \( \eta - \phi \) space. This small spacing is necessitated by the possible small radii of the picky jets.

The equations for the picky jet pileup subtracted 4-momentum (4.15 and 4.16) are used as a targets for calculating the subjet \( p_T \) cutoff parameter (Equation 4.14).

Our jet trimming procedure is as follows:

1. The original jet is decomposed into a set of subjets by running the kT clustering algorithm with size parameter 0.1 on its constituents.

2. If the jet comprises of only one subjet, no PU subtraction is performed.

3. If the jet comprises of more than one subjet, the subjets are ordered by decreasing \( p_T \). For each subjet rank \( j \) (\( j = 1 \) is the highest \( p_T \) subjet) a trimmed 4-momentum is defined as the sum momenta of all subjets up to and including \( j \):

\[
P^{(j)}_{\text{trim}} = \sum_{i \leq j} p_{\text{sub},i}.
\]

\[ (4.17) \]
4. The distance:

\[ |p_{\text{trim}}^{(j)} - \vec{p}_{\text{jet}}(\text{PU} \rightarrow 0)|^2 + |m_{\text{trim}}^{(j)} - m_{\text{jet}}(\text{PU} \rightarrow 0)|^2 \]  

(4.18)

is minimized to select the best ranks \( j_{\text{min}} \) of subjets to keep. This is equivalent to finding \( p_{T,\text{frac}} \).

Studies on the performance of this method can be found in [12]. This method has two obvious advantages. First, since only physically present particles can be subtracted, it somewhat guards against over-subtraction. Second, it produces a PU subtracted list of constituents. This list can be directly used for computation of other jet shape quantities without performing a dedicated PU subtraction computation each time. It is also important to note that b-tagging decisions must be evaluated prior to picky jet PU subtraction. This is because the framework does not account for the fact that these charged particles should not be subtracted. However, as mentioned in Section 4.5.1, CHS is performed a priori. Hence, all charged hadrons associated to pileup vertices have already been removed prior to jet construction.

For simulation we correct the picky jet \( p_T \) spectra to better match data. This is done by rescaling the \( p_T \) of the jets by constants. These constants depend on the sample type, if it's a b jet, and if it is the leading or a subleading light quark jet (for top quark events). These constants are all close to unity and were found by simple trial and error. The values that produce the best \( p_T \) spectra match between data and simulation are used. This method works because the discrepancy with data appears to be \( p_T \)-independent, but somewhat flavor-dependent. Hence, only a small number
of constant correction factors are required. The procedure used to correct the raw energy is somewhat analogous to the JES corrections of standard AK5 jets (equation 4.9). However, for picky jets there is an important conceptual difference. The goal is to obtain good agreement between data and simulation, not to correct the energy scale to truth level. All picky jet correction factors are summarized in Table 4.1. More detailed information on these procedures, as well as comparisons of picky jet spectra before and after the corrections can be found in section 3.1.2 of [13].

Table 4.1 : Summary of picky jet $p_T$ corrections.

<table>
<thead>
<tr>
<th>Event type</th>
<th>quark flavor</th>
<th>Uncorrected $p_T$ rank</th>
<th>picky jet $p_T$ correction</th>
</tr>
</thead>
<tbody>
<tr>
<td>Any</td>
<td>B</td>
<td>Any</td>
<td>0.98</td>
</tr>
<tr>
<td>$t\bar{t}$</td>
<td>non-B</td>
<td>1 – 2</td>
<td>0.97</td>
</tr>
<tr>
<td>$t\bar{t}$</td>
<td>non-B</td>
<td>$\geq$ 3</td>
<td>0.98</td>
</tr>
<tr>
<td>$Z \rightarrow \ell^+\ell^-$</td>
<td>non-B</td>
<td>Any</td>
<td>0.99</td>
</tr>
</tbody>
</table>

Additionally, a picky jet CSV shape correction study was performed. Because the CSV AK5 jet shape and CSV picky jet shape differ, we cannot just apply the AK5 corrections. Our method adjusts the simulation of the CSV shape by altering the CSV discriminator values on a jet by jet basis. The result is a closure based method to determine an additional systematic uncertainty. The study is thoroughly
documented in Appendix M of [13].

4.6 B Flavor Tagging

“Quark flavor tagging” or simply just “flavor tagging” identifies the flavor of a quark, a jet arises from. This analysis relies on being able to detect the two b quarks in our signal. Hence, b flavor tagging (b-tagging) is therefore extremely important in this analysis. The CMS provides multiple b-tagging algorithms that make use of one or both of these parameters for b-tagging [84]. The two main quantities used for b-tagging in this analysis are track impact parameters (IP) and secondary vertices (b-decay vertices). This introduction briefly goes over IPs and secondary vertices in the context of b-tagging.

The track IP can be used to distinguish the decay products of a b hadron from prompt tracks (tracks originating from the primary vertex). The general track pre-election for b jet algorithms which use impact parameters is:

- Transverse momentum of at least 1 GeV/c. This reduces poorly constructed tracks.
- At least eight hits total and a good $\chi^2/N$ fit ($N =$ degrees of freedom).
- At least two hits in the pixel system. The inner layers provide most discriminating power.
- Must be $\Delta R < 0.05$ around to the jet axis (the vector from the primary vertex in the direction of the jet momenta).
- The transverse distance (of closest approach) to the primary vertex must be smaller than 0.2 cm.
- The longitudinal distance (of closest approach) to the primary vertex must be smaller than 17 cm.
• The distance (of closest approach) to the jet axis must be less that $700 \mu m$ (an attempt to eliminate pileup).

• The point of closest approach must be withing 5 cm of the primary vertex.

The track IP significance ($S_{IP}$) alone can be used discriminate between the decay products of b and non b jets. This is done by ranking all tracks compromising a jet in terms of decreasing $S_{IP}$ and selecting the n-th highest $S_{IP}$ as the discriminator. The value of this discriminator is used to determine if the jet is tagged as a b jet or not. The Track Counting (TC) algorithm uses the highest $S_{IP}$ as the discriminator, the Track Counting High Efficiency (TCHE) algorithm uses the second highest, and Track Counting High Purity (TCHP) uses the third highest. TCHE and TCHP are necessary, as the TC tends to bias the first track to a high positive $S_{IP}$, yet the probability of multiple high positive $S_{IP}$ is small for light-parton jets.

Additionally, there are b-tagging algorithms that depend on combinations of IPs. For instance, jet probability (JP) algorithms and jet b probability (JBP) algorithms. These are similar to TC based algorithms in that they attempt to provide discriminator between b jets and non b jets. The JP algorithm estimates the likelihood that all tracks in a jet came from a simple primary vertex. The jet B probability (JBP) algorithm is a modification of the JP algorithm, even further specified for b-tagging. In average there are four reconstructed particles from b decays. JBP accounts for this by giving more weight to tracks, up to four, with the highest $S_{IP}$.

The secondary vertex in this analysis refers to the decay vertex of the b quark. The b quark arises from the primary vertex and decays at the secondary vertex.
Identification of the secondary vertex is greatly based on the properties of the b quark decay. B quarks have a life time of $\tau_b \sim 1.6$ ps, travel a length $c\tau_b$ of $\sim 500 \mu$m, and produce on average 5 charged particles per decay [111]. This travel length is within the errors of the primary vertex. Hence, the secondary vertex is expected to be slightly displaced from the primary vertex. The secondary vertex and its kinematic variables can be used to discriminate between b and non b jets. However, identification of a secondary vertex is non-trivial. As opposed to primary vertex identification, a secondary vertex is strictly identified by means of examining already established jets. Finding a secondary vertex is generally more difficult than finding a primary vertex. Stricter selection criteria are used when selecting the relevant tracks:

- Tracks must be within a cone of $\Delta R = 0.3$ around the jet axis.
- The maximal distance between the track and the jet axis is 0.2 cm.

Additionally, tracks must pass the “high-purity” working point [112]. Track purity is a particular track quality selection criteria designed to reduce fake rate (fraction of reconstructed tracks which are fake). The track purity criteria incorporates the normalized $\chi^2$ of the track fit, the track length, and impact parameter information in each iteration of the actual track reconstruction. High purity is the most stringent track purity working point.

The secondary vertex finding procedure is done using the adaptive vertex fit algorithm [113]. This algorithm was previously discussed in more detail in Section 4.1. The adaptive vertex algorithm is an iterative algorithm with three basic sequential steps. First find a the vertex position using tracks; second measure the compatibility
of the vertex and the tracks; and third remove tracks that fail a compatibility cut. The process is repeated until new vertices are no longer produced. The first iteration of the algorithm simply uses the interaction region as a constraint to identify and remove the prompt tracks in the jet. All other iterations produce secondary (decay) vertex candidates. Furthermore, secondary vertex candidates have to pass the following criteria to enhance the chance of identifying only b jets:

- Secondary vertices must share less than 65% of their associated tracks with the primary vertex.
- The significance of the radial distance between the two vertices has to exceed $3\sigma$.
- Secondary vertex candidates with a radial distance of more than 2.5 cm with respect to the primary vertex, with masses compatible with the mass of K0 or exceeding 6.5 GeV/$c^2$ are rejected. This reduces the contamination by vertices corresponding to the interactions of particles with the detector material and by decays of long-lived mesons.
- The flight direction (vector from the primary vertex to the secondary vertex, also known as the direction of flight) of each secondary vertex candidate has to be within a cone of $\Delta R < 0.5$ around the jet direction.
- The maximal distance between the track and the jet axis is 0.2 cm.

Simple Secondary Vortex (SSV) algorithms make use of the secondary vertices and use the significance of the flight distance of as a discriminating variable. The efficiency of SSV algorithms are limited by the reconstitution efficiency of the secondary vertices, about 65% [114]. Analogous to the TC algorithms, SSV algorithms come in variations based on purity working points. The High Efficiency Simple Secondary Vertex (SSVHE) algorithm requires every vertex to have have at least two associated tracks. The High Purity Simple Secondary Vertex (SSVHP) algorithm requires every
vertex to have at least three associated tracks. SSV algorithms would be ideal for this analysis, if they were not limited by the reconstruction efficiency of the secondary vertex. However, there exists another class of b-tagging algorithm, that combines the benefits of SSV algorithms and the JP algorithms. For instance, the combined secondary vertex (CSV) algorithm used in this analysis.

4.6.1 B-Tagging AK5 Jets

For the T2BW signal we identify AK5 jets that are likely to arise from b quarks by means of the CSV algorithm [84]. The CSV algorithm involves the use of secondary vertices as well as track impact parameters. One of the main features of the CSV algorithm is its ability to provide discrimination even when no secondary vertex is found. Hence, our b-tagging efficiency is not constrained by the efficiency of secondary vertex reconstruction. The CSV algorithm provides a "no vertex" case, which reverts to track-based variables that are combined in a way similar to that of the Jet Probability algorithm [111]. Additionally, the CSV algorithm provides a "pseudo vertex" case: if IP significance > 2, the computation of a subset of secondary vertex based quantities can be done even without an actual vertex fit. The CSV algorithm uses the following variables to calculate the likelihood ratios:

- The vertex category (real, pseudo, or no vertex)
- The 2D flight distance significance
- The vertex mass
- The number of tracks at the vertex
- The ratio of the energy carried by tracks at the vertex with respect to all tracks in the jet
- The pseudo-rapidity of the tracks at the vertex with respect to the jet axis
- The 2D IP significance of the first track that raises the invariant mass above the charm threshold of 1.5 GeV when subsequently summing up tracks ordered by decreasing IP significance
- The number of tracks in the jet
- The 3D signed IP significances for each track in the jet

These likelihood ratios can be used as discriminators between b, c, and lighter quark jets. Additionally, we use the standard CSV algorithm track constraint, only use tracks within a cone of radius 0.5 in $\eta - \phi$ centered on the jet axis. Our implementation of the CSV algorithm for b-tagging yields a single discriminator value for each jet. This analysis applies the CSV algorithm with either the standard tight ($D > 0.898$) or the medium ($D > 0.679$) discriminator working points. The tight working point corresponds to a misidentification probability for light-parton jets of about 0.1%. The medium working point corresponds to a misidentification probability for light-parton jets of about 1%.

This analysis uses Monte Carlo-based background predictions to improve the statistical precision of merely data-based background prediction methods. The Monte Carlo method produces a synthetic sample referred to as a “simulated sample”. In order for simulated samples to be better representative of actual data, simulated-to-data scale factors (SF) need to be applied for various CMS reconstruction processes. For all jets, simulated-to-data b-tagging related scale factors are parametrized by jet
$p_T$ and flavor. For non-B jets, simulated-to-data b-tagging related scale factors are parametrized by jet $p_T$ and $\eta$. The b-tagging related scale factors are applied using the prescription in [115]. Each simulated event is weighted as a function of the tagging status of every jet matched to a gen-jet with $p_T > 10$ GeV. A “gen-jet” (short for generated) is a jet in the simulation produced directly from modeling the physics processes without detector modeling and reconstruction. This is opposed to a normal jet in simulation or data, which is derived from detector observables. This jet matching is done so the aggregate simulated b-tagging efficiencies match data b-tagging efficiencies. Because these scale factors have some topology dependence, additional measurements are required. A $t\bar{t}$ based measurement [116] is used for processes involving production of one or more top quarks. A $\mu$+jets measurement [117] is used for all other processes.

**4.6.2 Quark/Gluon Discrimination for AK5 Jets**

We use a CMS quark/gluon likelihood discriminator $L_q$ [118] to differentiate between jets originating from top quark decay products and gluon jets produced in QCD or Z events. The discriminator is binned by event pile-up activity. The inputs are the number of charged particles in each jet, the number of neutral particles in each jet, and the “$p_T$ spread” of the jet constituents. The $p_T$ spread is defined as:

$$\sqrt{\frac{\sum_i p_{T,i}^2}{\sum_i p_{T,i}}}.$$  \hspace{1cm} (4.19)
It was found that the quark/gluon likelihood discriminator distribution has up to \( \sim 20\% \) differences between simulation and data. Hence, we need to correct for this discrepancy. We use a method that maps the intrinsic value \( L_q \) to a corrected value \( f(L_q) \) for each jet. This method is further described in Section 4.6.4. This method has the advantage of guaranteeing the differential yield predicted by the simulation remains unchanged. This is in stark contrast event reweighting techniques known to be prone to yield distortions (when derived as the product of jet weights for multiple jets in the event).

4.6.3 B-Tagging Picky Jets

This analysis uses the CSV algorithm to b-tag picky jets for our T2tt signal. The methodology is the same one used for our b-tagging of the AK5 jets for the T2bw signal. However, we replace the CSV algorithm track constraint. The CSV algorithm track constraint is to only use tracks within a cone of radius 0.5 in \( \eta - \phi \) centered on the jet axis. Instead, we require that the tracks must be associated to PF particles within the jet. This needs to be done because the picky jets vary in size.

As with the AK5 b-tagging, the output of the CSV algorithm gives a discriminator value for every jet. However, instead of just placing some cuts on these discriminators to identify a b jet, we use the CSV discriminator values as input in more sophisticated T2tt search region selection techniques (Section 5).

Furthermore, as with the AK5 b-tagging, we require a simulated-to-data correction
scheme. Since, we don’t strictly use the discriminators in the same sense we do in the AK5 B-tagging, we cannot just copy that scheme. Instead, we use a more comprehensive scheme based on an overall shape correction (Appendix M in [13]).

4.6.4 Quark/Gluon Shape Corrections

We correct the distributions of jet shape variables for specific types of jets. This is done by replacing the jet shape quantity $x$ (for a given jet) by a transformed quantity $f(x)$. Here the function $f$ yields a distribution $f(x)$ that is in better agreement between data and simulation than $x$. Ideally, this would be achieved by defining $f(x)$ as a cumulative distribution function (cdf) mapping $x$ in simulation to $x$ in data, in a well-matched data control region. However, the data samples usually contain some contamination by a process not represented in the simulated sample. To counter this contamination we use background subtraction. Background subtraction is performed using the simulation of $x$ distributions for all the flavors of jets that are predicted to be present in the background events. With this taken into account a reasonable correction can be performed. Once this is taken into account, the correction can be performed. Simulation of a jet produces a predicted value of $x$. $f(x)$ is then defined as the abscissa $x'$ in the data distribution that has equal cumulative probability. An example of this method is given in Figure 4.6. Here we map a gluon jet distribution of arbitrary quantity $x$ in simulation (blue curve) to a distribution of $x$ in data (red curve).
It’s been shown [119] that for quark/gluon discrimination, quark jets are decently modeled by the PYTHIA simulation [120] of jet fragmentation. However, PYTHIA/data comparisons and PYTHIA/Herwig++ [121] generator comparisons indicate that there should be less separation between gluon and light quark distributions or b-jet and light quark distributions. This indicates a small level of mis-modeling of gluon and b-jets. Hence, we use the cdf mapping method to derive a correction function for the quark/gluon discriminator for gluon and bottom jets. We assume the light quark jet distributions are sufficiently modeled by simulation.

The mapping function is obtained for gluon jets using a $Z \rightarrow \ell^+\ell^-$ control region. This process has a relatively high gluon jet fraction that increases with jet multiplicity. The function is parametrized by $p_T$. The contributions from the other flavors of jets are subtracted using simulation.

The mapping function is obtained for b jets using a $\geq 3$ jet control region. The function is parametrized by $p_T$. The contributions from the other flavors of jets are subtracted using simulation. Additionally, the corrections for the quark/gluon likelihood discriminator ($L_q$) (Section 4.6.2) distributions for tight b-tagged jets are derived. These use a single lepton (muon) CR. More information on these corrections can be found in [13].
4.7 Top Reconstruction

We do “top-reconstruction” for the T2tt signal. This utilizes the fact that the T2tt signal is kinematically constrained by two on-shell top quarks. The analysis makes use of the “Comprehensively Optimized Resonance Reconstruction Algorithm” (CORRAL) framework [12] for top reconstruction. CORRAL is not a “top-tagger”, but rather a “top-reconstructor” for events with a high chance of containing top quarks. The overall goal of CORRAL is the simultaneous identification of two top quarks.

Additionally, CORRAL uses a custom b jet likelihood BDT and a top-jet likelihood BDT. The corresponding discriminator values are used as inputs for some CORRAL top reconstruction processes. The specifics of these BDT are not given in this analysis, but can be found in Section 3.2.1 of [12].

4.7.1 Jet Seed

Top reconstruction starts off by using two BDTs to define “thin-seed jets” and “fat-seed jets” within our collection of picky-jets. The notion of thin-seed jet corresponds to a collimated jet, while the notion of fat-seed jet corresponds to a uncollimated (spread-out) jet. These sets are not necessarily disjoint. This difference is highlighted in Figure 4.3). The event preselection is:

- Three picky jets with an invariant mass around 170 GeV
- Each jet has $p_T (> 20 \ GeV)$ and $|\eta| (< 2.4)$
- Only one of the picky jets is b-tagged
The input parameters for the thin jet-seed BDT and the fat jet-seed BDT are the same. They are categorized as follows:

1. **Basic Jet Variables:**
   - Jet $p_T$
   - Rank of Jet $p_T$. Rank 1 is the jet with the maximum $p_T$ value.
   - Jet $|\eta|$

2. **b quark:**
   - CSV b-tagging discriminator
   - Rank of the CSV b-tagging discriminator. Rank 1 is the jet with the maximum CSV b-tagging discriminator value.

3. **Top:**
   - Top jet likelihood

4. **Distance Variables:**
   - $|\eta(\text{jet}) - \langle \eta_{p_T} \rangle|$. $\langle \eta_{p_T} \rangle$ is the $p_T$ weighted average value of $\eta$ for all jets in the event.
   - Rank of $|\eta(\text{jet}) - \langle \eta_{p_T} \rangle|$. The most central ($|\eta(\text{jet}) - \langle \eta_{p_T} \rangle| \to 0$) jets are ordered first.
   - $\langle \mathcal{N}(\Delta R) \rangle_{p_T}$. The average normalized distance between the jet of interest and all others, weighted by the $p_T$ of the others. $\mathcal{N}$ is the Gaussian with $\mu = 0$ and $\sigma = 1$. This is an isolation measure for jets. It is motivated by the fact that jets from top decays are more likely to be at the center of activity.
   - $\langle \exp(-\Delta R) \rangle_{p_T}$. Similar to $\langle \mathcal{N}(\Delta R) \rangle_{p_T}$, but with exponential fall-off. Probes a longer range of activity.
   - Rank of $\langle \exp(-\Delta R) \rangle_{p_T}$. Jets with smaller values are ordered first.

### 4.7.2 Top Candidate

The list of “seed-jets” is ordered in decreasing BDT discriminator quality, with the thin-seed jets going before the fat-seed jets. For every seed-jet, a list of all possible
“top candidates” are formed compromising the seed-jet and two lower $p_T$ jets. For each top candidate, the list of jets are reordered such that the one most likely to be a $b$ jet is first ($j_1$) and the remaining two are ordered by decreasing $p_T$ ($j_2, j_3$). “$j$” just stands for jet.

Top candidates (cand) are then refined into “thin-top” (t1) and “fat-top” (t2) candidates using two additional BDT’s. The notion of thin-top corresponds to a collimated top decay, while the notion of fat-top corresponds to a not collimated (spread-out) decay. Thin-tops are formed from thin seeded top candidates passing the thin-top BDT. The input parameters for the thin-top BDT are:

1. **Jet Seed:**
   - Seed jet quality value (BDT discriminator).
   - Quality rank of the thin-seed jet.

2. **Mass:**
   - $m_{123}$: The 3-jet invariant mass from summing the 4-momentum of $j_1$, $j_2$, and $j_3$.
   - $m_{12}/m_{123}$. $m_{12}$ is the 2-jet invariant mass by summing the 4-momenta of jets $j_1$ and $j_2$.
   - $\tan^{-1}(m_{13}/m_{12})$
   - $\tan^{-1}(m_{12}/m_{23})$
   - Deviation from the expected W mass:
     \[ \delta m_W = (m_{23} - \hat{m}_W)/\hat{\sigma}_W, \]  
     where $\hat{m}_W$ is the parameterized mean of the reconstructed W mass and $\hat{\sigma}_W$ is the spread of the reconstructed W mass.

3. **$p_T$:**
   - $p_T(j_1 + j_2 + j_3)/H_T$: The candidate $p_T$ relative to the scalar sum $p_T$ of all the jets in the event.
   - Rank of $p_T(j_1 + j_2 + j_3)/H_T$. Higher $p_T$ top candidates are ordered first.
• $\sum \frac{p_T^2}{\sum p_T}$. The sum is taken over all three jets that form the top candidate.

4. **b quark:**

• b jet likelihoods for j1, j2, and j3. This is important as wrong combinations can lead to two b jet coming from a top, which shouldn’t happen.

5. **Distance:**

• Maximum $\Delta R$ between all pairs of jets in the top candidate.

• Maximum $|\eta(jet) - \langle \eta_{p_T} \rangle|$ over all jets of the top candidate. Calorimetry and pile-up related issues in the forward region can be a spurious source of jets. These jets can form top candidates with high values of this variable.

• Rank of $|\eta(j1 + j2 + j3) - \langle \eta_{p_T} \rangle|$. The most central top candidates are ordered first.

6. **Area:**

• $N_{\text{enclosed}}$: Number of non-candidate jets contained in the circle ($C_{\text{enclosed}}$) enclosing the top candidate jets. A random combination of jets can also have a quasi-random number of other jets inside the candidate. However, true top candidates are collimated enough that this shouldn’t happen.

• Rank of $C_{\text{enclosed}}$. Rank 1 is the top candidate with the smallest radius.

• $\alpha_o$: The orientation of the top candidate in the $\eta - \phi$ plane. First the smallest-area ellipse that contains all jets in the top candidate is found. Second, the angle ($\alpha_o$) that the major axis of the ellipse forms with the $\phi$ axis is found.

If a thin seeded top candidate does not pass the thin-top BDT, it might still pass the fat-top BDT (be a fat-top). Additionally, fat seeded top candidates are run through the fat-top BDT to see if that are fat-tops. The input parameters of the fat-top BDT are mostly a subset of the thin-top BDT input variables. This is because there is less kinematic distinction between low $p_T$ top decays and combinatorial background.

The input parameters for the fat-top BDT are:

1. **Mass:**
• $m_{123}$: The 3-jet invariant mass from summing the 4-momentum of $j_1$, $j_2$, and $j_3$.
• $m_{12}/m_{123}$. $m_{12}$ is the 2-jet invariant mass by summing the 4-momenta of jets $j_1$ and $j_2$.
• $m_{13}/m_{123}$. $m_{13}$ is the 2-jet invariant mass by summing the 4-momenta of jets $j_1$ and $j_3$.
• $\tan^{-1}(m_{13}/m_{12})$
• Rank of $\delta m_W$. Less deviant candidates are ordered first.
• Deviation from the expected top quark mass:
\[
\delta m_t = \frac{(m_{123} - \hat{m}_t)}{\hat{\sigma}_t},
\]  
where $\hat{m}_t$ is the parameterized mean of the reconstructed top quark mass and $\hat{\sigma}_t$ is the spread of the reconstructed top quark mass.
• Rank of $\delta m_t$. Less deviant candidates are ordered first.

2. b quark:
• b jet likelihoods for $j_1$.

3. Top:
• The product of top jet likelihoods for all the top candidate jets. This is of value because the fat top selection criteria allows lower $p_T$ jets which have a higher non-top contamination rate.

4. Distance:
• Maximum $\Delta R$ between all pairs of jets in the top candidate.

5. Area:
• $N_{\text{enclosed}}$: Number of non-candidate jets contained in the circle ($C_{\text{enclosed}}$) enclosing the top candidate jets.
• $NE_{\text{enclosed}}$: Number of non-candidate jets contained in the eclipse ($C_{\text{enclosed}}$) enclosing the top candidate jets. Lack of collimation for low $p_T$ top decays means that the $\eta - \phi$ spread of decay products can be more elliptical than circular.
4.7.3 Top Pairs

Once we have a list of thin-tops and fat-tops, they are sorted in order of decreasing BDT quality with thin top coming first. A “top pair candidate” is any two candidates (t1, t1) or (t1, t2) that have no jets in common. These top candidate pairs (i, j) are then fed into another BDT to be ranked based on the following criteria:

1. **Number of top pairs:**
   - Number of top pairs in the event. Used to parameterize other cuts.
   - Number of top pairs in which top candidate j is used as the second candidate.

2. **Top quality:**
   - Rank of the quality of the top pair. This is equivalent to the rank of the top pair in a list created sequentially by trying to pair up quality-ranked top candidates i and j > i.
   - The individual quality ranks for the two top candidates i and j.
   - Rank of the quality of the top pair, but only within the list created from the same seed jet.
   - Thin-top and fat-top discriminator values for top candidate j.

3. **Mass:**
   - $m_{23}/m_{123}$ for top candidate i.
   - $m_{123}$ for top candidate j.
   - $\tan^{-1}(m_{13}/m_{12})$ for top candidate j.
   - $\delta m_W^{(j)}$. The W mass deviation for top candidate j.
   - $\delta m_t^{(i)}$. The t mass deviation for top candidate i.
   - Rank of $|\delta m_W^{(i)} + \delta m_W^{(j)}|$. Top pairs with smaller deviations are ordered first.
   - Rank of $|\delta m_t^{(i)} + \delta m_t^{(j)}|$. Top pairs with smaller deviations are ordered first.
   - $|\delta m_W^{(i)}|$, $|\delta m_t^{(i)}|$, and $|\delta m_t^{(j)}|$ ranks of the top pair. Top pairs with smaller deviations are ordered first.

4. **b quark:**
- Jet b-likelihood rank for j1 for both top candidate i and top candidate j.
- CSV b-tagging discriminator value for j1 for both top candidate i and top candidate j.
- Minimum CSV b-tagging discriminator rank between j2 and j3 of top candidate i and j.
- Minimum b-likelihood rank between j2 and j3 of top candidate i and j.

5. Top:

- The rank of the top pair. Top pairs are ordered by the largest product of top jet likelihoods using all six jets.

