

SEARCH FOR THE FLAVOR-CHANGING NEUTRAL CURRENT IN TOP
PAIR EVENTS WITH AN ASSOCIATED PHOTON USING 13 TEV
PROTON-PROTON COLLISION DATA COLLECTED WITH THE ATLAS
DETECTOR

by

JASON TYLER BARKELOO

A DISSERTATION

Presented to the Department of Physics
and the Graduate School of the University of Oregon
in partial fulfillment of the requirements
for the degree of
Doctor of Philosophy

March 2020

DISSERTATION APPROVAL PAGE

Student: Jason Tyler Barkeloo

Title: Search for the Flavor-Changing Neutral Current in Top Pair Events With an Associated Photon Using 13 TeV Proton-Proton Collision Data Collected With the ATLAS Detector

This dissertation has been accepted and approved in partial fulfillment of the requirements for the Doctor of Philosophy degree in the Department of Physics by:

David Strom	Chair
James Brau	Advisor
Spencer Chang	Core Member
Dev Sinha	Institutional Representative

and

Janet Woodruff-Borden	Vice Provost and Dean of the Graduate School
-----------------------	--

Original approval signatures are on file with the University of Oregon Graduate School.

Degree awarded March 2020

© 2020 Jason Tyler Barkeloo

This work is licensed under a Creative Commons

Attribution-NonCommercial-NoDerivs (United States) License.

DISSERTATION ABSTRACT

Jason Tyler Barkeloo

Doctor of Philosophy

Department of Physics

March 2020

Title: Search for the Flavor-Changing Neutral Current in Top Pair Events With an Associated Photon Using 13 TeV Proton-Proton Collision Data Collected With the ATLAS Detector

Abstract for FCNC here.

This dissertation includes previously published and unpublished co-authored material.

CURRICULUM VITAE

NAME OF AUTHOR: Jason Tyler Barkeloo

GRADUATE AND UNDERGRADUATE SCHOOLS ATTENDED:

University of Oregon, Eugene, Oregon
Miami University, Oxford, Ohio
Wittenberg University, Springfield, Ohio

DEGREES AWARDED:

Doctor of Philosophy, Physics, 2020, University of Oregon
Master of Science, Physics, 2012, Miami University
Bachelor of Science, Physics, 2010, Wittenberg University

PROFESSIONAL EXPERIENCE:

University of Oregon, Graduate Research Assistant, ATLAS Experiment, June 2014-Present

University of Oregon, Graduate Teaching Assistant, September 2012-June 2014

Miami University, Teaching and Research Assistant, August 2010 - May 2012

Wittenberg University, Undergraduate Researcher, August 2008 - May 2010

GRANTS, AWARDS AND HONORS:

Weiser Senior Teaching Assistant Award – University of Oregon – 2014

American Association of Physics Teachers Outstanding Teaching Assistant – Miami University – 2012

PUBLICATIONS:

J. Barkeloo, J. Brau, M. Breidenbach, R. Frey, D. Freytag, C. Gallagher, R. Herbst, M. Oriunno, B. Reese, A. Steinhebel, D. Strom. “A silicon-tungsten electromagnetic calorimeter with integrated electronics for the International Linear Collider” J. Phys.: Conf. Ser. **1162** 012016 (2019).

J. Kangara, A. Hachtel, M. C. Gillette, J. Barkeloo, E. Clements, S. Bali. “Design and construction of cost-effective fail-safe tapered amplifier systems for laser cooling and trapping experiments”, Am. J. Phys. **82**(8), 805 - 817 (2014).

A. Hachtel, J. Kleykamp, D. Kane, M. Marshall, B. Worth, J. Barkeloo, J. Kangara, J. Camenisch, M. Gillette, S. Bali. “An undergraduate lab on measurement of radiative broadening in atomic vapor”, Am. J. Phys. **81**(6), 471 (2013).

Additional ATLAS Collaboration publications can be found:
<http://inspirehep.net/search?p=exactauthor%3AJason.Barkeloo>

TABLE OF CONTENTS

Chapter	Page
I. INTRODUCTION	1
1.1. The Standard Model Top Quark	1
1.2. Searching for FCNC Top Quark Decays	2
II. THEORY	4
2.1. The Standard Model Particles	4
2.2. The Standard Model Interactions	6
2.3. Electroweak Symmetry Breaking	8
2.3.1. Gauge Boson Masses	9
2.3.2. Fermion Masses	10
2.3.3. The CKM Matrix	12
2.4. The Top Quark	13
2.4.1. Discovering The Top Quark	14
2.4.2. Production and Decay At Hadron Colliders	16
2.4.3. Beyond the Standard Model Top Quark Physics	20
2.5. The Flavor-Changing Neutral Current	21
2.5.1. The Standard Model Flavor Sector	21
2.5.2. The GIM Mechanism	23
2.5.3. New Physics With Enhancements to FCNCs	26

Chapter	Page
2.5.4. Current Measurements of Top FCNCs	29
III. THE LARGE HADRON COLLIDER AND THE ATLAS DETECTOR	34
3.1. The Large Hadron Collider	34
3.1.1. LHC Magnets	37
3.1.2. Luminosity	37
3.1.3. Pileup	39
3.2. The ATLAS Detector	41
3.2.1. Common Detector Variables and the ATLAS Coordinate System	43
3.2.2. Magnet Setup	44
3.2.3. Inner Detector	45
3.2.4. Electromagnetic and Hadronic Calorimeters	49
3.2.5. Muon System	52
3.2.6. Trigger and Data Acquisition	54
Level 1 Calorimeter	56
High Level Trigger	57
IV. SIMULATION AND RECONSTRUCTION	58
4.1. Simulation of pp Collisions	58
4.1.1. Matrix Element Calculation and Parton Distribution Functions	59
4.1.2. Parton Shower Calculation	61
4.1.3. Detector Simulation	61
4.1.4. Monte Carlo Generators Used for LHC Physics	62

Chapter		Page
4.2.	Creation of Flavor Changing Neutral Current Signal Events	62
4.2.1.	FCNC Events Produced With MadGraph5 aMC@NLO	63
4.3.	Object Reconstruction	67
4.3.1.	Electrons	68
4.3.2.	Muons	70
4.3.3.	Photons	71
4.3.4.	Jets	74
	B-Jets	76
4.3.5.	Missing Transverse Energy	78
	Neutrino Reconstruction	79
V.	SEARCH STRATEGY	80
5.1.	Major Backgrounds	80
5.1.1.	Overlap Removal	81
5.1.2.	Duplicate Event Removal	82
5.2.	Event Selection	82
5.3.	Event Classification: Neural Network Optimization	83
5.3.1.	Input Variables	84
5.3.2.	Architecture	90
	Training and Validation of Neural Networks	93
5.3.3.	Hidden Layer Studies	95
5.3.4.	B-Tagging Working Point Studies	99
5.4.	Initial Event Selection	102
5.4.1.	Triggers	103

Chapter	Page
5.5. Data and MC Pre-Selection Cuts	104
5.6. Background Evaluation: Control and Validation Regions	111
5.6.1. Backgrounds Without Photons	111
5.6.2. Background With Photons	116
W+ γ Control Region	116
$t\bar{t} + \gamma$ Control Region	120
5.7. Fake Rates	123
5.7.1. Electron \rightarrow Photon Fakes	123
5.7.2. Jet \rightarrow Lepton Fakes	128
5.7.3. Jet \rightarrow Photon Fakes	130
5.8. Signal Region	136
VI. ANALYSIS AND RESULTS	141
6.1. Systematic Uncertainties	141
6.1.1. Theoretical uncertainties	142
6.1.2. Experimental Uncertainties	143
6.1.3. Symmetrization, Smoothing, and Pruning of Systematic Uncertainties	145
Symmetrization	146
Smoothing	146
Pruning	147
6.2. Statistical Treatment of Results	147
6.3. Nuisance Parameters	151
6.4. Limit on Branching Ratio $t \rightarrow q\gamma$	160

Chapter	Page
VII. COMPLEMENTARY SEARCHES AND OUTLOOK	161
7.1. Comparison with Complementary Searches	161
7.2. Future Directions	161
7.3. Conclusion	161
 APPENDICES	
A. DERIVATION INFORMATION (TOPQ1)	162
B. COMPLETE LIST OF MONTE CARLO SAMPLES USED	163
C. ADDITIONAL PLOTS FROM NN STUDIES	168
C.1. Additional Shape Comparison Plots: μ +jets channel	168
C.2. Shape Comparison Plots: e +jets channel	171
D. SYSTEMATIC VARIATIONS FOR ELECTRON TO PHOTON FAKE RATE CALCULATION	176
E. SYSTEMATIC PULL PLOTS	180
REFERENCES CITED	181

LIST OF FIGURES

Figure	Page
2.1. Summary of $t\bar{t}$ production cross sections as a function of center of mass energy at Fermilab's Tevatron and CERN's LHC	15
2.2. Summary of single-top production cross sections as a function of center of mass energy at Fermilab's Tevatron and CERN's LHC	16
2.3. Leading order diagrams for $t\bar{t}$ production at hadron colliders	17
2.4. Leading order diagrams for single top quark production	18
2.5. Top quark decays in the Standard Model	18
2.6. Categorization of top quark pair decays in the Standard Model based on the decays of the W bosons	20
2.7. Box diagram of $K_L^0 \rightarrow \mu^+\mu^-$ through the exchange of W bosons	23
2.8. Box diagrams of $K_L^0 \rightarrow \mu^+\mu^-$ through the exchange of W bosons after the inclusion of the charm quark	24
2.9. An example loop diagram of a top quark decaying to a light quark and a photon	25
2.10. Flavor-changing neutral current top quark decays.	29
2.11. Flavor-changing limits from a variety of experiments in every channel $t \rightarrow uX(\text{top})$ and $t \rightarrow cX(\text{bottom})$	30
2.12. FCNC diagram for the search in single top production at the electron-proton collider HERA.	31
2.13. FCNC diagram for the search in single top production at the electron-positron collider, LEP.	31
2.14. Flavor-changing neutral current theoretical branching ratios and experimental limits (both ATLAS and CMS) updated through 2018 . . .	32
3.1. Map of LHC and the various detector experiments: ATLAS, CMS, LHCb, and ALICE located under the Franco-Swiss border near Geneva	35

Figure	Page
3.2. Schematic of the CERN accelerator complex.	36
3.3. Total integrated luminosity as a function of time delivered by the LHC(green), recorded (yellow) and declared good for physics analysis (blue) by the ATLAS detector throughout Run 2 consisting entirely of 13 TeV pp collisions.	39
3.4. Luminosity-weighted distribution of the mean number of interactions per bunch crossing for the entirety of Run 2 shown by individual years, 2015 (yellow), 2016(orange), 2017 (purple), 2018 (green), as well as an integrated total (blue).	40
3.5. A candidate dimuon event ($Z \rightarrow \mu^+ \mu^-$) with 28 reconstructed vertices collected in 2018 with the ATLAS detector.	40
3.6. A candidate dimuon event ($Z \rightarrow \mu^+ \mu^-$) with 65 reconstructed vertices collected in 2017 with the ATLAS detector.	41
3.7. Schematic of the ATLAS detector.	42
3.8. Coordinate system used in the ATLAS Collaboration.	44
3.9. Schematic of the windings of the ATLAS magnet.	45
3.10. Schematic of the ATLAS inner detector.	46
3.11. Unfolded transverse (left) and longitudinal (right) impact parameter resolutions measured with (Run 2) and without (Run 1) the IBL as a function of p_T	47
3.12. Blown up schematic of the ATLAS inner detector with more detail.	48
3.13. Schematic of the ATLAS hadronic and electromagnetic calorimeter systems.	49
3.14. Sketch of the accordian structure used in the LAr Calorimeter	50
3.15. Number of radiation lengths throughout the LAr calorimeter as a function of $ \eta $	51
3.16. Number of interaction lengths throughout the LAr calorimeter as a function of $ \eta $	52
3.17. Schematic of the ATLAS muon detector.	53

Figure	Page
3.18. Flow diagram of the ATLAS trigger and data acquisition system used in Run 2.	56
4.1. A pictoral view of the different steps for the creation of a MC event.	59
4.2. The bands are the momentum fraction, x , times the unpolarized parton distribution function obtained in NNLO NNPDF3.0 global analysis at scales $\mu^2 = 10 \text{ GeV}$ and $\mu^2 = 100 \text{ GeV}^2$	60
4.3. Normalized kinematics (p_T , m_t , and y_t) of individual top quarks produced by the model for each FCNC final state search and an official $t\bar{t}$ sample.	63
4.4. Normalized kinematics (p_T , $m_{t\bar{t}}$, and $y_{t\bar{t}}$) of the $t\bar{t}$ system produced by the model for each FCNC final state search and an official $t\bar{t}$ sample.	64
4.5. Normalized p_T of the electron and muons produced by the model for each FCNC final state search and an official $t\bar{t}$ sample.	64
4.6. Normalized p_T of the up, charm, and bottom quarks produced by the model for each FCNC final state search and an official $t\bar{t}$ sample.	65
4.7. Normalized p_T of the photons produced by the model for each FCNC final state search and an official $t\bar{t}$ sample, there is 0 contribution from the official $t\bar{t}$ sample.	65
4.8. Normalized p_T of the photons produced by the model using the Left-handed (LH) and Right-handed (RH) couplings	66
4.9. Cross section of a simulated ATLAS detector showing how various particles interact with ATLAS subsystems.	68
4.10. A sample parton-level event with many random soft jet objects, clustered with four different jets algorithms, illustrating the areas of the resulting hard jets. For k_T and Cambridge/Aachen the detailed shapes are in part determined by the specific set of ghosts used, and change when the ghosts are modified	75
4.11. Pictoral representation of an event with a b-jet showing the secondary vertex and impact parameter	77
4.12. The MV2c10 output for b, c, and light flavored jets in simulated $t\bar{t}$ and the background rejection as a function of the b-jet efficiency	78

Figure	Page
5.1. Normalized variables showing the shapes of neural network input variables for the μ +jets channel: γ_{iso} topo E_T cone40, γ_{pT} , $m_{q\gamma}$, $m_{l\gamma}$, m_{bW} , and $\Delta R_{j\gamma}$	87
5.2. Normalized variables showing the shapes of neural network input variables for the μ +jets channel: ΔR_{bl} , m_T^W , S_T , n_{jets} , χ_W^2 , and $p_T(q)$	88
5.3. Normalized variables showing the shapes of neural network input variables for the μ +jets channel: $\Delta R_{l\gamma}$, E (lepton), \cancel{E}_T , and $p_T(b)$	89
5.4. Pictoral representation of neural network architecture with 3 input variables, 2 hidden layers with 4 nodes each, and 1 output layer.	92
5.5. ROC Curves are shown for both search channels for a varying number of hidden layers. Orange lines correspond to one hidden layer, blue to 2 hidden layers and green to 3 hidden layers. The blue and green curves have near identical AUC values: 0.950 and 0.951 for the e +jets case and 0.962 for the μ +jets cases.	95
5.6. Accuracy plots for both channels for the 2 hidden layer neural network	96
5.7. Loss plots for both channels for the 2 hidden layer neural network	97
5.8. Normalized neural network output signal and background distribution plots are shown for both search channels for a varying number of hidden layers.	98
5.9. Significance plots for both channels for the 2 hidden layer neural network.	98
5.10. Accuracy and loss plots for the e +jets channel at 70%, 77%, and 85% b-tagging working points.	100
5.11. Accuracy and loss plots for the μ +jets channel at 70%, 77%, and 85% b-tagging working points.	100
5.12. Neural network output and significance plots for the e +jets channel at 70%, 77%, and 85% b-tagging working points.	101
5.13. Neural network output and significance plots for the μ +jets channel at 70%, 77%, and 85% b-tagging working points.	101
5.14. Photon p_T in the signal region pre-selection region. FCNC signal branching ratio is scaled to 1.5%.	105

Figure	Page
5.15. S_T and m_T^W in the signal region pre-selection region. FCNC signal branching ratio is scaled to 1.5%.	106
5.16. Top mass candidates in the signal region pre-selection: m_{Wb} and $m_{q\gamma}$. FCNC signal branching ratio is scaled to 1.5%.	107
5.17. Photon p_T (a), leading light jet p_T (b), lepton p_T (c), b-jet p_T (d), \cancel{E}_T (e), and n_{jets} (f) plots in the signal region pre-selection for the electron+jets channel. FCNC signal branching ratio is scaled to 1%.	108
5.18. Photon p_T (a), leading light jet p_T (b), lepton p_T (c), b-jet p_T (d), \cancel{E}_T (e), and n_{jets} (f) plots in the signal region pre-selection for the muon+jets channel. FCNC signal branching ratio is scaled to 1.5%.	109
5.19. Output of the Neural Network in the signal region pre-selection region. FCNC signal branching ratio is scaled to 1.5%.	110
5.20. S_T distributions in the 3(a,d), 4(b,e), and 5+(c,f) jets control and validation regions. The electron channel is shown on the top and the muon channel on the bottom, before scale factors are determined.	113
5.21. Event-level plots for the =4 jet validation region after scale factors have been applied in the electron channel. FCNC signal branching ratio is scaled to 1.5%.	114
5.22. Event-level plots for the =4 jet validation region after scale factors have been applied in the muon channel. FCNC signal branching ratio is scaled to 1.5%.	115
5.23. $W+\text{jets}+\gamma$ validation region plots for the electron channel. The FCNC signal sample is scaled to 1.5%.	118
5.24. $W+\text{jets}+\gamma$ validation region plots for the muon channel. The FCNC signal sample is scaled to 1.5%.	119
5.25. $t\bar{t}+\text{jets}+\gamma$ validation region plots for the electron channel. The FCNC signal sample is scaled to 1.5%.	121
5.26. $t\bar{t}+\text{jets}+\gamma$ validation region plots for the muon channel. The FCNC signal sample is scaled to 1.5%.	122
5.27. Probe p_T vs leading electron p_T in $Z \rightarrow e^+e^-$ Data Events	125
5.28. Probe p_T vs leading electron p_T in $Z \rightarrow e^+e^-$ MC Events	126