6. Area:

- The sum of area of the ellipses that enclose jets from each top candidate. The overlap is not double counted. This sum should not be overly large for the correct top combination.
- The intersection area between the two circles that enclose jets from each top candidate. For the correct combination of tops and signals where the two top quarks are produced more or less independently, this should not be very large.

7. Energy and $p_T$:

- $|\vec{p}_T^i + \vec{p}_T^j|/H_T$. $\vec{p}_T^i$ corresponds to top candidate i.
- $|H_T^i + H_T^j|/H_T$. $H_T^i$ corresponds to top candidate i.
- Rank of $H_T^i + H_T^j$. Top pair with higher $H_T$ are ranked first.

The BDT discriminator is used to choose the “top pair” out of the lists of all top candidate pairs.

We use CORRAL because it represents a substantial gain over existing top identification algorithms for T2tt. This analysis uses a CORRAL working point that reconstructs top pairs in signal events with 32% efficiency at a “purity” of 63%. The purity here is defined as the probability to obtain the correct top pair combination out of all possible combinations. In contract, the standard HEP Top Tagger (HTT) [122]
only has an efficiency of 4% - 9% and purity of 70% - 80% per top candidate. Figure 4.7 gives the comparison between CORRAL, HTT, and the CMS Top Tagger (CTT) \cite{123}. The left plot shows the efficiency of reconstructing at least one top as a function of $p_T$. The right plot shows the efficiency of reconstructing both tops as a function of $p_T$. A more detailed look at the performance of CORRAL in regards to other top taggers, can be found in Appendix K of \cite{13}.

4.8 Lepton Control Regions

There are two lepton control regions used in this analysis: the electron control region and the muon control region. Both the electron and muon control regions are used for Top, to identify W, and Z($\nu\bar{\nu}$) + jets backgrounds. This section goes over the electron and muon identification for their respected control regions.

4.8.1 Electron Identification

The selection criteria for the electron control region can be categorized as follows:

1. Trigger Criteria:
   - The event is required to pass the HLT_Ele27_WP80_v* (single electron) trigger.
   - The event must have at least one offline electron with $p_T \geq 30\text{ GeV}$ and $|\eta| \leq 2.4$.

2. Electron Criteria:
   - All electrons in event must have $p_T \geq 15\text{ GeV}$ and $|\eta| \leq 2.4$.
   - All electrons in event must pass the tight ID working point \cite{88}.
   - Combined Particle Flow Isolation: PFIsol $\leq 0.1 \times p_T$ ($\leq 0.07 \times p_T$ for electrons with $p_T \leq 20\text{ GeV}$ in the endcap).
The simulated-to-data scale factors used for our electron control sample are identification (ID) efficiency, Reconstruction (Reco) efficiency, and single electron trigger (Trig) efficiency. ID and Reco are provided by the electron/photon Physics Object Group ($\text{e}/\gamma$ POG) [124]. ID corrects for inefficiencies in electrons passing the quality criteria in the control sample selection. Reco corrects for inefficiencies in reconstruction of electrons starting from superclusters (Section 3.2.2). Trig corrects for the inefficiency of triggering on a desired event (the inefficiency of the trigger criteria). Trig was obtained by means of a tag-and-probe method. Information on tag-and-probe can be found in Appendix A.1.

4.8.2 Muon Identification

The selection criteria for the muon control region can be categorized as follows:

1. Trigger Criteria:
   
   - The event is required to pass the HLT_IsoMu24_eta2p1_v* (single muon) trigger.
   - The event must have at least one offline muon with $p_T \geq 28 \text{ GeV}$ and $|\eta| \leq 2.1$.

2. Muon Criteria:
   
   - All muons in event must have $p_T \geq 15 \text{ GeV}$ and $|\eta| \leq 2.4$.
   - All muons in event must pass the tight ID working point [87].
   - Combined Particle Flow Isolation: $\text{PFIso} \leq 0.12 \times p_T$.

The simulated-to-data scale factors used for our muon control sample are identification (ID) efficiency, Reconstruction (Reco) efficiency, and Trigger (Trig) efficiency.
These are provided by the muon Physics Object Group (mPOG) [125]. ID corrects for inefficiencies in muons passing the quality criteria in the control sample selection. Reco corrects for inefficiencies in reconstruction of muons starting from a well-reconstructed tracker track. Trig corrects for the inefficiency of triggering on a desired event (the inefficiency of the trigger criteria).

Additionally, we compare Z data to Z simulation, to determine the muon momentum scale and angular resolution. This is done by equalizing the Z invariant mass peak location and width. Without this equalization, the mass peaks in simulation would be displaced from the data mass peaks. Without this equalization, the mass widths in the simulation would be thinner than the mass widths in the data. This was implemented through the MuScleFit measurement method [126].

Lastly, we apply two residual correction factors for our muon control region: the \( H_T \) dependence residual correction (Appendix D.1 in [13]) and the mis-measured tails residual correction (Appendix D.2 in [13]). The \( H_T \) dependence residual correction arises from the fact that the simulated-to-data scale (ID, Iso, Trig) factors are measured from \( Z \rightarrow l^+l^- \). This makes them insufficient to probe the \( H_T \), high jet multiplicity, and rare effects related to b jets, in our \( t\bar{t} \) samples. Essentially, the MC is too efficient when selecting leptons in high \( H_T \) events. This problem is best demonstrated by the excess of simulated events at very high \( H_T \) in the single muon control region. This correlates to an inefficiency in the data arising for muons near b-tagged jets. The mis-measured tails residual correction is needed to correct the
simulation yield of a small population with mis-reconstructed muons that is present in data but not found in simulation.

4.9 Lepton Veto BDTs

The main background for this analysis is semi-leptonic \( t\bar{t} \) decay with an unidentified charged lepton in the decay chain. Since we confine ourselves to hadronic decays, this means being able to reject events with leptons coming either directly from the W decay or via an intermediate \( \tau \) decay, while simultaneously accepting events with leptons coming from hadron decay or fake leptons. Maximizing the separation between rejection and acceptance is done using a Multivariate Analysis (MVA) approach. More specifically, we use Boosted Decision Tree (BDT) algorithms to simultaneously accept and reject events. A more detailed explanation of BDTs in general can be found in Appendix A.3.2. The BDTs in this section are referred to as lepton “vetos”.

4.9.1 Electron Veto

Electrons are pre-selected requiring \( p_T > 5 \) GeV, \( |\eta| < 2.4 \), and to be identified by the loose POG electron ID [88]. The BDT input variables are restricted to only those well predicted by the method used to estimate the W and \( t\bar{t} \) backgrounds (Section 6). We initially consider all variables in the muon veto, as well as those used in the eg-POG BDT. That number is then restricted to only those variables that can be well predicted by the embedding method (Section 9 in [13]). Additionally, a few variables
found to have negligible impact have been removed: the $\eta$ width, the $\phi$ width, R9
($E_{3x3}/E_{\text{supercluster}}$), H/E (related to HCAL tower energy over ECAL seed cluster
energy), and $E_{1x5}/E_{5x5}$. The $\eta$ and $\phi$ width correspond to the differences (in $\eta$ and
$\phi$) between the supercluster and position of inner track extrapolated from interaction
vertex. “NxM” refers to a specific array of ECAL crystals. The 19 variables used for
the electron veto are categorized as follows:

1. **Basic Kinematic Variables**:
   - $p_T$

2. **Impact Variables**:
   - $d_{xy}$
   - $d_z$
   - SIP 3D

3. **Isolation Variables**:
   - Relative PFIso on charged (CHS subtracted)
   - Relative PFIso on neutral

4. **Jet Variables**: lepton = l, nearest jet to lepton = nearj
   - $\Delta R(l, \text{nearj})$
   - $p_T^l / p_T^{\text{nearj}}$
   - CSV b-tagger discriminator value of nearj

5. **Electron POG Variables**: SC = supercluster, calo refers to the ECAL, and
   the default electron track reconstruction is GSF
   - Fraction of energy lost by bremsstrahlung radiation [127]
   - $\chi^2$ (of both KF track and GSF track)
   - Number of track layers with measurements
   - $\Delta \eta(\text{SC}, \text{track})$ at vertex
   - $\Delta \eta(\text{SC}, \text{track})$ on calo surface
   - $\Delta \phi(\text{SC}, \text{track})$ at vertex
• Shower shape $\sigma_{\eta \eta}$: Measures the shower width (in terms of the crystal spacing) in the traverse extension along $\eta$. This largely unaffected by showering of electrons and photons in the tracker material. Hence, it’s a key variable in electron and photon identification [128].

- $E(\text{SC})/p(\text{track at vertex})$
- $1/E(\text{SC}) - 1/p(\text{track at vertex})$
- (electron cluster energy)/(momentum of calo surface)
- $E_{\text{preshower}}(\text{SC})/E(\text{SC})$

Figure 4.8 (Left) gives the output of the electron veto, events as a function of BDT discriminator value. Figure 4.8 (Right) overlays the receiver operating characteristic (ROC) curve of the electron veto with various other electron discrimination work points. This plot shows the efficiency for background as a function of the efficiency for signal. In both these figures, “signal” actually refers to the targets of the electron veto, the semi-leptonic $t\bar{t}$ decay with electrons coming either directly from the W decay or via an intermediate $\tau$ decay, while “background” refers to electrons originating from hadron decays (still true electrons) or from electron mis-identification of the hadrons themselves (fake electrons). The working point is chosen to obtain a 98% W electron tagging efficiency, with a corresponding $e(\text{background})$ of 5%. The corresponding electron veto discriminator cut is -0.5.

4.9.2 Muon Veto

Muons are pre-selected requiring $p_T > 5$ GeV, $|\eta| < 2.4$, and to be identified by the loose POG muon ID [87]. The BDT input variables are restricted to only those well predicted by the embedding method (Section 9 in [13]). The 8 variables used for
the muon veto are categorized as follows:

1. **Basic Kinematic Variables:**
   - \( p_T \)

2. **Impact Variables:**
   - \( d_{xy} \)
   - \( d_z \)
   - SIP 3D

3. **Isolation Variables:**
   - Relative PFIso on charged (CHS subtracted)
   - Relative PFIso on neutral

4. **Jet Variables:** lepton = l, nearest jet to lepton = nearj
   - \( \Delta R(l, \text{nearj}) \)
   - \( \frac{p_T^l}{p_T^{\text{nearj}}} \)
   - CSV b-tagger discriminator value of nearj

Figure 4.9 (Left) gives the output of the muon veto: events as a function of discriminator value. Figure 4.9 (Right) overlays the ROC curve of the muon veto with various other muon discrimination work points. The naming conventions are analogous to Figure 4.8. The working point is chosen to obtain a 98% W muon tagging efficiency, with a corresponding e(background) of 5%. The corresponding muon veto discriminator cut is -0.5.

### 4.9.3 Tau Veto

The main background after application of the electron and muon veto’s is semileptonic \( t\bar{t} \) decay with hadronically decaying \( \tau \)’s. Ideally we want to veto against these
hadronically decaying $\tau$’s. After application of a baseline selection ($E_T > 175$ GeV, $>= 5$ jets with $p_T > 30$ GeV and $|\eta| < 2.4$) on simulated $t\bar{t}$ events, less than 25% of the surviving hadronic $\tau$ decay into multiple charged tracks. The remaining decay into a single charged particle. This allows us to consider single tracks as a primary variable for our $\tau$ veto. This is expected to increase our $\tau$ veto efficiency. However, it creates challenges in keeping the signal efficiency high.

We define a preliminary $\tau$ candidate as a charged particle flow candidate with $p_T > 5$ GeV, $|\eta| < 2.4$, and transverse mass ($m_T$) < 68 GeV. $m_T$ is constructed from the $E_T$ and visible $\tau$ decay products. The transverse mass is constructed with the following three points in mind:

- Background Constraint: $W$ is decaying into $\tau$ and $\nu$, so we can place a constraint on consistency of the transverse mass ($m_T$) with that expected from $W$.

- Fake $\tau$ candidates: No $W$ constraint on the $m_T$. For fake $\tau$ the transverse mass distribution is determined from the track $p_T$ and $E_T$ (which need not be correlated).

- $\gamma p_T$ Addition: It is common to have an associated $\gamma$. This improves the $m_T$ resolution for $W$ decays while arbitrarily increasing $m_T$ for fake $\tau$ candidates, further improving the separation of signal and background.

The transverse mass is:

$$m_T = \sqrt{2 * p_T(\text{track + nearest } \gamma) * E_T * (1 - \cos(\Delta \phi))},$$

where the $p_T$ of the highest nearby $\gamma$ is added to the track $p_T$, and the nearest $\gamma$ must have $p_T > 0.5$ and $\Delta R(\text{track, nearest } \gamma) < 0.2$. The cut-off of 68 GeV was chosen in order to keep the fake rate for low $E_T$ signals below 10%. Figure 4.10 (Left) compares
the $m_T$ distributions (normalized number of events as a function of $m_T$) between signal charged tracks (signal: MStop = 620, LMSP = 40) and hadronic decaying $\tau$’s from $t\bar{t}$ (background). It demonstrates that cut-off removes most $t\bar{t}$ while having a small impact on the signal.

$\tau$ candidates are identified with a BDT trained on $t\bar{t}$ simulation for at particular T2tt “signal points”. A signal point corresponds to a particular combination of stop mass ($m_{\tilde{t}}$), the intermediate chargino mass ($m_{\tilde{\chi}_\pm}$)(when applicable), and the neutralino ($m_{\tilde{\chi}_0^1}$). The candidates come from the pool of preliminary $\tau$ candidates. The simulation includes the baseline selection as well as another preselection used to account for the isolation environment and track $p_T$ spectra: $\geq 5$ jets ($p_T > 30$ GeV) and $E_T > 175$ GeV. Additionally, since no event-wide quantities are used in the BDT, the dependence on signal point is negligible. The 14 variables used for the $\tau$ veto are categorized as follows:

1. **Track properties:**
   - Track $p_T$: The transverse energy of nearby particles is expected to be higher around high $p_T$ tracks, so the $p_T$ of the track is used. The steeply falling number of track as $p_T$ increases allows one to loosen isolation cuts as the $\tau$ candidate increases in $p_T$.
   - Track $|\eta|$.
   - Track $\Delta Z$: This is the distance to the primary vertex. It is used to suppress fake $\tau$’s.

2. **Local activity variables:**
   - Charged isolation: The sum $p_T$ of charged particles within various cones ($\Delta R < 0.1, 0.2, 0.3, 0.4$).
   - Total isolation: The sum $p_T$ of all particles within various cones ($\Delta R < 0.1, 0.2, 0.3, 0.4$).
• $\Delta R$ to nearest track: The $\eta - \phi$ distance to the closest track with $p_T > 1$ GeV.

3. Jet Variables:

• $\Delta R$ to jet axis: If the jet containing the $\tau$ candidate is within acceptance ($p_T > 20$ GeV, $|\eta| < 2.4$), the distance between the jet axis and the candidate can differ for real $\tau$’s as compared to fakes.

• Jet CSV b-Tag discriminator: The CSV b-Tag discriminate of that jet is provided as input to the BDT, if the jet containing the $\tau$ candidate is within acceptance ($p_T > 20$ GeV, $|\eta| < 2.4$).

Figure 4.10 (Top Right) gives the output of the tau (MStop = 620, MLSP = 40): Normalized number of events as a function of maximum tau discriminator value. Figure 4.10 (Bottom) shows the ROC curve of the tau veto. The discriminator cut is chosen to keep the overall signal fake rate below 10%. We consider all $\tau$ candidates with a discriminator value greater than 0.77 as isolated. The result in a veto of approximately 62% of single lepton, hadronic $\tau$ events where the $\tau$ track is within the selection criteria. Additionally, 3.7% of events from a signal point are lost.
Figure 4.4: Picky Jets by profile. Top Left: sub jet s1 (higher $p_T$) and s2 (lower $p_T$). Top Right: 1D $p_T$ profile of top left. Bottom right: Gaussian kernel of top right indicating the jet should be split. Bottom left: An example of a jet scenario that shouldn’t split. [11]
Figure 4.5: The efficiency of jet reconstruction algorithms (before PU corrections). The left is the more collimated top quark in SUSY stop pair production. The right is the less collimated top quark in SUSY stop pair production. The SUSY stop pair production uses a stop mass of 600 GeV and LSP mass of 50 GeV. [12]

Figure 4.6: Cumulative distribution function (cdf) mapping method for correcting jet shapes. [13].
Figure 4.7: The Top tagging efficiencies for constructing at least one top (left) and constructing both tops (right) for CORRAL, CTT, and HTT. [12]
Figure 4.8: Electron Veto. Left: BDT discriminator for signal (leptons from $W$ decays in $t\bar{t}$, either coming directly from the $W$ or via an intermediate $\tau$ decay) and for background (true leptons coming from hadron decays or charged hadrons misidentified as leptons). Right: ROC curve for BDT compared with reference working points of other veto algorithms available in CMS. The working point of the BDT has been chosen to be 98% efficient on signal and 5% on background, corresponding to a discriminator cut of -0.5. [13]
Figure 4.9 : Muon Veto [13]. Left: BDT discriminator for signal (leptons from $W$ decays in $ttbar$, either coming directly from the $W$ or via an intermediate $\tau$ decay) and for background (true leptons coming from hadron decays or charged hadrons mis-identified as leptons). Right: ROC curve for BDT compared with reference working points of other veto algorithms available in CMS. The working point of the BDT has been chosen to be 98% efficient on signal and 5% on background, corresponding to a discriminator cut of -0.5. [13]
Figure 4.10: Tau Veto. Top Left: The $m_T$ distribution constructed for using all charged tracks in signal and for hadronic $\tau$’s in $t\bar{t}$ events. The veto is set for $m_T < 68\text{GeV}$. Top Right: The maximum $\tau$ veto discriminator value for all $\tau$ candidates in an event. Bottom: The event level $t\bar{t}$ veto efficiency as a function of signal efficiency for stop and LSP masses of 620 GeV and 40 GeV, respectively. [13]
Chapter 5

The Signal

This section details the methods used to identify our signals (T2bW and T2tt) and the methodology used to pick our search regions. The overall goal is to find a method to separate the signal from the non-signal. We use a set of preselection and event selection Boosted Decision Trees (BDT) to accomplish this identification.

5.1 Event Selection

Our signal BDT involves a twofold selection process. First a loose BDT preselection and then a tight BDT discriminator selection. The CMS dataset we search for our SUSY signals (Table B.2) was obtained with the “online preselection trigger” (Appendix B). Additionally, we use “offline preselection” requirements (Appendix B) on this dataset.

5.1.1 BDT Preselection

Our BDTs requires additional preselection criteria. However, they are different between T2bW and T2tt, the main difference being the T2tt signal uses the CORRAL top tagger. More information on CORRAL can be found in subsection 4.7. Both BDT preselections include all the lepton vetoes as well as kinematic cuts above the
corresponding trigger turn-on. Unless otherwise noted, all cuts are placed on standard AK5 PF jets.

The T2bW search region BDT preselection cuts are defined as:

- \( \geq 5 \) jets \( (p_T > 30 \text{ GeV}, |\eta| < 2.4) \)
- \( \geq 1 \) Tight CSV b-Tagged jet \( (p_T > 30 \text{ GeV}, |\eta| < 2.4) \)

The T2tt search region BDT preselection cuts are defined as:

- \( \geq 1 \) CORRAL reconstructed top pair
- \( \geq 1 \) Tight b-Tagged picky jet \( (p_T > 30 \text{ GeV}, |\eta| < 2.4) \)

### 5.1.2 BDT Input Variables

After BDT preselection is BDT constructing. The first part in BDT construction is the selection of BDT input variables. This is done in two basic steps. First, a selection of \( \sim 100 \) variables based on efficiency curves of signal against \( t\bar{t} \) and \( Z \rightarrow \nu\bar{\nu} \) is selected. Second, a subset of these variables is chosen by looking at the point by point maximum significance for different combinations of variables. All jets are standard AK5 jets and are required to be in acceptance \( (p_T > 30 \text{ GeV}, |\eta| < 2.4) \), unless noted otherwise.

The 14 T2bW BDT input variables can be categorized as follows:

1. **Basic**
   - \( E_T \)
   - Number of AK5 jets
• $\Delta \phi(E_T, \text{jet}3)$
• Number of medium CSV b-Tagged jets

2. Quark likelihood: These come from the CMS quark/gluon likelihood calculation [118]. $L_q(\text{jet}_i)$ is the quark likelihood for jet$_i$.

• Leading quark likelihood of all jets in the event: Max[$L_q(\text{jet}_i)$], i is over all jets.
• Second leading quark likelihood of all jets in the event: 2nd-Max[$L_q(\text{jet}_i)$], i is over all jets.
• Product of all quark likelihoods:

$$\prod_{i=1}^{n} L_q(\text{jet}_i). \quad (5.1)$$


• $H_T$(Along)/$H_T$(Away):

$$H_T(\text{Along}) = \sum p_T(\text{jet}_i), \ i \in |\Delta \phi(E_T, \text{jet}_i)| < \pi/2, \quad (5.2)$$

$$H_T(\text{Away}) = \sum p_T(\text{jet}_j), \ j \in |\Delta \phi(E_T, \text{jet}_j)| > \pi/2 \quad (5.3)$$

• RMS[$p_T(\text{jet}_i)$], i is over all jets.
• RMS[$\Delta \phi(E_T, \text{jet}_i)$], i is over all jets.
• Invariant mass of the highest $p_T$ pair of all jets uniquely paired to their nearest neighbor. The $p_T$ of jet pairs is designed to single out possible W decay to two quark jet pairs

4. b jet: RMS = root mean square.

• b jet invariant mass (two leading CSV medium b-tagged jets)
• b jet transverse mass:

$$m_T(b) = \sqrt{2 * p_T(\text{jet}) * E_T(1 - \cos(\Delta \phi))} \quad (5.4)$$

, jet is the nearest CSV medium b-tagged jet to $E_T$
• RMS[|$\Delta \eta$(b jet, jet$_i$)], b jet is leading CSV medium tagged b jet, i is over all jets but the b jet.

The 24 T2tt BDT input variables can be categorized as follows:
1. Basic

- $E_T$
- Number of AK5 jets

2. Quark likelihood: This comes from the CMS quark/gluon likelihood calculation \[118\]. $L_q(jet_i)$ is the quark likelihood for $jet_i$.

- Product of all quark likelihoods:

\[
\prod_{i=1}^n L_q(jet_i).
\] (5.5)

3. b jet

- Picky b jet transverse mass:

\[
m_T(b) = \sqrt{2 \times p_T(jet) \times E_T(1 - \cos(\Delta \phi))},
\] (5.6)

where jet is the nearest picky medium b-tagged jet to $E_T$.

4. Activity variables: These two variables are computed with a kernel density estimate \[129, 130\] of the $p_T$ distribution of each picky jet in the event. From the sum of the kernel densities, these two quantities are calculated.

- Peak $|\eta|$ location activity in the event
- Distance in $|\eta|$ between two peaks in event activity

5. CORRAL variables: These refer to CORRAL top pair reconstruction \[12\]. All top-tag related quantities refer to the leading quality top pair candidate. “Top1” (“Top2”) refers to the top candidate in the leading pair with the highest (lowest) top discriminator value.

- Number of working point 98 (WP98) CORRAL top pairs \[12\]
- top pair mass
- top2($p_T$)/top1($p_T$)
- Ellipse size: Total area covered by the top pair in $\eta - \phi$ space. The area is measured as the ellipse containing the top candidate.
- $\Delta \phi(E_T, \text{top2})$
- top2 thin quality
- top2 fat quality
- $p_T$ of all 6 top jets that make up the top pair
• Min[ΔR(jet_i, jet_k)], i, k ∈ top1
• Min[ΔR(jet_i, jet_k)], i, k ∈ top2
• Min[Δφ(∑E_T, jet_i)], i ∈ top2
• Maximum picky b jet discriminator value in each of the two top candidates of the top pair.

5.1.3 BDT Training

The SUSY models that predict our signals contain unknown mass parameters. In T2bW we have three unknown masses: intermediate chargino (m_χ±), stop (m_˜t), and the neutralino (m_˜χ^0_1). In T2tt we have two unknown masses: stop (m_˜t) and the neutralino (m_˜χ^0_1). Hence, each signal is really a “signal space”, best parametrized by these unknown masses. One particular set of mass values is referred to as a “signal point” within the signal space. Additionally, we focus on a reduced T2bW signal space in which:

\[ m_χ^± = x \cdot m_˜t + (1 - x) \cdot m_˜χ^0_1, \]  \hspace{1cm} (5.7)

where we consider three fractions for x: 0.25, 0.50, and 0.75. x is referred to as the mass fraction or mass split parameter.

There is a continuum of signal points, hence an infinite set of possible signal points we could use to train our BDTs on, “training points”. BDT training studies, with different tuning parameters, found that the BDTs are largely insensitive to fine tunings. Furthermore, BDTs trained with minor difference in training signal points are found to only produce minor differences in the discrimination over a wide range of signal points. Hence, we select a few training signal points that span a wide range
(Table 5.1) of signal space. We select 30 T2bW signal training points and 18 T2tt signal training points. BDTs are trained individually for each of these signal training points. However, since the T2bW signal is lacking in statistics, the T2bW input signal points are actually the weighted sum of four neighboring points in signal space.

Each BDT is trained against simulated background samples (Tables B.16, B.14) and the appropriate simulated signal samples (Tables B.9, B.9). In order to combat over-fitting, the simulated signal and the simulated background samples were cut during training. These samples are run through our offline preselection requirements (Appendix B) and the BDT preselection criteria, before being used.

5.2 Search Region Selection

A BDT trained on a signal training point is known to have discriminating power for neighboring points too. Hence, every BDT (and training point) corresponds to a “search region”. A search region is a region of signal points where the BDT has discriminating power.

We have to eventually calculate the backgrounds in each search region we use. This alone makes it extremely impracticable to use all 30 T2bW search regions and all 18 T2tt search regions (Table 5.1). Additionally, the actual range of the certain search regions may be redundant. The goal is then to pick a few specific search regions that maximize the number of expected excluded signal points in the signal space. In
order to accomplish this we define the significance as:

\[
N_S \frac{N_S}{\sqrt{N_S + N_B + (E_S * N_S)^2 + (E_B * N_B)^2 + 3/2}}.
\] (5.8)

Here

- \(N_S\) is the number of signal points.
- \(N_B\) is the number of background points.
- \(E_B\) is the estimated relative systematic uncertainty on the background prediction. All backgrounds except for \(t\bar{t}\) are assumed to have an uncertainty of 100%.
- \(E_S\) is the estimated relative signal uncertainty. The signal uncertainties are taken from the estimated uncertainties in some example search regions \(O(15\%)\).

For both T2bW and T2tt, we use the following algorithm:

1. Collect the BDTs all trained on different signal points: 30 for T2bW and 18 for T2tt.
2. For each BDT, we use 200 different discriminator cuts to collect the 200 unique search regions. We use a discriminator granularity of 0.1.
3. Find the combination of \(N\) (\(N\) starts at 0) search regions out of all (6000 for T2bW, 3600 for T2tt) possible search regions that as a group exclude the most number of signal level at a 95% confidence level. If multiple search regions exclude the same number of points, the one with the most aggregate significance is chosen.
4. Repeat step (3) with \(N = N + 1\) until the gain of adding a new search region is negligible.
We choose 5 final T2bW search regions and 4 final T2tt search regions. Search region are parametrized by mass variables $m_{\tilde{t}}$, $m_{\tilde{\chi}_1^0}$, and $x$ (T2bW only). The name of these regions, training points, and discriminator cuts, are shown in Table 5.2.