Figure	Page
5.29. 2-Dimensional scale factors derived using the $Z \rightarrow e^+e^-$ events for unconverted and converted photon types shown with statistical uncertainties	128
5.30. Pictoral representation of the regions used for the ABCD method in this search	132
5.31. Regions used for the ABCD method shown for both converted and unconverted photon types. MC events are shown in the gradient histogram and data events are shown with the black points. All final signal region cuts are included except the neural network cut to give enough statistics in the regions for the ABCD method to be a reasonable estimate of the fake rate scale factor.	134
5.32. Photon p_T (a), lepton p_T (b), leading light jet p_T (c), $\Delta R_{l\gamma}$ (d), $\gamma_{\text{topoetcone40}}$ (e), and b-jet p_T (f) distributions in the final signal region with scale factors applied for the electron channel	137
5.33. FCNC top candidate mass (a), Standard Model top candidate mass (b), \cancel{E}_T (c), N_{jets} (d), m_T^W (e), and S_T (f) distributions in the final signal region with scale factors applied for the electron channel	138
5.34. Photon p_T (a), lepton p_T (b), leading light jet p_T (c), $\Delta R_{l\gamma}$ (d), $\gamma_{\text{topoetcone40}}$ (e), and b-jet p_T (f) distributions in the final signal region with scale factors applied for the muon channel	139
5.35. FCNC top candidate mass (a), Standard Model top candidate mass (b), \cancel{E}_T (c), N_{jets} (d), m_T^W (e), and S_T (f) distributions in the final signal region with scale factors applied for the muon channel	140
6.1. Nuisance parameters after pruning for $e+\text{jets}$ channel	152
6.2. Pull values for the various nuisance parameters considered in the fit for the $e+\text{jets}$ channel.	153
6.3. Bin by bin normalization γ factors used in each region for the $e+\text{jets}$ channel.	154
6.4. Negative-log likelihood (goodness of fit) as a function of signal strength using data in all regions for the $e+\text{jets}$ channel.	154
6.5. Correlation matrix with at least one coefficient above 20% for $e+\text{jets}$ channel	155
6.6. Nuisance parameters after pruning for $\mu+\text{jets}$ channel	156

Figure	Page
6.7. Pull values for the various nuisance parameters considered in the fit for the μ +jets channel.	157
6.8. Bin by bin normalization γ factors used in each region for the μ +jets channel.	158
6.9. Negative-log likelihood (goodness of fit) as a function of signal strength using data in all regions for the μ +jets channel.	158
6.10. Correlation matrix with at least one coefficient above 20% for μ +jets channel	159
C.1. Normalized variables showing the shapes of neural network input variables for the μ +jets channel: [E (bjet), η_b , ΔR_{jb} , E (light jet), light jet η , and χ^2_ν the total χ^2 fit value	169
C.2. Normalized variables showing the shapes of neural network input variables for the μ +jets channel: [lepton p_T , lepton η , lepton isolation , chi^2_{SM} the bW χ^2 value from neutrino reconstrucion ,photon E, and photon η	170
C.3. Normalized variables showing the shapes of neural network input variables for the e +jets channel: $\Delta R_{l\gamma}$, E (lepton), \not{E}_T , and $p_T(b)$	171
C.4. Normalized variables showing the shapes of neural network input variables for the e +jets channel: γ_{iso} topo E_T ccone40, γ_{p_T} , $m_{q\gamma}$, $m_{l\gamma}$, m_{bW} , and $\Delta R_{j\gamma}$	172
C.5. Normalized variables showing the shapes of neural network input variables for the e +jets channel: ΔR_{bl} , m_T^W , S_T , n_{jets} , χ^2_W , and $p_T(q)$	173
C.6. Normalized variables showing the shapes of neural network input variables for the e +jets channel: [E (bjet), η_b , ΔR_{jb} , E (light jet), light jet η , and χ^2_ν the total χ^2 fit value	174
C.7. Normalized variables showing the shapes of neural network input variables for the e +jets channel: [lepton p_T , lepton η , lepton isolation , chi^2_{SM} the bW χ^2 value from neutrino reconstrucion, photon E, and photon η	175
D.1. 2-Dimensional scale factors derived using the $Z \rightarrow e^+e^-$ events within a 5 GeV mass window around the Z boson (2.5 GeV on each side) for unconverted and converted photon types shown with statistical uncertainties.	177
D.2. 2-Dimensional scale factors derived using the $Z \rightarrow e^+e^-$ events within a 15 GeV mass window around the Z boson (7.5 GeV on each side)	

Figure	Page
for unconverted and converted photon types shown with statistical uncertainties.	178
D.3. 2-Dimensional scale factors derived using the $Z \rightarrow e^+e^-$ events within a 20 GeV mass window around the Z boson (10 GeV on each side) for unconverted and converted photon types shown with statistical uncertainties.	179

LIST OF TABLES

Table	Page
2.1. Particles of the Standard Model	6
2.2. Summary of the Standard Model field contents and their representations in the gauge group, labeled by their dimension ($SU(3)_C$ and $SU(2)_L$) and weak hypercharge ($U(1)_Y$)	7
2.3. Summary of final states of a W boson, when accounting for color permutations and the CKM matrix	19
2.4. Expected branching ratios for various flavor-changing neutral current processes in the Standard Model and multiple theories that predict enhancements to the branching ratio. Two-Higgs Double Models with flavor-violating Yukawa couplings (2HDM), quark single models (QS), minimal supersymmetry models with 1TeV squarks and gluinos (MSSM), R-parity violating supersymmetry models (RPV), and extra-dimensional models (XD).	26
4.1. Photon identification variables used for <i>loose</i> and <i>tight</i> photon identification	73
5.1. Separation of normalized variables between signal and background in the e+jets and μ +jets channels for the variables used as input to the final neural network.	85
5.2. Branching ratio values with a significance of 2 after neural network optimization	99
5.3. The electron trigger requirements in the event selections	103
5.4. The muon trigger requirements in the event selections	104
5.5. Derived $t\bar{t}$ and W+jets scale factors for QCD multijet backgrounds . . .	114
5.6. Fit result W+jets+ γ normalization scale factors including both statistical and systematic uncertainties	117
5.7. Fit result $t\bar{t} + \gamma$ normalization scale factors including both statistical and systematic uncertainties	120

Table	Page
5.8. Monte Carlo hadronic photon origin event yields in each region for converted and unconverted photons, including statistical errors	133
5.9. Closure test results and the resulting correction factor values, θ_{MC} for hadronic fake estimates including statistical errors	133
5.10. Hadron scale factors determined in data for both channels and both photon types using the ABCD method. Statistical and systematic uncertainties are shown.	135
5.11. Signal region event yields after all scale factors and cuts are applied . .	136
A.1. TOPQ1 Object Definitions	162
B.1. List of MC16a (r-tag 9364) samples used in this search for the $t\bar{t} + V$ and diboson processes. All samples are available and used for MC16d (r-tag 10201) and MC16e (r-tag 10724) unless otherwise stated.	163
B.2. List of MC16a (r-tag 9364) samples used in this search for the signal sample, $t\bar{t}$, and single-top processes. All samples are available and used for MC16d (r-tag 10201) and MC16e (r-tag 10724) unless otherwise stated.	164
B.3. List of MC16a (r-tag 9364) samples used in this search for the W+jets background. All samples are available and used for MC16d (r-tag 10201) and MC16e (r-tag 10724). All processes are simulated for up to 2 partons at NLO and 4 partons at LO.	165
B.4. List of MC16a (r-tag 9364) samples used in this search for the Z+jets background. All samples are available and used for MC16d (r-tag 10201) and MC16e (r-tag 10724). All processes are simulated for up to 2 partons at NLO and 4 partons at LO.	166
B.5. List of MC16a (r-tag 9364) samples used in this search for the W+jets+ γ and Z+jets+ γ backgrounds. All samples are available and used for MC16d (r-tag 10201) and MC16e (r-tag 10724). All processes are simulated for up to 1 parton at NLO and 3 partons at LO.	167

CHAPTER I

INTRODUCTION

The Standard Model of particle physics has proven itself an exceptional and resilient theory since the combination of the electromagnetic and weak interactions in 1961[1]. Further theoretical work combined the Higgs Mechanism[2, 3] with the electroweak theory[4, 5]. The resiliency of this theoretical model has been tested to further degrees of accuracy over the decades with one of the most recent test being the experimental confirmation of the Higgs Boson in 2012[6, 7] using the Large Hadron Collider (LHC). Further precision measurements are ongoing at various experiments at the LHC, including the ATLAS experiment.

However, the Standard Model is known to have flaws and disagreements with nature. For example, the Standard Model predicts massless neutrinos which is in conflict with experimental observation of neutrino flavor oscillation and does not provide an explanation for dark matter particles or their interactions with currently known particles. While these large gaps in the Standard Model are well known, every precision measurement made has yet to give any significant new hints toward physics beyond the Standard Model. One new pathway to look for these hints at the LHC is through top quark decays.

1.1. The Standard Model Top Quark

The top quark was first observed at Fermilab's Tevatron in 1995[8] but the increase in energy and amount of data at the LHC has produced orders of magnitude more top quarks than previously seen, opening up a pathway to precision measurements of the properties of the top quark. The top quark is the heaviest

fundamental particle with a mass of 172.51 ± 0.27 (stat) ± 0.42 (syst)[9]. This large mass also means that the top quark lifetime is very short ($5 * 10^{-25}$ s) and decays before it can hadronize. This allows the study of its branching ratios and decay modes directly. The Standard Model predicts that the top quark decays through the charged current mode nearly 100% of the time, $t \rightarrow qW$ ($q = b, s, d$)[10]. The Standard Model also predicts a rare branching ratio of the top quark through a flavor changing neutral current (FCNC) process, to a neutral boson (photon, Z boson, Higgs Boson, or gluon) and up type quark with a heavily suppressed branching ratio on the order of 10^{-14} [11].

1.2. Searching for FCNC Top Quark Decays

Precision measurements are an important litmus test for the Standard Model. Predicted branching ratios for FCNC processes in top quark decays are far beyond the experimental reach of the LHC and any observation of these decay modes would be a sure sign of new physics. Branching ratios are an important measurement due to a litany of theories for new physics beyond the Standard Model (BSM). These BSM theories such as Minimal Supersymmetric models[12], R-parity-violating Supersymmetric models[13] and two Higgs doublet models[14] predict enhancements in the top sector by many orders of magnitude. Even a null result of a search will set an upper limit on the branching ratio that can assist in ruling out future physical models based on their amount of large top sector enrichment.

This dissertation presents a search for top FCNCs using the entire Run 2 dataset at the LHC, containing combined 2015-2018 datasets taken by the ATLAS experiment totaling 139 fb^{-1} of integrated luminosity taken at $\sqrt{s} = 13 \text{ TeV}$. This analysis looks for an excess of events coming from top quark pair produced events where one top

quark decays to the most likely decay mode (a bottom quark and W boson) and the other to an up type quark (up or charm) and a photon. Chapter II presents a theoretical background for the Standard Model with a closer view on the usual extensions to include the FCNC vertices. Following this, Chapter III will discuss the LHC and the ATLAS experiment used in the creation of the dataset used in the analysis. In Chapter IV the special signal simulation requirements will be presented as well as the common background event simulation methodology. The search strategy including the creation of signal, control and validation regions and the training of a neural network will be examined in Chapter V. Chapter VI will discuss the results and the conclusions drawn from these results will be presented in Chapter VII.

CHAPTER II

THEORY

In this chapter a theoretical background will be presented on the Standard Model of particle physics with special attention paid to the top quark's properties and decays. This will include discussion of all of the fundamental particles and their interactions through the fundamental forces of nature: electromagnetism and the strong and weak nuclear forces.

2.1. The Standard Model Particles

The Standard Model of particle physics is the cornerstone of our understanding of the basic building blocks of nature and their interactions. Typical matter is made up of atoms consisting of electrons around an inner nucleus of protons and neutrons. These protons and neutrons are made up of a collection of up and down type quarks along with sea quarks and gluons. Protons consist of two up type quarks and a down type quark, while neutrons contain two down type quarks and a single up type quark.

Within the Standard Model all matter (quarks and leptons) is made up of fermions (spin- $\frac{1}{2}$ particles). In addition to the up and down type quarks and the electron the remaining fermions can be described as additional 'generations' that are similar, each consisting of two quarks and a lepton. Every generation consists of quarks with electric charge $+\frac{2}{3}$ and $-\frac{1}{3}$ and a lepton with charge -1 along with an electrically neutral lepton called a neutrino. All of these fermions are listed in Table 2.1.

While we can observe the leptons in nature, free quarks do not exist. We can observe quarks only as part of a bound state called a hadron. If a hadron consists of a

quark-antiquark pair ($q\bar{q}$) it is called a meson. Sets of three quarks or three antiquarks (qqq or $\bar{q}\bar{q}\bar{q}$) are called baryons. Protons and neutrons are baryonic matter.

In addition to the fermions the Standard Model contains gauge, or vector, bosons which are spin-1 particles that carry the fundamental forces of nature. The electromagnetic force is mediated by the massless, charge neutral photon (γ). The nuclear weak force is carried by two massive bosons: the electrically neutral Z boson and the charged (± 1) W boson. These bosons together dictate electroweak interactions within the Standard Model. The remaining force, the nuclear strong force, is carried through the gluon (g). Gluons are massless and chargeless but carry color, an analog of electric charge in the electroweak interaction. The last remaining piece of the Standard Model is the scalar (spin-0) Higgs boson. This Higgs boson is a massive electrically neutral boson that is responsible for giving mass to the massive fundamental particles. All of these bosons are also shown in Table 2.1.

	Particle	Spin	Charge	Mass
Quarks				
u type	u			$2.4^{+0.6}_{-0.4}$ MeV
	c	$\frac{1}{2}$	$\frac{2}{3}$	1.28 ± 0.03 GeV
	t			173.1 ± 0.6 GeV
d type	d			$4.7^{+0.5}_{-0.4}$ MeV
	s	$\frac{1}{2}$	$-\frac{1}{3}$	96^{+8}_{-4} MeV
	b			$4.18^{+0.04}_{-0.03}$ GeV
Leptons				
e doublet	e	$\frac{1}{2}$	-1	$0.5109989461 \pm 0.000000003$ MeV
	ν_e		0	< 2 eV
μ doublet	μ	$\frac{1}{2}$	-1	$105.6583745 \pm 0.0000024$ MeV
	ν_μ		0	< 2 eV
τ doublet	τ	$\frac{1}{2}$	-1	1776.86 ± 0.12 GeV
	ν_τ		0	< 2eV
Bosons				
Vector	γ	1	0	$< 10^{-18}$ eV
	g	1	0	0
	W	1	\pm	80.385 ± 0.0015 GeV
	Z	1	0	91.1876 ± 0.0021 GeV
Scalar	H	0	0	$125.09 \pm 0.21 \pm 0.11$ GeV

TABLE 2.1. Particles of the Standard Model [10]

2.2. The Standard Model Interactions

In addition to describing these particles the Standard Model also describes the ways in which these particles are capable of interacting. All of the particles described in the previous section are included in the following theory. The Standard Model Lagrangian is simply a function of fields and their derivatives taken at one point in spacetime, x^μ .

$$\mathcal{L}[\phi_i(x), \partial_\mu \phi_i(x)]$$

The Standard Model is defined having a gauge symmetry

$$G_{SM} = SU(3)_C \times SU(2)_L \times U(1)_Y$$

Name	Field Components	$(SU(3)_C, SU(2)_L, U(1)_Y)$	Comments
Spin-1/2 Quarks			
Q	$(u_L \ d_L)$ u_R d_R	$(3,2,\frac{1}{6})$ $(3,1,\frac{2}{3})$ $(3,1, -\frac{1}{3})$	x3 generations
Spin-1/2 Leptons			
L	$(\nu_L \ e_L)$ e_R	$(1,2,-\frac{1}{2})$ $(1,1,-1)$	x3 generations
Spin-0 Higgs			
ϕ	$(\phi^+ \ \phi^0)$	$(1,2,\frac{1}{2})$	
Spin-1 Gauge Bosons			
Gluons W B	$G^{1,...,8}$ $(W^1 \ W^2 \ W^3)$ B^0	$(8,1,0)$ $(1,3,0)$ $(1,1,0)$	

TABLE 2.2. Summary of the Standard Model field contents and their representations in the gauge group, labeled by their dimension ($SU(3)_C$ and $SU(2)_L$) and weak hypercharge ($U(1)_Y$)

where all of the particle content described in Section 2.1 is described under this symmetry in Table 2.2[15]. The eight color-anticolor combinations of spin-1 gluons are associated with $SU(3)_C$ where the C denotes 'color' quantum numbers of the gauge group. Any particle that carries color will interact with gluons via the strong nuclear interaction. The three spin-1 gauge bosons W_a , $a = 1, 2, 3$ conduct the weak isospin $SU(2)_L$ symmetry where the L denotes that only left-handed chiral fermions transform with respect to this symmetry. The left-handed fermions are $SU(2)$ doublets while the right-handed components are singlets. The remaining spin-1 gauge boson, B, is associated with $U(1)^Y$ weak hypercharge.

The W^\pm , Z , and photon (γ) are produced by spontaneous symmetry breaking $SU(2)_L \times U(1)_Y$ implied by the Higgs potential discussed Section 2.3.

One general way of expressing the Standard Model Lagrangian is

$$\mathcal{L}_{SM} = \mathcal{L}_{kinetic} + \mathcal{L}_{Higgs} + \mathcal{L}_{Yukawa}$$

where we have a kinetic term for the bosons and fermions, a Higgs term which includes the Higgs kinetic term and potential, and a term that dictates the Yukawa interactions. This can be expanded out

$$\mathcal{L} = \left(-\frac{1}{4} F_{\mu\nu} F^{\mu\nu} + i\bar{\psi} \not{D} \psi + \text{h.c.} \right) + (|D_\mu \phi|^2 + V(\phi)) + (\psi_i Y_{ij} \psi_j \phi + \text{h.c.})$$

All of the gauge field strengths are included within $F_{\mu\nu}$. To maintain gauge invariance for the kinetic terms the derivative must be replaced by the covariant derivative

$$\partial^\mu \rightarrow D^\mu = \partial^\mu + ig_s G_a^\mu L_a + ig W_a^\mu T_a + ig' B^\mu Y$$

The eight gluon fields are described by G_a^μ , the three weak interaction boson fields are described by W_a^μ , and the hypercharge boson field is B^μ . Here g_s , g , and g' are the gauge coupling constants, and L_a , T_a , and Y are the generators for $SU(3)_C$, $SU(2)_L$, and $U(1)_Y$ respectively. The generators are described by the Gell-Man matrices for L_a ($\frac{1}{2}\lambda_a$ for triplets, 0 for singlets), the Pauli matrices for T_a ($\frac{1}{2}\sigma_a$ for doublets, 0 for singlets), and the $U(1)_Y$ charges for Y .

2.3. Electroweak Symmetry Breaking

The spin-1 particles within the Standard Model are massless; however we know that the real bosons made up of these states are massive (i.e. the W^\pm and Z). This comes about from a spontaneous symmetry breaking of $SU(2)_L \times U(1)_Y$ by a particle

whose ground state is not invariant under the symmetry. The scalar Higgs field is capable of doing this while at the same time not breaking the $SU(3)_C$ symmetry leaving the gluons massless because ϕ is an $SU(3)_C$ singlet.

Expanding the Yukawa term of the Lagrangian we see that the covariant derivative of the scalar field becomes

$$D_\mu \phi = \partial_\mu \phi - \frac{i}{2} (g W_\mu^a \frac{\sigma^a}{2} + g' B_\mu) \phi$$

and the potential

$$V(\phi) = \lambda (\phi^\dagger \phi - \frac{\mu^2}{2\lambda})^2 = \lambda (\phi^\dagger \phi)^2 - \mu^2 \phi^\dagger \phi + \frac{\mu^4}{4\lambda}$$

It follows that if $\frac{\mu^2}{\lambda} > 0$ then this potential has a non-zero vacuum expectation value $v^2 \equiv \frac{\mu^2}{\lambda}$. The condition $\lambda > 0$ is a requirement on the vacuum stability of the model such that $\mu^2 < 0$ is the requirement for spontaneous symmetry breaking. Therefore we can make a gauge transformation such that

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

2.3.1. Gauge Boson Masses

The mass eigenstates of the spin-1 bosons can be recovered from this by writing out the mass part of the gauge boson kinetic term $|D^\mu \phi|^2$

$$\mathcal{L}_{mass} = -\frac{1}{8} \begin{pmatrix} 0 & v \end{pmatrix} \begin{pmatrix} gW_3 + g'B & g(W_1 - iW_2) \\ g(W_1 + iW_2) & -gW_3 + g'B \end{pmatrix}^2 \begin{pmatrix} 0 \\ v \end{pmatrix}$$

and if we define the weak mixing angle as $\tan\theta_W \equiv \frac{g'}{g}$ and the mass eigenstates of the bosons

$$W^\pm \equiv \frac{1}{\sqrt{2}}(W_1 \mp iW_2)$$

$$Z \equiv \cos\theta_W W_3 - \sin\theta_W B$$

$$A \equiv \sin\theta_W W_3 + \cos\theta_W B$$

then the mass matrix can be diagonalized to give

$$\mathcal{L}_{mass} = -\frac{1}{4}g^2 v^2 W^+ W^- - \frac{1}{8}(g^2 + g'^2)v^2 Z^2$$

which gives us the masses of our bosons: $M_W^2 = \frac{1}{4}g^2 v^2$, $M_Z^2 = \frac{1}{4}(g^2 + g'^2)v^2$, and $M_A^2 = 0$. Since the field A is massless there is an unbroken $U(1)$ that is identified as the photon and $U(1)_{EM}$.

2.3.2. Fermion Masses

Fermions acquire their masses from the Yukawa term in the Lagrangian. If we transform into the mass basis, also using the replacement of the scalar field corresponding to a physical Higgs boson

$$\phi = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + H(x) \end{pmatrix}$$

then we can write the $SU(2)_L$ quark doublets into their components

$$Q_{Li} = \begin{pmatrix} U_{Li} \\ D_{Li} \end{pmatrix}$$

The Yukawa term can be broken down into a baryonic and leptonic portion. The lepton portion leads to charged lepton masses after the Higgs acquires a vacuum expectation value after electroweak symmetry breaking.

$$-\mathcal{L}_{\text{Yukawa}}^{\text{leptons}} = Y_{ij}^e \overline{L}_{Li} \phi E_{Rj}$$

This leads to three physical parameters, which are generally chosen to be the three charged lepton masses (electron, muon, tau). The baryonic term leads to quark masses and flavors. Ten physical parameters arise from the baryonic Yukawa interactions

$$-\mathcal{L}_{\text{Yukawa}}^{\text{quarks}} = Y_{ij}^d \overline{Q}_{Li} \phi D_{Rj} + Y_{ij}^u \overline{Q}_{Li} \tilde{\phi} U_{Rj} + \text{h.c.}$$

Expanding this in terms of the above $SU(2)_L$ components

$$-\mathcal{L} = (M_d)_{ij} \overline{D}_{Li} D_{Rj} + (M_u)_{ij} \overline{U}_{Li} U_{Rj} + \text{h.c.}$$

such that $M_q = \frac{v}{\sqrt{2}} Y^q$. Unitary matricesies (V_{qL} and V_{qR}) that take us to the mass basis can always be found.

$$V_{qL} M_q V_{qR}^\dagger = M_q^{\text{diag}}, \text{ for } q=u,d$$

The quark mass eigenstates that follow from this are:

$$q_{Li} = (V_{qL})_{ij} q_{Lj}^I \text{ and } q_{Ri} = (V_{qR})_{ij} q_{Rj}^I$$

being transformed from the interaction basis I .

2.3.3. The CKM Matrix

In the interaction basis interactions between quarks and W^\pm come from the $\bar{\psi}D\psi$ kinetic term in the Lagrangian. We can explicitly write this out in both the interaction and mass basis such that:

$$\text{Interaction Basis: } -\mathcal{L}_{W^\pm}^q = \frac{g}{2} \overline{Q}_{Li} \gamma^\mu W_\mu^a \tau^a Q_{Li} + \text{h.c.}$$

$$\text{Mass Basis: } -\mathcal{L}_{W^\pm}^q = \frac{g}{\sqrt{2}} \begin{pmatrix} \bar{u}_L & \bar{c}_L & \bar{t}_L \end{pmatrix} \gamma^\mu (V_{uL} V_{dL}^\dagger)_{ij} W_\mu^+ \begin{pmatrix} d_L \\ s_L \\ b_L \end{pmatrix} + \text{h.c.}$$

The term $V = V_{uL} V_{dL}^\dagger$ in the previous equation is a unitary 3×3 matrix known as the Cabibbo-Kobayashi-Maskawa (CKM) matrix [16, 17] that describes quark mixing.

The CKM matrix can be expressed simply:

$$V_{CKM} = \begin{bmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{bmatrix}$$

The components of the CKM matrix are not all arbitrary. This can be expressed in the Wolfenstein parametrization[18] to show that there are only four parameters, three real and one complex. The Wolfenstein parameterization is an expansion in $V_{us} = \lambda = 0.2257^{+0.009}_{-0.0010}$. Where $A = 0.814^{+0.021}_{-0.022}$, $\rho = 0.135^{+0.031}_{-0.016}$, and $\eta = 0.349^{+0.015}_{-0.017}$ [10].

$$V_{CKM}^{\text{Wolfenstein}} = \begin{bmatrix} 1 - \frac{1}{2}\lambda^2 & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1 - \frac{1}{2}\lambda^2 & A\lambda^2 \\ A\lambda^3(1 - \rho - i\eta) & -A\lambda^2 & 1 \end{bmatrix}$$

This formulation demonstrates quite clearly that the CKM matrix is almost diagonal but has small off-diagonal terms. This means that interactions between quarks is strongest within a generation i.e. up and down quarks, charm and strange quarks, and top and bottom quarks experience the largest degree of mixing between each other. A top quark directly decays to a bottom quark and W boson almost all of the time ($\approx 99.83\%$ of the time). While the mixing does not prohibit decaying to a strange or down quark instead of a bottom, it is a significantly rarer event ($\approx 0.16\%$ and $\approx 0.01\%$, respectively). The values for all of these parameters have been measured experimentally and are collected by the Particle Data Group[10]:

$$V_{CKM} = \begin{bmatrix} 0.97420 \pm 0.00021 & 0.2243 \pm 0.0005 & 0.00394 \pm 0.00036 \\ 0.218 \pm 0.004 & 0.997 \pm 0.017 & 0.0422 \pm 0.0008 \\ 0.0081 \pm 0.0005 & 0.0394 \pm 0.0023 & 1.019 \pm 0.025 \end{bmatrix}$$

2.4. The Top Quark

This dissertation will pay particular interest to the top quark, the heaviest of all of the fundamental particles, over 40 times heavier than its generational partner the bottom quark and 40% larger than the Higgs boson the next heaviest particle. One of its most interesting properties, due to its large mass, is that it has an extremely small lifetime ($\approx 5 \times 10^{-25}s$). This lifetime is orders of magnitude shorter than the characteristic time over which the strong nuclear interaction takes place ($\approx 10^{-23}s$), meaning that it decays before it can hadronize by combining with another quark to form a mesonic or baryonic system. This allows us to observe the decays of the top quark directly as opposed to the decay products of a top quark system. The top quark was first proposed as an explanation of CP violation in kaon decays by

Kobayashi and Maskawa in 1973[17] which predicted a third generation of quarks. At the time there was little evidence for this. The later discovery of the charm quark, that was predicted to exist due to a suppression caused by the GIM mechanism[19], and formed the foundation of Kobayashi and Maskawa's work provided further evidence. The discovery of the charm filled out a complete 2 generation Standard Model (up, down, strange, and charm quarks and the electron and muon). However the further observation of the tau lepton meant that the symmetry between lepton and quark generations would be broken without the existence of a third generation of quarks, the top and bottom, providing soft evidence of a third generation of quarks. Within just a few years, in 1977, the bottom was also discovered[20] and the theoretical predictions of the Standard Model at the time heavily favored the existence of a 6th quark, the 3rd and heaviest up-type quark.

2.4.1. Discovering The Top Quark

Almost 20 years went by after the discovery of the bottom quark and many experiments came up empty handed even while discovering the W and Z bosons (the Super Proton Synchrotron at CERN) before direct evidence for the top quark was observed at Fermilab's Tevatron in both the CDF and D0 experiments in 1995[8, 21]. While the Tevatron only operated at a maximum energy of 1.96 TeV during its lifetime, further properties of the top quark could not be probed in detail due to a lack of statistics. Not until the LHC began operation at a center of mass energy of 7 TeV in 2010 were enough top quarks produced to study the details of its interactions.

The amount of top quarks produced scale up by the energy as seen in Figure 2.1 as well as the integrated luminosity, or number of events produced, of the accelerator

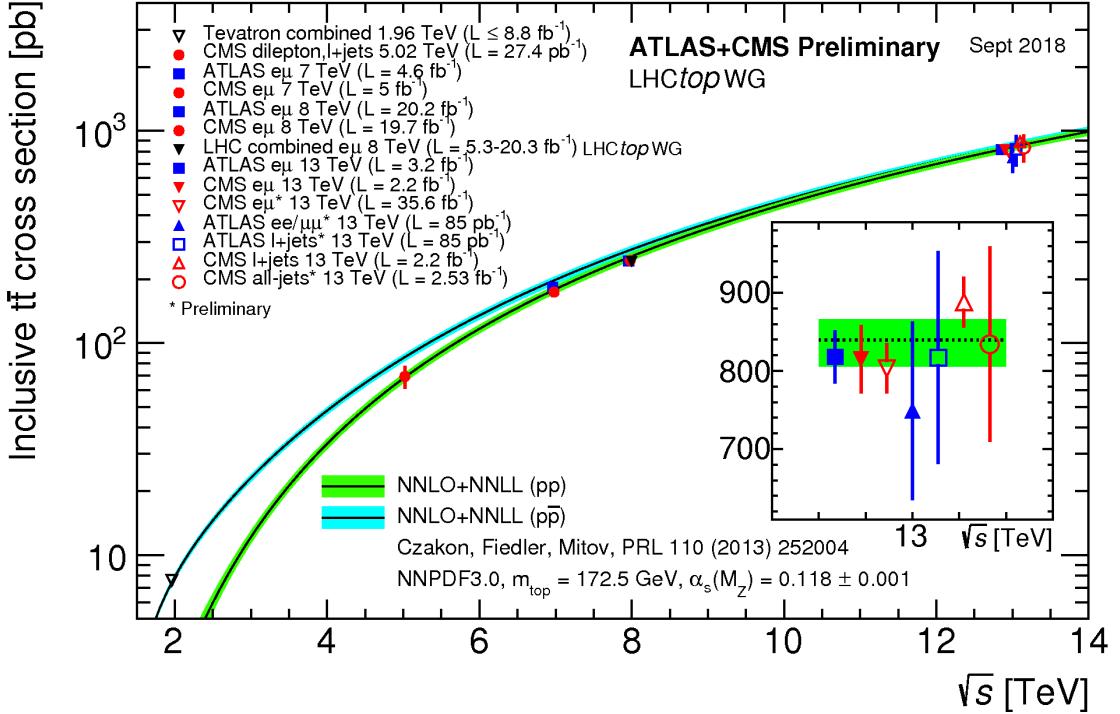


FIGURE 2.1. Summary of $t\bar{t}$ production cross sections as a function of center of mass energy at Fermilab’s Tevatron and CERN’s LHC. For a top mass of 172.5 GeV and center of mass energy of 13 TeV the central value cross section is calculated to be 831 pb [22]

and detector setup. The LHC has significantly more of both than the Tevatron did. Throughout the entirety of Run 2 at the LHC there are expected to be more than 115,000 top pair events produced within the ATLAS detector, allowing physicist to probe the details of the top quark better than ever before.