We also quantify how important each variable is in calculation of the search region (Table 5.4 and 5.4). This is done through calculation of a “variable ranking” using “The Toolkit for Multivariate Data Analysis with ROOT” (TMVA) software package [131]. Each branch (corresponding to a variable) in each tree get a weight dependent on how many events entered that specific branch. The weights for each tree are then normalized so the sum of all weights equals 100%. The aggregate BDT variable weights are computed by adding up all of the individual tree weights for each variable. These weights are further modulated by the trees overall weight in the BDT. These weights are finally normalized to add up 100%.
<table>
<thead>
<tr>
<th>$M(\tilde{t})$</th>
<th>$M(\tilde{\chi}_0^0)$</th>
<th>$x$</th>
<th>$M(\tilde{t})$</th>
<th>$M(\tilde{\chi}_0^0)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>500, 525</td>
<td>25, 50</td>
<td>0.25</td>
<td>300</td>
<td>25</td>
</tr>
<tr>
<td>700, 725</td>
<td>25, 50</td>
<td>0.25</td>
<td>425</td>
<td>25</td>
</tr>
<tr>
<td>400, 425</td>
<td>75, 100</td>
<td>0.25</td>
<td>550</td>
<td>25</td>
</tr>
<tr>
<td>600, 625</td>
<td>75, 100</td>
<td>0.25</td>
<td>650</td>
<td>25</td>
</tr>
<tr>
<td>450, 475</td>
<td>75, 100</td>
<td>0.25</td>
<td>800</td>
<td>25</td>
</tr>
<tr>
<td>750, 775</td>
<td>75, 100</td>
<td>0.25</td>
<td>300</td>
<td>75</td>
</tr>
<tr>
<td>300, 325</td>
<td>125, 150</td>
<td>0.25</td>
<td>425</td>
<td>75</td>
</tr>
<tr>
<td>400, 425</td>
<td>175, 200</td>
<td>0.25</td>
<td>350</td>
<td>125</td>
</tr>
<tr>
<td>550, 575</td>
<td>175, 200</td>
<td>0.25</td>
<td>450</td>
<td>125</td>
</tr>
<tr>
<td>700, 725</td>
<td>175, 200</td>
<td>0.25</td>
<td>575</td>
<td>125</td>
</tr>
<tr>
<td>650, 675</td>
<td>275, 300</td>
<td>0.25</td>
<td>650</td>
<td>125</td>
</tr>
<tr>
<td>350, 375</td>
<td>25, 50</td>
<td>0.50</td>
<td>700</td>
<td>125</td>
</tr>
<tr>
<td>600, 625</td>
<td>25, 50</td>
<td>0.50</td>
<td>450</td>
<td>250</td>
</tr>
<tr>
<td>800, 825</td>
<td>25, 50</td>
<td>0.50</td>
<td>550</td>
<td>250</td>
</tr>
<tr>
<td>500, 525</td>
<td>75, 100</td>
<td>0.50</td>
<td>675</td>
<td>250</td>
</tr>
<tr>
<td>700, 725</td>
<td>75, 100</td>
<td>0.50</td>
<td>800</td>
<td>250</td>
</tr>
<tr>
<td>550, 575</td>
<td>125, 150</td>
<td>0.50</td>
<td>550</td>
<td>350</td>
</tr>
<tr>
<td>400, 425</td>
<td>175, 200</td>
<td>0.50</td>
<td>550</td>
<td>350</td>
</tr>
<tr>
<td>700, 725</td>
<td>175, 200</td>
<td>0.50</td>
<td>550</td>
<td>350</td>
</tr>
<tr>
<td>450, 475</td>
<td>225, 250</td>
<td>0.50</td>
<td></td>
<td></td>
</tr>
<tr>
<td>600, 625</td>
<td>225, 250</td>
<td>0.50</td>
<td></td>
<td></td>
</tr>
<tr>
<td>750, 775</td>
<td>225, 250</td>
<td>0.50</td>
<td></td>
<td></td>
</tr>
<tr>
<td>400, 425</td>
<td>25, 50</td>
<td>0.75</td>
<td></td>
<td></td>
</tr>
<tr>
<td>550, 575</td>
<td>25, 50</td>
<td>0.75</td>
<td></td>
<td></td>
</tr>
<tr>
<td>800, 825</td>
<td>25, 50</td>
<td>0.75</td>
<td></td>
<td></td>
</tr>
<tr>
<td>350, 375</td>
<td>75, 100</td>
<td>0.75</td>
<td></td>
<td></td>
</tr>
<tr>
<td>650, 675</td>
<td>75, 100</td>
<td>0.75</td>
<td></td>
<td></td>
</tr>
<tr>
<td>500, 525</td>
<td>125, 150</td>
<td>0.75</td>
<td></td>
<td></td>
</tr>
<tr>
<td>500, 525</td>
<td>175, 200</td>
<td>0.75</td>
<td></td>
<td></td>
</tr>
<tr>
<td>650, 675</td>
<td>225, 250</td>
<td>0.75</td>
<td></td>
<td></td>
</tr>
<tr>
<td>450, 475</td>
<td>225, 250</td>
<td>0.75</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 5.1: List of input signal points used for training the search region boosted decision trees
<table>
<thead>
<tr>
<th>Search Region Name</th>
<th>Abbreviation</th>
<th>Training Signal Point</th>
<th>Discriminator Cut</th>
<th>Signal efficiency [%]</th>
</tr>
</thead>
<tbody>
<tr>
<td>T2bW Low X</td>
<td>T2bW_LX</td>
<td>$M(\tilde{t}) = 550, 555$ GeV, $M(\tilde{\chi}_0^1) = 145, 200$ GeV, $x = 0.25$</td>
<td>$&gt; 0.94$</td>
<td>25</td>
</tr>
<tr>
<td>T2bW Medium X, High Mass</td>
<td>T2bW_MXHM</td>
<td>$M(\tilde{t}) = 550, 555$ GeV, $M(\tilde{\chi}_0^1) = 125, 150$ GeV, $x = 0.50$</td>
<td>$&gt; 0.92$</td>
<td>14</td>
</tr>
<tr>
<td>T2bW High X, High Mass</td>
<td>T2bW_HXHM</td>
<td>$M(\tilde{t}) = 400, 425$ GeV, $M(\tilde{\chi}_0^1) = 25, 50$ GeV, $x = 0.75$</td>
<td>$&gt; 0.82$</td>
<td>10</td>
</tr>
<tr>
<td>T2bW Very High Mass</td>
<td>T2bW_VHM</td>
<td>$M(\tilde{t}) = 550, 555$ GeV, $M(\tilde{\chi}_0^1) = 25, 50$ GeV, $x = 0.75$</td>
<td>$&gt; 0.93$</td>
<td>12</td>
</tr>
<tr>
<td>T2tt Low Mass</td>
<td>T2tt_LM</td>
<td>$M(\tilde{t}) = 300$ GeV, $M(\tilde{\chi}_0^1) = 25$ GeV</td>
<td>$&gt; 0.79$</td>
<td>8</td>
</tr>
<tr>
<td>T2tt Medium Mass</td>
<td>T2tt_MM</td>
<td>$M(\tilde{t}) = 425$ GeV, $M(\tilde{\chi}_0^1) = 75$ GeV</td>
<td>$&gt; 0.83$</td>
<td>16</td>
</tr>
<tr>
<td>T2tt High Mass</td>
<td>T2tt_HM</td>
<td>$M(\tilde{t}) = 550$ GeV, $M(\tilde{\chi}_0^1) = 25$ GeV</td>
<td>$&gt; 0.92$</td>
<td>25</td>
</tr>
<tr>
<td>T2tt Very High Mass</td>
<td>T2tt_VHM</td>
<td>$M(\tilde{t}) = 675$ GeV, $M(\tilde{\chi}_0^1) = 250$ GeV</td>
<td>$&gt; 0.95$</td>
<td>19</td>
</tr>
</tbody>
</table>

Table 5.2 : Search Region Definitions

<table>
<thead>
<tr>
<th>Variable</th>
<th>T2bW_LX</th>
<th>T2bW_LM</th>
<th>T2bW_MXHM</th>
<th>T2bW_HXHM</th>
<th>T2bW_VHM</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\Delta \phi(\not{E}_T, \text{jet}3)$</td>
<td>8.8%</td>
<td>9.8%</td>
<td>8.8%</td>
<td>9.6%</td>
<td>7.7%</td>
</tr>
<tr>
<td>b jet invariant mass</td>
<td>4.6%</td>
<td>6.0%</td>
<td>5.7%</td>
<td>5.2%</td>
<td>5.1%</td>
</tr>
<tr>
<td>Product of all quark likelihoods</td>
<td>5.6%</td>
<td>5.7%</td>
<td>5.9%</td>
<td>5.6%</td>
<td>6.0%</td>
</tr>
<tr>
<td>Leading W candidate invariant mass</td>
<td>6.3%</td>
<td>4.3%</td>
<td>5.5%</td>
<td>4.5%</td>
<td>6.1%</td>
</tr>
<tr>
<td>$H_T(\text{Along})/H_T(\text{Away})$</td>
<td>8.0%</td>
<td>7.7%</td>
<td>7.3%</td>
<td>7.2%</td>
<td>7.8%</td>
</tr>
<tr>
<td>Leading quark likelihood</td>
<td>7.1%</td>
<td>8.6%</td>
<td>6.5%</td>
<td>7.4%</td>
<td>6.1%</td>
</tr>
<tr>
<td>$\not{E}_T$</td>
<td>7.1%</td>
<td>6.7%</td>
<td>7.5%</td>
<td>8.2%</td>
<td>8.5%</td>
</tr>
<tr>
<td>b jet transverse mass</td>
<td>8.5%</td>
<td>7.5%</td>
<td>7.8%</td>
<td>8.8%</td>
<td>8.3%</td>
</tr>
<tr>
<td>Number of AK5 jets</td>
<td>6.8%</td>
<td>7.3%</td>
<td>8.8%</td>
<td>7.3%</td>
<td>10.6%</td>
</tr>
<tr>
<td>Number of medium CSV b-Tagged jets</td>
<td>6.0%</td>
<td>4.9%</td>
<td>5.6%</td>
<td>6.3%</td>
<td>5.3%</td>
</tr>
<tr>
<td>RMS[$p_T(\text{jet},)$]</td>
<td>7.1%</td>
<td>5.4%</td>
<td>6.9%</td>
<td>5.4%</td>
<td>5.7%</td>
</tr>
<tr>
<td>RMS[$\Delta \eta(\text{b jet, jet})$]</td>
<td>7.9%</td>
<td>9.2%</td>
<td>8.7%</td>
<td>8.7%</td>
<td>7.6%</td>
</tr>
<tr>
<td>RMS[$\Delta \phi(\not{E}_T, \text{jet})$]</td>
<td>7.3%</td>
<td>8.1%</td>
<td>7.7%</td>
<td>7.4%</td>
<td>7.1%</td>
</tr>
<tr>
<td>Second leading quark likelihood</td>
<td>8.8%</td>
<td>9.0%</td>
<td>7.3%</td>
<td>8.4%</td>
<td>8.1%</td>
</tr>
</tbody>
</table>

Table 5.3 : T2bW variable rank
<table>
<thead>
<tr>
<th>Variable</th>
<th>T2tt_LM</th>
<th>T2tt_MM</th>
<th>T2tt_HM</th>
<th>T2tt_VHM</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\Delta \phi(\eta, \text{top2})$</td>
<td>6.4%</td>
<td>5.3%</td>
<td>5.4%</td>
<td>5.0%</td>
</tr>
<tr>
<td>Distance in $</td>
<td>\eta</td>
<td>$ between two peaks</td>
<td>5.2%</td>
<td>4.5%</td>
</tr>
<tr>
<td>Peak $</td>
<td>\eta</td>
<td>$ location</td>
<td>6.0%</td>
<td>4.3%</td>
</tr>
<tr>
<td>Product of all quark likelihoods</td>
<td>5.3%</td>
<td>3.6%</td>
<td>3.7%</td>
<td>3.8%</td>
</tr>
<tr>
<td>$E_T^\gamma$</td>
<td>3.1%</td>
<td>6.1%</td>
<td>5.9%</td>
<td>5.7%</td>
</tr>
<tr>
<td>Min[$\Delta \phi(\eta, \text{jet}_i)$], $i \in \text{top2}$</td>
<td>4.5%</td>
<td>4.8%</td>
<td>4.4%</td>
<td>4.6%</td>
</tr>
<tr>
<td>Min[$\Delta R(\text{jet}_i, \text{jet}_k)$], $i, k \in \text{top1}$</td>
<td>5.6%</td>
<td>4.7%</td>
<td>4.2%</td>
<td>4.1%</td>
</tr>
<tr>
<td>Min[$\Delta R(\text{jet}_i, \text{jet}_k)$], $i, k \in \text{top2}$</td>
<td>5.4%</td>
<td>4.2%</td>
<td>4.3%</td>
<td>4.5%</td>
</tr>
<tr>
<td>Picky b jet transverse mass</td>
<td>4.6%</td>
<td>5.3%</td>
<td>5.0%</td>
<td>4.9%</td>
</tr>
<tr>
<td>Number of AK5 jets</td>
<td>4.7%</td>
<td>6.4%</td>
<td>7.5%</td>
<td>7.7%</td>
</tr>
<tr>
<td>Number of WP98 CORRAL top pairs</td>
<td>4.4%</td>
<td>4.4%</td>
<td>4.5%</td>
<td>4.2%</td>
</tr>
<tr>
<td>Top1, jet1 $p_T$</td>
<td>3.1%</td>
<td>3.5%</td>
<td>4.3%</td>
<td>3.8%</td>
</tr>
<tr>
<td>Top1, jet2 $p_T$</td>
<td>3.3%</td>
<td>3.2%</td>
<td>2.4%</td>
<td>3.0%</td>
</tr>
<tr>
<td>Top1, jet3 $p_T$</td>
<td>3.2%</td>
<td>3.3%</td>
<td>4.0%</td>
<td>3.9%</td>
</tr>
<tr>
<td>Maximum picky b jet disc in top1</td>
<td>3.0%</td>
<td>3.0%</td>
<td>3.1%</td>
<td>3.0%</td>
</tr>
<tr>
<td>Total area covered by the top pair in $\eta - \phi$ space</td>
<td>4.4%</td>
<td>4.1%</td>
<td>2.9%</td>
<td>3.5%</td>
</tr>
<tr>
<td>Top pair mass</td>
<td>3.3%</td>
<td>3.1%</td>
<td>3.0%</td>
<td>2.9%</td>
</tr>
<tr>
<td>Top2 fat quality</td>
<td>5.4%</td>
<td>5.0%</td>
<td>3.4%</td>
<td>3.4%</td>
</tr>
<tr>
<td>Top2, jet1 $p_T$</td>
<td>2.9%</td>
<td>4.3%</td>
<td>5.9%</td>
<td>6.7%</td>
</tr>
<tr>
<td>Top2, jet2 $p_T$</td>
<td>2.6%</td>
<td>2.6%</td>
<td>3.0%</td>
<td>2.8%</td>
</tr>
<tr>
<td>Top2, jet3 $p_T$</td>
<td>2.4%</td>
<td>2.9%</td>
<td>2.6%</td>
<td>2.9%</td>
</tr>
<tr>
<td>Maximum picky b jet disc in top2</td>
<td>2.7%</td>
<td>3.2%</td>
<td>3.4%</td>
<td>3.3%</td>
</tr>
<tr>
<td>Top2($p_T$)/Top1($p_T$)</td>
<td>3.6%</td>
<td>3.4%</td>
<td>4.1%</td>
<td>3.4%</td>
</tr>
<tr>
<td>Top2 thin quality</td>
<td>4.9%</td>
<td>4.8%</td>
<td>4.6%</td>
<td>4.6%</td>
</tr>
</tbody>
</table>

Table 5.4: T2tt variable rank
Chapter 6

Top, W, and Z($\rightarrow \nu \bar{\nu}$) + jets backgrounds

We need to be able to predict the background we would expect in our search regions. For our top and electroweak backgrounds we use “MC sample reweighting”. The MC sample reweighting methodology is described in Section A.2. The general idea for MC sample reweighting of non-QCD backgrounds is to use events where the W/Z decay into visible leptons to arrive at set of correction factors for the simulation. These factors are then applied verbatim to simulated events where the W/Z decay hadronically, W/Z decay into non-reconstructed leptons, or the Z decay invisibly.

MC sample reweighting involves defining particular process specific control regions (CR) where we believe the simulation should agree with data. We use three process specific control regions for our MC reweighting: $t\bar{t}$, W, and Z. The lack of a top control region is discussed in Section 6.5.3. Additionally, We use a multi-smearing (Section A.2.2) of 50 correction samplings for all non-QCD processes.

For simulation, the process specific control regions are derived from the simulated $t\bar{t}$ dataset (Table B.14), the simulated W dataset (Table B.11), and the simulated Z dataset (Table B.10). In data, the process specific control regions are derived from single leptonic control region. The leptonic control region is based on the single electron dataset (Table B.3) and single muon datasets (Table B.4). This is discussed
in section 6.1. The MC reweighting technique ultimately provides scale factors sets 
(SFS) that correspond to particular classes of physics corrections. These SFS are 
derived through the $t\bar{t}$, W, and Z control regions and discussed in Section 6.2.

### 6.1 Leptonic Control Regions

The single lepton control region is data taken on a single electron or single muon 
trigger. Events must have at least one tight-IDed lepton offline. To standardize our 
selection, we use physics object “standard” definitions provided by the POGs. This 
way we can use standard correction factors given by the electron and muon Physics 
Object Groups (POGs). The standard criteria for various physics objects is as follows:

- **Electron:**
  - $p_T \geq 15 \text{ GeV}$, $|\eta| \leq 2.4$
  - Pass tight ID working point [38]
  - Combined PFIso $\leq 0.1 \times p_T$
  - Combined PFIso $\leq 0.7 \times p_T$, for $p_T \leq 20 \text{ GeV}$ electrons in the ECAL endcap

- **Muon:**
  - $p_T \geq 15 \text{ GeV}$, $|\eta| \leq 2.4$
  - Pass tight ID working point [40]
  - Combined PFIso $\leq 0.12 \times p_T$

- **Jet:**
  - $p_T \geq 30 \text{ GeV}$
  - Comprised of PF candidates
  - Clustered using anti-kT jet algorithm, size parameter 0.5
  - Unless otherwise noted, all jets disambiguated of leptons before clustering: 
    remove all medium-IDed electrons ($p_T \geq 15 \text{ GeV}$) and tight-IDed muons 
    ($p_T \geq 10 \text{ GeV}$) from the input list of PF candidates.
• **b-tag:**
  - \( p_T \geq 30 \text{ GeV}, |\eta| \leq 2.4 \)
  - CSV b tagging algorithm discriminator value is \( > 0.898 \) for a “tight” (bT) working point and \( > 0.679 \) for a “medium” (bM) working point.
  - b-tagging is disambiguated of leptons by using only the jet constituents as input to various b-tagging algorithms.

The event selection for the leptonic control region is:

• Passes the HLT\_IsoMu24\_eta2p1\_v* trigger and has an offline muon with \( p_T \geq 28 \text{ GeV}, |\eta| \leq 2.1 \), or passes the HLT\_Ele27\_WP80\_v* trigger has an offline electron with \( p_T \geq 30 \text{ GeV}, |\eta| \leq 2.4 \)

• \( \geq 2 \) lepton disambiguated jets

• \( \Delta R(l_i, l_j) \geq 0.05 \) between all pairs of leptons \( l_i \) and \( l_j \)

• If there are exactly two reconstructed leptons of the same flavor in the event, then \( m(ll) \geq 56 \text{ GeV} \).

Additionally, we define a “relaxed” criteria for QCD events. We don’t get enough QCD events with the standard criteria. This is because MC does not adequately model the rate of leptons from non-prompt sources like punch-through hadrons, anomalous fragmentation of jets into highly isolated particles, mis-reconstruction, etc. The relaxed criteria is as follows:

• **Electron:**
  - Same \( p_T \) and \( |\eta| \) as standard electron criteria
  - Pass loose ID working point [38]
  - Combined PF\text{Iso} \leq 2 \times p_T

• **Muon:**
  - Same \( p_T \) and \( |\eta| \) as standard muon criteria
  - Pass loose ID working point [40]
  - Combined PF\text{Iso} \leq 2 \times p_T
• Jet:
  - Pass all standard jet criteria
  - All jets within $\Delta R < 0.3$ of electrons or muons are removed from jet list

6.2 SFS for Kinematic Corrections

A “Scale factor set” (SFS) is a set of SF that represent a particular class of simulated background correction. SF are derived through MC reweighting. More general information on the MC reweighting procedure can be found in the appendix A.2. Every SFS has:

• A background process being corrected for.
• A control region (CR) where the SF’s are derived.
• A set of $m$ truth-level quantities used to parametrized the SF.
• A set of $n$ observed quantities used to define $n$-dimensional histogram yields in data and simulation.

In order to accurately map a correction back to the physics process responsible, every SFS is based on only one process. All other processes are then treated as backgrounds to this process. These faux backgrounds are then subtracted from the data yields (subtract the simulation yields) before we begin unfolding. We derive SFS consecutively. We reweigh the simulation after a SFS is derived. We use the reweighed simulation to derive the next SFS. We define the following process specific leptonic control regions for our SFSs:

• $t\bar{t}$
– Exactly one lepton, which is a muon. We chose the muon, as opposed to electron, because the muon events are less contaminated with poorly modeled QCD multijet production events.
– $\geq 3$ jets with $p_T \geq 30$ GeV and $|\eta| \leq 2.4$
– $\geq 1$ bT CSV b-tagged jet
– $\geq 2$ bM CSV b-tagged jet

• **W**

– Exactly one lepton, which is a muon. We chose the muon, as opposed to electron, because the muon events are less contaminated with poorly modeled QCD multijet production events.
– $m_T(W) \geq 40$ GeV
– bM b-tag veto for non b-specific SFS

• **Z**

– Exactly two leptons, same flavor
– $80$ GeV $< m(ll) < 100$ GeV
– Medium b-tag veto for non b-specific SFS

The SFS along with the corresponding CR, observables, and simulated truth parameterization are listed in Table 6.1. Here “N.” stands for “number of”. Additionally, unless stated otherwise, all jets used for counting and $H_T$ (scalar sum $p_T$) computations have $p_T \geq 30$ GeV and $|\eta| \leq 2.4$. The simulated kinematics of well-known top quark/W/Z boson decays, should be fully parametrized by the 4-momentum of the top quarks/W/Z bosons. Hence, we omit jets coming from decay products of the top quark/W/Z bosons when doing jet counting and $H_T$ computations.

### 6.2.1 Effect of MC-Reweighting on Control Distributions

This section shows selected data-simulation agreements (closures) before and after the SFS have been applied. Additional closures can be found at Section 6.3 of [13].
These closure checks are done to make sure our MC weighting methodology produces satisfactory results. The results are satisfactory and in all cases improve upon the non-corrected simulation. The labeling convention is as follows:

- **SM**: The ratio of the simulation to data, uncorrected simulation prediction.
- **× SF**: The ratio of the simulation to data, fully corrected simulation prediction
- **× SF \(t\bar{t}\) etc**: The fractional composition of the corrected simulation
- **\(j70\)**: \(p_T \geq 70\) GeV, \(|\eta| \leq 2.4\) jet multiplicity. All other jet counting (i.e. when not stated) uses a \(p_T \geq 30\) GeV threshold.
- **bM**: The medium CSV b-tagged jet.
- **bT**: The tight CSV b-tagged jet.
- **\(\geq 1\) bT bM**: At least one tight-tagged plus one medium-tagged jet, i.e. \(\geq 2\) medium b-tagged jets in total
- **Zll100**: \(Z \rightarrow \ell^+\ell^-\) selection, with \(p_T(\ell\ell) \geq 100\) GeV.

Figures 6.1 to 6.3 show the simulation to data ratios as a function of jet multiplicity, for various b-tag multiplicity requirements. “lepFake” corresponds to an additional QCD production process scale factor set. These account for residual tensions in the \(m_T(W) \leq 30\) GeV regions per reconstructed jet multiplicity bin. This is a crude modeling and not expected to be predicted well. This is shown to give a idea of how much QCD contamination can be present. The maximum expected QCD contamination for the exactly 2 jet, 1 tight b-tag, \(m_T(W) \geq 40\) GeV single muon region, is \(\sim 5\%\). For all other cases the contamination is \(\sim 1\%\). The agreement between simulation and data agreement after reweighting is \(10\%\) or better in most regions.
Figure 6.1: Simulated over data yields as a function of jet multiplicity ($p_T \geq 30 \text{ GeV}, |\eta| \leq 2.4$) for various event selections as stated in the titles. [13].
Figure 6.2: Simulated over data yields as a function of jet multiplicity \((p_T \geq 30 \text{ GeV, } |\eta| \leq 2.4)\) for various event selections as stated in the titles. [13].
Figure 6.3: Simulated over data yields as a function of jet multiplicity ($p_T \geq 30\text{ GeV}, |\eta| \leq 2.4$) for various event selections as stated in the titles. [13].
6.3 Closure Test for Leptonic Control Regions

We perform a closure test (observing the agreement between simulation and data) as a way to check our methodology, as well as to arrive at closure yield bias and closure uncertainty. Ideally, we want to see corrected simulated in much better agreement with data, than uncorrected simulated samples.

6.3.1 Closure Test in the Single Lepton Control Region

For each search region discriminator, we check the closure in the single leptonic control region, between data and simulation processes. We do this as a way to check our methodology, as well as to arrive at closure yield bias and closure uncertainty. Figures 6.4 and 6.5 show these closures for T2bW and T2tt respectively. The labeling convention is as follows:

- **SM**: The ratio of the simulation to data, with a subset of corrections applied divided by the final prediction with all corrections applied. Corrections are listed in Table 6.2. The residual lepton correction come from the fact that MC is too efficient when selecting lepton in high $H_T$ regions with b jets. The corresponding study for this correction is given in Appendix D in [13].

- $\times SF t\bar{t}$ etc: The fractional composition of the corrected simulation

- **SR**: The whole search region

- **SRxN**: A region N times larger (in discriminator range) then the SR. If $D_{T2bw} \geq d$, then the bin “$SR \times n$” contains events with $D_{T2bw} \geq 1 - n \times (1 - d)$.

- **Errors**: Error bars are statistical only and mostly dominated by uncertainties on the data yield.

We see that corrected simulation is generally $\sim 20\%$ or better (where statistically significant). The closure uncertainty is directly taken from the uncertainty in the
Figure 6.4: Closure in the single lepton control region, where the lepton has been removed from all quantities (jets and \( \vec{E}_T \)). Observed over predicted yields as a function of the BDT discriminator values for various T2bW search regions. [13]
Figure 6.5: Closure in the single lepton control region, where the lepton has been removed from all quantities (jets and $E_T$). Observed over predicted yields as a function of the BDT discriminator values for various T2tt search regions. [13]
search region. However, the yields in the search region are too statistically limited to provide a search region dependent bias correction. Hence, we use a more statistically plentiful region, SR $\times 2$, to get a search region dependent bias correction. This is reasonable as SR $\times 2$ is statistically compatible with the search region in all regions. The closure bias and uncertainties used in the final prediction are listed Tables 6.3.

Although not used in this analysis, the same 1-variable closures studies performed in $Z \rightarrow \ell^+\ell^-$ control region (Section 6.3.2) were performed in the single lepton control region. The results of those studies are presented here for completeness (Figure 6.6).

Figure 6.6 : Worst variable closure in the single lepton control region, where the lepton has been removed from all quantities (jets and $E_T$). Left: T2bW variables. Right: T2tt variables. [13]
6.3.2 Closure Test in the $Z \to \ell^+\ell^-$ Control Region

The closure in the $Z \to \ell^+\ell^-$ control region is given in Figures 6.7 and 6.8. Only four bins per search region can be made with decent statistical precision. Additionally, there are insufficient data events to be able to probe the search region itself. Regardless, data and corrected simulation is generally within 20% agreement or better. In order better understand the bias we relax the preselection. Figures 6.9 and 6.10 show the relaxed preselection plots. We observe no large scale bias in yield as a function of search region discriminator.

Due to the low statistics we arrive at an alternative closure metric. The 1-variable ($x$) closure of $x$ is defined as:

$$\text{closure}(x) \equiv \sum_i f_{i}^{\text{SR}}(x) \frac{n_{i}^{\text{data}}(x)}{n_{i}^{\text{MC}}(x)},$$

where the sum is over all bins in the $x$ histogram, $f_{i}^{\text{SR}}(x)$ is the fractional search region contribution for bin $i$, $n_{i}^{\text{data}}(x)$ is the data yield in bin $i$, and $n_{i}^{\text{MC}}(x)$ is the simulated yield in bin $i$. An example of this closure uncertainly for $H_{T}(\text{Along})/H_{T}(\text{Away})$ is given in Figure 6.11. The closure systematic uncertainty for the $Z$ prediction, is obtained from the worst 1-variable closure. The worse is one with maximal deviation of:

$$\frac{|\text{closure}(x) - 1|}{\sigma_{\text{stat}[\text{closure}(x)]}}.$$

This is taken per search region and for the tightest MVA discriminator cut preselection possible. Figure 6.12, Table 6.4, and Table 6.5 give these worst closures.
Figure 6.7: Closure in the $Z \rightarrow \ell^+\ell^-$ control region, where the lepton has been removed from all quantities (jets and $E_T$). Observed over predicted yields as a function of the BDT discriminator values for various T2bW search regions. [13]
Figure 6.8: Closure in the $Z \rightarrow \ell^+\ell^-$ control region, where the lepton has been removed from all quantities (jets and $E_T$). Observed over predicted yields as a function of the BDT discriminator values for various T2tt search regions. [13]
Figure 6.9: Closure in the $Z \rightarrow \ell^+\ell^-$ control region, after a relaxed preselection, where the lepton has been removed from all quantities (jets and $E_T$). Observed over predicted yields as a function of the BDT discriminator values for various T2bW search regions. [13]
Figure 6.10: Closure in the $Z \rightarrow \ell^+ \ell^-$ control region, after a relaxed preselection, where the lepton has been removed from all quantities (jets and $E_T$). Observed over predicted yields as a function of the BDT discriminator values for various T2tt search regions. [13]
Figure 6.11: Observed and predicted yields as a function of $H_T^{\text{along}} / H_T^{\text{away}}$ for events in the $Z \rightarrow \ell^+ \ell^-$ control region and various T2bW search regions.
Figure 6.12 : Worst variable closure in the $Z \rightarrow \ell^+\ell^-$ control region, where the lepton has been removed from all quantities (jets and $E_T$). Left: T2bW variables. Right: T2tt variables. [13]

6.4 $E_T$ Triggered Control Region

An additional $E_T$ triggered control region study was performed. This is done as an independent check for Z+jets and W+jets processes. In simulation, the $E_T$ triggered sample is made using simulated events after applying the detector effect corrections listed in Section A.2.2, and the truth-level kinematics reweighting in Section 6.2. In data, the $E_T$ triggered region is derived from the $H_T$ dataset given in Table B.5. The selection for $E_T$ triggered search region is:

- $E_T \geq 175$ GeV
- $\geq 2$ jets with $p_T \geq 70$ GeV and $|\eta| \leq 2.4$
- QCD rejection cuts: $\Delta \phi(E_T, j1 - 2) \geq 0.5$ and $\Delta \phi(E_T, j3) \geq 0.3$
- Lepton vetos applied

For both Z+jets and W+jets processes discrepancies of $\sim 5\%$ in the event counts
relative to those predicted were observed. These discrepancies were taken as an additional uncertainty in the event counts for these background processes. These uncertainties can be found in Table 6.10 and Table 6.9, for Z and $t\bar{t}+W$ respectively.

6.5 Prediction in Search Regions

The rest of this section gives the predicted simulation count and uncertainties in each of our SUSY search regions for our single top, $t\bar{t}$, W, and Z+jets backgrounds.

6.5.1 Uncertainties

There are six types of uncertainties are of relevance for our top and EWK backgrounds:

1. **Simulated sample size**: Statistical uncertainties due to the amount of simulation surviving various search region requirements and the number of smeared events (for the $E_T$ corrections) (Section A.2.2) simulated per input event.

2. **Lepton Veto BDT selection SF**: Uncertainties on the efficiency of prompt leptons to pass the lepton BDT selection criteria. These studies are documented in Appendix H of [13].

3. **Kinematics reweighting statistics**: Statistical uncertainty of the leptonic control regions used to derive the event kinematics corrections (Section 6.2). As the same events can be reused for different scale factor sets (SFS), this is computed in a properly correlated way via a bootstrapping procedure [132].

   - Generate 40 pseudo-datasets from the lepton-triggered collider data, by weighting each event by a random number drawn from a Poisson distribution with mean 1.
   - Each pseudo-dataset gives a different SFS
   - For each search region the propagated statistical uncertainty is the spread (RMS) of predicted yields, over the ensemble of predictions from applying the different SFS.
4. **Closure systematics**: The closure level obtained in leptonic control regions for the corrections applied to the MC-based prediction (\(E_T\), jet shape, kinematics reweighting), is used as a systematic uncertainty as described in Section 6.3.

5. **Closure in \(E_T\) region**: An additional systematic uncertainty of 5\% for W and Z processes due to non-closure observed in the \(E_T\) triggered region and not present in the lepton triggered control regions. Discussed in Section 6.4.

6. **MC modeling**: See Section 6.5.3. For the single top background only, which has less than 1\% contribution in any search region, a 50\% systematic uncertainty is assumed on the simulated yields since there is no adequate control region to derive correction factors nor closure levels for this prediction.

### 6.5.2 Impact

The various corrections effecting the total predicted yields of the top and EWK processes in the search region, including the \(E_T\) triggered corrections, can be found in Table 6.6. The numbers in this table are the ratio of yields. This list is cumulative, in that each correction is done sequentially from top to bottom. The numerator is the yield after the particular correction in applied. The denominator is the yield in the previous step. The denominator of the first row is the simulation yields with only MC pileup reweighting (Section B.3). “\(E_T\) corrections” are the \(E_T\) scale and resolution corrections.

The impact of the kinematic reweighting scale factors sets on the total predicted yields of the top and EWK processes is given in Table 6.7. The numbers in this table are the ratio of yields. The numerator is the yield if all but that particular scale factor set is applied. The denominator is the yield when all corrections and scale factor sets are applied.
6.5.3 Top Background

Unfortunately, the limited amount of simulated single top statistics entering our search regions makes it impossible to define an adequate single top control region. This makes it impossible to derive kinematic correction factors or closure levels. The estimated yield of the single top is negligible. This is shown in Table 6.8. We do not incorporate single top into our backgrounds.

6.5.4 EWK backgrounds

The predicted $t\bar{t}$ and $W$ background yields are given in Table 6.9. The predicted $Z$ yields are given in Table 6.10.
<table>
<thead>
<tr>
<th>SFS</th>
<th>Process/CR</th>
<th>Observable</th>
<th>Truth parameterization</th>
</tr>
</thead>
<tbody>
<tr>
<td>ttSpectrum</td>
<td>$t\bar{t}$</td>
<td>$p_T$ of the vectorial sum of the lepton, $E_T$, and nearest (in $\Delta R$) medium b-tagged jet to the lepton $p_T$ of the other medium b-tagged jet, $p_T(l + E_T)$</td>
<td>$p_T$ of leading top quark $p_T$ of second top quark</td>
</tr>
<tr>
<td>ttNJ</td>
<td>$t\bar{t}$</td>
<td>$N. \ p_T \geq 30 \ \text{GeV jets}$ $N. \ p_T \geq 70 \ \text{GeV jets}$ $p_T(l + E_T)$</td>
<td>$N. \ p_T \geq 30 \ \text{GeV jets}$ $N. \ p_T \geq 70 \ \text{GeV jets}$ $p_T(l + E_T)$</td>
</tr>
<tr>
<td>zRecoil</td>
<td>$Z$</td>
<td>$N. \ p_T$ jets $H_T$ $p_T(l)$</td>
<td>$H_T$ $p_T(l)$</td>
</tr>
<tr>
<td>zNJ</td>
<td>$Z$ (b vetoed)</td>
<td>$N. \ p_T \geq 30 \ \text{GeV jets}$ $N. \ p_T \geq 70 \ \text{GeV jets}$ $p_T(l)$</td>
<td>$N. \ p_T \geq 30 \ \text{GeV jets}$ $N. \ p_T \geq 70 \ \text{GeV jets}$ $p_T(l)$</td>
</tr>
<tr>
<td>wRecoil</td>
<td>W (b vetoed)</td>
<td>$H_T$ $p_T(l + E_T)$</td>
<td>$H_T$ $p_T(l + E_T)$</td>
</tr>
<tr>
<td>wNb</td>
<td>W</td>
<td>$N. \ p_T$ jets $p_T(l + E_T)$ $p_T(l + E_T)$</td>
<td>$p_T(l + E_T) p_T(l + E_T)$</td>
</tr>
<tr>
<td>wNJ70</td>
<td>W</td>
<td>$N. \ p_T \geq 30 \ \text{GeV jets}$ $N. \ p_T \geq 70 \ \text{GeV jets}$ $p_T(l + E_T)$</td>
<td>$N. \ p_T \geq 30 \ \text{GeV jets}$ $N. \ p_T \geq 70 \ \text{GeV jets}$ $p_T(l + E_T)$</td>
</tr>
</tbody>
</table>

Table 6.1: Scale factor sets
Table 6.2: Corrections applied to the simulation in the lepton closure tests.

<table>
<thead>
<tr>
<th>Correction</th>
<th>Included in &quot;SM&quot;</th>
<th>Included in the full prediction</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pileup</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>MET</td>
<td>no</td>
<td>yes</td>
</tr>
<tr>
<td>ak5 JEC</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>ak5 b-Tag</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>picky JEC</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>picky b-Tag</td>
<td>no</td>
<td>yes</td>
</tr>
<tr>
<td>Q/G likelihood</td>
<td>no</td>
<td>yes</td>
</tr>
<tr>
<td>Std. Lepton trigger</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>Std. Lepton ID/ISO</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>Residual Lepton ID</td>
<td>no</td>
<td>yes</td>
</tr>
<tr>
<td>Kinematic correction</td>
<td>no</td>
<td>yes</td>
</tr>
</tbody>
</table>
Table 6.3: Systematic uncertainty and correction due to closure of the MC reweighting prediction, for the $t\bar{t}$ and $W$ processes.

<table>
<thead>
<tr>
<th>Search Region</th>
<th>Bias Correction</th>
<th>Closure uncertainty (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>T2bW_LX</td>
<td>1.01</td>
<td>0.25</td>
</tr>
<tr>
<td>T2bW_LM</td>
<td>1.09</td>
<td>0.10</td>
</tr>
<tr>
<td>T2bW_MXHM</td>
<td>1.20</td>
<td>0.23</td>
</tr>
<tr>
<td>T2bW_HXHM</td>
<td>1.20</td>
<td>0.20</td>
</tr>
<tr>
<td>T2bW_VHM</td>
<td>1.21</td>
<td>0.33</td>
</tr>
<tr>
<td>T2tt_LM</td>
<td>1.04</td>
<td>0.08</td>
</tr>
<tr>
<td>T2tt_MM</td>
<td>1.18</td>
<td>0.15</td>
</tr>
<tr>
<td>T2tt_HM</td>
<td>1.26</td>
<td>0.23</td>
</tr>
<tr>
<td>T2tt_VHM</td>
<td>1.35</td>
<td>0.31</td>
</tr>
</tbody>
</table>
Table 6.4: For T2bW search regions, the worst variable closure in the $Z \to \ell^+\ell^-$ control region, where the lepton has been removed from all quantities (jets and $E_T$).

<table>
<thead>
<tr>
<th>Search Region</th>
<th>Cut</th>
<th>Variable</th>
<th>Closure</th>
</tr>
</thead>
<tbody>
<tr>
<td>T2bW_LX</td>
<td>presel</td>
<td>$\Delta\phi(E_T,jet3)$</td>
<td>1.50 ± 0.42</td>
</tr>
<tr>
<td>T2bW_LX</td>
<td>$\geq -0.5$</td>
<td>num medium b-tags</td>
<td>0.61 ± 0.27</td>
</tr>
<tr>
<td>T2bW_LM</td>
<td>presel</td>
<td>leading q likelihood</td>
<td>1.32 ± 0.26</td>
</tr>
<tr>
<td>T2bW_LM</td>
<td>$\geq -0.5$</td>
<td>$E_T$</td>
<td>1.32 ± 0.31</td>
</tr>
<tr>
<td>T2bW_MXHM</td>
<td>presel</td>
<td>product q likelihood</td>
<td>0.65 ± 0.16</td>
</tr>
<tr>
<td>T2bW_MXHM</td>
<td>$\geq -0.5$</td>
<td>$\Delta\phi(E_T,j3)$</td>
<td>0.70 ± 0.21</td>
</tr>
<tr>
<td>T2bW_HXHM</td>
<td>presel</td>
<td>$m_T(b)$</td>
<td>0.79 ± 0.18</td>
</tr>
<tr>
<td>T2bW_HXHM</td>
<td>$\geq -0.5$</td>
<td>leading q likelihood</td>
<td>1.32 ± 0.28</td>
</tr>
<tr>
<td>T2bW_VHM</td>
<td>presel</td>
<td>$\Delta\phi(E_T,jet3)$</td>
<td>0.79 ± 0.16</td>
</tr>
<tr>
<td>T2bW_VHM</td>
<td>$\geq -0.5$</td>
<td>$\Delta\phi(E_T,jet3)$</td>
<td>0.75 ± 0.20</td>
</tr>
</tbody>
</table>
Table 6.5: For T2tt search regions, the worst variable closure in the $Z \rightarrow \ell^+ \ell^-$ control region, where the lepton has been removed from all quantities (jets and $E_T$).

<table>
<thead>
<tr>
<th>Search Region</th>
<th>Cut</th>
<th>Variable</th>
<th>Closure</th>
</tr>
</thead>
<tbody>
<tr>
<td>T2tt_LM</td>
<td>presel</td>
<td>$\Delta \phi(\vec{E}_T, \text{top2})$</td>
<td>$1.24 \pm 0.10$</td>
</tr>
<tr>
<td>T2tt_LM</td>
<td>$\geq -0.5$</td>
<td>product q likelihood</td>
<td>$1.22 \pm 0.21$</td>
</tr>
<tr>
<td>T2tt_LM</td>
<td>$\geq 0$</td>
<td>top2 fat quality</td>
<td>$1.16 \pm 0.22$</td>
</tr>
<tr>
<td>T2tt_MM</td>
<td>presel</td>
<td>$\text{top2}(p_T) / \text{top1}(p_T)$</td>
<td>$1.18 \pm 0.10$</td>
</tr>
<tr>
<td>T2tt_MM</td>
<td>$\geq -0.5$</td>
<td>$\text{top2}(p_T) / \text{top1}(p_T)$</td>
<td>$1.28 \pm 0.14$</td>
</tr>
<tr>
<td>T2tt_MM</td>
<td>$\geq 0$</td>
<td>Max b-jet disc</td>
<td>$1.32 \pm 0.22$</td>
</tr>
<tr>
<td>T2tt_HM</td>
<td>presel</td>
<td>ellipse size</td>
<td>$1.14 \pm 0.10$</td>
</tr>
<tr>
<td>T2tt_HM</td>
<td>$\geq -0.5$</td>
<td>top2 jet3 $p_T$</td>
<td>$1.39 \pm 0.17$</td>
</tr>
<tr>
<td>T2tt_HM</td>
<td>$\geq 0$</td>
<td>top2 jet3 $p_T$</td>
<td>$1.30 \pm 0.25$</td>
</tr>
<tr>
<td>T2tt_VHM</td>
<td>presel</td>
<td>top2 fat quality</td>
<td>$1.38 \pm 0.22$</td>
</tr>
<tr>
<td>T2tt_VHM</td>
<td>$\geq -0.5$</td>
<td>$\eta$ peak location</td>
<td>$1.39 \pm 0.16$</td>
</tr>
<tr>
<td>T2tt_VHM</td>
<td>$\geq 0$</td>
<td>top2 thin quality</td>
<td>$0.74 \pm 0.21$</td>
</tr>
</tbody>
</table>
Table 6.6: Impact of the various corrections on the total predicted yields of top and EWK processes. Picky jet CSV discriminator shape corrections have no impact on T2bW search regions, which only utilizes anti-\(k_T\) jets.

<table>
<thead>
<tr>
<th>Search region</th>
<th>T2bW</th>
<th>T2tt</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>LX</td>
<td>LM</td>
</tr>
<tr>
<td>(E_T) corrections</td>
<td>1.27</td>
<td>0.97</td>
</tr>
<tr>
<td>q/g likelihood &amp; picky jet spectra</td>
<td>1.12</td>
<td>1.17</td>
</tr>
<tr>
<td>Picky jet CSV shape</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Kinematics reweighting</td>
<td>1.06</td>
<td>1.05</td>
</tr>
<tr>
<td>Lepton Veto BDT selection SF</td>
<td>1.07</td>
<td>1.04</td>
</tr>
</tbody>
</table>

Table 6.7: Impact of various kinematics reweighting scale factor sets on the total predicted yields of top and EWK processes.

<table>
<thead>
<tr>
<th>Search region</th>
<th>T2bW</th>
<th>T2tt</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>LX</td>
<td>LM</td>
</tr>
<tr>
<td>ttSpectrum</td>
<td>1.07</td>
<td>1.03</td>
</tr>
<tr>
<td>ttNJ</td>
<td>0.96</td>
<td>0.97</td>
</tr>
<tr>
<td>zRecoil</td>
<td>1.03</td>
<td>1.01</td>
</tr>
<tr>
<td>ZNJ</td>
<td>0.99</td>
<td>0.97</td>
</tr>
<tr>
<td>ZNb</td>
<td>0.95</td>
<td>0.97</td>
</tr>
<tr>
<td>ZHTb</td>
<td>1.01</td>
<td>1.01</td>
</tr>
<tr>
<td>wRecoil</td>
<td>1.02</td>
<td>1.01</td>
</tr>
<tr>
<td>wNb</td>
<td>0.98</td>
<td>1.00</td>
</tr>
<tr>
<td>wNJ70</td>
<td>1.00</td>
<td>1.00</td>
</tr>
</tbody>
</table>
Table 6.8 : MC predicted single top yields and uncertainties (in events).

<table>
<thead>
<tr>
<th>Search region</th>
<th>T2bW</th>
<th>T2tt</th>
</tr>
</thead>
<tbody>
<tr>
<td>Single top yield</td>
<td>0.0002</td>
<td>0.0011</td>
</tr>
<tr>
<td>MC statistics</td>
<td>0.0007</td>
<td>0.0021</td>
</tr>
<tr>
<td>Lepton Veto BDT selection SF</td>
<td>0.0012</td>
<td>0.0040</td>
</tr>
<tr>
<td>MC modeling</td>
<td>0.0001</td>
<td>0.0006</td>
</tr>
<tr>
<td>Total uncertainty (yield)</td>
<td>0.0014</td>
<td>0.0046</td>
</tr>
<tr>
<td>Total uncertainty (%)</td>
<td>806.61</td>
<td>410.14</td>
</tr>
</tbody>
</table>

Table 6.9 : MC reweighting predicted t\bar{t} and W yields and uncertainties (in events).

<table>
<thead>
<tr>
<th>Search region</th>
<th>T2bW</th>
<th>T2tt</th>
</tr>
</thead>
<tbody>
<tr>
<td>t\bar{t} + W yield</td>
<td>6.41</td>
<td>30.35</td>
</tr>
<tr>
<td>W fraction</td>
<td>0.17</td>
<td>0.05</td>
</tr>
<tr>
<td>MC statistics</td>
<td>0.84</td>
<td>1.65</td>
</tr>
<tr>
<td>Lepton Veto BDT selection SF</td>
<td>1.03</td>
<td>2.24</td>
</tr>
<tr>
<td>Kinematics reweighting stat.</td>
<td>0.20</td>
<td>0.44</td>
</tr>
<tr>
<td>Closure (1\ell control region)</td>
<td>1.59</td>
<td>2.78</td>
</tr>
<tr>
<td>Closure (E_T region)</td>
<td>0.05</td>
<td>0.08</td>
</tr>
<tr>
<td>Total uncertainty (yield)</td>
<td>2.08</td>
<td>3.96</td>
</tr>
<tr>
<td>Total uncertainty (%)</td>
<td>32.5</td>
<td>13.0</td>
</tr>
</tbody>
</table>
Table 6.10 : MC reweighting predicted $Z \rightarrow \nu \bar{\nu}$ yields and uncertainties (in events).

<table>
<thead>
<tr>
<th>Search region</th>
<th>T2bW</th>
<th>T2tt</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>LX</td>
<td>LM</td>
</tr>
<tr>
<td>$Z \rightarrow \nu \bar{\nu}$ yield</td>
<td>1.88</td>
<td>4.57</td>
</tr>
<tr>
<td>MC statistics</td>
<td>0.23</td>
<td>0.46</td>
</tr>
<tr>
<td>Kinematics reweighting stat.</td>
<td>0.51</td>
<td>0.62</td>
</tr>
<tr>
<td>Closure ($Z \rightarrow \ell^+\ell^-$ control region)</td>
<td>0.73</td>
<td>1.46</td>
</tr>
<tr>
<td>Closure ($E_T$ region)</td>
<td>0.09</td>
<td>0.23</td>
</tr>
<tr>
<td>Total uncertainty (yield)</td>
<td>0.93</td>
<td>1.67</td>
</tr>
<tr>
<td>Total uncertainty (%)</td>
<td>49.3</td>
<td>36.6</td>
</tr>
</tbody>
</table>
Chapter 7

QCD Multijet background

For the QCD multi-jet background we use MC sample reweighting as in Section 6. The section briefly touches on the process and provides the results. A more detailed breakdown of the QCD multi-jet sample reweighting can be found in [13, 132].

As with the top, W, and Z(\nu\bar{\nu}) + jets MC sample reweighting, the goal is to find a suitable control region from we can be used to derive simulation yields and uncertainties. This is referred to as the Low $E_T$ control region. However, this is just the start of the QCD prediction. The full method uses a total of 6 BDT control regions, in addition to the Low $E_T$ baseline region.

As opposed to EWK backgrounds entering the search region due to $E_T$ from neutrinos, QCD enter the search region from heavy flavor decay or a severe jet mis-measurement. Therefore, a large component of the QCD prediction involves correcting for these effects in the simulated samples while preserving all correlations between $E_T$ and jets in the event which is not necessary in the other prediction. There are four $E_T$ corrections:

1. **Jet core energy resolution correction**: Corrects a bias in the $E_T$ distribution at moderate $E_T$ values and correcting the $E_T$ magnitude at high $E_T$ values. This
corresponds to correcting a jet “core”. The core should ideally follow a gaussian
distribution. The core is known to be wider in data relative to simulation.

2. **Unclustered energy correction** : The unclustered energy is not properly
simulated in the simulation. This correction smears $E_T$ proportionally to the
magnitude of the unclustered energy in the detector with the aim of reducing
biases in the $\Delta \phi(b, E_T)$ distributions.

3. **Residual $\Delta \phi(b, E_T)$ correction** : This corrects a small bias in the correlation
between $E_T$ and jets in b-tagged events with high $E_T$. It is a tool to estimate
systematics.

4. **Jet tail resolution correction** : This corrects jet resolution tail discrepancies
between data and simulation. Correct the total normalization of events at high
$E_T$ while correcting the relative fractions of different $E_T$ sources in this region.

The QCD predictions are then extrapolation into the search region and the determi-
nation of uncertainties on the predictions is different than in EWK. This methodology
is detailed in [13,132]. The results of the QCD prediction are given in Table 7.1.
Table 7.1: Results of the QCD prediction.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>T2bW_LX</td>
<td>0.5112 ± 0.2060</td>
<td>+0.0137</td>
<td>±0.1697</td>
<td>+0.5840, −0.2063</td>
<td>0.3049 ± 0.2626</td>
</tr>
<tr>
<td>T2bW_LM</td>
<td>0.0729 ± 0.0612</td>
<td>+0.1091</td>
<td>±0.0646</td>
<td>+0.5416, −0.0729</td>
<td>0.0000 ± 0.0000</td>
</tr>
<tr>
<td>T2bW_MXHM</td>
<td>0.0965 ± 0.0824</td>
<td>+0.0336</td>
<td>±0.0808</td>
<td>+0.0725, −0.0965</td>
<td>0.1223 ± 0.1137</td>
</tr>
<tr>
<td>T2bW_HXHM</td>
<td>0.0012 ± 0.0012</td>
<td>+0.0198</td>
<td>±0.0014</td>
<td>+0.0055, −0.0012</td>
<td>0.0000 ± 0.0000</td>
</tr>
<tr>
<td>T2bW_VHM</td>
<td>0.0000 ± 0.0000</td>
<td>+0.0134</td>
<td>±0.0000</td>
<td>+0.0146, −0.0000</td>
<td>0.0000 ± 0.0000</td>
</tr>
<tr>
<td>T2tt_LM</td>
<td>0.3314 ± 0.2652</td>
<td>+0.0000</td>
<td>±0.1559</td>
<td>+1.4775, −0.3314</td>
<td>0.0000 ± 0.0000</td>
</tr>
<tr>
<td>T2tt_MM</td>
<td>0.0032 ± 0.0035</td>
<td>+0.1083</td>
<td>±0.0023</td>
<td>+0.2247, −0.0030</td>
<td>0.0002 ± 0.0002</td>
</tr>
<tr>
<td>T2tt_HM</td>
<td>0.0001 ± 0.0001</td>
<td>+0.0178</td>
<td>±0.0000</td>
<td>+0.0691, −0.0001</td>
<td>0.0002 ± 0.0002</td>
</tr>
<tr>
<td>T2tt_VHM</td>
<td>0.0000 ± 0.0000</td>
<td>+0.0178</td>
<td>±0.0000</td>
<td>+0.0057, −0.0000</td>
<td>0.0000 ± 0.0000</td>
</tr>
</tbody>
</table>
High mass $\tilde{t}$ produce boosted decay products, hence they can produce events with high $E_T$. Unfortunately, $t\bar{t}Z$ (with $Z \to \nu\bar{\nu}$) has the same signature, making it an irreducible background for this search. Hence, we need to understand the $t\bar{t}Z$ background.

Because $t\bar{t}Z$ has such a low production cross section ($\sim 242$ fb [56]) compared to the other SM process, it’s not really possible to find control regions enriched in $t\bar{t}Z$ with statistics high enough for MC reweighting. Hence, we cannot use the standard MC reweighting technique. In lieu of MC reweighting, we use MC simulated events to obtain the predictions of the yields in the SUSY search regions. Additionally, we derive the following three systematic uncertainties:

1. **MC Generation Uncertainty**: Compare our $t\bar{t}Z$ simulated yields to those given by a different generator, in the SUSY search regions

2. **Cross section Uncertainty**: Use a $t\bar{t}Z$ enriched control region to validate the inclusive NLO cross section prediction on data

3. **Data-Simulation Kinematic Differences Uncertainty**: Use the maximum kinematic closure uncertainty of the $t\bar{t} + W$ and $Z \to \nu\bar{\nu}$ predictions as an additional systematic
8.1 Yields and MC Generation Uncertainty

The are two different $t\bar{t}Z$ simulated samples used are this analysis (Table B.16). These are referred to the MCatLNO and Madgraph samples. Information on how these samples were generated is given in Appendix B.3. We use the MCatLNO for the calculation of our yields, because it is expected to better describe the kinematics at low jet multiplicity with respect to the Madgraph. The yields and NLO Uncertainty are given in Table 8.1. The MC generation uncertainty is:

\[
\frac{|\text{MCatNLO} - \text{Madgraph}|}{\text{MCatNLO}}. \tag{8.1}
\]

8.2 Cross Section Uncertainty

We derive the uncertainty on the $t\bar{t}Z$ cross section by starting with the triple lepton control region. The data sample corresponding to this control region is given in Table B.6. The triggers and luminosity for this dataset are summarized in Table B.7.

We are looking for a subset of the triple lepton control region, a $t\bar{t}Z$ control region. This is a region where the Z decays into two opposite sign, same flavor leptons and the top-pair decays semi-leptonically. However, this region has a leading background of di-bosons (20% contribution). Since, we have no dedicated way to measure di-bosons, we must estimate this contribution from simulation. Additionally, we need to factor in the uncertainty of the di-boson cross section in our measurement for the $t\bar{t}Z$ cross section. Hence, we choose a region where the di-bosons dominates relative to the
Table 8.1: Simulated yields for $t\bar{t}Z$ for both the Madgraph and MCatNLO samples.