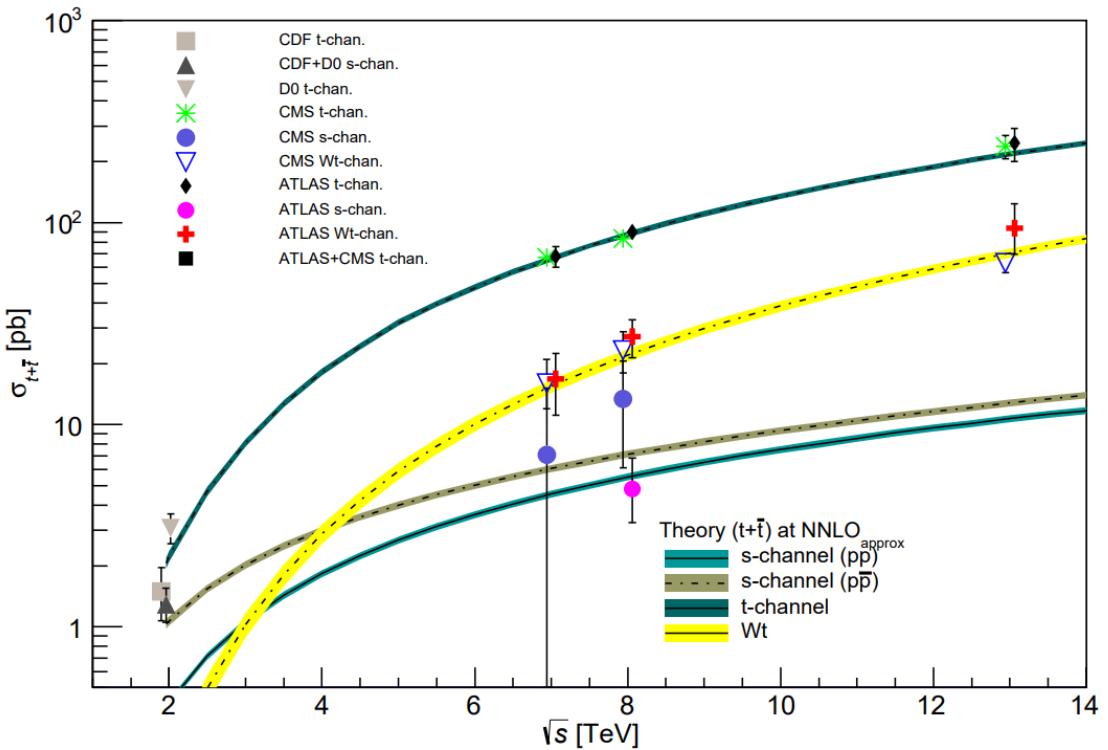


FIGURE 2.2. Summary of single-top production cross sections as a function of center of mass energy at Fermilab’s Tevatron and CERN’s LHC[22]

2.4.2. Production and Decay At Hadron Colliders

There are multiple ways to produce top quarks at the LHC. The most prevalent production mechanism of top quarks is through producing top/anti-top quark pairs. This can be done, to leading order, either by quark-antiquark annihilation (Figure 2.3a) or gluon gluon fusion (Figure 2.3b-d). At the Tevatron, a proton anti-proton collider, the leading diagram was quark-antiquark annihilation because of the significantly larger amount of antiquarks in the collisions. At the LHC the major production mechanism is gluon-gluon fusion, $\approx 90\%$, while quark-antiquark annihilation accounts for only $\approx 10\%$ of top quark pair production. This is a strong interaction process and is therefore quite common. The cross sections are shown in Figure 2.1.

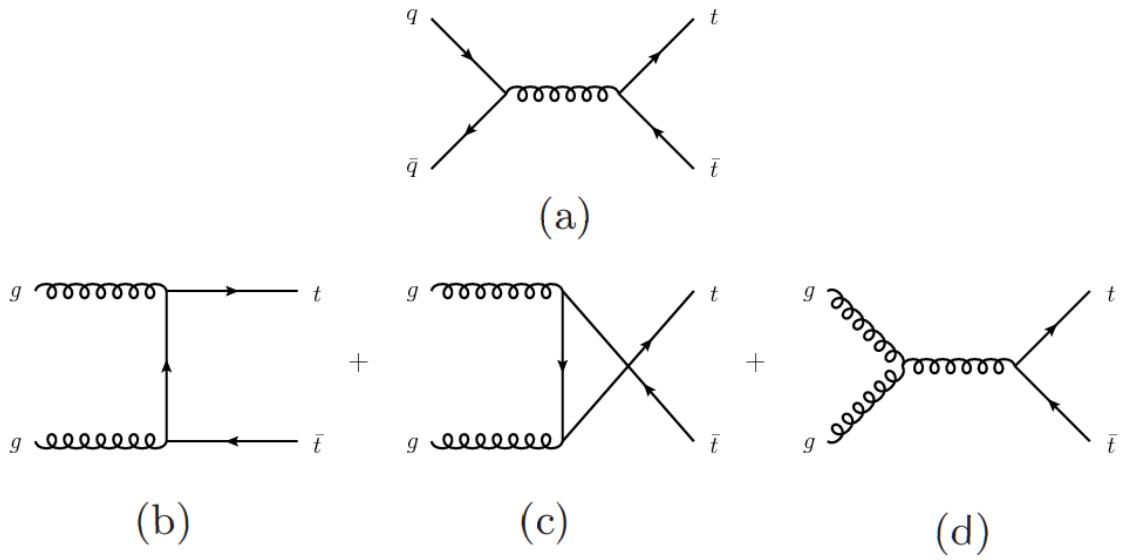


FIGURE 2.3. Leading order diagrams for the production of top/anti-top quark pairs at hadron colliders. Quark-antiquark annihilation diagram in (a) while (b)-(d) show various gluon-gluon fusion diagrams

Single top quarks can also be produced through weak interactions, which is less common, and the leading diagrams are shown in Figure 2.4 and their cross sections as measured at the LHC and Tevatron are shown in Figure 2.2. Comparing the leading production mechanisms at 13 TeV, it can be seen that $t\bar{t}$ production is about a factor of 4 larger than single top production. This means there are about 8x as many top quarks looking at pair produced events, as two are produced per event while giving an additional experimental handle in looking for an invariant mass of final state products around the top quark mass.

Because the top quark decays before hadronization, we can study the branching ratios to various products directly. Figure 2.5 shows ways the top quark is allowed to decay in the Standard Model. Figure 2.5a shows the most likely decays where the top quark decays to a down-type quark and a W boson. The branching ratio of these decays is proportional to the square of the corresponding matrix element in the CKM

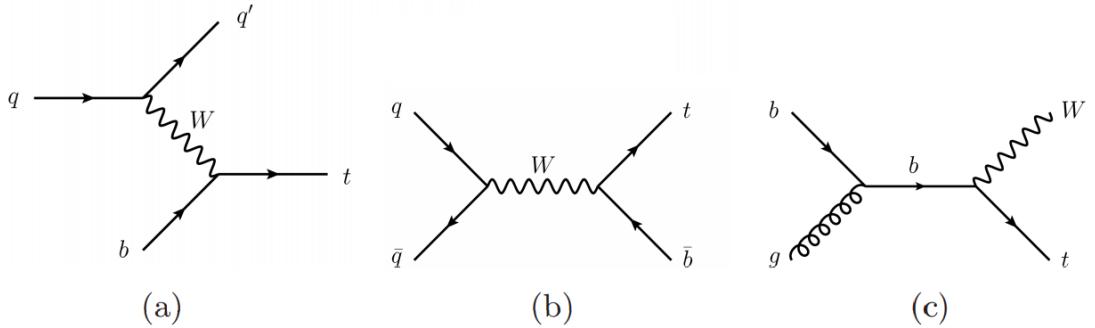


FIGURE 2.4. Leading order diagrams for the production of single top quarks at hadron colliders. The t-channel diagram is shown in (a), the s-channel in (b), and production in association with a W-boson is shown in (c)

matrix as shown in Section 2.3.3. The sum of these branching ratios is unity within standard error bars such that the implication of the diagrams, corresponding to the flavor-changing neutral current decays, shown in Figure 2.5b, are highly suppressed which will be explored further in Section 2.5.

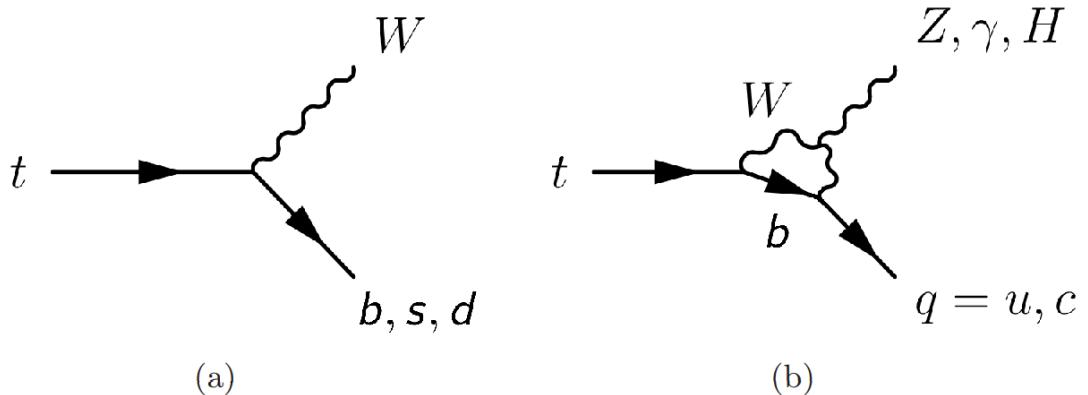


FIGURE 2.5. Top quark decays in the Standard Model

Since the matrix element V_{tb} in the CKM matrix is essentially unity each top usually decays to a b quark and a W boson. The final state of the top pair events are then typically categorized by the decay of the W bosons. The W boson can decay leptonically to a lepton (electron, muon, or tau) and its associated

Decay Mode	States	Decay Mode	States	Decay Mode	States
$W^+ \rightarrow e^+ \nu_e$	1	$W^+ \rightarrow u\bar{d}$	$\times 3 V_{ud} ^2$	$W^+ \rightarrow c\bar{d}$	$\times 3 V_{cd} ^2$
$W^+ \rightarrow \mu^+ \nu_\mu$	1	$W^+ \rightarrow u\bar{s}$	$\times 3 V_{us} ^2$	$W^+ \rightarrow c\bar{s}$	$\times 3 V_{cs} ^2$
$W^+ \rightarrow \tau^+ \nu_\tau$	1	$W^+ \rightarrow u\bar{b}$	$\times 3 V_{ub} ^2$	$W^+ \rightarrow c\bar{b}$	$\times 3 V_{cb} ^2$

TABLE 2.3. Summary of final states of a W^+ boson, when accounting for color permutations and the CKM matrix. The table holds true if you flip all particles to their antiparticles

neutrino or hadronically to quarks. This means top pair events are described as "all-hadronic" when both W bosons decay hadronically, "leptonic" if both W bosons decay leptonically, or "semi-leptonic" if one W boson decays hadronically and the other leptonically. The ratios of these events is shown in Figure 2.6. However, because of how the tau decays and interacts, these events are typically treated separately. Only the electron and muons are considered as the leptons in the "semi-leptonic" final state of a $t\bar{t}$ event. A quick counting (neglecting final state particle masses) of the final states of W bosons is shown in Table 2.3. The unitarity condition of the CKM matrix implies that $|V_{ud}|^2 + |V_{us}|^2 + |V_{ub}|^2 = 1$. The phase space of the final product then has 1 state for each of the leptons and approximately 3 for each of the up type quarks so a single top quark can be expected to decay into a lepton $\approx \frac{3}{9}$ of the time or any combination of quarks the other $\approx \frac{6}{9}$ of the time. This holds when looking at the branching ratios of top quark pairs in Figure 2.6. The end result of this is that the regions of special interest for this dissertation, the electron and muon semi-leptonic final states, occur $\approx 30\%$ of the time. While this is not the largest selection of the final state branching ratios the presence of a lepton makes it easier to look for these final states.

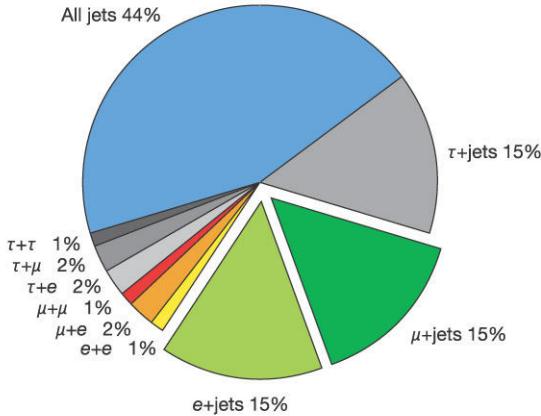


FIGURE 2.6. Categorization of top quark pair decays in the Standard Model based on the decays of the W bosons[23]

2.4.3. Beyond the Standard Model Top Quark Physics

Many questions remain open with the intricacies of the Standard Model involving top quarks. The mass of the top quark is at a mass scale similar to the W, Z, and Higgs bosons. Does this imply the top plays a special role in the mechanism of electroweak symmetry breaking? After the discovery of the Higgs: why do the top quark and Higgs boson have the exact masses they do which allows the electroweak potential to be stable up to very high energy scales as well as allowing the Universe to sit in a meta-stable state[24]? Because of its large mass the top quark is the main destabilizer of the Higgs potential. It may be possible that new physics models will predict phenomena that can be more easily found in the properties of the top quark due to changes in the Higgs sector. Many new physics models will present most dramatically in top quarks as radiative corrections to new massive particles will affect tops first before presenting in deviations in properties of the other, significantly lighter, fermions. For example, a new Higgs like particle would most likely couple most strongly to the top quark. Now that an unprecedented number of top quarks

events are accessible, the predictions of the Standard Model can be tested with much greater precision. Top quark production modes, decay modes, couplings, and various other properties are now being measured and providing limits on the phase spaces of any new physical models that exist or ruling out entire classes of models that have yet to be written down. Arguably the properties of top quark are one of the most likely places that will point to a new understanding of our Universe.

2.5. The Flavor-Changing Neutral Current

The idea of "flavor" in the Standard Model refers to copies of the $SU(3)_C \times U(1)_{EM}$ representation, shown previously in Table 2.2. Specifically it will be used in the discussion of the generational change of one quark into another through some interaction. As an example, Figure 2.5b shows a flavor-changing neutral current interaction.

2.5.1. The Standard Model Flavor Sector

Quarks that interact with W^\pm bosons have interactions that stem from the kinetic term of the Standard Model Lagrangian. If this is written out explicitly in the mass eigenstates the interaction looks like:

$$-\frac{g}{\sqrt{2}} \begin{pmatrix} \bar{u}_L & \bar{c}_L & \bar{t}_L \end{pmatrix} \gamma^\mu W_\mu^+ V \begin{pmatrix} d_L \\ s_L \\ b_L \end{pmatrix} + \text{h.c.}$$

where the W interacts directly to change the flavor of quarks from an down(up)-type to an up(down)-type quark. The CKM matrix can be thought of as a rotation between the mass and interaction basis and the fact that the CKM matrix is non-

diagonal means that the W boson interacts with quarks of different generations. The small off-diagonal elements of the CKM matrix mean that this generational change is a significantly smaller effect than interactions that change flavors within a generation, e.g. a top quark going to a bottom quark and W boson is much more likely than a top quark decaying to a down or strange quark and a W boson. Going two generations away from the diagonal the mixing from the CKM matrix is even smaller. The interaction with a W boson between two quarks is the only interaction vertex in the Standard Model that allows for both flavor and generation changes.

As opposed to this flavor changing charged current interaction the FCNC is an interaction between neutral gauge bosons and fermions. These FCNC processes involve either up or down type quarks or involve charged or neutral leptons. The flavor of the fermion is changed but the electric charge is conserved because it interacts with a neutral boson as opposed to the W^\pm like in the charged current interaction. FCNC interactions are forbidden at tree-level in the Standard Model but can occur via higher order processes such as loops (as shown in Figure 2.5b).

There are four neutral bosons in the Standard Model: the gluon, photon, Higgs boson, and Z boson. Each of these can mediate FCNC interactions but all are forbidden at tree-level. The gluon and photon correspond to exact gauge symmetries and have diagonal, flavor universal couplings since their interactions with fermions come through the kinetic terms. The ramification of this is that they only interact with fermions of the same flavor. The Standard Model Higgs cannot couple to fermions of different flavor since the Standard Model fermions are chiral and the Higgs couplings to fermions align with the fermion mass matrix. In the Standard Model there is only a single Higgs doublet and the only source of the fermionic masses is the Higgs vacuum expectation value. The Z boson can only connect to quarks from the

same type (up or down). When you move from the interaction to the mass eigenstates the rotation matrices only include terms such as $U_{uL}U_{uL}^\dagger = 1$ as opposed to the CKM matrix terms ($U_{uL}U_{dL}^\dagger$) which means these couplings are also flavor universal. The Standard Model FCNC suppression is built in through these means as opposed to generation-changing charged current processes. The charged current processes rely on the CKM parameters which are free parameters of the Standard Model and as such are measured and "put in." These FCNCs are suppressed multiple ways in the Standard Model.

2.5.2. The GIM Mechanism

Historically the original Cabibbo model of particle physics only had three quarks: the up, down, and strange. Studies of Kaon decay during the late 1960's suggested there were no neutral current interactions in the Standard Model at the time. The decay $K^+ \rightarrow \mu^+\nu_\mu$ was observed but the process $K_L^0 \rightarrow \mu^+\mu^-$ which was predicted was not observed. Even in the absence of a tree level decay the K_L^0 decay was still predicted via a box diagram through the exchange of W bosons shown in Figure 2.7.

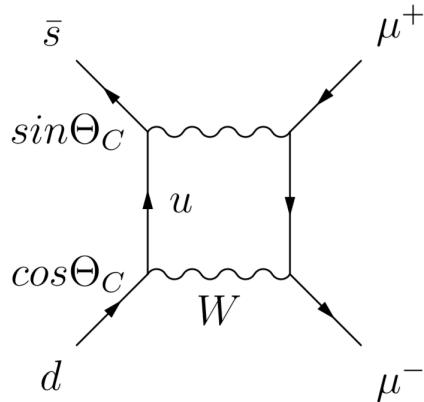


FIGURE 2.7. Box diagram of $K_L^0 \rightarrow \mu^+\mu^-$ through the exchange of W bosons

Interactions at the time were thought to have strangeness quantum number interactions that changed strangeness in the following way:

$$\Delta S = 0 : u\bar{u} + d\bar{d}\cos^2\Theta_C + s\bar{s}\sin^2\Theta_C$$

$$\Delta S = 1 : (s\bar{d} + d\bar{s})\sin\Theta_C\cos\Theta_C$$

The non-observation of the predicted decay led Glashow, Iliopoulos, and Maiani to predict the existence of a fourth quark, the charm, in 1970[19]. The addition of the charm led to two quark doublets and an almost perfect cancellation between the box diagrams:

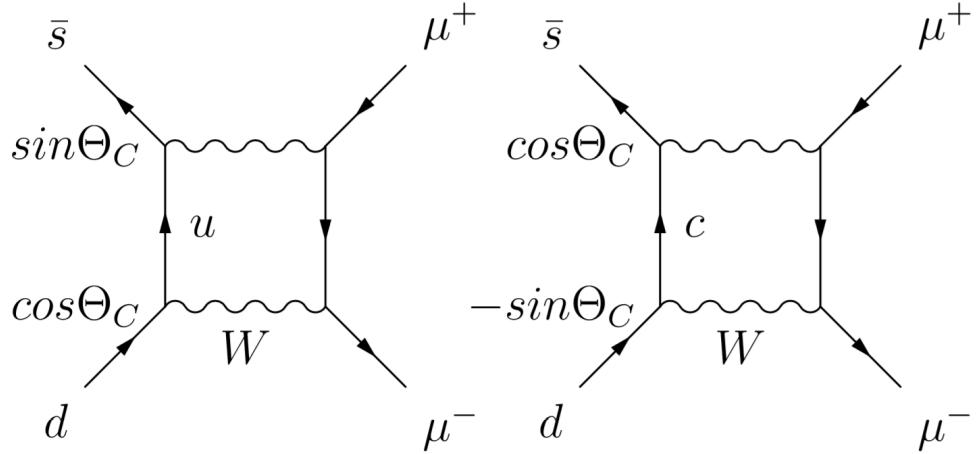


FIGURE 2.8. Box diagrams of $K_L^0 \rightarrow \mu^+\mu^-$ through the exchange of W bosons after the inclusion of the charm quark

These box diagrams mean that the strangeness change in interactions can be rewritten as

$$\Delta S = 0 : u\bar{u} + c\bar{c} + (d\bar{d} + s\bar{s})\cos^2\Theta_C + (s\bar{s} + d\bar{d})\sin^2\Theta_C$$

$$\Delta S = 1 : (s\bar{d} + d\bar{s} - d\bar{s} - s\bar{d})\sin\Theta_C\cos\Theta_C$$

The addition of the charm means that, in the approximation $m_c = m_u$, the $\Delta S = 1$ terms cancel exactly, as the new box diagrams show in Figure 2.8. The FCNC interactions in top quark decays are suppressed through this mechanism as well, with the further inclusion of the bottom and top quarks. They are further suppressed by being proportional to the quark mixing of off-diagonal elements in the CKM matrix, which are significantly less than 1. A loop diagram for top quark FCNC is shown in Figure 2.9. This loop process is very rare. The Standard Model branching ratio for these top FCNC interactions are shown along predicted enhancements from a variety of models of physics beyond the Standard Model in Table 2.4.

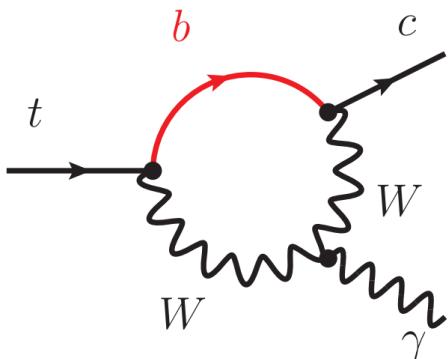


FIGURE 2.9. An example loop diagram of a top quark decaying to a light quark and a photon

Process	SM	2HDM	QS	MSSM	RPV	XD
$t \rightarrow u\gamma$	$4 * 10^{-16}$	—	$\leq 4 * 10^{-8}$	$\leq 10^{-8}$	$\leq 10^{-9}$	—
$t \rightarrow c\gamma$	$5 * 10^{-14}$	$\leq 10^{-7}$	$\leq 4 * 10^{-8}$	$\leq 10^{-8}$	$\leq 10^{-9}$	$\leq 10^{-9}$
$t \rightarrow uZ$	$7 * 10^{-17}$	—	$\leq 6 * 10^{-4}$	$\leq 10^{-7}$	$\leq 10^{-6}$	—
$t \rightarrow cZ$	$1 * 10^{-14}$	$\leq 10^{-6}$	$\leq 6 * 10^{-4}$	$\leq 10^{-7}$	$\leq 10^{-6}$	$\leq 10^{-5}$
$t \rightarrow ug$	$4 * 10^{-14}$	—	$\leq 9 * 10^{-7}$	$\leq 10^{-7}$	$\leq 10^{-6}$	—
$t \rightarrow cg$	$5 * 10^{-12}$	$\leq 10^{-4}$	$\leq 9 * 10^{-7}$	$\leq 10^{-7}$	$\leq 10^{-6}$	$\leq 10^{-10}$
$t \rightarrow uH$	$2 * 10^{-17}$	$\leq 6 * 10^{-6}$	—	$\leq 10^{-5}$	$\leq 10^{-9}$	—
$t \rightarrow cH$	$3 * 10^{-15}$	$\leq 2 * 10^{-3}$	—	$\leq 10^{-5}$	$\leq 10^{-9}$	$\leq 10^{-4}$

TABLE 2.4. Expected branching ratios for various flavor-changing neutral current processes in the Standard Model[11] and multiple theories that predict enhancements to the branching ratio. Two-Higgs Double Models with flavor-violating Yukawa couplings (2HDM) [11, 25], quark single models (QS) [26, 27], minimal supersymmetry models with 1TeV squarks and gluinos (MSSM) [12], R-parity violating supersymmetry models (RPV) [28], and extra-dimensional models (XD) [29].

2.5.3. New Physics With Enhancements to FCNCs

Various theoretical models that include physics beyond what is included in the Standard Model are proposed to solve problems that exist with the Standard Model or an explanation of known phenomena that are not in agreement with the Standard Model. Various models seek to solve different problems, e.g., providing a dark matter candidate or fixing the naturalness problem of the Standard Model resulting from an unexpectedly high amount of fine-tuning from loop corrections to the Higgs mass. Top quark FCNCs in the Standard Model are currently so far from experimental reach (≈ 10 orders of magnitude) that they are impossible to observe even with major improvements to the accelerator and detector technologies. Table 2.4 also shows a variety of theories beyond the Standard Model which predict large enhancements to FCNC top couplings. For the most part these enhancements come from terms that have very heavy particles moving in the loops. Therefore searching for FCNCs

with top quarks provides a particularly good handle to study models of new physics and rule out the phase spaces of these models by lowering the expected limit. An explanation of the various models explored in Table 2.4 follows.

Two-Higgs-Doublet Models (2HDM): 2HDM are a simple extension of the Standard Model which contain two Higgs doublets instead of the one currently contained in the Standard Model. This leads to a much richer phenomenology in the Higgs sector with two CP even neutral Higgs bosons (h and a heavier H), a CP odd pseudoscalar A , and two charged Higgs bosons H^\pm . The currently discovered Higgs boson can be mapped to either h or H depending on various limits in the model of choice. These models can typically be described by an additional six parameters: the four Higgs masses (m_h, m_H, m_A, m_{H^\pm}), the ratio of the two vacuum expectation values and a mixing angle that diagonalizes the mass matrix of the neutral CP even Higgs bosons. Many supersymmetric models predict the existence of an extra Higgs doublet. Some of these models also attempt to explain the baryon asymmetry of the Universe[30]. 2HDM Models predict very large enhancements to FCNC interactions due to an extension of the electroweak symmetry breaking sectors. Some of these models (type III 2HDM, and models of minimal flavor violation) include tree level FCNCs[31] which is why the enhancement brings the branching ratio up to an observable level.

Quark Singlet Models (QS): QS involve an extension to the Standard Model in the form of an extra vector-like quark singlet that couples strongly to the top quark, typically in the form of a top-partner quark. These heavier t' quarks could explain the fine-tuning of the Higgs boson mass through cancellation of some or all of the top loop diagrams present in the radiative corrections to the Higgs mass. These models

generally imply then that the CKM matrix is no longer unitary and tree level FCNCs are allowed which offers a great enhancement to potential branching ratios[26, 27].

Minimal Supersymmetric Models (MSSM): Supersymmetric models where every Standard Model particle has a super particle partner typically aim to solve multiple problems with the Standard Model at once. In general the lightest supersymmetric particle, which is stable, provides a good dark matter candidate. MSSM models have super partner quarks (squarks) and super partner gluons (gluinos) on a mass scale of ≈ 1 TeV. Top FCNCs can occur through loop diagrams still, as they do in the Standard Model, but the loop is enhanced as it includes the supersymmetric particle of the top quark (the heavier stop quark)[12, 32].

R-Parity Violating Supersymmetric Models (RPV): Another supersymmetric model R parity is no longer conserved because all superpartners are odd under the parity. FCNCs can also occur at the one loop level in these models in loops which no longer conserve baryon or lepton number.[28]

Extra-dimensional Models (XD): XD Models are Randall-Sundrum models that describe the Universe as a warped-geometry higher dimensional space where elementary particles are localized on a (3+1)-dimensional brane. These models offer a potential solution to the hierarchy problem of the Standard Model by adding in a mechanism to explain the difference between the typical scales over which FCNCs take place (the electroweak scale) and the Planck scale. In these models FCNCs exist due to flavor-violating couplings between Standard Model fermions and Kaluza-Klein excitations of the gauge bosons in the Standard Model[29]. Due to their overlap in the extra dimension with Kaluza-Klein guage modes the flavor-violating couplings will be largest in the top sector.

2.5.4. Current Measurements of Top FCNCs

All four of the neutral boson mediated FCNC channels can be searched for individually and model independently. Each channel will have its own signature and various advantages and disadvantages in performing the search. Each of the potential tree-level diagrams is shown in Figure 2.10, all of which are forbidden at tree level in the Standard Model.

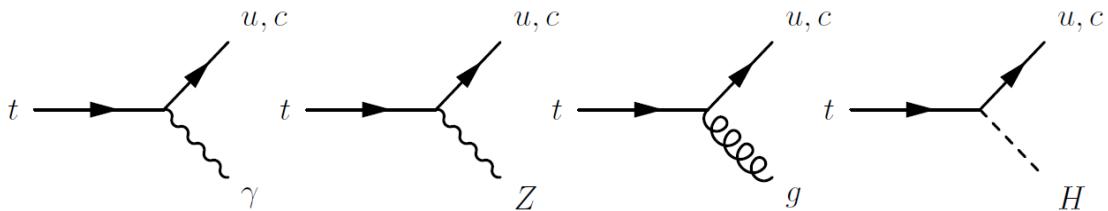


FIGURE 2.10. Flavor-changing neutral current top quark decays.

The FCNC channels involving all of the various neutral bosons can be searched for with the ATLAS detector in both single-top production modes as well as the decay mode through $t\bar{t}$ events. All of the production channels provide a sharp separation between $u \rightarrow tX$ and $c \rightarrow tX$ but have fewer events than searching in the decay mode using $t\bar{t}$ events. Some final states of these various FCNC events such as those involving the decay mode search for $t \rightarrow qZ$ exploit the tri-lepton final state ($Z \rightarrow ll$, with the other top decaying leptonically $t \rightarrow bW \rightarrow bl\nu$). The FCNC process involving the Higgs boson has the advantage of being able to look in a wide range of potential final states of the Higgs and can be successfully tackled using a variety of different methods.

Limits have been set on these FCNC processes at various experiments in the past. The electron-proton collider HERA at DESY, the electron-proton collider LEP at CERN, and the proton-proton colliders the Tevatron at Fermilab and the LHC at

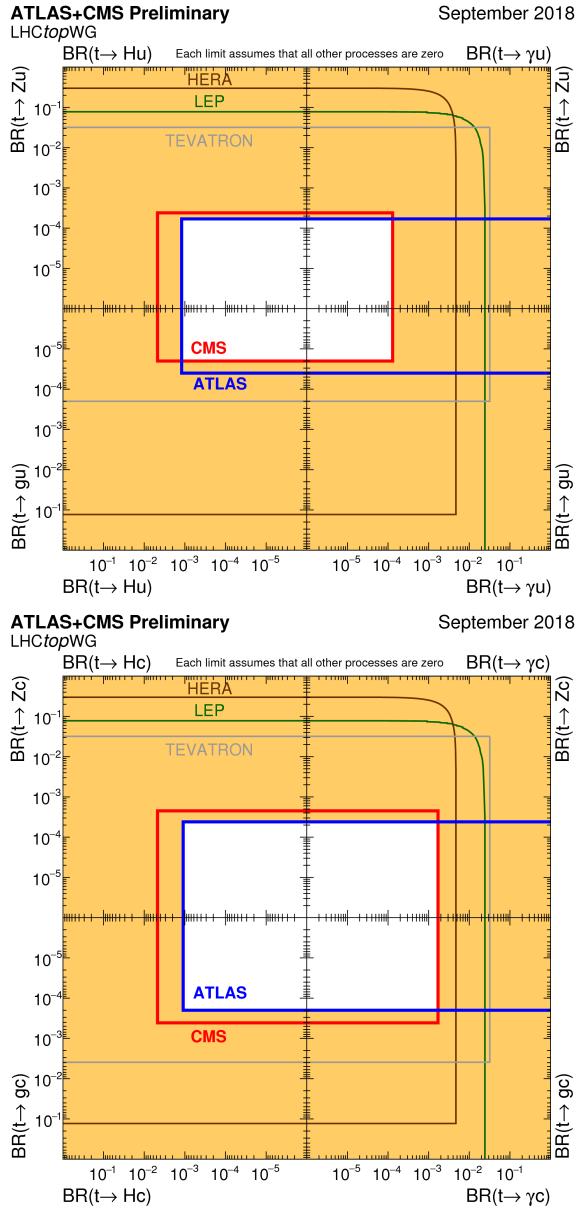


FIGURE 2.11. Flavor-changing limits from a variety of experiments in every channel $t \rightarrow uX$ (top) and $t \rightarrow cX$ (bottom) [22]

CERN have all had experiments searching for the FCNC processes. The collected limits of these experiments are shown in Figure 2.11. Due to the energy of these early colliders 209 GeV center of mass energy for LEP and 318 GeV for HERA only production modes could be searched for as they are below the production threshold

for $t\bar{t}$ pairs. The diagrams searched for at these experiments are shown in Figure 2.12 (HERA) and Figure 2.13 (LEP).

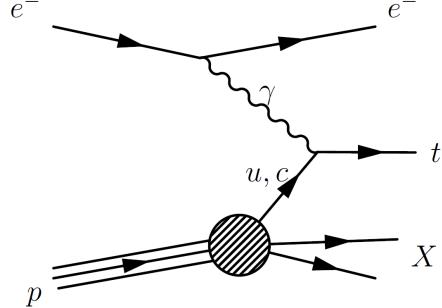


FIGURE 2.12. FCNC diagram for the search in single top production at the electron-proton collider HERA.

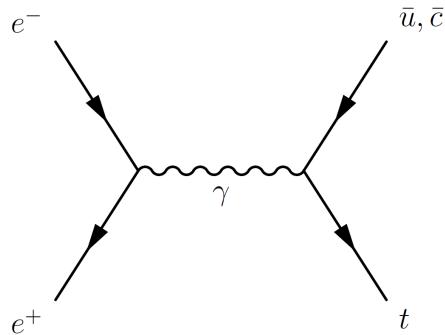


FIGURE 2.13. FCNC diagram for the search in single top production at the electron-positron collider, LEP.

Shown in Figure 2.14 are the Standard Model theoretical predictions, various beyond the Standard Model predictions (as discussed in Section 2.5.3), and experimental limits for all processes $t \rightarrow Xq$ where q is an up-type quark and X is any neutral boson.