<table>
<thead>
<tr>
<th>Search Region</th>
<th>MCatNLO</th>
<th>Madgraph</th>
<th>Uncertainty</th>
</tr>
</thead>
<tbody>
<tr>
<td>$T2bW_{LX}$</td>
<td>0.59 ± 0.07</td>
<td>0.60 ± 0.09</td>
<td>3%</td>
</tr>
<tr>
<td>$T2bW_{LM}$</td>
<td>2.46 ± 0.15</td>
<td>2.36 ± 0.18</td>
<td>4%</td>
</tr>
<tr>
<td>$T2bW_{MXHM}$</td>
<td>0.83 ± 0.09</td>
<td>0.94 ± 0.13</td>
<td>12%</td>
</tr>
<tr>
<td>$T2bW_{HXHM}$</td>
<td>1.72 ± 0.14</td>
<td>1.55 ± 0.15</td>
<td>10%</td>
</tr>
<tr>
<td>$T2bW_{VHM}$</td>
<td>0.62 ± 0.08</td>
<td>0.64 ± 0.10</td>
<td>4%</td>
</tr>
<tr>
<td>$T2tt_{LM}$</td>
<td>1.34 ± 0.11</td>
<td>1.44 ± 0.15</td>
<td>7%</td>
</tr>
<tr>
<td>$T2tt_{MM}$</td>
<td>2.66 ± 0.18</td>
<td>3.08 ± 0.24</td>
<td>16%</td>
</tr>
<tr>
<td>$T2tt_{HM}$</td>
<td>1.62 ± 0.15</td>
<td>1.86 ± 0.19</td>
<td>15%</td>
</tr>
<tr>
<td>$T2tt_{VHM}$</td>
<td>0.99 ± 0.11</td>
<td>1.25 ± 0.16</td>
<td>26%</td>
</tr>
</tbody>
</table>

t$\bar{t}$Z contribution by inverting the b-tag requirement. We define the $t\bar{t}$Z and di-boson control regions:

- **$t\bar{t}$Z Control Region:**
  - 3 leptons (muon or electron)
  - Muons must pass tight working point and isolation
  - Electrons must pass medium working point and isolation
  - Lepton $p_T$ thresholds: $p_T > 20$ GeV, 10 GeV, 10 GeV. Referenced analysis indicate that this leads to a 100% trigger efficiency
  - OSSF Mass: $\geq 80$ GeV, $< 100$ GeV. OSSF = Opposite sign, same flavor lepton pair whose invariant mass is closest to that of a Z boson
  - $\geq 3$ AK5 jets($p_T > 30$ GeV, $|\eta| < 2.4$)
  - $\geq 3$ picky jets($p_T > 20$ GeV, $|\eta| < 2.4$)
– ≥ 1 picky medium b-tagged jets (p_T > 30 GeV, |\eta| < 2.4)

• Di-Boson Control Region:
  – 3 leptons (muon or electron)
  – Muons must pass tight working point and isolation
  – Electrons must pass medium working point and isolation
  – Lepton p_T thresholds: p_T > 20 GeV, 10 GeV, 10 GeV. Referenced analysis indicate that this leads to a 100% trigger efficiency
  – OSSF Mass: ≥ 80 GeV, < 100 GeV. OSSF = Opposite sign, same flavor lepton pair whose invariant mass is closest to that of a Z boson
  – ≥ 3 AK5 jets (p_T > 30 GeV, |\eta| < 2.4)
  – ≥ 3 picky jets (p_T > 20 GeV, |\eta| < 2.4)
  – ≥ 0 picky medium b-tagged jets (p_T > 30 GeV, |\eta| < 2.4)

The resulting yields in ttZ and di-boson control region, of the number of picky jets (p_T > 20 GeV, |\eta| < 2.4), are shown in Figure 8.1.

Figure 8.1: Yields in ttZ related control regions. Left: Simulated and data yields in the ttZ control region as a function of the number of picky jets with p_T > 20 GeV and |\eta| < 2.4. Right: Simulated and data yields in the di-boson control region as a function of the number of picky jets with p_T > 20 GeV and |\eta| < 2.4. [13]
The relationship between data and simulation in the control regions can be written as follows:

\[
\text{Data}_{\text{reg } 1} - \text{Other}_{\text{reg } 1} = SF_{t\bar{t}Z} \cdot t\bar{t}Z_{\text{reg } 1} + SF_{VV} \cdot VV_{\text{reg } 1}, \tag{8.2}
\]

\[
\text{Data}_{\text{reg } 2} - \text{Other}_{\text{reg } 2} = SF_{t\bar{t}Z} \cdot t\bar{t}Z_{\text{reg } 2} + SF_{VV} \cdot VV_{\text{reg } 2}, \tag{8.3}
\]

where reg 1 is the t\bar{t}Z control region, reg 2 is the di-boson control region, Other is all processes not t\bar{t}Z or di-boson, SF_{t\bar{t}Z} is the t\bar{t}Z Scale Factor, and SF_{VV} is the di-boson Scale Factor. Knowing the yields, we solve for the scale factors. The cross section uncertainty is taken as the uncertainty on the t\bar{t}Z Scale Factor. The results are shown in Table 8.2.

Table 8.2: Yields for data and simulation in both the t\bar{t}Z and di-boson control regions. The uncertainty on \( SF_{t\bar{t}Z} \) is taken as the cross section systematic uncertainty.

<table>
<thead>
<tr>
<th></th>
<th>Data</th>
<th>t\bar{t}Z</th>
<th>SF_{t\bar{t}Z}</th>
<th>Di-boson</th>
<th>SF_{VV}</th>
<th>Other SM</th>
</tr>
</thead>
<tbody>
<tr>
<td>t\bar{t}Z Reg.</td>
<td>15.00 ± 3.87</td>
<td>13.60 ± 0.41</td>
<td>0.69 ± 0.31</td>
<td>1.93 ± 0.14</td>
<td>0.80 ± 0.36</td>
<td>4.09 ± 0.78</td>
</tr>
<tr>
<td>Di-boson Reg.</td>
<td>15.00 ± 3.87</td>
<td>2.37 ± 0.17</td>
<td>0.69 ± 0.31</td>
<td>11.79 ± 0.34</td>
<td>0.80 ± 0.36</td>
<td>4.00 ± 0.72</td>
</tr>
</tbody>
</table>
8.3 Kinematic Uncertainty

The data-simulation kinematic uncertainty is usually derived via a re-weighted MC prediction like we do for the Z and \(t \bar{t}\) backgrounds. However, since our \(t \bar{t}Z\) control region is low in statistics, that method does not work. In lieu of re-weighting, we use the maximum closure uncertainty, in the lepton control regions, between the \(t \bar{t} + W\) and \(Z \rightarrow \nu \bar{\nu}\) background predictions (Section 6.3). We also apply a sanity check to verify that closure uncertainties in the lepton control regions are applicable to the \(t \bar{t}Z\) prediction. This is done by applying the \(t \bar{t}\) kinematic scale factors to the \(t \bar{t}Z\) system, ignoring the Z. This checks the cases where the bulk of the \(t \bar{t}Z\) kinematics lies in a region of phases space that we observe is very discrepant for \(t \bar{t}\). We see this effect is insignificant compared to the closures. Table 8.3 summarized the data-simulation kinematic uncertainty.

8.4 Prediction in the Search Regions

A summary of the final \(t \bar{t}Z\) prediction and all uncertainties, in all search regions, is given in Table 8.4.
Table 8.3: Comparison of data-simulation kinematic uncertainties. The first column shows the change of $t\bar{t}Z$ yields after applying the $t\bar{t}$ kinematic SFs. The second column shows the kinematic closures derived in Section 6.3 for $t\bar{t} + W$. The third column show the kinematic closures derived in Section 6.3 for $Z$. Finally, the last column is the kinematics uncertainty that we apply to the $t\bar{t}Z$ sample.

<table>
<thead>
<tr>
<th>Search Region</th>
<th>$t\bar{t}Z$ reweighting</th>
<th>$t\bar{t} + W$ closure</th>
<th>$Z$ closure</th>
<th>Kinematics uncertainty</th>
</tr>
</thead>
<tbody>
<tr>
<td>$T2bW_{LX}$</td>
<td>2%</td>
<td>25%</td>
<td>39%</td>
<td>39%</td>
</tr>
<tr>
<td>$T2bW_{LM}$</td>
<td>4%</td>
<td>10%</td>
<td>32%</td>
<td>32%</td>
</tr>
<tr>
<td>$T2bW_{MXHM}$</td>
<td>2%</td>
<td>23%</td>
<td>30%</td>
<td>30%</td>
</tr>
<tr>
<td>$T2bW_{HXHM}$</td>
<td>1%</td>
<td>20%</td>
<td>32%</td>
<td>32%</td>
</tr>
<tr>
<td>$T2bW_{VHM}$</td>
<td>2%</td>
<td>33%</td>
<td>25%</td>
<td>33%</td>
</tr>
<tr>
<td>$T2tt_{LM}$</td>
<td>2%</td>
<td>8%</td>
<td>16%</td>
<td>16%</td>
</tr>
<tr>
<td>$T2tt_{MM}$</td>
<td>3%</td>
<td>15%</td>
<td>32%</td>
<td>32%</td>
</tr>
<tr>
<td>$T2tt_{HM}$</td>
<td>7%</td>
<td>23%</td>
<td>30%</td>
<td>30%</td>
</tr>
<tr>
<td>$T2tt_{VHM}$</td>
<td>8%</td>
<td>31%</td>
<td>26%</td>
<td>31%</td>
</tr>
</tbody>
</table>
Table 8.4: $t\bar{t}Z$ prediction and uncertainties.

<table>
<thead>
<tr>
<th>Search Region</th>
<th>$t\bar{t}Z$ Prediction</th>
<th>Cross Sect.</th>
<th>NLO</th>
<th>Kinematics</th>
<th>Total Syst.</th>
<th>Total Unc.</th>
</tr>
</thead>
<tbody>
<tr>
<td>T2bW_LX</td>
<td>0.59 ± 11%</td>
<td>31%</td>
<td>3%</td>
<td>39%</td>
<td>50%</td>
<td>51%</td>
</tr>
<tr>
<td>T2bW_LM</td>
<td>2.46 ± 6%</td>
<td>31%</td>
<td>4%</td>
<td>32%</td>
<td>45%</td>
<td>45%</td>
</tr>
<tr>
<td>T2bW_MxHM</td>
<td>0.83 ± 11%</td>
<td>31%</td>
<td>12%</td>
<td>30%</td>
<td>45%</td>
<td>46%</td>
</tr>
<tr>
<td>T2bW_HxHM</td>
<td>1.72 ± 8%</td>
<td>31%</td>
<td>10%</td>
<td>32%</td>
<td>46%</td>
<td>46%</td>
</tr>
<tr>
<td>T2bW_VHM</td>
<td>0.62 ± 14%</td>
<td>31%</td>
<td>4%</td>
<td>33%</td>
<td>45%</td>
<td>47%</td>
</tr>
<tr>
<td>T2tt_LM</td>
<td>1.34 ± 8%</td>
<td>31%</td>
<td>7%</td>
<td>16%</td>
<td>36%</td>
<td>37%</td>
</tr>
<tr>
<td>T2tt_MM</td>
<td>2.66 ± 7%</td>
<td>31%</td>
<td>16%</td>
<td>32%</td>
<td>47%</td>
<td>48%</td>
</tr>
<tr>
<td>T2tt_HM</td>
<td>1.62 ± 9%</td>
<td>31%</td>
<td>15%</td>
<td>30%</td>
<td>46%</td>
<td>47%</td>
</tr>
<tr>
<td>T2tt_VHM</td>
<td>0.99 ± 11%</td>
<td>31%</td>
<td>26%</td>
<td>31%</td>
<td>51%</td>
<td>52%</td>
</tr>
</tbody>
</table>
Chapter 9

The Results And Interpretation

This chapter provides the results and interpretation obtained from this analysis.

9.1 Data Yields

Figures 9.1 and 9.2 show the 18.9 fb$^{-1}$ data and simulated yields for the every search region as a function of BDT discriminator. These are shown for completeness. The vertical red line indicate the particular cut values we use to define our search regions. The blue bar indicates the total simulation uncertainty (systematic and statistical). Some systematic uncertainties are only defined for the search region, so must be approximated here:

1. All QCD systematic are replaced by a flat 50% systematic uncertainty.
2. The propagated QCD prediction statistical uncertainty (taking into account multi-smearing) is approximated.
3. The kinematic closure uncertainty on the $Z \rightarrow \nu\bar{\nu}$ background in the search region: $t\bar{t} + W$ uncertainty multiplied by $Z$(uncertainty) / $(t\bar{t} + W)$(uncertainty).
4. No systematic uncertainty is placed on the single top yield, as we disregard the single top yields.
5. The systematic uncertainty to cover the statistical uncertainty of the kinematic scale factors, is not included.

The bottom planes of these figures show the corresponding ratios. The “data” line is the data yield divided by the prediction yield. The solid colored lines are the
proportions of the particular simulation processes over the total prediction. The blue band is the relative systematic uncertainty on the prediction. The “std. MC” line is the simulation without any corrections divided by the prediction. The “M(\bar{t})” line corresponds to that particular simulated signal. We observe no statistically significant deviations. Figures 9.4 and 9.3 show the corresponding plots sans lepton veto. These regions have higher statistics. They are useful to check the t\bar{t} + W kinematic closure. Simulation and data agree within 20% for all our discriminators.

The data and simulation yields in our search regions, for the analogous 18.9 fb\(^{-1}\) study, are given in Table 9.1 and 9.2. No statistically significant deviations are found. The 19.6 fb\(^{-1}\) data and simulation yields in our search region are given in Table 9.3 and 9.4. Still, no statistically significant deviations are observed. The addition of 657 pb\(^{-1}\) of data, which correspond to an additional 39549036 events in Dataset D of Table B.2, produce very little change in our search regions. The changes are summarized in Tables 9.5 and 9.6. At most one extra event is observed.

<table>
<thead>
<tr>
<th></th>
<th>T2tt_LM</th>
<th>T2tt_MM</th>
<th>T2tt_HM</th>
<th>T2tt_VHM</th>
</tr>
</thead>
<tbody>
<tr>
<td>tt, W+jets, and single t</td>
<td>19.8 ± 3.2</td>
<td>8.64 ± 1.81</td>
<td>3.21 ± 1.02</td>
<td>1.00 ± 0.53</td>
</tr>
<tr>
<td>Z+jets</td>
<td>0.69 ± 0.23</td>
<td>2.30 ± 0.90</td>
<td>1.92 ± 0.84</td>
<td>0.59 ± 0.28</td>
</tr>
<tr>
<td>t\bar{t}Z</td>
<td>1.34 ± 0.49</td>
<td>2.66 ± 1.27</td>
<td>1.62 ± 0.75</td>
<td>0.99 ± 0.52</td>
</tr>
<tr>
<td>QCD multijet</td>
<td>0.91 ± 0.58</td>
<td>0.17 ± 0.07</td>
<td>0.04 ± 0.02</td>
<td>0.01 ± 0.01</td>
</tr>
<tr>
<td>All SM backgrounds</td>
<td>22.7 ± 3.3</td>
<td>13.8 ± 2.4</td>
<td>6.8 ± 1.5</td>
<td>2.6 ± 0.8</td>
</tr>
<tr>
<td>Observed data</td>
<td>16</td>
<td>18</td>
<td>7</td>
<td>2</td>
</tr>
</tbody>
</table>

Table 9.1 : T2tt Search Region Yields for 18.9 fb\(^{-1}\) study
Figure 9.1: T2bW yields with all lepton vetos applied. The bottom panes give the ratio of observed over predicted yields. The error on the data line is the error on the ratio of the data statistical uncertainty and simulation statistical uncertainty. The blue error band is the relative systematic error on the prediction yield. [13].
Figure 9.2: T2tt yields with all lepton vetos applied. The bottom panes give the ratio of observed over predicted yields. The error on the data line is the error on the ratio of the data statistical uncertainty and simulation statistical uncertainty. The blue error band is the relative systematic error on the prediction yield. [13].
Figure 9.3: T2bW yields with no lepton vetos applied. The bottom panes give the ratio of observed over predicted yields. The error on the data line is the error on the ratio of the data statistical uncertainty and simulation statistical uncertainty. The blue error band is the relative systematic error on the prediction yield. [13].
Figure 9.4: T2tt yields with no lepton vetos applied. The bottom panes give the ratio of observed over predicted yields. The error on the data line is the error on the ratio of the data statistical uncertainty and simulation statistical uncertainty. The blue error band is the relative systematic error on the prediction yield. [13].
### Table 9.2: T2bW Search Region Yields for 18.9 fb$^{-1}$ study

<table>
<thead>
<tr>
<th></th>
<th>T2bW_LX</th>
<th>T2bW_LM</th>
<th>T2bW_MXHM</th>
<th>T2bW_HXHM</th>
<th>T2bW_VHM</th>
</tr>
</thead>
<tbody>
<tr>
<td>$t\bar{t}$, $W$+jets, and single $t$</td>
<td>6.41 ± 2.08</td>
<td>30.4 ± 4.0</td>
<td>3.41 ± 1.05</td>
<td>12.1 ± 2.8</td>
<td>2.00 ± 0.78</td>
</tr>
<tr>
<td>Z+jets</td>
<td>1.88 ± 0.93</td>
<td>4.57 ± 1.67</td>
<td>1.66 ± 0.72</td>
<td>1.77 ± 0.73</td>
<td>1.24 ± 0.54</td>
</tr>
<tr>
<td>$t\bar{t}$Z</td>
<td>0.59 ± 0.30</td>
<td>2.46 ± 1.11</td>
<td>0.83 ± 0.39</td>
<td>1.72 ± 0.79</td>
<td>0.62 ± 0.29</td>
</tr>
<tr>
<td>QCD multijet</td>
<td>0.71 ± 0.35</td>
<td>0.36 ± 0.19</td>
<td>0.10 ± 0.12</td>
<td>0.01 ± 0.01</td>
<td>0.01 ± 0.01</td>
</tr>
<tr>
<td>All SM backgrounds</td>
<td>9.6 ± 2.3</td>
<td>37.7 ± 4.4</td>
<td>6.0 ± 1.3</td>
<td>15.6 ± 3.0</td>
<td>3.9 ± 1.0</td>
</tr>
<tr>
<td>Observed data</td>
<td>12</td>
<td>47</td>
<td>6</td>
<td>14</td>
<td>4</td>
</tr>
</tbody>
</table>

### Table 9.3: T2tt Search Region Yields for 19.6 fb$^{-1}$ study

<table>
<thead>
<tr>
<th></th>
<th>T2tt_LM</th>
<th>T2tt_MM</th>
<th>T2tt_HM</th>
<th>T2tt_VHM</th>
</tr>
</thead>
<tbody>
<tr>
<td>$t\bar{t}$, $W$+jets, and single $t$</td>
<td>20.43 ± 3.35</td>
<td>8.93 ± 1.87</td>
<td>3.32 ± 1.05</td>
<td>1.03 ± 0.55</td>
</tr>
<tr>
<td>Z+jets</td>
<td>0.71 ± 0.24</td>
<td>2.38 ± 0.93</td>
<td>1.99 ± 0.87</td>
<td>0.61 ± 0.29</td>
</tr>
<tr>
<td>$t\bar{t}$Z</td>
<td>1.39 ± 0.51</td>
<td>2.75 ± 1.31</td>
<td>1.68 ± 0.78</td>
<td>1.02 ± 0.54</td>
</tr>
<tr>
<td>QCD multijet</td>
<td>0.94 ± 0.60</td>
<td>0.18 ± 0.07</td>
<td>0.04 ± 0.02</td>
<td>0.01 ± 0.01</td>
</tr>
<tr>
<td>All SM backgrounds</td>
<td>23.5 ± 3.5</td>
<td>14.3 ± 2.5</td>
<td>7.0 ± 1.6</td>
<td>2.7 ± 0.8</td>
</tr>
<tr>
<td>Observed data</td>
<td>16</td>
<td>19</td>
<td>8</td>
<td>3</td>
</tr>
</tbody>
</table>

### 9.2 Signal Efficiencies

We use the full search region selection on simulated signal (T2bW and T2tt) samples B.3.1 as well as additional corrections (Section 9.3) to arrive at expected signal yields in our search regions. Up until now, unless otherwise indicated, the MC procedure referred to “Full Simulation” (GEANT4 [133]). It is the standard used to model the full CMS detector. However, the simulate signal samples use “Fast Simulation”. Fast simulation takes into account detector response in a pragmatic and simplified way. It is designed to produce simulation 100 times faster than full simulation, but at the cost of accuracy for novel analyses such as this one. We use fast simulation, because it is faster to create. More information on the differences between fast simulation and full simulation can be found at [134].
<table>
<thead>
<tr>
<th>tt, W+jets, and single t</th>
<th>T2bW_LX</th>
<th>T2bW_LM</th>
<th>T2bW_MXHM</th>
<th>T2bW_HXHM</th>
<th>T2bW_VHM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Z+jets</td>
<td>6.63 ± 2.15</td>
<td>31.4 ± 4.14</td>
<td>3.53 ± 1.09</td>
<td>12.5 ± 2.90</td>
<td>2.07 ± 0.81</td>
</tr>
<tr>
<td>ttZ</td>
<td>1.94 ± 0.96</td>
<td>4.73 ± 1.73</td>
<td>1.72 ± 0.74</td>
<td>1.83 ± 0.75</td>
<td>1.28 ± 0.56</td>
</tr>
<tr>
<td>QCD multijet</td>
<td>0.61 ± 0.31</td>
<td>2.54 ± 1.15</td>
<td>0.86 ± 0.40</td>
<td>1.78 ± 0.82</td>
<td>0.64 ± 0.30</td>
</tr>
<tr>
<td>All SM backgrounds</td>
<td>9.9 ± 2.4</td>
<td>39.0 ± 4.6</td>
<td>6.2 ± 1.4</td>
<td>16.1 ± 3.1</td>
<td>4.0 ± 1.0</td>
</tr>
<tr>
<td>Observed data</td>
<td>12</td>
<td>48</td>
<td>6</td>
<td>15</td>
<td>4</td>
</tr>
</tbody>
</table>

Table 9.4: T2bW Search Region Yields for 19.6 fb$^{-1}$ study

<table>
<thead>
<tr>
<th>Observed data</th>
<th>T2tt_LM</th>
<th>T2tt_MM</th>
<th>T2tt_HM</th>
<th>T2tt_VHM</th>
</tr>
</thead>
<tbody>
<tr>
<td>18.9 fb$^{-1}$ study</td>
<td>16</td>
<td>18</td>
<td>7</td>
<td>2</td>
</tr>
<tr>
<td>19.6 fb$^{-1}$ study</td>
<td>16</td>
<td>19</td>
<td>8</td>
<td>3</td>
</tr>
<tr>
<td>Difference</td>
<td>0</td>
<td>+1</td>
<td>+1</td>
<td>+1</td>
</tr>
</tbody>
</table>

Table 9.5: Summary of T2tt Search Region Yields

The absolute counts of simulated signal samples that make it through the search regions selection (including correction) are not actually of interest. The acceptance efficiency ($\epsilon$) is the quantity of interest. For every signal (including distinction for different mass fractions $x$) and search region, we calculate ($\epsilon$) as a function of $m_{\tilde{\chi}^0_1}$ and $m_{\tilde{t}}$. ($\epsilon$) is the ratio of passing signal events ($P_{\text{pass}}$) to the total amount of signal events ($P_{\text{alt}}$):

$$\epsilon = \frac{P_{\text{pass}}}{P_{\text{alt}}}$$  \hspace{1cm} (9.1)

The efficiency in the 4 different T2tt regions is shown in Figure 9.5. The T2bW signal adds the extra mass splitting parameter $x$ (0.25, 0.50, 0.75). Each T2bW search region was trained on one of these three values of $x$ (Table 5.2). Additionally, the T2bW simulated signal was generated for only these three values of $x$. However,
just because a particular simulated signal event was generated with $x$ does not mean it will only make it into a search region trained on that $x$. For the simulated T2bw signal with $x = 0.25$, the efficiency in the 5 different T2bW search regions is shown in Figure 9.6. Figures 9.7 and 9.8 correspond to $x = 0.50$ and 0.75, respectively. While the efficiency is generally very low, it is non-zero.

### 9.3 Signal Uncertainties

The signal efficiencies were subject to several unique corrections. More information on the study of these corrections can be found in Section 10 of [13]. Additional plots on the corresponding uncertainties can be found in Appendix F of [13]. A great many of these corrections and uncertainties account for differences in physics reconstruction processes between data, full simulation, and fast simulation. Uncertainties are assessed as a function of $m_{\tilde{t}}$, $\tilde{\chi}_1^0$, and $x$. A summary of corrections and uncertainties is as follows:

1. **Parton Distribution Function (PDF)**: This is the uncertainty associated with modeling the parton distribution functions. It is evaluated following the recommendation of the PDF4LHC group [135]. The PDF sets CTEQ10 [136], MSTW2008 [137], and NNPDF2.3 [138], are considered.

2. **Initial state radiation (ISR)**: Studies of the hadronic recoil of heavy boosted systems [139] suggest the MadGraph $Z+$ jets simulation samples overestimate

<table>
<thead>
<tr>
<th>Observed data</th>
<th>T2bW_LX</th>
<th>T2bW_LM</th>
<th>T2bW_MXHM</th>
<th>T2bW_HXHM</th>
<th>T2bW_VHWM</th>
</tr>
</thead>
<tbody>
<tr>
<td>18.9 fb$^{-1}$ study</td>
<td>12</td>
<td>47</td>
<td>6</td>
<td>14</td>
<td>4</td>
</tr>
<tr>
<td>19.6 fb$^{-1}$ study</td>
<td>12</td>
<td>48</td>
<td>6</td>
<td>15</td>
<td>4</td>
</tr>
<tr>
<td>Difference</td>
<td>0</td>
<td>+1</td>
<td>0</td>
<td>+1</td>
<td>0</td>
</tr>
</tbody>
</table>

Table 9.6: Summary of Search Region Yields
Figure 9.5: Simulated $\tilde{t}\tilde{t}\rightarrow tt\tilde{\chi}_1^\pm\tilde{\chi}_1^0$ Signal Efficiencies. Top Left: LM. Top Right: MM. Bottom Left: HM. Bottom Right: VHM.
Figure 9.6: Simulated $\tilde{t}\tilde{t} \rightarrow \tilde{\chi}^+\tilde{\chi}^- b\bar{b}$ ($x = 0.25$) Signal Efficiencies. Top Left: LX. Top Right: LM. Middle Left: MXHM. Middle Right: VHM. Bottom: HXHM.
Efficiency

Figure 9.7: Simulated $\tilde{t}\tilde{t} \rightarrow \tilde{\chi}^+\tilde{\chi}^-bb$ ($x = 0.50$) Signal Efficiencies. Top Left: LX. Top Right: LM. Middle Left: MXHM. Middle Right: VHM. Bottom: HXHM
Efficiency

0
0.002
0.004
0.006
0.008
0.01
0.012
0.014
0.016
0.018

$[\text{GeV}]$

$t\sim m$

200 300 400 500 600 700 800

$[\text{GeV}]$

1
0
$\chi \sim m$

50 100 150 200 250 300 350 400

$x = 0.75$

$\chi \sim + + W \rightarrow +$

$\chi \sim +$

$\chi \sim b + \rightarrow t\sim ; t\sim t\sim \rightarrow pp$

Figure 9.8: Simulated $\bar{t}t \rightarrow \tilde{\chi}^+ \tilde{\chi}^- b\bar{b}$ ($x = 0.75$) Signal Efficiencies. Top Left: LX. Top Right: LM. Middle Left: MXHM. Middle Right: VHM. Bottom: HXHM.
the transverse boost of the $Z$ boson, especially at high $p_T$. Similar behavior was observed in $t\bar{t}$ and $WZ$ systems. It is attributed to a mis-modeling of the initial state radiation. The recipe to counter this effect is described in [140]. It is applied to both the $t\bar{t} \rightarrow t\tilde{\tau}_1^0\tilde{\chi}_1^0$ and $t\bar{t} \rightarrow \tilde{\tau}_1^+\tilde{\tau}_1^- b\bar{b}$ samples.