Since the publication of Figure 2.14 the ATLAS experiment has published a result in the production mode of the $t \rightarrow q\gamma$ vertex using 81 fb^{-1} of data (LHC data runs between 2015-2017). Upper limits have been set on $t \rightarrow u\gamma$ left-handed (right-handed) branching ratio of 2.8×10^{-5} (6.1×10^{-5}) and upper limits on $t \rightarrow c\gamma$

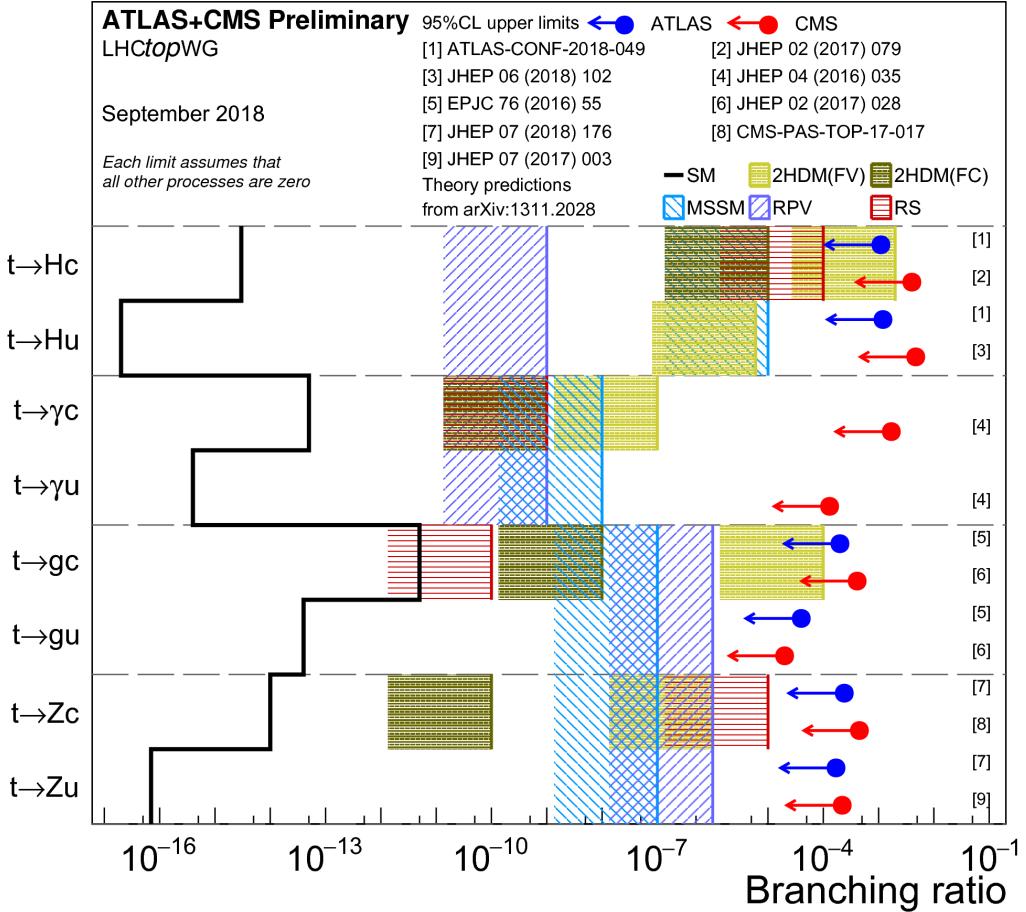


FIGURE 2.14. Flavor-changing neutral current theoretical branching ratios and experimental limits (both ATLAS and CMS) updated through 2018 [22]

left-handed (right-handed) branching ratio of 22×10^{-5} (18×10^{-5}) [33]. Production mode searches offer better reach for $t \rightarrow u\gamma$ over $t \rightarrow c\gamma$ due to the higher prevalence of up quarks in the parton distribution function of the colliding protons.

This dissertation presents a search for the process $t \rightarrow q\gamma$ in the decay channel using top quark pair events. The process $t\bar{t} \rightarrow bWq\gamma \rightarrow bl\nu q\gamma$ will be searched for using the ATLAS detector. This final state ($bl\nu q\gamma$) is a straightforward channel in that it has one of each type of object that can be reconstructed with ATLAS. There are 2 jets from the quarks but one is a b-jet that has qualities that can distinguish it from a normal quark jet which will be discussed in Section 4.3.4.1. In addition to the

jets the final state also has a charged lepton (an electron or muon), a photon, and a neutrino. The reconstruction of these objects is discussed in Chapter IV. The FCNC process is a unique process where a top quark decays directly into an up-type quark and a photon. The jet and photon we measure are required to be reconstructed into a top quark by measuring the invariant mass requirement which gives an excellent handle for separating signal from background.

CHAPTER III

THE LARGE HADRON COLLIDER AND THE ATLAS DETECTOR

This chapter describes the experimental details of the collider complex at the LHC and specifically the ATLAS detector used to produce, collect, and measure various particle properties. The subsystems of the ATLAS detector are described in detail.

3.1. The Large Hadron Collider

The LHC is the world's largest and most energetic particle accelerator. As a hadron collider the LHC collides particles made up of quarks, typically proton-proton collisions. Protons, as opposed to electrons/positrons at a previous collider such as LEP, have much higher mass and a significantly smaller amount of energy loss during acceleration due to synchrotron radiation (which scales as $\frac{1}{m^4}$). Due to this the LHC is able to reach a much higher center of mass energy using the same circular ring used by LEP, though this higher energy comes at a cost. Due to hadrons being made up of constituent partons (quarks and gluons), not all of which interact in any given collision, the particles that do not take place in the hard interaction are left over and create a 'messier' environment in the detectors. This is opposed to lepton colliders, where all of the energy that goes into the collision is present in the final state particles coming from the interaction point. The implication of this is that at hadron colliders the momentum along the beam axis cannot be known, only momentum in the transverse plane of a collision is known due to conservation of momentum.

The LHC is housed in a 27 km ring running beneath the Franco-Swiss border near Geneva, Switzerland and accelerates beams of protons (ions) to a center of mass

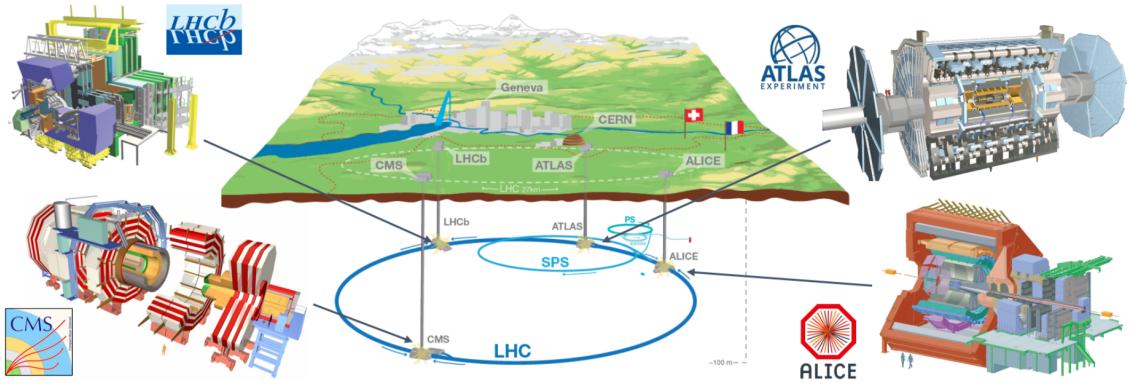


FIGURE 3.1. Map of LHC and the various detector experiments: ATLAS, CMS, LHCb, and ALICE located under the Franco-Swiss border near Geneva[34]

energy of 13 TeV (5 TeV) using two counterpropagating circular beams around the ring. The particles are then collided at one of the four primary interaction points, each of which house a dedicated detector as shown in Figure 3.1.

In addition to the LHC beam line the accelerator uses a series of smaller accelerators to increase the energy of the particles before being introduced into the LHC. This accelerator complex is detailed in Figure 3.2. The start of the accelerator chain, and source of LHC protons, is the Linear Accelerator 2 (LINAC 2, purple) where hydrogen gas is placed inside of an electric field that separates the protons and electrons. The remaining protons are passed through radiofrequency (RF) cavities and accelerated to 50 MeV using electric fields which oscillate at a frequency specific to the distance between any two RF cavities.

After leaving LINAC 2 the protons are injected into the Proton Synchrotron Booster (BOOSTER, light purple) and accelerated to 1.4 GeV before being passed to the Proton Synchrotron (PS, magenta) in two batches with a separation of 1.2 seconds. The PS accelerates the protons to 25 GeV to be injected into the Super Proton Synchrotron (SPS, blue) in a series of four batches separated by 3.6 seconds and are accelerated to 450 GeV. The SPS is the second largest accelerator in the

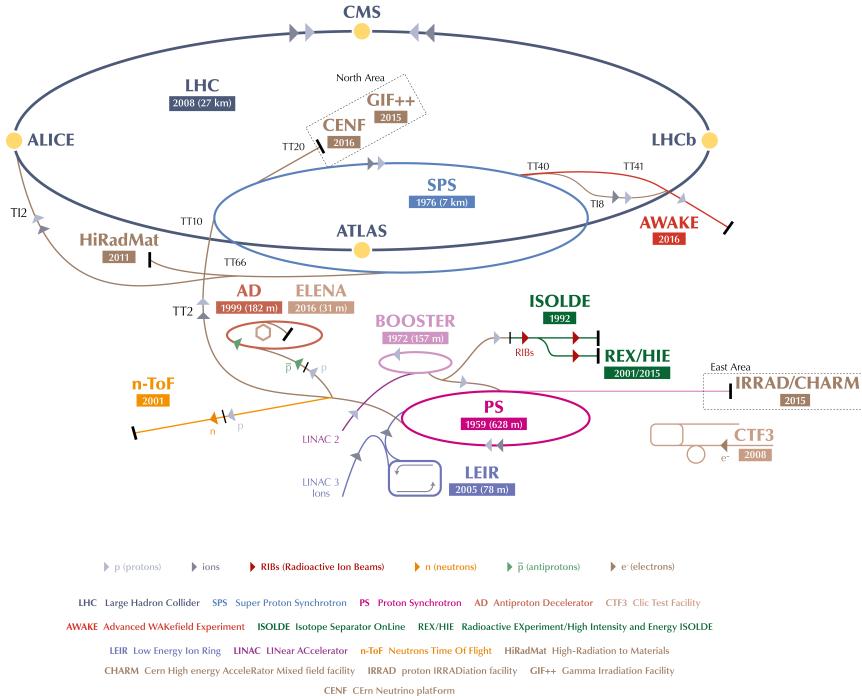


FIGURE 3.2. Schematic of the CERN accelerator complex.[35]

complex. After reaching the 450 GeV of the SPS the particles are split and injected into the LHC in opposing directions where they are further accelerated to a collision energy of 6.5 TeV per beam leading to a center of mass energy of 13 TeV for the LHC during Run-2.

The first proton-proton collisions were produced in the LHC in 2008 at the injection energy of the SPS, $\sqrt{s} = 900$ GeV. During testing a faulty electrical connection caused a magnet quench, or a sudden loss of superconductivity, to occur. This broke the nearby magnets and caused a delay in operations until late 2009 when LHC Run-1 began at a collision energy of $\sqrt{s} = 7$ TeV and later raised to $\sqrt{s} = 8$ TeV in 2012 to complete Run-1. Various upgrade and repairs on the LHC occurred

throughout the long shutdown between 2012-2015 where the center of mass energy was increased to the LHC Run-2 energy of $\sqrt{s} = 13$ TeV.

3.1.1. LHC Magnets

The energies achieved in the collisions are only possible due to the LHC magnets that bend and focus the colliding particles. The LHC uses the most powerful magnet technology that can be produced on an industrial scale. There are 1232 superconducting dipole magnets each being 15m in length, weighing over 35 tons, and producing uniform magnetic fields of up to 8.4 T. The niobium-titanium cables must be cooled to 1.9 K and operate with a current of 11,800 A. Of these 1232 magnets 1104 are used to bend the particles around the ring and the remaining 128 are used in the beam dump. To achieve the same center of mass energy using standard non-superconducting magnets the 27 km LHC would instead have to be upwards of 120 km long.

Since the bunches of particles are charged they will naturally diverge while traveling if not focused. To correct for this an additional 392 quadrupole magnets, 5-7m in length, are used to focus the beam. These quadrupoles are used in pairs: one which focuses in the horizontal plane and defocuses in the vertical plane and the other which focuses in the vertical plane and defocuses in the horizontal plane. Together these magnets keep the beam squeezed to a usable size. All of these magnets have two apertures, one for each of the counter-propagating beams.

3.1.2. Luminosity

The amount of data collected at collider experiments is determined not only by the center of mass energy of colliding particles but also the rate of events produced.

This rate is called the luminosity and can be determined by the square of the number of particles in each bunch (since any one in one bunch can interact with any one in the other), the time between bunches, and the cross section of the bunch (or probability of a collision).

$$\mathcal{L} = \frac{1}{\sigma} \frac{dN}{dt}$$

For any given proton-proton pair the luminosity can be expressed as:

$$\mathcal{L} = \frac{1}{4\pi\sigma_x\sigma_y}$$

and can be expanded for the whole beam with the inclusion of the number of protons per bunch (N_1 and N_2), the number of bunches (N_b), and the frequency at which the bunches overlap (f) to:

$$\mathcal{L} = \frac{N_1 N_2 N_b f}{4\pi\sigma_x\sigma_y}$$

which can be iterated over the running time of the LHC (the total time with beams of proper size and energy propagating through the LHC) giving the total delivered luminosity. This total integrated luminosity as a function of time during LHC Run-2 is shown in Figure 3.3. This luminosity value can be multiplied by the probability, or cross section, of any particular final state to obtain the number of times that final state is produced with a given luminosity.

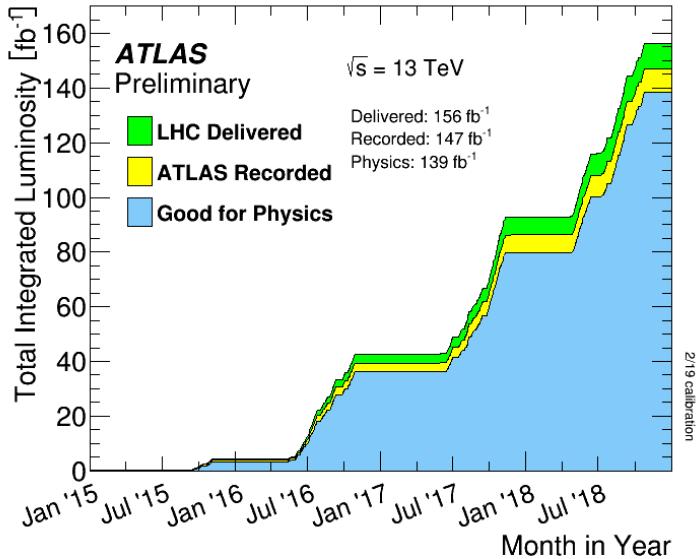


FIGURE 3.3. Total integrated luminosity as a function of time delivered by the LHC(green), recorded (yellow) and declared good for physics analysis (blue) by the ATLAS detector throughout Run 2 consisting entirely of 13 TeV pp collisions. [36]

3.1.3. Pileup

Increasing the luminosity is very beneficial for increasing the statistics needed when searching for rare events but it brings additional challenges as well. Most interactions at any given detector are not hard-scatter events that correspond to potentially interesting physics cases but are instead soft collisions which create noise in the various detector experiments. The LHC works hard to deliver as much data to the experiments as possible and delivers bunches of protons at a time. It is possible for multiple pairs of protons to undergo these soft inelastic collisions at a time. The average number of interactions per bunch crossing, or pileup $\langle \mu \rangle$, for Run-2 was 33.7, shown in Figure 3.4. The pileup must be accounted for when separating the tracks and energy deposited from an interesting hard-scatter event from the other soft collisions which occur at nearly the same time within a detector. The difficulty of separating out one event from another can be seen in Figure 3.5 where there are 28 reconstructed

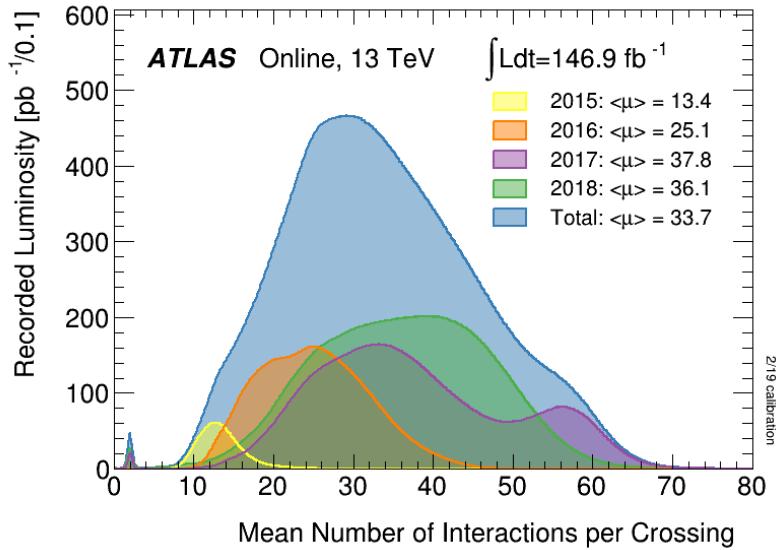


FIGURE 3.4. Luminosity-weighted distribution of the mean number of interactions per bunch crossing for the entirety of Run 2 shown by individual years, 2015 (yellow), 2016 (orange), 2017 (purple), 2018 (green), as well as an integrated total (blue).[36]

verticies. An extreme case of 65 reconstructed verticies is also shown in Figure 3.6. As the LHC will continue to operate at higher and higher luminosities in the future, the amount of pileup that will need to be dealt with will continue to increase.

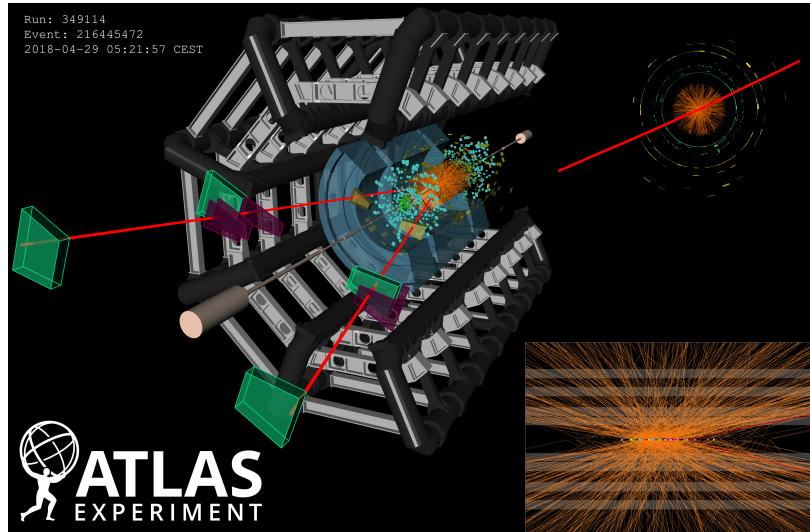


FIGURE 3.5. A candidate dimuon event ($Z \rightarrow \mu^+\mu^-$) with 28 reconstructed verticies collected in 2018 with the ATLAS detector.[36]

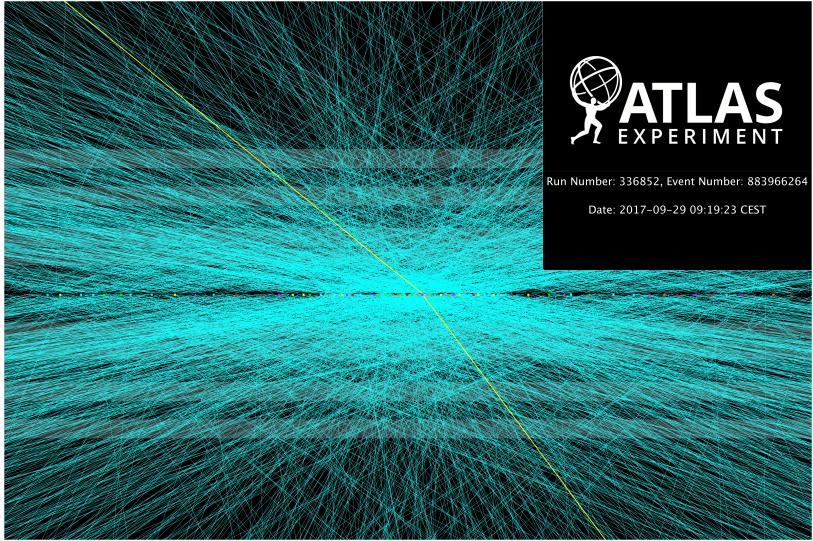


FIGURE 3.6. A candidate dimuon event ($Z \rightarrow \mu^+ \mu^-$) with 65 reconstructed verticies collected in 2017 with the ATLAS detector.[36]

3.2. The ATLAS Detector

The ATLAS detector, depicted in Figure 3.7, is one of the two general-purpose detectors at the LHC. It is the largest detector of its kind ever built at 46 meters in length, 25 meters in diameter, weighing 7000 tons, and containing around 3000 kilometers of cables[37]. Around the interaction points within the detector the ATLAS detector covers nearly the entire solid angle and is nominally symmetric. ATLAS is built up of a variety of concentric subsystems, which will be discussed throughout this section, each with a specialized task and optimized for the measurement of different particle signatures. The primary subsystems used to measure particle trajectories and momenta accurately are the inner detector (Section 3.2.3), the hadronic and electromagnetic calorimeters (Section 3.2.4), and the muon system (Section 3.2.5). The inner detector measures the paths of charged particles, called tracks. The electromagnetic and hadronic calorimeters measure the energy of charged and neutral

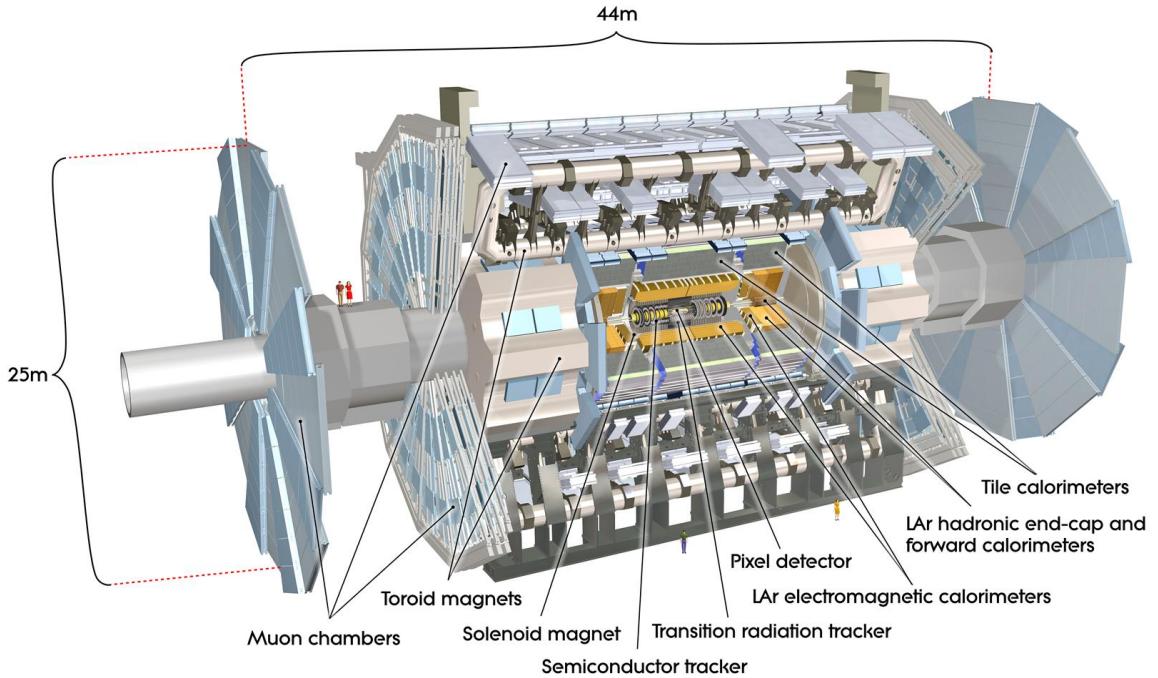


FIGURE 3.7. Schematic of the ATLAS detector.[37]

particles. The muon system measures the momenta of minimum ionizing particles (MIPs).

In addition to the various detectors and calorimeters the ATLAS detector has a magnet system (Section 3.2.2) that bends charged particles in the detector, allowing for a measurement of their charge and momentum and distinguishing them from neutral particles. Between the inner detector and the calorimeters is a solenoid which provides an axial magnetic field. Between the calorimeters and the muon system is a toroidal magnet, from which ATLAS got its original acronym (**A** Toroidal **L**H_C **A**pparatu**S**).

3.2.1. Common Detector Variables and the ATLAS Coordinate System

The ATLAS detector uses a right-handed coordinate system with the origin at the interaction point. In a Cartesian coordinate system the z-axis is defined to be along the beam pipe (positive towards LHC Point 8) while the x-axis points toward the center of the LHC ring which means the positive y-axis points upwards as shown in Figure 3.8. In practice coordinates used are a modified polar coordinate system. In the transverse (xy-)plane to the beam line the azimuthal angle, ϕ , is measured around the beam axis and radius, r , are used. Away from the transverse plane the pseudorapidity, η , is defined by the polar angle (from the y-axis), θ , to be $\eta = -\ln[\tan(\theta/2)]$. Differences in η are Lorentz invariant under longitudinal boosts such that the differences in the rest frames of colliding particles are not important for massless particles. Since the particles typically present in the ATLAS detector are highly energetic, and therefore have a large boost, the pseudorapidity is a good estimate of the true rapidity of the particles. Massless particles are also produced uniformly in η and not in θ which is why η is preferred.

The distance between any two objects within the ATLAS detector can be described geometrically by the variable $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$. Another common variable used is the missing transverse energy, E_T^{miss} . The information known about the missing energy is limited to the transverse plane because the momenta of the colliding particles is unknown along the beamline.

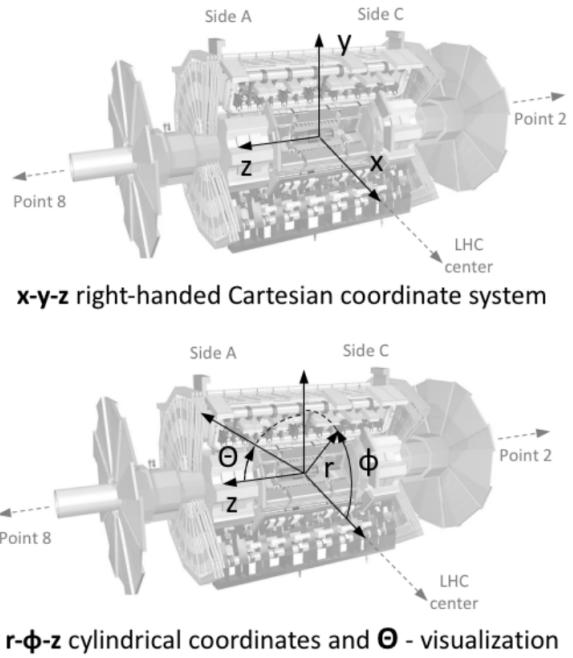


FIGURE 3.8. Coordinate system used in the ATLAS Collaboration.[34]

3.2.2. Magnet Setup

The ATLAS detector has two magnet systems of note. The first is the superconducting solenoid that surrounds the inner detector with a magnetic field aligned with the beam axis. The solenoid has a magnetic field of 2T that makes the tracking of charged particles possible with the inner detector. This magnet is a thin single layer coil, which is imperative to minimize the amount of material in front of the calorimeters.

The toroid system consist of two parts, the end-cap and the barrel magnets. The windings of these magnets is shown in Figure 3.9. Each of these magnets consists of eight superconducting air-core coils, together weighing 830 tons. The end-cap coils are interleaved with the barrel coils. A peak magnetic field strength of 3.9T (4.1T) is achieved in the barrel (end-cap) toroid which assists in the track and momentum

measurement of high energy muons as they leave the ATLAS detector. The barrel toroids cover a range of $|\eta| < 1.4$ while the endcap toroids cover the range $1.6 < |\eta| < 2.7$. The remaining region is covered by a combination of the field of the two sets of toroids.

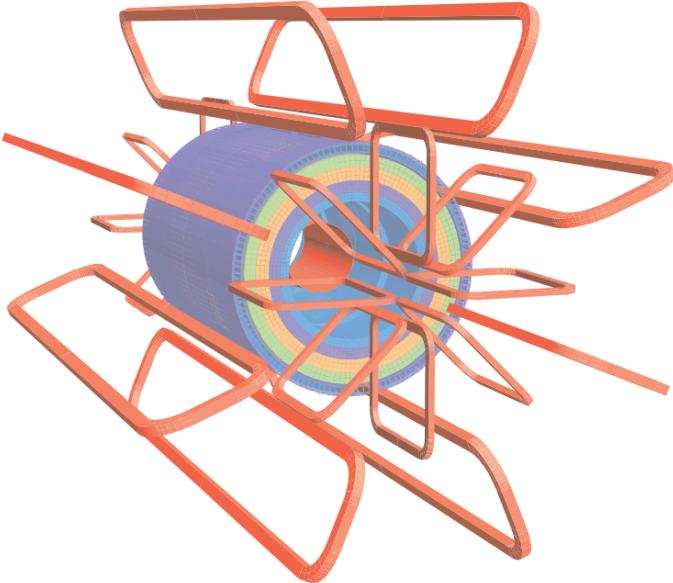


FIGURE 3.9. Schematic of the windings of the ATLAS magnet.[37]

3.2.3. Inner Detector

The inner detector sits inside the solenoid magnet and is used to reconstruct charged particle tracks as they bend due to the magnetic field. The inner detector is made up of four distinct parts. The Insertable B-Layer (IBL) [38], the Pixel Detectors, the Semiconductor Tracker (SCT), and the Transition Radiation Tracker (TRT)[39]. The inner detector provides complete coverage for charged particle tracking, extending to $|\eta| < 2.5$. Momentum resolution as well as primary and secondary vertex measurements are done using the inner detector. Secondary vertices are important for identifying of particles with delayed decays such as bottom quarks (Section 4.3.4.1),

charm quarks, and tau leptons. A schematic of the inner detector can be seen in Figure 3.12.

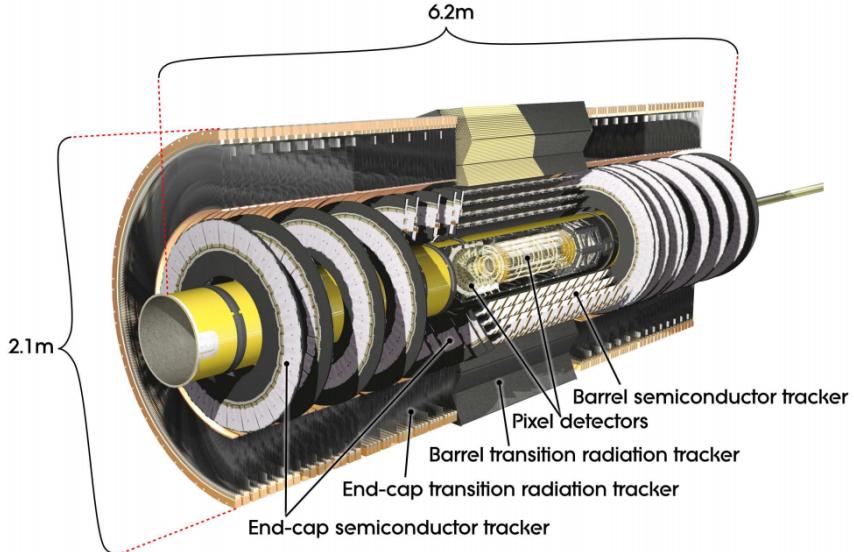


FIGURE 3.10. Schematic of the ATLAS inner detector.[37]

The IBL was added during Long Shutdown I (2016) of the LHC and is closer to the interaction point than the innermost layer in Run 1. This required adding a smaller beam pipe (reduction in radius from 29 mm to 25mm) but was able to improve the resolution of vertexes and thus the reconstruction of events involving bottom quark decays as well as allowing for charm quark decays to be classified better than ever before. This improvement is shown in Figure 3.11 in a study of the impact parameter resolution. An improvement of up to 40% is seen with the inclusion of the IBL in the low p_T region. The IBL functions as a fourth layer of the Pixel Detector and uses planar sensors (similar to the Pixel Detector) as well as 3D sensors allowing electrons to interact with the bulk of the sensor as opposed to just the surface.

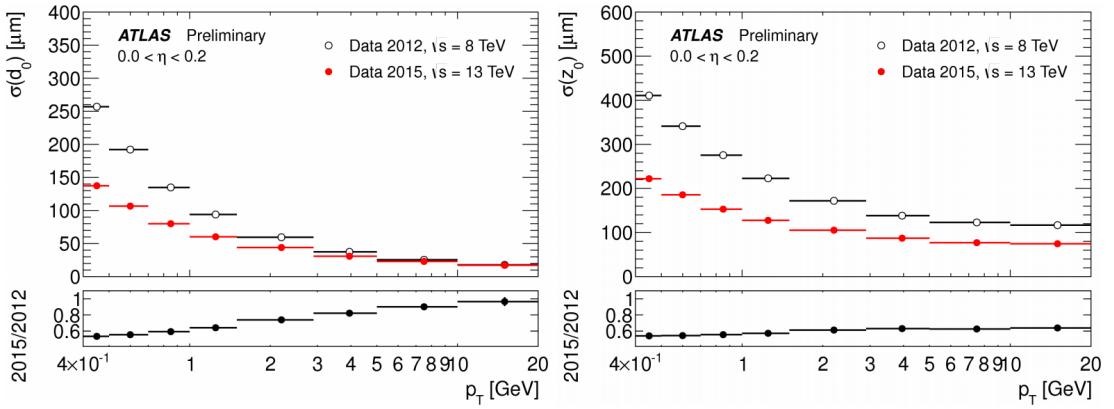


FIGURE 3.11. Unfolded transverse (left) and longitudinal (right) impact parameter resolution measured with (Run 2) and without (Run 1) the IBL as a function of p_T .[40]

The next layer from the beam pipe, as detailed in Figure 3.12, is the Pixel Detector which is a series of high granularity silicon pixel detectors which measure a position when a charged particle passes through them. These silicon pixels are n-doped silicon wafers biased with a high voltage that allow for the creation of electron hole pairs. The electron then drifts toward the electrode which creates the position signal in the readout electronics. In addition to the IBL there are three more cylindrical layers which are designed to ensure single pixel isolation and minimize leakage. The pixels are $50 \times 400 \mu\text{m}^2$. For complete coverage to the cylindrical system, endcaps are placed on each side of the central barrel. These endcaps consist of four wheels that have trapezoid shaped silicon pixels. The three barrel layers consist of 67 million pixels and the endcaps total an additional 13 million pixels. After the Pixel Detector is the Semiconductor Tracker (SCT) which is also made up of barrel and endcap detectors. The barrel SCT is four cylindrical layers of silicon microstrip trackers where the endcaps are nine discs on each side of the barrel made up of either silicon or gallium arsenide semiconductors. The SCT contains over 60m^2 of silicon detectors with over 6 million readout channels.