3. **Jet Energy Scale (JES)** (Section 4.5): The uncertainties associated with JES corrections must be accounted for. These correction factors and uncertainties for “standard” jets have been measured in data as described in [101]. The jet energy scale (JES) uncertainties are propagated by scaling the momenta of all jets. More information on this procedure is given in Section 10.3.3 of [13].

4. **Jet Energy Resolution (JER)** (Section 7): The core width of the jet energy resolution distribution is about 10% wider in data than simulation [101]. The uncertainty in the jet energy resolution is minimal (few %) for most of the signal masses.

5. **B-tagging** (Section 4.6): The scale factors and their associated uncertainties for the CVS b-tagging algorithm between the fast simulation, full simulation, and data, were studied extensively. For T2bW, the fast simulation to full simulation is studied in Appendix L of [13] and full simulation to data is given in [141]. For T2tt, fast simulation, full simulation, and data, are studied in Appendix L of [13]. Additionally, for T2tt picky jet b-tagging an additional uncertainty is calculated. This reflects the dependence of b-tagging scale factors on the relative sample composition between data and simulation. This is done through CSV shape comparison studies. Studies conclude a conservative value of 4% uncertainty is appropriate for the relative sample composition.

6. **CORRAL** (Section 4.7): The performance of CORRAL top reconstruction between full simulation and data is studied to extract the corresponding uncertainties. The top reconstruction efficiency as a function of the top $p_T$ is studied between the full simulation and data. The studies performed show the agreement observed is better than 5%. Therefore, a conservative value of 5% is used as the corresponding uncertainty.

7. **Fast Simulation Substructure**: The variables utilizing the sub-structure based picky jets and CORRAL top-tagging are not expected to be represented in fast simulation as they are in full simulation. Hence, the top reconstruction efficiency as a function of the top $p_T$ of the generated top is parameterize and studied between the two simulation types. The difference in the performance between full and fast simulation is compared to obtain an associated uncertainty.

8. **Statistical Uncertainty**: This is the statistical uncertainty on the simulated signal. It can be quite large, especially for high cross section T2bW signals.
9. **Luminosity Uncertainty**: A luminosity uncertainty of 2.6% is applied to the simulated signal yield.

Uncertainties are summed in quadrature to obtain the overall signal efficiency uncertainty. Figure 9.9 shows uncertainty for $\bar{t}t \rightarrow t\tilde{\chi}^0_1\tilde{\chi}^0_1$. The uncertainties for $\bar{t}t \rightarrow \tilde{\chi}^\pm\tilde{\chi}^0 -\bar{b}b$ ($x = 0.25, 0.50, 0.75$) are given in Figures 9.10, 9.11, and 9.12, respectively.

Figure 9.9: Simulated $\bar{t}t \rightarrow t\tilde{\chi}^0_1\tilde{\chi}^0_1$ Signal Uncertainties on Efficiency. Top Left: LM. Top Right: MM. Bottom Left: HM. Bottom Right: VHM.
Figure 9.10 : Simulated $\tilde{t}\tilde{t} \rightarrow \tilde{\chi}^+\tilde{\chi}^-\tilde{b}\tilde{b}$ ($x = 0.25$) Signal Uncertainties on Efficiency. Top Left: LX. Top Right: LM. Middle Left: MXHM. Middle Right: VHM. Bottom: HXHM.
Relative systematic uncertainty

\[ t \sim m \]

\[ \chi \sim m_0 \]

\[ x = 0.50 \]

\[ \chi \sim + \]

\[ W \to + \]

\[ \chi \sim b + \to t \]

\[ t \to pp \]

Figure 9.11: Simulated $\tilde{t} \tilde{t} \to \tilde{\chi}^+ \tilde{\chi}^- \bar{b} \bar{b}$ ($x = 0.50$) Signal Uncertainties on Efficiency. Top Left: LX. Top Right: LM. Middle Left: MXHM. Middle Right: VHM. Bottom: HXHM.
Figure 9.12: Simulated $\tilde{t}\tilde{t} \rightarrow \tilde{\chi}^0\tilde{\chi}^\pm b\bar{b}$ ($x = 0.75$) Signal Uncertainties on Efficiency. Top Left: LX. Top Right: LM. Middle Left: MXHM. Middle Right: VHM. Bottom: HXHM
9.4 Upper Limits on Data

The number of observed events in each search region ($N_{\text{obs}}$) was not found to be in significant excess of the expected background. Hence, in-lieu of discovery, limits are set.

The signal regions are parameterized by $m_{\tilde{t}}$, $m_{\tilde{\chi}_0^0}$, and $x$ (for T2bW only). For each point in this parameter space, the corresponding uncertainty on the signal efficiency is $\delta_\epsilon(m_{\tilde{t}}, m_{\tilde{\chi}_0^0}, x)$. The number of observed events in each signal region is $N_{\text{obs}}$. Combining this information estimates the number of observed events (with uncertainty) for every signal point:

$$N_{\text{obs}} \rightarrow N_{\text{obs}} \pm \delta_\epsilon(m_{\tilde{t}}, m_{\tilde{\chi}_0^0}, x).$$  \hspace{1cm} (9.2)

When setting limits, there are two general hypotheses that can be used to explain observing $N_{\text{obs}}$ [142]:

1. Null hypothesis ($H_0$): There is signal and background in the observed data: $N_{\text{obs}} = N_s + N_b$

2. Alternative hypothesis ($H_1$): There is only background in the observed data: $N_{\text{obs}} = N_b$

where $N_s$ represents an unknown number of signal events and $N_b$ represents the known background. $H_0$ is tested against $H_1$. The general goal being to exclude $H_0$ for some set of signal points with at least 95% confidence level (CL). CL is defined as 1 - p-value. This analysis uses the modified-frequentist CL$_s$ method [143] with a one-sided profile likelihood ratio test statistic [142] to accomplish something akin to this.
The statistical uncertainties of the observed number of events are modeled as poisson distributions. Systematic uncertainties of the background predictions are assumed multiplicative. They are modeled with lognormal distributions. The CL$_s$ method can be summarized with the help of two basic conditional probabilities:

\[
\text{CL}_{s+b} = P(\leq N_{\text{obs}} | N_s + N_b),
\]

\[
\text{CL}_b = P(\leq N_{\text{obs}} | N_b).
\]

CL$_{s+b}$ is the probability of $N_s + N_b$ fluctuating to $N_{\text{obs}}$ or less. CL$_b$ is the probability of $N_b$ fluctuating to $N_{\text{obs}}$ or less. CL$_{s+b}$ corresponds to the probability to get a result less compatible with $H_0$ than $H_1$. CL$_b$ corresponds to the probability to get a result less compatible with $H_1$ than $H_0$. However, the CL$_s$ method does not actually use CL$_{s+b}$ to calculate the 95% CL upper limit with respect to $H_0$. Instead CL$_s$ is used:

\[
\text{CL}_s = \frac{\text{CL}_{s+b}}{\text{CL}_b}.
\]

Strictly speaking CL$_s$ is not a probability, however it is treated as such. Furthermore, CL$_s$ provides a more conservative limit:

\[
\text{CL}_s \geq \text{CL}_{s+b}.
\]

The upper limit on $N_s$ at 95% confidence level using the CL$_s$ metric ($\text{CL}_s \leq 0.05$) is referred to as 95% CL$_s$(N$_{\text{obs}}$). Additionally, 95% CL$_s$(N$_{\text{exp}}$), 95% CL$_s$(N$_{\text{exp-std.}}$),
and 95% CL$_s$(N$_{\text{exp+std.}}$) are calculated. These are best understood in comparison to N$_{\text{obs}}$. N$_{\text{obs}}$ corresponds to a single data set. It is merely a single point on a distribution of all possible data sets that could have been arrived at. N$_{\text{exp}}$ corresponds to the median value of this N$_{\text{obs}}$ distribution. N$_{\text{exp}}$ characterizes the sensitivity of the experiment itself. N$_{\text{exp+std.}}$ and N$_{\text{exp-std.}}$ correspond to ±1 standard deviation (std.) from the median, respectively. The method for calculating 95% CL$_s$(N$_{\text{exp}}$) utilizes a data set corresponding to the median value of N$_{\text{obs}}$ called the “Asimov” dataset. 95% CL$_s$(N$_{\text{exp+std.}}$) and 95% CL$_s$(N$_{\text{exp-std.}}$) use analogous datasets representing the 0.16 and 0.84 quantile of the N$_{\text{obs}}$ distribution, respectively.

Figure 9.13 shows 95% CL$_s$(N$_{\text{obs}}$) for the $\bar{t}t \rightarrow t\tilde{\ell}^+\tilde{\chi}_1^0\tilde{\chi}_1^0$ signal. Figures 9.14, 9.15, and 9.16 show 95% CL$_s$(N$_{\text{obs}}$) for the $\bar{t}t \rightarrow \tilde{\chi}^+\tilde{\chi}^- b\bar{b}$ ($x = 0.25, 0.50, 0.75$) signals, respectively. Figures 9.17, 9.18, 9.19, and 9.20, are the corresponding plots for 95% CL$_s$(N$_{\text{exp}}$). Figures 9.21, 9.22, 9.23, and 9.24, are the corresponding plots for 95% CL$_s$(N$_{\text{exp-std.}}$). Figures 9.25, 9.26, 9.27, and 9.28, are the corresponding plots for 95% CL$_s$(N$_{\text{exp+std.}}$). No meaningful deviations or anomalies are observed in any of these upper limits.

For every search region, for every point in the parameter space, 95% CL$_s$(σ$_{\text{obs}}$), 95% CL$_s$(σ$_{\text{exp}}$), 95% CL$_s$(σ$_{\text{exp-std.}}$), and 95% CL$_s$(σ$_{\text{exp+std.}}$) are calculated. These are the cross section upper limits corresponding to 95% CL$_s$(N$_{\text{obs}}$), 95% CL$_s$(N$_{\text{exp}}$), 95% CL$_s$(N$_{\text{exp-std.}}$), and 95% CL$_s$(N$_{\text{exp+std.}}$), respectively. They are calculated as follows:
Figure 9.13: 95% CL$_{s}$ limit on N$_{\text{obs}}$ for $t\bar{t} \rightarrow t\bar{t}\chi_{0}^{0}\chi_{0}^{0}$. Top Left: LM. Top Right: MM. Bottom Left: HM. Bottom Right: VHM.
Figure 9.14: 95% CL$_s$ limit on $N_{obs}$ for $\tilde{t}\tilde{t}\rightarrow \tilde{\chi}^+\tilde{\chi}^-b\bar{b}$ ($x = 0.25$). Top Left: LX. Top Right: LM. Middle Left: MXHM. Middle Right: VHM. Bottom: HXHM.
Figure 9.15: 95% CL$_s$ limit on N$_{obs}$ for $\tilde{t}\bar{t} \rightarrow \tilde{\chi}^+\tilde{\chi}^- b\bar{b}$ ($x = 0.50$). Top Left: LX. Top Right: LM. Middle Left: MXHM. Middle Right: VHM. Bottom: HXHM
Figure 9.16: 95% CL$_s$ limit on $N_{obs}$ for $\bar{t}t \rightarrow \tilde{\chi}^+\tilde{\chi}^- b\bar{b}$ ($x = 0.75$). Top Left: LX. Top Right: LM. Middle Left: MXHM. Middle Right: VHM. Bottom: HXHM.
Figure 9.17: 95% CL$_{S}$ limit on $N_{\text{exp}}$ for $\tilde{t}\tilde{\tau} \rightarrow \tilde{t}\tilde{\tau}_{1}^0\tilde{\chi}_{1}^0$. Top Left: LM. Top Right: MM. Bottom Left: HM. Bottom Right: VHM.
Figure 9.18: 95% CL_{exp} limit on N_{exp} for \tilde{t} \tilde{\bar{t}} \rightarrow \tilde{\chi}^{\pm} \tilde{\chi}^{-} b \bar{b} (x = 0.25). Top Left: LX. Top Right: LM. Middle Left: MXHM. Middle Right: VHM. Bottom: HXHM
Figure 9.19: 95% CL_{exp} limit on N_{exp} for \tilde{t}\tilde{t} \rightarrow \tilde{\chi}^+\tilde{\chi}^- b\bar{b} (x = 0.50). Top Left: LX. Top Right: LM. Middle Left: MXHM. Middle Right: VHM. Bottom: HXHM.
Figure 9.20: 95% CL_{s} limit on N_{exp} for \tilde{t} \tilde{t} \rightarrow \tilde{\chi}^{+} \tilde{\chi}^{-} \tilde{b} \tilde{b} (x = 0.75). Top Left: LX. Top Right: LM. Middle Left: MXHM. Middle Right: VHM. Bottom: HXHM
Figure 9.21: 95% CLs limit on $N_{\text{exp-std.}}$ for $\bar{t}t \rightarrow \bar{t}t\bar{\chi}_1^0\bar{\chi}_1^0$. Top Left: LM. Top Right: MM. Bottom Left: HM. Bottom Right: VHM.
Figure 9.22: 95% CL$_s$ limit on $N_{\text{exp-std.}}$ for $\tilde{t}\tilde{t} \rightarrow b + \tilde{\chi}^-; \tilde{\chi}^- \rightarrow W^+ + \nu$ ($x = 0.25$). Top Left: LX. Top Right: LM. Middle Left: MXHM. Middle Right: VHM. Bottom: HXHM
Figure 9.23: 95% CL$_{s}$ limit on $N_{\text{exp-std.}}$ for $\bar{t}t \rightarrow \tilde{\chi}^{\pm} \tilde{\chi}^{-} b \bar{b}$ ($x = 0.50$). Top Left: LX. Top Right: LM. Middle Left: MXHM. Middle Right: VHM. Bottom: HXHM.
Figure 9.24: 95% CL$_s$ limit on $N_{\text{exp-std.}}$ for $\bar{t}t \rightarrow \tilde{\chi}^+ \tilde{\chi}^- b \bar{b}$ ($x = 0.75$). Top Left: LX. Top Right: LM. Middle Left: MXHM. Middle Right: VHM. Bottom: HXHM.
Figure 9.25: 95% CLs limit on $N_{\text{exp+std.}}$ for $t\bar{t} \rightarrow t\bar{t}\chi_1^-\chi_1^0$. Top Left: LM. Top Right: MM. Bottom Left: HM. Bottom Right: VHM.
Figure 9.26: 95% CL$_s$ limit on N$_{\text{exp+std.}}$ for $\tilde{t}\tilde{t} \rightarrow b + \tilde{\chi}^\pm; \tilde{\chi}^0 \rightarrow W^\pm + \tilde{\chi}^0_1$ ($x = 0.25$). Top Left: LX. Top Right: LM. Middle Left: MXHM. Middle Right: VHM. Bottom: HXHM
Figure 9.27: 95% CL$_s$ limit on $N_{\text{exp+std.}}$ for $\tilde{t}\tilde{t} \rightarrow \tilde{\chi}^+\tilde{\chi}^-b\bar{b}$ ($x = 0.50$). Top Left: LX. Top Right: LM. Middle Left: MXHM. Middle Right: VHM. Bottom: HXHM.
Figure 9.28: 95% CL$_s$ limit on $N_{\text{exp+std.}}$ for $\tilde{t}\tilde{t} \rightarrow \tilde{\chi}^+ \tilde{\chi}^- QQ$ ($x = 0.75$). Top Left: LX. Top Right: LM. Middle Left: MXHM. Middle Right: VHM. Bottom: HXHM.
95% $\text{CL}_s(\sigma) = \frac{95\% \text{ CL}_s(N)}{\epsilon L_{\text{int}}}$, \hspace{1cm} (9.7)

where $\epsilon$ is the signal efficiency given in Section 9.2. Figure 9.29 shows 95% $\text{CL}_s(\sigma_{\text{obs}})$ for the $\tilde{t}\tilde{\bar{t}} \rightarrow t\bar{t}\tilde{\chi}_1^0\tilde{\chi}_1^0$ signal. Figures 9.30, 9.31, and 9.32 show 95% $\text{CL}_s(\sigma_{\text{obs}})$ for the $\tilde{t}\tilde{\bar{t}} \rightarrow \tilde{\chi}^+\tilde{\chi}^- b\bar{b}$ ($x = 0.25, 0.50, 0.75$) signals, respectively. Figures 9.33, 9.34, 9.35, and 9.36, are the corresponding plots for 95% $\text{CL}_s(\sigma_{\text{exp}})$. Figures 9.37, 9.38, 9.39, and 9.40, are the corresponding plots for 95% $\text{CL}_s(\sigma_{\text{exp-std.}})$. Figures 9.41, 9.42, 9.43, and 9.44, are the corresponding plots for 95% $\text{CL}_s(\sigma_{\text{exp+std.}})$. As with 95% $\text{CL}_s(N_s)$, no meaningful deviations or anomalies are observed in any of these upper limits.

9.5 Best Regions and Quantities

The search regions overlap, so any given signal point may correspond to multiple search regions. This poses a serious problem. Calculating a particular quantity of interest ($\epsilon$, $\delta_\epsilon$, $\sigma_{\text{exp}}$, $\sigma_{\text{exp-std.}}$, and $\sigma_{\text{exp+std.}}$) for some signal point ($m_{\tilde{t}}$, $m_{\tilde{\chi}^0}$, $x$), depends on which search region you take it from. Therefore a “best” search region is defined for every signal point. Quantities of interest for a signal point are taken from this best search region. They are referred to as “best” quantities. The best regions used in this analysis are the ones that produce the strongest expected limit on the signal production cross section (lowest 95% $\text{CL}_s(\sigma_{\text{exp}})$). Figure 9.45 displays the best search regions in T2tt and T2bW mass planes. The color corresponds to the particular search region used. The best expected efficiencies are given in Figure 9.46. The best
Figure 9.29: 95% CL$_{s}$ limit on $\sigma_{\text{obs}}$ for $\tilde{t}\tilde{t} \rightarrow t\tilde{\chi}_1^0\tilde{\chi}_1^0$. Top Left: LM. Top Right: MM. Bottom Left: HM. Bottom Right: VHM.
Figure 9.30: 95% CLs limit on $\sigma_{\text{obs}}$ for $t\bar{t} \rightarrow \tilde{\chi}^+ \tilde{\chi}^- b\bar{b}$ ($x = 0.25$). Top Left: LX. Top Right: LM. Middle Left: MXHM. Middle Right: VHM. Bottom: HXHM.
Figure 9.31: 95% CL$_s$ limit on $\sigma_{\text{obs}}$ for $t\bar{t} \rightarrow \tilde{t}\tilde{t}$ ($x = 0.50$). Top Left: LX. Top Right: LM. Middle Left: MXHM. Middle Right: VHM. Bottom: HXHM
Figure 9.32: 95% CL$_s$ limit on $\sigma_{\text{obs}}$ for $t\bar{t} \rightarrow \tilde{\chi}^+\tilde{\chi}^-b\bar{b}$ ($x = 0.75$). Top Left: LX. Top Right: LM. Middle Left: MXHM. Middle Right: VHM. Bottom: HXHM.
Figure 9.33: 95% CL limit on $\sigma_{\exp}$ for $\tilde{t}\tilde{t} \rightarrow t\tilde{\chi}_1^0\tilde{\chi}_1^0$. Top Left: LM. Top Right: MM. Bottom Left: HM. Bottom Right: VHM.
Figure 9.34: 95% CLs limit on $\sigma_{\text{exp}}$ for $\bar{t} t \rightarrow \tilde{\chi}^+ \tilde{\chi}^- b \bar{b}$ ($x = 0.25$). Top Left: LX. Top Right: LM. Middle Left: MXHM. Middle Right: VHM. Bottom: HXHM
Figure 9.35: 95% CL$_s$ limit on $\sigma_{\exp}$ for $\tilde{t}\tilde{t} \rightarrow \tilde{\chi}^+\tilde{\chi}^0_{1/2}$ ($x = 0.50$). Top Left: LX. Top Right: LM. Middle Left: MXHM. Middle Right: VHM. Bottom: HXHM.
Figure 9.36: 95% CL$_s$ limit on $\sigma_{\text{exp}}$ for $t\bar{t} \rightarrow \tilde{\chi}^+ \tilde{\chi}^- b\bar{b}$ ($x = 0.75$). Top Left: LX. Top Right: LM. Middle Left: MXHM. Middle Right: VHM. Bottom: HXHM.
Figure 9.37: 95% CL$_s$ limit on $\sigma_{\text{exp-std.}}$ for $\bar{t}t \rightarrow \bar{t}t\tilde{\chi}^0_1\tilde{\chi}^0_1$. Top Left: LM. Top Right: MM. Bottom Left: HM. Bottom Right: VHM.
Figure 9.38: 95% CL$_s$ limit on $\sigma$ for $\tilde{t}\tilde{t}\rightarrow b+\tilde{\chi}^+\tilde{\chi}^-b\bar{b}$ ($x = 0.25$). Top Left: LX. Top Right: LM. Middle Left: MXHM. Middle Right: VHM. Bottom: HXHM
Figure 9.39: 95% CL$_s$ limit on $\sigma_{\text{exp-std.}}$ for $\tilde{t}\bar{t}\rightarrow \tilde{\chi}^+\tilde{\chi}^-b\bar{b} \ (x = 0.50)$. Top Left: LX. Top Right: LM. Middle Left: MXHM. Middle Right: VHM. Bottom: HXHM.
Figure 9.40: 95% CL$_{s}$ limit on $\sigma_{\text{exp-std.}}$ for $\tilde{t}\tilde{t} \to b + \tilde{\chi}^+\tilde{\chi}^- b\bar{b}$ $(x = 0.75)$. Top Left: LX. Top Right: LM. Middle Left: MXHM. Middle Right: VHM. Bottom: HXHM.
Figure 9.41: 95% CLa limit on $\sigma_{\text{exp+std.}}$ for $\bar{t}t \rightarrow \bar{t}t\tilde{\chi}_0^0\tilde{\chi}_0^0$. Top Left: LM. Top Right: MM. Bottom Left: HM. Bottom Right: VHM.
Figure 9.42: 95% CL$_s$ limit on $\sigma_{\text{exp+std.}}$ for $\tilde{t}\tilde{t}\rightarrow\tilde{\chi}^+\tilde{\chi}^-bb\ (x = 0.25)$. Top Left: LX. Top Right: LM. Middle Left: MXHM. Middle Right: VHM. Bottom: HXHM
Figure 9.43 : 95\% CL$_{s}$ limit on $\sigma_{\text{exp+std.}}$ for $\tilde{t}\tilde{t} \rightarrow \tilde{\chi}^+\tilde{\chi}^- b\bar{b}$ ($x = 0.50$). Top Left: LX. Top Right: LM. Middle Left: MXHM. Middle Right: VHM. Bottom: HXHM
Figure 9.44 : 95% CL$_s$ limit on $\sigma_{\exp+\text{std.}}$ for $t\bar{t} \rightarrow \tilde{\chi}^+ \tilde{\chi}^- b\bar{b}$ ($x = 0.75$). Top Left: LX. Top Right: LM. Middle Left: MXHM. Middle Right: VHM. Bottom: HXHM.
expected uncertainties are given in Figure 9.47. The best 95% CLs limit on $\sigma_{\text{obs}}$, $\sigma_{\text{exp}}$, $\sigma_{\text{exp-std.}}$, and $\sigma_{\text{exp-std.}}$ are given in Figures 9.48, 9.49, 9.50, and 9.50, respectively. Nothing unordinary is observed.

9.6 Limit calculation

The best 95% CLs upper limits on $\sigma$ are normalized to $\sigma_{\text{SUSY}}$. $\sigma_{\text{SUSY}}$ is the theoretical $t\bar{t}$ production cross section as a function of $t$ mass (Figure 2.8). If the normalized cross section is less than unity, then that point in parameter space is excluded. The 95% CLs upper limit on $\sigma$ shouldn’t be less than $\sigma_{\text{SUSY}}$. In practice, regions in parameter space excluded with at least 95% CLs are usually continuous and can be described by the “exclusion lines” ($\sigma/\sigma_{\text{SUSY}} = 1$) encompassing those regions. Additionally, the distributions of 95% CLs($\sigma_{\text{obs}}$) around those exclusion lines are used to calculate the corresponding quantities at ±1 standard deviations. These quantities are referred to as $\sigma_{\text{obs-std.}}$ and $\sigma_{\text{obs+std.}}$. They are analogous to $\sigma_{\text{exp-std.}}$ and $\sigma_{\text{exp+std.}}$, respectively. Figures 9.52, 9.53, and 9.54, show the best normalized 95% CLs limits on $\sigma_{\text{obs}}$, $\sigma_{\text{obs-std.}}$, and $\sigma_{\text{obs+std.}}$, respectively. Figures 9.55, 9.56, and 9.57, show the best normalized 95% CLs limits on $\sigma_{\text{exp}}$, $\sigma_{\text{exp-std.}}$, and $\sigma_{\text{exp+std.}}$, respectively. The exclusion line is superimposed on all these normalized $\sigma$ figures. Nothing unordinary is observed and results are consistent with the 18.9 fb$^{-1}$ study [13,17].

For every signal region, all exclusion lines are superimposed on the best 95% CLs limit on $\sigma_{\text{obs}}$. This is shown in Figure 9.58 and summarizes the results of this analysis.
Figure 9.45: Best Search Regions. Top Left: $\bar{t}t \rightarrow \bar{t}t \tilde{x}^0 \tilde{x}^0$. Top Right: $\bar{t}t \rightarrow \tilde{\chi}^+ \tilde{\chi}^- b \bar{b}$ ($x = 0.25$). Bottom Left: $\bar{t}t \rightarrow \tilde{\chi}^+ \tilde{\chi}^- b \bar{b}$ ($x = 0.50$). Bottom Right: $\bar{t}t \rightarrow \tilde{\chi}^+ \tilde{\chi}^- b \bar{b}$ ($x = 0.75$). For $\bar{t}t \rightarrow \bar{t}t \tilde{\chi}_1^0 \tilde{\chi}_1^0$: 1-LM, 2-MM, 3-HM, and 4-VHM. For $\bar{t}t \rightarrow \tilde{\chi}^+ \tilde{\chi}^- bb$: 1-LX, 2-LM, 3-MXHM, 4-VHM, and 5-HXHM.
Figure 9.46: Best Signal Efficiencies. Top Left: $\tilde{t}\tilde{t} \rightarrow t\tilde{t}_{\chi^0_{1}}$. Top Right: $\tilde{t}\tilde{t} \rightarrow \tilde{\chi}^+\tilde{\chi}^-bb \ (x = 0.25)$. Bottom Left: $\tilde{t}\tilde{t} \rightarrow \tilde{\chi}^+\tilde{\chi}^-bb \ (x = 0.50)$. Bottom Right: $\tilde{t}\tilde{t} \rightarrow \tilde{\chi}^+\tilde{\chi}^-bb \ (x = 0.75)$. 
Figure 9.47: Best Signal Uncertainties. Top Left: $t\bar{t} \rightarrow t\bar{t}\tilde{\chi}^0\tilde{\chi}^0$. Top Right: $t\bar{t} \rightarrow \tilde{\chi}^0\tilde{\chi}^0$ ($x = 0.25$). Bottom Left: $t\bar{t} \rightarrow \tilde{\chi}^+\tilde{\chi}^-b\bar{b}$ ($x = 0.50$). Bottom Right: $t\bar{t} \rightarrow \tilde{\chi}^+\tilde{\chi}^-b\bar{b}$ ($x = 0.75$).
Figure 9.48: Best 95% CL$_s$ limit on $\sigma_{\text{obs}}$. Top Left: $\bar{t}t \rightarrow \bar{t}t \chi^0_1 \chi^0_1$. Top Right: $\bar{t}t \rightarrow \bar{t}t \chi^+ \chi^- b\bar{b}$ ($x = 0.25$). Bottom Left: $\bar{t}t \rightarrow \chi^+ \chi^- b\bar{b}$ ($x = 0.50$). Bottom Right: $\bar{t}t \rightarrow \chi^+ \chi^- b\bar{b}$ ($x = 0.75$).
Figure 9.49: Best 95% CL_s limit on $\sigma_{\exp}$. Top Left: $\tilde{t}\tilde{t} \to t\bar{t}\chi^{0}\chi^{0}$. Top Right: $\tilde{t}\tilde{t} \to \tilde{t}\bar{t} + \tilde{W}^{+}\tilde{W}^{-} + \chi^{0}\chi^{0} (x = 0.25)$. Bottom Left: $\tilde{t}\tilde{t} \to \tilde{t}\bar{t} + \tilde{W}^{+}\tilde{W}^{-} + \chi^{0}\chi^{0} (x = 0.50)$. Bottom Right: $\tilde{t}\tilde{t} \to \tilde{t}\bar{t} + \tilde{W}^{+}\tilde{W}^{-} + \chi^{0}\chi^{0} (x = 0.75)$. 

\[ m_{t} [GeV] \]

\[ m_{\tilde{t}} [GeV] \]
Figure 9.50: Best 95% CLs limit on $\sigma_{\text{exp-std.}}$. Top Left: $\tilde{t}\tilde{t} \rightarrow t\bar{t}\tilde{\chi}_1^0\tilde{\chi}_1^0$. Top Right: $\tilde{t}\tilde{t} \rightarrow \tilde{t}\tilde{t}\chi^+\chi^- b\bar{b} \ (x=0.25)$. Bottom Left: $tt \rightarrow \tilde{t}\tilde{t}\chi^+\chi^- b\bar{b} \ (x=0.50)$. Bottom Right: $tt \rightarrow \tilde{t}\tilde{t}\chi^+\chi^- b\bar{b} \ (x=0.75)$.
Figure 9.51: Best 95% CLs limit on $\sigma_{\text{exp+std.}}$. Top Left: $\bar{t}t \rightarrow t\bar{t}\tilde{\chi}_1^0\tilde{\chi}_0^0$. Top Right: $\bar{t}t \rightarrow \tilde{t}\tilde{t}\rightarrow b\bar{b}\tilde{\chi}_+\tilde{\chi}_- \chi_1^0 \chi_1^0 \chi_1^0 (x = 0.25)$. Bottom Left: $\bar{t}t \rightarrow \tilde{t}\tilde{t}\rightarrow \tilde{t}\tilde{t}\rightarrow b\bar{b}\tilde{\chi}_+\tilde{\chi}_- \chi_1^0 \chi_1^0 \chi_1^0 (x = 0.50)$. Bottom Right: $\bar{t}t \rightarrow \tilde{t}\tilde{t}\rightarrow \tilde{t}\tilde{t}\rightarrow b\bar{b}\tilde{\chi}_+\tilde{\chi}_- \chi_1^0 \chi_1^0 \chi_1^0 (x = 0.75)$. 
Figure 9.52: Best 95% CL_s limit on $\sigma_{\text{obs}}/\sigma_{\text{SUSY}}$. Top Left: $\tilde{t}\tilde{t} \rightarrow t\bar{t}\tilde{\chi}^0_1\tilde{\chi}^0_1$. Top Right: $\tilde{t}\tilde{t} \rightarrow \tilde{t}\tilde{t} \tilde{\chi}^+\tilde{\chi}^- b\bar{b}$ ($x = 0.25$). Bottom Left: $\tilde{t}\tilde{t} \rightarrow \tilde{t}\tilde{t} \tilde{\chi}^+\tilde{\chi}^- b\bar{b}$ ($x = 0.50$). Bottom Right: $\tilde{t}\tilde{t} \rightarrow \tilde{t}\tilde{t} \tilde{\chi}^+\tilde{\chi}^- b\bar{b}$ ($x = 0.75$).
Figure 9.53: Best 95% CLs limit on $\sigma_{\text{obs-std.}}/\sigma_{\text{SUSY}}$. Top Left: $\bar{t}t \rightarrow t\bar{t}\chi_1^0\chi_1^0$. Top Right: $\bar{t}t \rightarrow \tilde{\chi}^+\tilde{\chi}^- b\bar{b}$ ($x = 0.25$). Bottom Left: $\bar{t}t \rightarrow \tilde{\chi}^+\tilde{\chi}^- b\bar{b}$ ($x = 0.50$). Bottom Right: $\bar{t}t \rightarrow \tilde{\chi}^+\tilde{\chi}^- b\bar{b}$ ($x = 0.75$).
Figure 9.54: Best 95% CLs limit on $\sigma_{\text{obs+std.}}/\sigma_{\text{SUSY}}$. Top Left: $\tilde{t}\tilde{t} \rightarrow t\tilde{t}\tilde{\chi}^0_1$. Top Right: $\tilde{t}\tilde{t} \rightarrow \tilde{\chi}^+\tilde{\chi}^-bb$ ($x = 0.25$). Bottom Left: $\tilde{t}\tilde{t} \rightarrow \tilde{\chi}^+\tilde{\chi}^-bb$ ($x = 0.50$). Bottom Right: $\tilde{t}\tilde{t} \rightarrow \tilde{\chi}^+\tilde{\chi}^-bb$ ($x = 0.75$).
Figure 9.55: Best 95% CLs limit on $\sigma_{\text{exp}}/\sigma_{\text{SUSY}}$. Top Left: $\bar{t}t \rightarrow t\bar{t}\chi^0_1\chi^0_1$. Top Right: $\bar{t}t \rightarrow \tilde{t}\tilde{t}+\tilde{b}\tilde{b} (x=0.25)$. Bottom Left: $\bar{t}t \rightarrow \tilde{t}\tilde{t}+\tilde{b}\tilde{b} (x=0.50)$. Bottom Right: $\bar{t}t \rightarrow \tilde{t}\tilde{t}+\tilde{b}\tilde{b} (x=0.75)$.
Figure 9.56: Best 95% CL$_s$ limit on $\sigma_{\text{exp-std.}}/\sigma_{\text{SUSY}}$. Top Left: $\bar{t}t \rightarrow \bar{t}t\chi^0_1\chi^0_1$. Top Right: $\bar{t}t \rightarrow \tilde{\chi}^\pm\tilde{\chi}^- bb$ ($x = 0.25$). Bottom Left: $\bar{t}t \rightarrow \tilde{\chi}^+\tilde{\chi}^- bb$ ($x = 0.50$). Bottom Right: $\bar{t}t \rightarrow \tilde{\chi}^+\tilde{\chi}^- bb$ ($x = 0.75$).
Figure 9.57: Best 95% CLs limit on $\sigma_{\text{exp+std.}}/\sigma_{\text{SUSY}}$. Top Left: $\bar{t}t \rightarrow \bar{t}t\tilde{\chi}_1^0\tilde{\chi}_1^0$. Top Right: $\bar{t}t \rightarrow \tilde{\chi}_+\tilde{\chi}^-bb$ ($x = 0.25$). Bottom Left: $\bar{t}t \rightarrow \tilde{\chi}_+\tilde{\chi}^-bb$ ($x = 0.50$). Bottom Right: $\bar{t}t \rightarrow \tilde{\chi}_+\tilde{\chi}_-bb$ ($x = 0.75$).
The expected (observed) exclusions line is in red (black). The red (black) dashed lines are the expected (observed) ±1 std. variations. Figure 9.59 gives the corresponding plots for the 18.9 fb$^{-1}$ study [13]. Unfortunately, no significant deviations are observed between this 19.6 fb$^{-1}$ study and the 18.9 fb$^{-1}$ study. The excluded regions are taken from the expected exclusion lines. For $\tilde{t}\tilde{t} \rightarrow t\tilde{t}\tilde{\chi}_1^0\tilde{\chi}_1^0$, the 95% CL$_s$ limit excludes regions $m_\tilde{t} < 755$ GeV for $m_{\tilde{\chi}_1^0} < 200$ GeV. For $\tilde{t}\tilde{t} \rightarrow \tilde{\chi}^+\tilde{\chi}^-b\bar{b}$ scenarios with $m_\tilde{t} < 625$ GeV are excluded. This is the same result as the 18.9 fb$^{-1}$ study, as it is essentially categorizing the experiment itself. The observed exclusions regions are generally smaller than those of the 18.9 fb$^{-1}$ study. This can be explained by smaller than expected potential $\tilde{t}\tilde{t} \rightarrow t\tilde{t}\tilde{\chi}_1^0\tilde{\chi}_1^0$ and $\tilde{t}\tilde{t} \rightarrow \tilde{\chi}^+\tilde{\chi}^-b\bar{b}$ signals in the additional data.
Figure 9.58: For the 19.6 fb⁻¹ study, the 95% CLσ limit on the $\tilde{t}\tilde{t}$ production cross section and expected exclusion areas in the $m_{\tilde{t}} - m_{\chi^0}$ plane. Top Left: $\tilde{t}\tilde{t} \to t\bar{t}\chi^0\bar{\chi}^0$. Top Right: $\tilde{t}\tilde{t} \to \tilde{t}\tilde{t}\chi^+\chi^- b\bar{b}$ ($x = 0.25$). Bottom Left: $\tilde{t}\tilde{t} \to \tilde{t}\tilde{t}\chi^+\chi^- b\bar{b}$ ($x = 0.50$). Bottom Right: $\tilde{t}\tilde{t} \to \tilde{t}\tilde{t}\chi^+\chi^- b\bar{b}$ ($x = 0.75$).
Figure 9.59: For the 18.9 fb$^{-1}$ study, the 95% CL$S$ limit on the $\tilde{t}\tilde{t}$ production cross section and expected exclusion areas in the $m_\tilde{t} - m_{\chi^0}$ plane. Top Left: $\tilde{t}\tilde{t} \rightarrow t\bar{t}\tilde{\chi}_1^0\tilde{\chi}_1^0$. Top Right: $\tilde{t}\tilde{t} \rightarrow \tilde{\chi}_+^0\tilde{\chi}_-^0b\bar{b} (x = 0.25)$. Bottom Left: $tt \rightarrow \tilde{\chi}_-^0\tilde{\chi}_-^0b\bar{b} (x = 0.50)$. Bottom Right: $tt \rightarrow \tilde{\chi}_+^0\tilde{\chi}_-^0b\bar{b} (x = 0.75)$. Here LSP on the y-axis corresponds to $\tilde{\chi}_1^0$. [13].
Appendix A

Methods

A.1 Tag and Probe

"Tag and Probe" (T&P) [144] is a generalized methodology for measuring the efficiency of a user defined criteria on a particular particle type (electron, muon, etc). T&P works through exploitation of mass resonances (J/Ψ, Z, etc) which decay into two particles of said particle type. T&P can be summarized for every event as follows:

1. **Find “tags”:** tags are well identified particles. Tags pass a “tight” identification criteria.

2. **Find “probes”:** probes are not well identified particles. The probes only passes a “loose” identification criteria. Probes are selected such that a valid T&P combination can be formed. These must have invariant mass consistent with the mass resonance.

3. **Find “passing probes”:** a probe that passes the user defined criteria.

The efficiency ($\epsilon$) of the user defined criteria is simply the ratio of all the passing probes ($P_{\text{pass}}$) to the total amount of probes ($P_{\text{all}}$):

$$\epsilon = \frac{P_{\text{pass}}}{P_{\text{all}}}$$  \hspace{1cm} (A.1)
A.2 MC Reweighting

At its simplest, MC reweighting is reweighting of simulated events in order for the simulated events to be better representative of the data. Reweighting is done by means of scale factors. These scale factors are obtained by comparing the data events and simulated events in specific control regions.

The electroweak and QCD multijets background process in this analysis share a common event reweighting method. This method accounts for the observed discrepancies between data events and simulated events for various SM processes. Simulated events are generally generated in a two step MC process, physics process generation and detector response to those physics processes. Each of these MC steps need to be corrected for. Hence, the MC corrections fall into two categories:

1. Modeling of process kinematics: $p_T$ spectrum of top quarks/W/Z bosons, multiplicity of radiated jets, b production cross section, etc.

2. Detector reconstruction effects: lepton/b-tag/trigger efficiencies, jet resolution, etc.

While it is tempting to combine all needed corrections into a few scale factors, that strategy is extremely error prone (especially for complex analysis) and difficulty to judge the validity of. This analysis chooses to take MC corrections on a step-by-step basis. Attempting to physically identify meaningful effect when the simulation needs to be corrected for and checking the relevant closures (data-simulation agreement as the result of the correction) for each correction. This methodology allows us to justify each MC correction as we go and it provides quantitative measures for the validity of
A.2.1 Process Kinematics

MC modeling of the SM (or beyond the SM) is no trivial matter. Many SM processes require phenomenological models (hard scale, showering, hadronization, etc.) and/or finite order approximation in Matrix Element calculation (QCD, etc), in order to get actual results. Generally speaking, the more complicated the physics process trying to be modeled, the more difficult it is to generate that corresponding simulated process. Essentially, our ability to accurately model these processes are limited by our computational ability, our use of phenomenological models, and the incompleteness of the SM itself.

The MC SM processes are corrected for by weighting the MC by the appropriate scale factors (SF). These SF are derived by making sure certain distributions in a particular control region agree between simulated events and data events. Additionally, these SF are parametrized by truth level quantities. These are quantities corresponding to the “generated particles” (gen objects) in the MC SM generation process. This is opposed to the parameterization done in reconstruction reweighting, using quantities from the detector reconstruction. By using the gen objects we keep certain SM constraints that might be distorted through detector reconstruction.

The general relationship between data events, simulated events, and scale factor
The data yield in control or search region $i$, $U_i$, is given by

$$U_i = \sum_j (R_j^i SF^j), \quad (A.2)$$

where $U_i$ is the data yield in control or search region $i$, $R_j^i$ is the simulated yield in control region $i$ for process $j$ (entries of the response matrix), and $SF^j$ is the scale factor for process $j$.

For example, say we wanted to derive $SF$'s for the SM processes $t\bar{t}$, $W$, $Z(p_T < 200 \text{ GeV})$, $Z(p_T \geq 200 \text{ GeV})$, and other (everything else). This would be done by comparing simulated events to data events in five corresponding control regions A, B, C, D, and E. This could be written as follows:

$$
\begin{bmatrix}
U^A \\
U^B \\
U^C \\
U^D \\
U^E
\end{bmatrix} =
\begin{bmatrix}
R^A_{t\bar{t}} & R^A_W & R^A_Z(\text{LOW}) & R^A_Z(\text{HIGH}) & R^A_{\text{other}} \\
R^B_{t\bar{t}} & R^B_W & R^B_Z(\text{LOW}) & R^B_Z(\text{HIGH}) & R^B_{\text{other}} \\
R^C_{t\bar{t}} & R^C_W & R^C_Z(\text{LOW}) & R^C_Z(\text{HIGH}) & R^C_{\text{other}} \\
R^D_{t\bar{t}} & R^D_W & R^D_Z(\text{LOW}) & R^D_Z(\text{HIGH}) & R^D_{\text{other}} \\
R^E_{t\bar{t}} & R^E_W & R^E_Z(\text{LOW}) & R^E_Z(\text{HIGH}) & R^E_{\text{other}}
\end{bmatrix}
\begin{bmatrix}
SF^{t\bar{t}} \\
SF^W \\
SF^{Z(\text{LOW})} \\
SF^{Z(\text{HIGH})} \\
SF^{\text{other}}
\end{bmatrix}, \quad (A.3)
$$

where LOW means $p_T < 200 \text{ GeV}$ and HIGH means $p_T \geq 200 \text{ GeV}$. At first glance we may be tempted to merely do a matrix inversion and solve for the SF’s:

$$SF^j = \sum_i ((R^{-1})_j^i U^i). \quad (A.4)$$

However, matrix inversion actually leads to chaotic oscillations in the SF. This makes the SF unusable and meaningless. Matrix inversion, fundamentally misconstrues
what the variables physical represent. The yields aren’t constants arising from the underlying physics. The yields are actually random numbers sampled from poisson distributions (where the mean is the cross section times efficiency). Hence, even if the simulation was modeled perfectly (all SF = 1):

\[ U^i \neq \sum_j R^i_j. \]  

The poisson distribution represents the underlying physics. The yields only represent the particular random experiment we did. If we did the experiment again, we would get another yield. The SF’s we want are not for our random experiment yield, but for the mean of the poisson distribution that produced that yield.

Instead of matrix inversion, we do “stable unfolding” [145, 146] to solve for the SF’s. Unfolding refers to the problem of estimating probability distributions where no parametric form is available and data are subject to additional random fluctuations (due to limited resolution). STable refers to the stability of results (they aren’t chaotic like the matrix inversion). The particular type of unfolding we use is called “iterative bayesian unfolding”. This is because our unfolding procedure uses iteration as well as elements of bayesian analysis. Starting with probability theory:

\[ SF^j = \frac{1}{\epsilon_j} \sum_i \left( P(C_j|E_i) U^i \right), \]

where \( C_j \) is the cause (true yield in process \( j \)), \( E_i \) is the effect (found yield in region \( i \)), and:
\[ \epsilon_j = \sum_i R_{ij} \]  
(A.7)

Using Bayes theorem we can evaluate the conditional probability:

\[ P(C_j|E_i) = \frac{P(E_i|C_j) P(C_j)}{P(E_i)} \]  
(A.8)

where:

\[ P(E_i) = \sum_j (P(E_i|C_j) P(C_j)). \]  
(A.9)

Hence,

\[ P(C_j|E_i) = \frac{P(E_i|C_j) P(C_j)}{\sum_j (P(E_i|C_j) P(C_j))}. \]  
(A.10)

We also know that

\[ P(C_j) = SF^j, \]  
(A.11)

\[ P(E_i|C_j) = R_{ij}. \]  
(A.12)

Hence,

\[ SF^j = \frac{1}{\epsilon_j} \sum_i \left( \frac{R_{ij}^j SF^j}{\sum_k (R_{ik}^j SF^k)} U^i \right). \]  
(A.13)

This equation is made iterative by making the input SF’s the current iteration \([N]\) and the output SF’s the next iteration \([N+1]\):

\[ SF^j[N+1] = SF^j[N] \frac{1}{\epsilon_j} \sum_i \left( \frac{R_{ij}^j U^i}{\sum_k (R_{ik}^j SF^k[N])} \right). \]  
(A.14)
Additionally, we set all initial $[0]$ scale factors to unity:

$$SF^j[0] = 1.$$  \hspace{1cm} (A.15)

It was found that our scale factors quickly converge with this method, so we only require four iterations at most in this analysis.

### A.2.2 Reconstruction Effects

When possible, reconstruction effects are accounted for through means of CMS standardized scale factors. These are given throughout this thesis. These often come from different CMS analyses or CMS POGs. When not possible, custom measurements are performed. Reconstruction effects include custom lepton BDT selection efficiency measurements see (Section 6.5.1) as well as muon/electron trigger, identification, and reconstruction measurements.

The remainder of this subsection will focus on the technique of “multi-smearing”. Many $E_T$ corrections applied to simulated events use random number generator in their input. This can pose a fundamental problem in the way we calculate efficiencies and corrections.

If our MC correction is dependent on random numbers, then whether a particular event passes a cut or enters a control/search region could be dependent on that random distribution. Generally, this is not a problem we need to worry about if we have high statistics or cases when the correction is going to be sampled many times. However, in the case of low simulated event statistics control regions or final
predictions where the efficiency of a single event entering the region is an important component of the total event yield, the problem needs to be addressed. In these cases, we want to perform the correction many times and establish the efficiency from an ensemble of corrections. “Multi-smearing” was developed to do just that in low statistic regions. In multi-smearing, for every individual simulated event we produce a multitude of corrected events.

We define the yield $N_{j,i}$ of the event “$i$” entering a search region or histogram bin “$j$” as

$$N_{j,i} = w_i \epsilon_{j,i}, \quad \text{(A.16)}$$

where

- $w_i =$ Weight for that simulated event (including all appropriate scale factors and 1 to 1 corrections)
- $\epsilon_{j,i} = R_{j,i}/N_s :$ efficiency for the event entering that region
- $R_{j,i} =$ Number of times event $i$ enters region $j$ out of a total of $N_s$ corrections

The uncertainty of event $i$ entering a search region or histogram bin $j$ is

$$E_{j,i} = \sqrt{(w_i \epsilon_{j,i})^2 + w_i^2 \epsilon_{j,i} (1 - \epsilon_{j,i})/N_s}. \quad \text{(A.17)}$$

The first term is just the uncertainty we would expect, while the second term is the binomial error associated with the efficiency. This analysis uses a multi-smearing of 50 correction samplings for all QCD prediction, 10 for establishing QCD systematics, and 50 for all non-QCD processes.
A.3 Machine Learning

Machine learning (ML) is a sub-field of computer science. ML arose from the study of pattern recognition as well as computational learning theory in artificial intelligence (AI). AI can be thought of as the simulation of human intelligence processes by computers. ML was perhaps best described in 1959 by Arthur Samuel as a “Field of study that gives computers the ability to learn without being explicitly programmed” [147].

ML focuses on computer algorithms that can learn from and make predictions on data. These algorithms generally work by constructing a model from a given set of inputs in order to make “data-driven” predictions or decisions. This is opposed to non-ML algorithms that strictly follow static programming instructions. In practice, the line between ML and non-ML algorithms can be contextual.

The two general classes of “data” we talk about in ML are unlabeled data and labeled data. Here a “label” or “tag” \( (l_i) \) is just some extra variable that provides further information on a particular piece of data. For instance, a collection of random pictures could be described as unlabeled data, while a collection of random pictures with tags of content could be described as labeled data. This schema is of course relative to the type of ML investigation being done. ML includes unsupervised learning, supervised learning, and reinforcement learning. Additionally, we refer to the term “classifier” as a algorithm designed to classify things.

Unsupervised learning is used to draw inferences from unlabeled data. Unsuper-
vised learning is essentially about “hidden description”. This is classifying the data according to it’s inherent structure. Ideally we want something analogous to a feature matrix $D = D_{ji}$. The goal of unsupervised learning is to create a classifier, because no classifier is known. Unsupervised learning uses unlabeled dataset to derive algorithms we hope to be classifiers. The ability of these algorithms to be used as classifiers can validated through labeled “test datasets”. The general workflow of an unsupervised learning problem, with example, is given below.

- **Goal is to find a classifier**: Need to find classifiers for a collection of unlabeled data
  
  1. **Build algorithm based on unlabeled input data**: Build algorithm based on unlabeled dataset. Eventually, we might get a algorithm that seems to select what roughly corresponds to a label. The algorithm may have discovered the label class, without being explicitly given it. Hence, unsupervised.
  
  2. **Test algorithm on a set of labeled test data. Algorithm becomes a classifier if it classifies**: Test our algorithm on data with different labels. Including data that contains the label we discovered.
  
  3. **If the classifier is satisfactory, deploy the classifier**

Supervised learning is used to draw inferences from labeled data. Supervised learning is essentially about “function approximation”. This means finding a function $(g)$ that maps labels and features:

$$g(D_j) \approx l_j,$$  \hspace{1cm} \text{(A.18)}

such that we will be able to predict the label associated with a new feature row $D_{j\prime}$ by $g(D_{j\prime})$. The goal of supervised learning is to sort data. Supervised learning uses
a known “training dataset” with labels. The training dataset is basically examples of what we want our algorithm to accomplish (known \((D_j,l_j)\) pairs). Supervised leaning incorporates these examples to derive a more general mapping between data and labels. As with unsupervised learning, we can validated our algorithms with “test datasets”. Once we have an adequate mapping we can use it to label unlabeled data. The general workflow of an supervised learning problem, with example, is given below.

- **Goal is to classify the data:** Need to classify a collection of data labeled by type
  1. *Build classifier that map data to labels, by means of a labeled training set of data:* Build classifiers that maps the data (input) to the type of label (output). At this point we’d likely have a classifier for many different labels. The algorithm is not deriving new labels, it is merely using the ones given to it. Hence, supervised.
  2. *Test classifier on a set of test data:* Test our classifiers on labeled data
  3. *If the classifier is satisfactory, deploy the classifier*

Supervised learning includes two algorithm categories, classification and regression. In classification the algorithm produces categorical response values. In Regression the algorithm produces continuous-response values. This analysis uses supervised learning, particularly boosted decision trees via classification. More information on the fascinating field of machine learning can be found at [148].

### A.3.1 Decision Tree Basics

A decision tree is a flowchart like algorithm that uses a tree-like graph to build classification or regression models. At its base form the decision tree is a collection
of conditional expressions, which determine how data should be split based upon input variables. These are called “decisions nodes” and they split data (binarily) into “decision branches”. Decision branches often lead to other decision nodes, which splits the data even further. This splitting process continues until some threshold is reached or some pre-specified level of deepness (how many time splitting can occur) is reached. The first decision node in called a “root node” and the final node is called the “leaf node”. The leaf nodes for the classifiers we use only have two possible outcomes, +1 for signal and -1 for background, depending on content. A visual representation of a general decision tree is presented in Figure A.1.

![Simple Decision Tree](image)

Figure A.1 : Simple Decision Tree
We are dealing with huge datasets, hence our decision trees are best modeled via statistical machine learning:

\[ Y = f(X) + \epsilon, \]

where \( Y \) is the output response (predicted values), \( X \) is the input values (predictors), \( f \) is the unknown relationship, and \( \epsilon \) is a predictor independent random error.

Our decision trees fully separate the predictor space, \( R_j \), into disjoint regions (Figure A.2). When the decision tree is run on input, the “predicted value” is simply the mean value of the training output that fell into this region. The maximum prediction accuracy on our decision trees is then given by the minimization of the residual square sum (RSS):

\[ \sum_{j=1}^{J} \left[ \sum_{j \in 1} (y_i - \hat{y}_{R_j})^2 \right], \]

where \( y_i \) is the observation for \( Y \) and \( \hat{y}_i \) is prediction for \( Y \) based on some \( x_i \).

Hence, the goal is now to find a set of \( R_j \) that sufficiently minimizes RSS. In practice this is not feasible as it requires comparing the results from an infinite number of possible \( R_j \) collections. Hence, we use a method called “top-down greedy splitting” (TGS). TGS does not attempt to optimize the RSS of the whole tree, but attempts to optimize the RSS for each step (each decision). TGS starts from the top of the tree and at each step divides the predictor space as to minimize RSS. The TGS method can be done relatively quickly and it produces good prediction on our training set. However, a good predictor for the training set doesn’t necessarily translate to a good
predictor for actual data. The model could “overfit”. Overfitting is just memorizing the training data as opposed to discovering some underlying relationship. The tree could be too complex for it’s own good.

In order to combat over-fitting we use “pruning”. Pruning removes sections of the decision tree that provide little classification power. Here we are trying to minimize:

\[ |T| \sum_{m=1}^{[T]} \sum_{i: x_i \in R_m} (y_i - \hat{y}_{R_m})^2 + \alpha |T|, \]  

(A.21)

where \( T \) is a subtree of \( T_0 \) (original tree), \(|T|\) is the number of leave in the sub-tree \( T \), and \( \alpha \) is a non-negative tuning parameter. When \( \alpha = 0 \), the sub-tree is equal to the
original tree. As $\alpha$ increases, the minimization corresponds to some sub-tress with less leaves. When $\alpha = \infty$, the sub-tree is just the tree with one leaf node. We determine $\alpha$ by the k-fold cross-validation method. The training data is partitioned into k equal sized sub-samples: 1 validation sample k-1 training samples. We then construct a set of trees and prune (generating a series of $T_i$ and $\alpha_i$). The mean square error is then computed for the validation set. This whole process is actually repeated k times (the cross validation), rotating which sample used as the validation set. This produces k estimates of the mean square error. This result is averaged for each value of $\alpha$. $\alpha$ is then picked such that it minimizes this average error. The sub-tree corresponding to this alpha is the chosen as the result of our pruning. However, even with pruning our decision trees with suffer from high variance. We therefore use Boosted Decisions trees as described in the next section.

A.3.2 Boosted Decision Trees

Decision trees usually suffer from high variance. This can be seen if we half the training set and produces two trees. Typically these trees might sufficiently differ. In order to combat such a problem we use “Boosted Decision Trees” (BDTs). Instead of fitting all our data to a single tree, we sequentially construct a collection of small trees. The idea is to combine many weak classifiers into a strong classifier. Each tree is based on information of previously constructed smaller trees.

We work with a collection of residuals $r_i$ and small sub-trees $\hat{f}_b$. Residuals here
corresponding to events the tree failed to identify properly. In each iteration, \( \hat{f}_b \) is fit to training data \((X, r)\) and has \(d\) internal nodes. This is the training data with extra weight given to the residuals. The tree and residuals are updated accordingly:

\[
\hat{f}(x)^{\text{new}} = \hat{f}(x)^{\text{old}} + \lambda \hat{f}_b(x),
\]

\[
r_i^{\text{new}} = r_i^{\text{old}} - \lambda \hat{f}_b(x_i),
\]

where \(\lambda\) is the “shrinkage parameter”. The shrinkage parameter is a small positive number that controls the rate of the BDT learning. The final tree is just the weighted sum of all the trees. Each tree can be quite small: \(d = 1\) (“decision stump”) is not uncommon. We used AdaBoost [149] for all our BDTs.
Appendix B

Samples

B.1 Preselection

The online preselection for all search samples is:

- $\mathcal{E}_T > 80\text{Gev}$
- $\geq 2$ jets ($p_T > 50$, $|\eta| < 2.4$)

The offline preselection for all search samples is:

- No isolated leptons: Suppress background with genuine $\mathcal{E}_T$ coming from W decays.
- $\mathcal{E}_T > 175$
- $\geq 2$ jets ($p_T > 70$, $|\eta| < 2.4$), where the online selection is fully efficient
- $Min[|\Delta\phi(\mathcal{E}_T,jet1)|,|\Delta\phi(\mathcal{E}_T,jet2)|] > 0.5$: Suppress rare QCD multijet events with severely mismeasured high $p_T$ jets.
- $|\Delta\phi(\mathcal{E}_T,jet3)| > 0.3$: suppress rare QCD multijet events with severely mismeasured high $p_T$ jets

B.2 The Data Samples

The data sets, triggers, and luminosities, where we looked for our signals (T2bW and T2tt), are listed in Table B.1 and Table B.2. These are referred to as our signal sample: CMS Run 1 Data and CMS Run 1 Parked Data. The CMS Run 1 Data was previously used for the [13] analysis. The CMS Run 1 Parked Data is novel to this
analysis in that it includes and extra 657 pb$^{-1}$ of data. The MET Run D Dataset is a subset of the METParked Run D Dataset.

The single electron, single muon, and $H_T$ dataset are given in Tables B.3, B.4, and B.5 respectively. The $H_T$ dataset was collected with prescaled triggers, so no luminosity is given.

<table>
<thead>
<tr>
<th>Dataset</th>
<th>trigger</th>
<th>$L$(fb$^{-1}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>/MET/Run2012A-13Jul2012-v1/AOD</td>
<td>HLT_DiCentralPFJet50_PFMET80_v*</td>
<td>0.797</td>
</tr>
<tr>
<td>/MET/Run2012A-recover-06Aug2012-v1/AOD</td>
<td>HLT_DiCentralPFJet50_PFMET80_v*</td>
<td>0.081</td>
</tr>
<tr>
<td>/MET/Run2012B-13Jul2012-v1/AOD</td>
<td>HLT_DiCentralPFJet50_PFMET80_v*</td>
<td>4.411</td>
</tr>
<tr>
<td>/MET/Run2012C-24Aug2012-v1/AOD</td>
<td>HLT_DiCentralPFNoPUJet50_PFMETORPFMETNoMu80_v*</td>
<td>0.474</td>
</tr>
<tr>
<td>/MET/Run2012C-PromptReco-v2/AOD</td>
<td>HLT_DiCentralPFNoPUJet50_PFMETORPFMETNoMu80_v*</td>
<td>6.330</td>
</tr>
<tr>
<td>/MET/Run2012C-EcalRecover 11Dec2012-v1/AOD</td>
<td>HLT_DiCentralPFNoPUJet50_PFMETORPFMETNoMu80_v*</td>
<td>0.133</td>
</tr>
<tr>
<td>/MET/Run2012D-PromptReco-v1/AOD</td>
<td>HLT_DiCentralPFNoPUJet50_PFMETORPFMETNoMu80_v*</td>
<td>6.712</td>
</tr>
<tr>
<td>Total Luminosity</td>
<td></td>
<td>18.948</td>
</tr>
</tbody>
</table>

Table B.1 : CMS: Run 1 - Dataset

<table>
<thead>
<tr>
<th>Dataset</th>
<th>trigger</th>
<th>$L$(fb$^{-1}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>/MET/Run2012A-13Jul2012-v1/AOD</td>
<td>HLT_DiCentralPFJet50_PFMET80_v*</td>
<td>0.797</td>
</tr>
<tr>
<td>/MET/Run2012A-recover-06Aug2012-v1/AOD</td>
<td>HLT_DiCentralPFJet50_PFMET80_v*</td>
<td>0.081</td>
</tr>
<tr>
<td>/MET/Run2012B-13Jul2012-v1/AOD</td>
<td>HLT_DiCentralPFJet50_PFMET80_v*</td>
<td>4.411</td>
</tr>
<tr>
<td>/MET/Run2012C-24Aug2012-v1/AOD</td>
<td>HLT_DiCentralPFNoPUJet50_PFMETORPFMETNoMu80_v*</td>
<td>0.474</td>
</tr>
<tr>
<td>/MET/Run2012C-PromptReco-v2/AOD</td>
<td>HLT_DiCentralPFNoPUJet50_PFMETORPFMETNoMu80_v*</td>
<td>6.330</td>
</tr>
<tr>
<td>/MET/Run2012C-EcalRecover 11Dec2012-v1/AOD</td>
<td>HLT_DiCentralPFNoPUJet50_PFMETORPFMETNoMu80_v*</td>
<td>0.133</td>
</tr>
<tr>
<td>/METParked/Run2012D-22Jan2013-v1/AOD</td>
<td>HLT_DiCentralPFNoPUJet50_PFMETORPFMETNoMu80_v*</td>
<td>7.969</td>
</tr>
<tr>
<td>Total Luminosity</td>
<td></td>
<td>19.395</td>
</tr>
</tbody>
</table>

Table B.2 : CMS: Run 1 - Parked Dataset

<table>
<thead>
<tr>
<th>Dataset</th>
<th>trigger</th>
<th>$L$(fb$^{-1}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>/SingleElectron/Run2012A-13Jul2012-v1/AOD</td>
<td>HLT_Ele27_WP80_v*</td>
<td>0.796</td>
</tr>
<tr>
<td>/SingleElectron/Run2012A-recover-06Aug2012-v1/AOD</td>
<td>HLT_Ele27_WP80_v*</td>
<td>0.081</td>
</tr>
<tr>
<td>/SingleElectron/Run2012C-24Aug2012-v1/AOD</td>
<td>HLT_Ele27_WP80_v*</td>
<td>0.474</td>
</tr>
<tr>
<td>/SingleElectron/Run2012C-PromptReco-v2/AOD</td>
<td>HLT_Ele27_WP80_v*</td>
<td>0.133</td>
</tr>
<tr>
<td>Total Luminosity</td>
<td></td>
<td>18.937</td>
</tr>
</tbody>
</table>

Table B.3 : Single Electron Dataset
<table>
<thead>
<tr>
<th>Dataset</th>
<th>trigger</th>
<th>$L$(fb$^{-1}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>/SingleMu/Run2012A-13Jul2012-v1/AOD</td>
<td>HLT_IsoMu24_eta2p1_v*</td>
<td>0.796</td>
</tr>
<tr>
<td>/SingleMu/Run2012A-recover-06Aug2012-v1/AOD</td>
<td>HLT_IsoMu24_eta2p1_v*</td>
<td>0.081</td>
</tr>
<tr>
<td>/SingleMu/Run2012B-13Jul2012-v1/AOD</td>
<td>HLT_IsoMu24_eta2p1_v*</td>
<td>4.411</td>
</tr>
<tr>
<td>/SingleMu/Run2012C-24Aug2012-v1/AOD</td>
<td>HLT_IsoMu24_eta2p1_v*</td>
<td>0.474</td>
</tr>
<tr>
<td>/SingleMu/Run2012C-PromptReco-v2/AOD</td>
<td>HLT_IsoMu24_eta2p1_v*</td>
<td>0.133</td>
</tr>
<tr>
<td>/SingleMu/Run2012D-PromptReco-v1/AOD</td>
<td>HLT_IsoMu24_eta2p1_v*</td>
<td>6.712</td>
</tr>
<tr>
<td>Total Luminosity</td>
<td></td>
<td>18.937</td>
</tr>
</tbody>
</table>

Table B.4 : Single Muon Dataset

<table>
<thead>
<tr>
<th>Dataset</th>
</tr>
</thead>
<tbody>
<tr>
<td>/HT/Run2012A-13Jul2012-v1/AOD</td>
</tr>
<tr>
<td>/HT/Run2012A-recover-06Aug2012-v1/AOD</td>
</tr>
<tr>
<td>/JetHT/Run2012B-13Jul2012-v1/AOD</td>
</tr>
<tr>
<td>/JetHT/Run2012C-24Aug2012-v2/AOD</td>
</tr>
<tr>
<td>/JetHT/Run2012C-EcalRecover 11Dec2012-v1/AOD</td>
</tr>
<tr>
<td>/JetHT/Run2012C-PromptReco-v2/AOD</td>
</tr>
<tr>
<td>/JetHT/Run2012D-PromptReco-v1/AOD</td>
</tr>
</tbody>
</table>

Table B.5 : $H_T$ Dataset

### B.3 The Simulated Samples

There are several simulated samples used in this analysis. These are generated by various ‘Monte Carlo” (MC) programs. MC refers a broad class of computational algorithms that make use of repeated random sampling to obtain numerical results. We use MC programs to generate synthetic data. We use simulated signal samples as well as simulated background related samples.

Additionally, all simulated samples are reweighed for pileup using the standard CMS procedure [150] when applicable. First, the distribution of the average number of interactions per bunch crossing is obtained for the corresponding dataset. Second, each event is weighted such that the distribution of the number of generated
interactions per event in simulation matches the corresponding distribution in data.

Furthermore, all simulated samples are corrected for integrated luminosity when applicable. This means weighting the simulated events so that it gives numbers we would expect at a particular dataset luminosity. To make the integrated luminosity of the simulated sample match the data, simulated samples are weighted as follows:

\[
weight = \frac{\mathcal{L}_{data} \times \sigma}{N_{MC}}, \quad (B.1)
\]

where \(\mathcal{L}_{data}\) is the integrated luminosity of the data, \(N_{MC}\) is the number of events in the simulated sample, and \(\sigma\) is the cross section of the simulated sample process.
Table B.7 : Luminosity and trigger used to collect the triple lepton control regions

<table>
<thead>
<tr>
<th>Trigger</th>
<th>Lep. Collected</th>
<th>L (fb$^{-1}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>HLT_Ele17_CaloIdT_CaloIsoVL_TkIdVL_TkIsoVL_Ele8_CaloIdT_CaloIsoVL_TkIdVL_TkIsoVL</td>
<td>$eee$</td>
<td>19.034</td>
</tr>
<tr>
<td>HLT_Mu17_Ele8_CaloIdT_CaloIsoVL_TkIdVL_TkIsoVL</td>
<td>$e\mu\mu$</td>
<td>19.034</td>
</tr>
<tr>
<td>HLT_Mu8_Ele17_CaloIdT_CaloIsoVL_TkIdVL_TkIsoVL</td>
<td>$e\mu\mu$</td>
<td>19.034</td>
</tr>
<tr>
<td>HLT_Mu17_Mu8</td>
<td>$\mu\mu\mu$</td>
<td>19.034</td>
</tr>
</tbody>
</table>

B.3.1 The Simulated Signal Samples

Simulated signal samples were produced using the MADGRAPH (v5.1.3.30) event generator [151] with the CTEQ6L [152] parton distribution functions (PDFs). The PYTHIA (version 6.4.26) [120] generator is used to perform the parton showering. The MC CMS fast simulation package [134] is then used to model detector response and produce the final simulated signal samples.

For the T2tt simulated sample (Table B.8):

- The mass of the stop squark ($m_{\tilde{t}}$) is varied from 200 to 1000 GeV
- The mass of the LSP mass ($m_{\tilde{\chi}^0_1}$) is varied from 0 to 700 GeV
- The masses are varied in steps of 25 GeV

For the T2bW simulated sample (Table B.9):

- The mass of the stop squark ($m_{\tilde{t}}$) is varied from 200 to 1000 GeV
- The mass of the LSP mass ($m_{\tilde{\chi}^0_1}$) is varied from 0 to 550 GeV
- The masses are varied in steps of 25 GeV
- The chargino mass ($\tilde{\chi}^\pm$) is defined through the fraction “x” of the stop mass:
  $$\tilde{\chi}^\pm = x m_{\tilde{t}} + (1 - x) m_{\tilde{\chi}^0_1}$$
Table B.8 : T2tt simulated dataset

- x = 0.25, 0.50, and 0.75.

Table B.9 : T2bW simulated dataset

B.3.2 The Simulated Background Related Samples

Simulated background related samples for Z, W, diboson, τtW, t, t̄t, ttZ, and multijet QCD, are used in this analysis. This analysis uses the following MC programs to generate these simulated events: MADGRAPH (version 5.1.3.30) [151], POWHEG (version 1.0 r1380) [153–157], PYTHIA (version 6.4.26) [120], MC@NLO (version 3.41) [158,159], TAUOLA (version 27.121.5) [160], and HERWIG (version 6.5.20) [121].

The Z (Table B.10) and W (Table B.11) samples are generated, with up to three additional partons, by MADGRAPH. Diboson (Table B.17) and τtW (Table B.12) simulated samples are generated by MADGRAPH. t (Table B.13) and t̄t (Table B.14)
samples are generated with POWHEG. Two Multijet QCD samples are generated (Table B.15): one with PYTHIA and one with MADGRAPH. Two $t\bar{t}Z$ samples are generated (Table B.16): one with MC@NLO and the other with MADGRAPH. MC@NLO is used because it is impossible to define a control sample for $t\bar{t}Z$ (it is too similar to the signal). MADGRAPH is used to calculate the $t\bar{t}Z$ systematic uncertainties. $\tau$ decays are done via TAUOLA.

As with the signal simulated samples, PYTHIA generation is used to perform the parton showing. Except for the MC@NLO $t\bar{t}Z$ sample which uses HERWIG. The GEANT4 [133] full simulation package is then used to model detector response and produce the final background samples.

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Cross Section (pb)</th>
<th>Number of Events</th>
</tr>
</thead>
<tbody>
<tr>
<td>/ZJetsToNuNu_HT-100</td>
<td></td>
<td></td>
</tr>
<tr>
<td>/ZJetsToNuNu_HT-200</td>
<td></td>
<td></td>
</tr>
<tr>
<td>/ZJetsToNuNu_HT-200</td>
<td></td>
<td></td>
</tr>
<tr>
<td>/ZJetsToNuNu_HT-300</td>
<td></td>
<td></td>
</tr>
<tr>
<td>/ZJetsToNuNu_HT-400</td>
<td></td>
<td></td>
</tr>
<tr>
<td>/ZJetsToNuNu_HT-500</td>
<td></td>
<td></td>
</tr>
<tr>
<td>/ZJetsToNuNu_HT-600</td>
<td></td>
<td></td>
</tr>
<tr>
<td>/DYJetsToLL_HT-200</td>
<td></td>
<td></td>
</tr>
<tr>
<td>/DYJetsToLL_HT-400</td>
<td></td>
<td></td>
</tr>
<tr>
<td>/DY1JetsToLL_M-50</td>
<td></td>
<td></td>
</tr>
<tr>
<td>/DY2JetsToLL_M-50</td>
<td></td>
<td></td>
</tr>
<tr>
<td>/DY3JetsToLL_M-50</td>
<td></td>
<td></td>
</tr>
<tr>
<td>/DY4JetsToLL_M-50</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table B.10 : Z simulated dataset ( $Z \rightarrow \nu\bar{\nu}$, $Z \rightarrow l^+l^-$ )

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Cross Section (pb)</th>
<th>Number of Events</th>
</tr>
</thead>
<tbody>
<tr>
<td>/WJetsToLNu_HT-100</td>
<td></td>
<td></td>
</tr>
<tr>
<td>/WJetsToLNu_HT-200</td>
<td></td>
<td></td>
</tr>
<tr>
<td>/WJetsToLNu_HT-300</td>
<td></td>
<td></td>
</tr>
<tr>
<td>/WJetsToLNu_HT-400</td>
<td></td>
<td></td>
</tr>
<tr>
<td>/WJetsToLNu_HT-500</td>
<td></td>
<td></td>
</tr>
<tr>
<td>/WJetsToLNu_HT-600</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table B.11 : W simulated dataset
Table B.12 : ttW simulated dataset

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Cross Section (pb)</th>
<th>Number of Events</th>
</tr>
</thead>
<tbody>
<tr>
<td>/TTWJets_8TeV-madgraph/Summer12_DR53X-PU_S10_START53_V7A-v1/AODSIM</td>
<td>0.232</td>
<td>$1.90 \times 10^4$</td>
</tr>
</tbody>
</table>

Table B.13 : t simulated dataset

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Cross Section (pb)</th>
<th>Number of Events</th>
</tr>
</thead>
<tbody>
<tr>
<td>/Tt-channel_4e2tauStar_8TeV-powheg-tauola/Summer12_DR53X-PU_S10_START53_V7A-v1/AODSIM</td>
<td>56.4</td>
<td>$3.8 \times 10^6$</td>
</tr>
<tr>
<td>/Tt-channel_4e2tauStar_8TeV-powheg-tauola/Summer12_DR53X-PU_S10_START53_V7A-v1/AODSIM</td>
<td>30.7</td>
<td>$1.9 \times 10^6$</td>
</tr>
<tr>
<td>/TtToDilepton_tW-channel_8TeV-powheg-tauola/Summer12_DR53X-PU_S10_START53_V7A-v1/AODSIM</td>
<td>1.18</td>
<td>$3.0 \times 10^6$</td>
</tr>
<tr>
<td>/TtToDilepton_tW-channel_8TeV-powheg-tauola/Summer12_DR53X-PU_S10_START53_V7A-v1/AODSIM</td>
<td>1.18</td>
<td>$3.0 \times 10^6$</td>
</tr>
</tbody>
</table>

B.4 Trigger Efficiency

The trigger efficiency in data is measured by means of a muon control sample. We require at least 5 central jets with $p_T > 30$ GeV and $|\eta| < 2.4$. Furthermore, at least 2 of these central jets must have $p_T > 70$ GeV. The trigger efficiency as a function of offline $E_T$ for all 2012 runs combined is shown in figure B.1. For our lowest offline cut of $E_T > 175$ GeV, the trigger is $\sim 95\%$ efficient. For our preselection region, the trigger efficiency is found to be $96.6\%$ with an uncertainty of $0.4\%$ for data (97.4\% with an uncertainty of 0.1\% for simulation).

The trigger efficiency in our search regions is determined with a simulated $t \bar{t}$ dataset. All preselection requirements are applied. The muon and electron veto are not applied. Our signals (T2bW and T2tt) are very similar in topology to the subset of $t \bar{t}$ events with these selection criteria. The HLT triggers in the run C/D data sets (HLT_DiCentralPFNoPUJet50_PFMETORPFMETNoMu80_v*) are not present in the simulation. Hence, the HLT trigger used is that of run A/B (HLT_DiCentralPFJet50_PFMET80_v*). Of all events passing the offline selection cuts, $\gtrsim 99\%$ also passed the trigger. The efficiency is plotted in Figure B.2.
Table B.14 : tt simulated dataset

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Cross Section (pb)</th>
<th>Number of Events</th>
</tr>
</thead>
<tbody>
<tr>
<td>/TTJets/MassiveBinDecay/TuneZ2star_8TeV-madgraph-tauola/Summer12_DR53X-PU_S10_START53_V7A-v1/AODSIM</td>
<td>234</td>
<td>6.9 × 10^6</td>
</tr>
<tr>
<td>/TTJets/MassiveBinDecay/TuneZ2star_8TeV-madgraph-tauola/Summer12_DR53X-PU_S10_START53_V7A-v2/AODSIM</td>
<td>234</td>
<td>28 × 10^6</td>
</tr>
<tr>
<td>/TTJets/MassiveBinDecay/TuneZ2star_8TeV-powheg-hepth/Summer12_DR53X-PU_S10_START53_V18-v1/AODSIM</td>
<td>244</td>
<td>22 × 10^6</td>
</tr>
</tbody>
</table>

Table B.15 : Multijet QCD simulated dataset

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Cross Section (pb)</th>
<th>Number of Events</th>
</tr>
</thead>
<tbody>
<tr>
<td>/QCD/HT-250to500_TuneZ2star_8TeV-madgraph-pythiaAODSIM/Summer12_DR53X-PU_S10_START53_V7A-v1/AODSIM</td>
<td>3.91 × 10^3</td>
<td>4.1 × 10^6</td>
</tr>
<tr>
<td>/QCD/HT-500to1000_TuneZ2star_8TeV-madgraph-pythiaAODSIM/Summer12_DR53X-PU_S10_START53_V7A-v1/AODSIM</td>
<td>3.83 × 10^3</td>
<td>4.1 × 10^6</td>
</tr>
<tr>
<td>/QCD/HT-1000to2000_TuneZ2star_8TeV-madgraph-pythiaAODSIM/Summer12_DR53X-PU_S10_START53_V7A-v1/AODSIM</td>
<td>3.81 × 10^3</td>
<td>4.1 × 10^6</td>
</tr>
</tbody>
</table>

Table B.16 : ttZ simulated dataset

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Cross Section (pb)</th>
<th>Number of Events</th>
</tr>
</thead>
<tbody>
<tr>
<td>/TTZJets18TeV-madgraph-pythiaAODSIM/Summer12_DR53X-PU_S10_START53_V7A-v1/AODSIM</td>
<td>0.206</td>
<td>2.10 × 10^3</td>
</tr>
<tr>
<td>/ttbarZ_18TeV-Madspin-xMcatNLO-herwig/Summer12_DR53X-PU_S10_START53_V19-v1/AODSIM</td>
<td>0.206</td>
<td>1.40 × 10^3</td>
</tr>
</tbody>
</table>

Table B.17 : Diboson simulated dataset

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Cross Section (pb)</th>
<th>Number of Events</th>
</tr>
</thead>
<tbody>
<tr>
<td>/WWJetsTo2L2Nu_TuneZ2star_8TeV-madgraph-tauola/Summer12_DR53X-PU_S10_START53_V7A-v1/AODSIM</td>
<td>6.02</td>
<td>1.9 × 10^6</td>
</tr>
<tr>
<td>/WZJetsTo2L2Q_TuneZ2star_8TeV-madgraph-tauola/Summer12_DR53X-PU_S10_START53_V7A-v1/AODSIM</td>
<td>2.46</td>
<td>3.2 × 10^6</td>
</tr>
<tr>
<td>/WZJetsTo2LNu_TuneZ2star_8TeV-madgraph-tauola/Summer12_DR53X-PU_S10_START53_V7A-v1/AODSIM</td>
<td>1.10</td>
<td>2.9 × 10^6</td>
</tr>
<tr>
<td>/ZZJetsTo2L2Nu_TuneZ2star_8TeV-madgraph-tauola/Summer12_DR53X-PU_S10_START53_V7A-v1/AODSIM</td>
<td>0.358</td>
<td>9.5 × 10^5</td>
</tr>
<tr>
<td>/ZZJetsTo2L2Q_TuneZ2star_8TeV-madgraph-tauola/Summer12_DR53X-PU_S10_START53_V7A-v1/AODSIM</td>
<td>1.34</td>
<td>1.9 × 10^6</td>
</tr>
<tr>
<td>/ZZJetsTo2L_LH_TuneZ2star_8TeV-madgraph-tauola/Summer12_DR53X-PU_S10_START53_V7A-v1/AODSIM</td>
<td>0.213</td>
<td>4.8 × 10^5</td>
</tr>
</tbody>
</table>
Figure B.1: Search trigger efficiency as a function of offline $E_T$. [13].
Figure B.2: The Trigger turn-on curves as a function of BDT (MVA) discriminator values. Top is T2bw. Bottom is T2tt. [13].
Bibliography


SUSYSMSSummaryPlots8TeV#SUSY_theory_cross_sections_plots.


[38] H. Haber and G. Kane, “The Search for Supersymmetry: Probing Physics
Beyond the Standard Model”,  Phys. Reports 117 (1987) 75,


[40] S. Dawson, E. Eichten, and C. Quigg, “Search for Supersymmetric Particles in
Hadron-Hadron Collisions”,  Phys. Rev. D31 (1985) 1581,


doi:10.1142/9789812839657_0001, 10.1142/9789814307505_0001,


[44] G. R. Farrar and P. Fayet, “Phenomenology of the Production, Decay, and
Detection of New Hadronic States Associated with Supersymmetry”,  Phys.


[55] https://twiki.cern.ch/twiki/bin/view/LHCPhysics/TtbarNNLO.


[60] K. Hanke, “Past and present operation of the CERN PS Booster”,


[68] CMS Collaboration, “Description and performance of track and


[70] https://twiki.cern.ch/twiki/bin/view/CMSPublic/SWGuideEcalRecoClustering.


[79] CMS Collaboration, “Performance of CMS muon reconstruction in $pp$ collision
events at $\sqrt{s} = 7$ TeV”, JINST 7 (2012) P10002,


Volume 2: Data Acquisition and High-Level Trigger. CMS trigger and

[82] K. Rose, “Deterministic annealing for clustering, compression, classification,
regression, and related optimization problems”, Proceedings of the IEEE 86
(1998), no. 11, 2210–2239, doi:10.1109/5.726788.

[84] CMS Collaboration, “b-Jet Identification in the CMS Experiment”,


[85] CMS Collaboration, F. Beaudette, “The CMS Particle Flow Algorithm”, in


[87] https://twiki.cern.ch/twiki/bin/view/CMSPublic/SWGuideMuonId.

[88] https:

//twiki.cern.ch/twiki/bin/view/CMSPublic/WorkBookMetAnalysis#Type_0_Correction.

[89] W. Adam, R. Frhwirth, A. Strandlie, and T. Todorov, “Reconstruction of electrons with the Gaussian-sum filter in the CMS tracker at LHC”,


[92] https://twiki.cern.ch/twiki/bin/view/CMSPublic/

WorkBookMetAnalysis#Type_0_Correction.
[93] CMS Collaboration, “MET performance in 8 TeV data”,


doi:10.1016/0550-3213(93)90166-M.


[99] M. Cacciari and G. P. Salam, “Pileup subtraction using jet areas”,


[112] CMS Collaboration, “Description and performance of track and


[115] https://twiki.cern.ch/twiki/bin/view/CMS/BTagSFMethods#Extension_to_multiple_operating.


(2012).


[124] https://twiki.cern.ch/twiki/bin/view/Main/EGammaScaleFactors2012#2012_8_TeV_data_53X.

[125] https://twiki.cern.ch/twiki/bin/view/CMS/MuonReferenceEffs.


274
[128] https://twiki.cern.ch/twiki/bin/view/CMSPublic/
SWGuideEgammaShowerShape.
Function”, Ann. Math. Statist. 27 (09, 1956) 832–837,
Mullin, V. Pavlunin and R. Rossin, “Scalar Top Quark Search with Jets and
Missing Transverse Momentum in pp Collisions at

√

s = 7 TeV”,

Phys. Conf. Ser. 331 (2011) 032049,




[140] https://twiki.cern.ch/twiki/bin/viewauth/CMS/SMST2ccMadgraph8TeV.


[150] https://twiki.cern.ch/twiki/bin/view/CMS/PileupInformation.