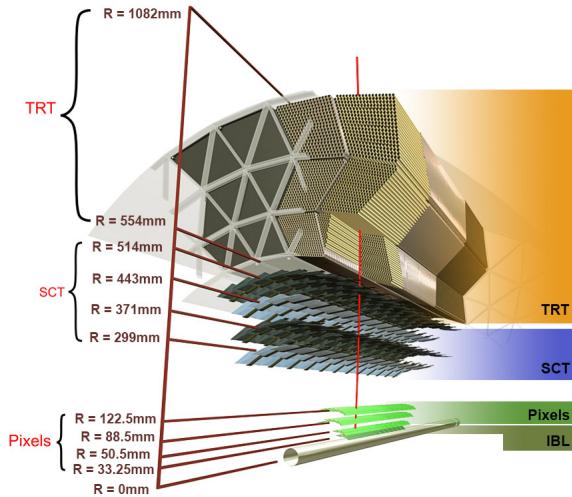


FIGURE 3.12. Blown up schematic of the ATLAS inner detector with more detail.[41]

The final part of the Inner Detector is the Transition Radiation Tracker (TRT) [42]. The TRT is a straw detector surrounding the SCT. Every straw is a 4 mm in diameter Kapton tube with a 0.03 mm diameter gold-plated tungsten wire in its center. In the barrel region there are 50,000 straws that are each 144 cm long and an additional 250,000 straws in both endcaps which are 39 cm in length. Each straw is filled with an active gas mixture made up of mostly Xenon or Argon.

When charged particles traverse across the TRT straws they ionize the active gas mixture and produce ionization clusters. The amount of clusters created depends on how far the charged particle traveled through the TRT (5-6 clusters per mm). The straw walls are held at a high negative voltage such that the primary electrons are accelerated toward the gold-plated tungsten wire anode creating more ionization by liberating more electrons from the active gas and producing a detectable signal which is amplified and read out. Transition radiation occurs when a particle makes a transition between materials with different dielectric constants and the energy radiated is directly proportional to the Lorentz factor of the particle. This allows for an excellent discrimination between electrons and charged pions.

3.2.4. Electromagnetic and Hadronic Calorimeters

While the inner detector focuses on tracking the charged particles as they pass through the detector the ATLAS Calorimeter system is designed to absorb and measure the energy of neutral and charged particles. The exceptions to this are muons which are able to penetrate through the calorimeters into the muon and neutrinos which do not interact at all within the ATLAS detector. The calorimeters can be broken down into two major systems, the Liquid Argon (LAr) calorimeter[43] and the tile calorimeter (TileCal)[44]. Both of these systems are shown in Figure 3.13.

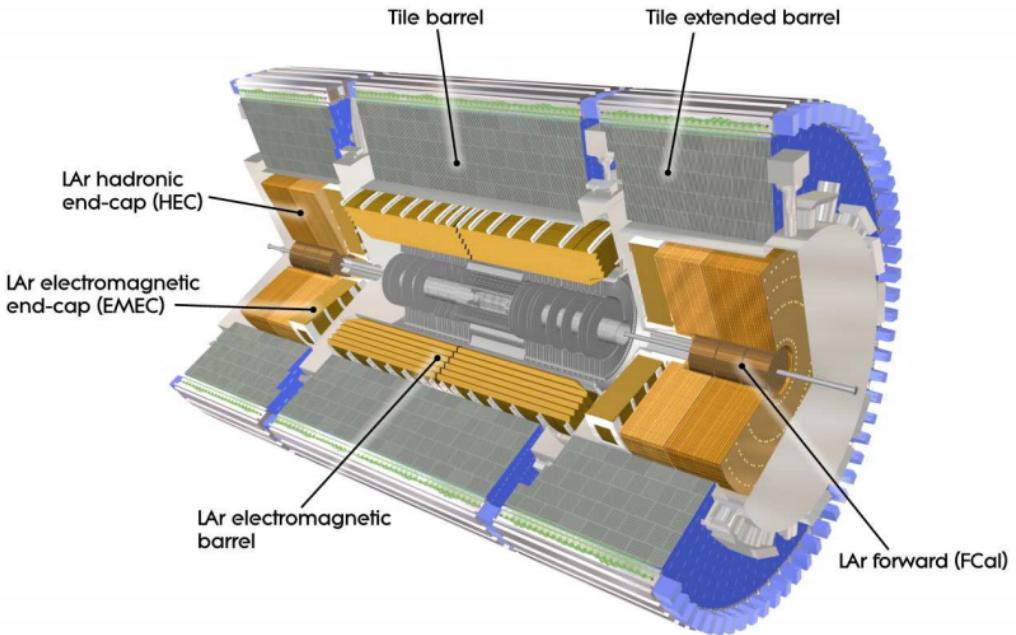


FIGURE 3.13. Schematic of the ATLAS hadronic and electromagnetic calorimeter systems.[37]

The LAr calorimeter is a sampling calorimeter. Sampling calorimeters use alternating layers of a dense absorbing material and an active material to measure the signal produced by showering particles. The LAr calorimeter uses lead as the absorbing material and liquid argon measured with copper-tungsten sensors as the

active layer. The layers in the LAr calorimeter are arranged in an "accordian-shaped" geometry shown in Figure 3.14 to provide complete azimuthal coverage. This allows for the electromagnetic energy resolution to be uniform in the azimuthal direction. Sampling calorimeters do not directly measure the entire energy of the particle, only the interactions that occur in the active layers. The stochastic nature of the processes being measured means that large fluctuations can occur while measuring electromagnetic showers. These fluctuations mean that sampling calorimeters must account for sampling statistics as opposed to other types of calorimeters where the entirety of the energy is absorbed with an active layer, such as scintillators.

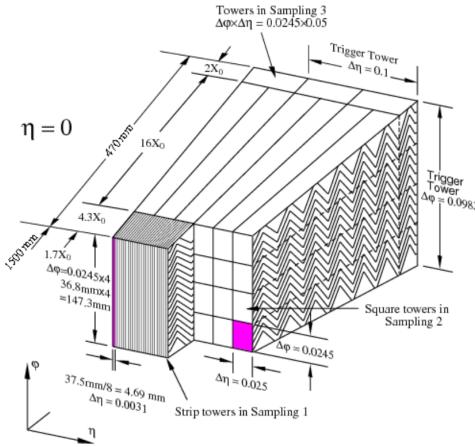


FIGURE 3.14. Sketch of the accordian structure used in the LAr Calorimeter.[43]

For sampling calorimeters it is important to know the ratio $\frac{E_{\text{visible}}}{E_{\text{deposited}}}$ so that the energy of a particle can be reconstructed based on only the energy measured by the active layers. This ratio must be measured with test beams where the original beam energy is known precisely. Sampling calorimeters allow for the complete detection of electromagnetic showers. Because there is a large amount of material to traverse through, all of the energy can be deposited within the detector. The amount of material traversed by each particle is an important aspect as it includes not only the

active material and absorber but also the support structures and cables that can play a role in particle interactions. The thickness of a material passed through is typically measured in radiation lengths (X_0), where an electron passing through one radiation length will lose $1/e$ of its energy to bremsstrahlung. The amount of radiation lengths in the LAr calorimeter is shown in Figure 3.15.

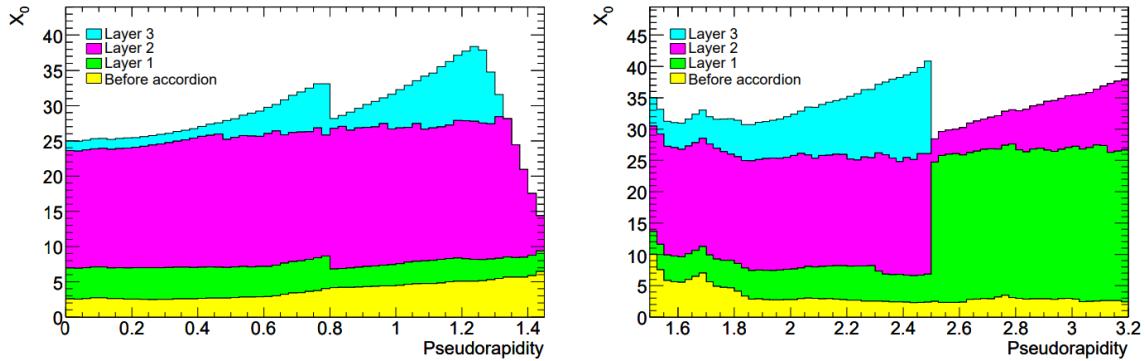


FIGURE 3.15. Number of radiation lengths throughout the LAr calorimeter as a function of $|\eta|$ [37]

Forward from the barrel there are two electromagnetic endcap (EMEC) wheels with a similar accordion structure to the modules in the barrel that cover ranges $1.4 < |\eta| < 2.5$ and $2.5 < |\eta| < 3.2$. Outside of the EMEC wheels is the LAr hadronic endcap (HEC) with a simpler parallel plate structure. The last part of the LAr calorimeter is the LAr forward calorimeter (FCal). Due to the FCal's proximity to the beamline the particle flux is very high so a dense calorimeter is used to avoid losing energy into other pieces of the detector. The FCal is made up of three layers: the first is copper and the others are tungsten.

The remaining calorimeter system is the TileCal which is primarily responsible for hadronic calorimetry in the central region $|\eta| < 1.7$. TileCal is also a sampling calorimeter with iron plate absorbers and plastic scintillating tiles. The scintillating tiles are placed orthogonal to the beamline and readout using wavelength shifting

fibers connected to photomultiplier tubes on the outside of the system. TileCal has a fixed central barrel and two extended barrel sections as shown in Figure 3.13. The extended barrel sections can be moved. The total nuclear interaction length of the TileCal is 7.4λ , where λ is the mean distance a hadronic particle will travel before experiencing an inelastic interaction with the material it is traveling through. The total interaction length for each section of the calorimeter is shown in Figure 3.16.

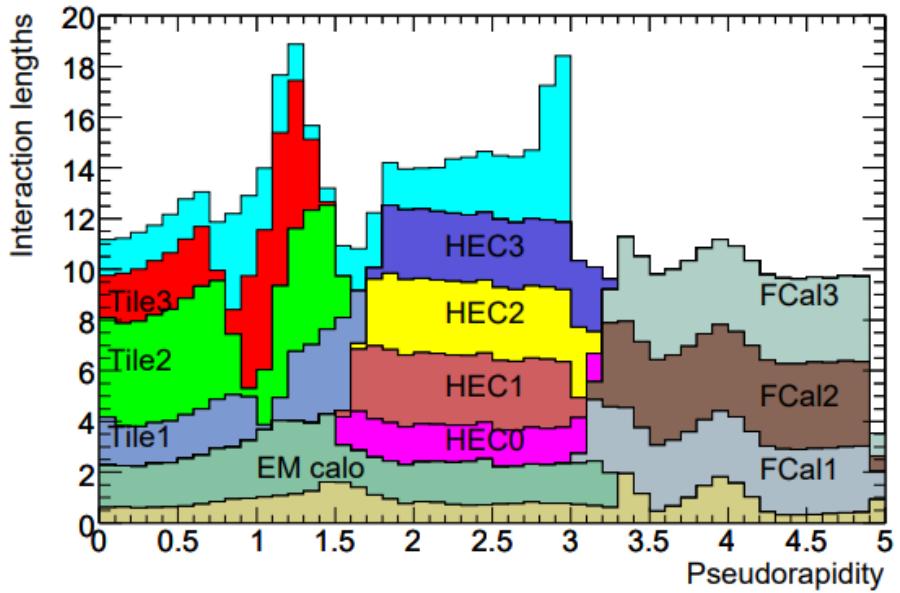


FIGURE 3.16. Number of interaction lengths throughout the LAr calorimeter as a function of $|\eta|$

3.2.5. Muon System

The final and outermost subdetector of the ATLAS detector is the muon spectrometer, which measures the momentum of muons. Different technologies are used in the barrel and endcap regions for both measurement and triggering (deciding which events to keep when only a small fraction of events can be recorded). For the barrel region, $|\eta| < 2.7$, three layers of Monitored Drift Tubes (MDT) are used for

precision energy and tracking measurements and Resistive Plate Chambers (RPC) for triggering. In the forward region, $2.0 < |\eta| < 2.7$, where the flux is higher, Cathode Strip Chambers (CSC) are used for energy and position measurements, and Thin Gap Chambers (TGC) are used for triggering. These systems, shown in Figure 3.17, are aided by the magnetic field created by the toroid system discussed in Section 3.2.2.

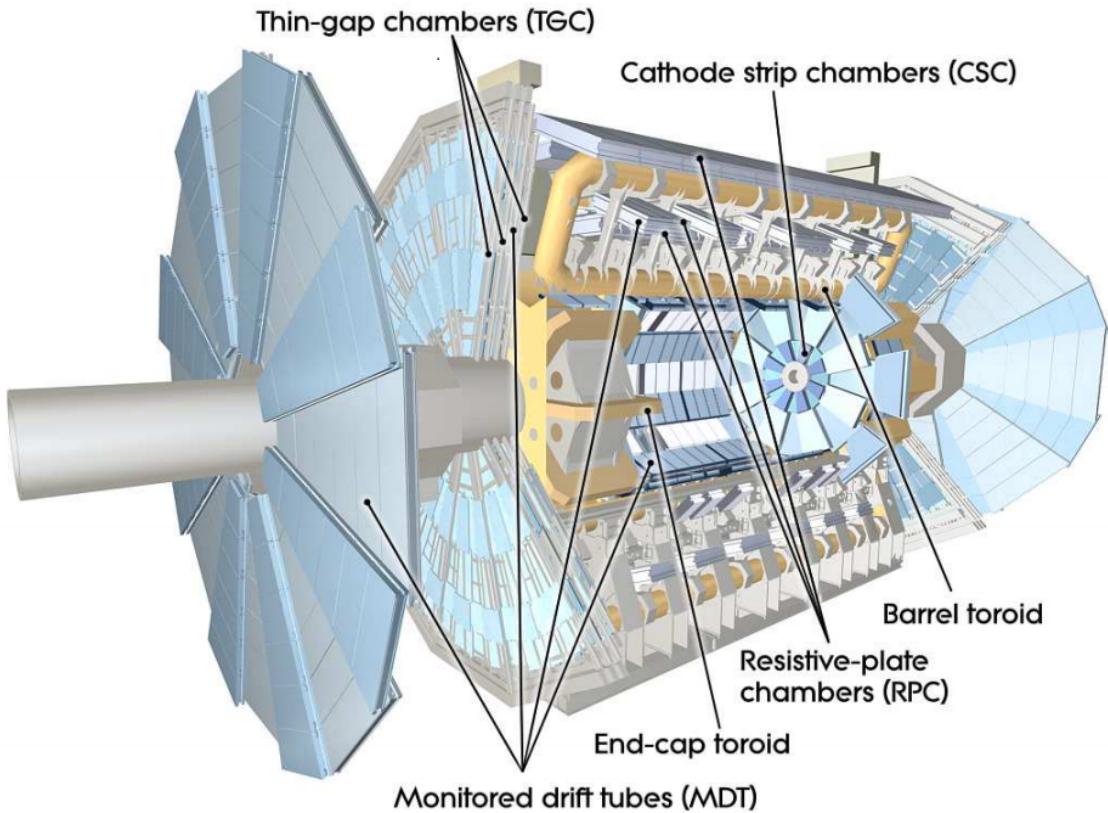


FIGURE 3.17. Schematic of the ATLAS muon detector.[37]

MDTs are arranged in chambers of up to 6 layers of aluminum tubes ranging in length from one to six meters. Each tube is 30 mm in diameter and contains a sense wire 50 μm in diameter. The chambers are arranged with a support spacer in between layers of MDTs that have a built-in optical sensor to monitor the drift tubes (hence the name) for deformations. This ensures that the precision of measurements does

not change over time. The MDTs are only used in the barrel and not in the forward region because they are inappropriate in areas with high rates, in this case a high flux of muons.

For the forward region CSCs are used. CSCs consist of arrays of positively charged wires crossed with negatively charged strips within a gas. As muons pass through they knock electrons from atoms in the gas which go toward the anode wires. Since the strips and wires are perpendicular to each other two position coordinates are read out. CSCs have the benefit of giving acceptable one and two-track resolution in a high flux environment.

The trigger system for muons in the barrel region uses RPCs which are parallel plates with opposite charges separated by a gaseous volume. A muon passing through an RPC knocks electrons from the gas which cause an avalanche of electrons that get picked up by the external metal strips rather than by the electrode. The pattern of metal strips that gets hit gives a quick measurement of the muon momentum which is used by the trigger to make the immediate decision about the event. The endcap muon trigger relies on TGCs. TGCs are anode wires with graphite cathodes in between thin layers of fiberglass laminate. Similarly to why CSCs are used over MDRs in the forward region TGCs have excellent timing resolution and can handle the high flux of muons in the forward regions.

3.2.6. Trigger and Data Acquisition

The amount of data the LHC is capable of producing is staggering, and the ATLAS trigger system is required to reduce the enormous amount of data produced to a reasonable amount while keeping the most interesting events. The LHC provides collisions at a rate of 40MHz. Every event saved to tape requires about 1.6MB of

space [45]. To keep all of the data produced 64TB/s would need to be saved or 230PB of data for a 12-hour run or 400EB of data per year (150 days of uptime).

In order to reduce this to a manageable level the ATLAS trigger system uses a two level trigger system. A hardware trigger, Level 1 or L1 trigger, is used to lower the rate from 40MHz to between 75 and 100kHz which is sent to the next level of the trigger system, the High Level Trigger (HLT). Another factor of 50 in rate reduction is achieved by this software based HLT to reduce the rate below 2kHz. A flowchart of the ATLAS trigger and data acquisition system is shown in Figure 3.18. When combined with partial event readouts and the <2kHz full event readout the total bandwidth requirement is around 3GB/s to be written. This means that only around 0.004% of data is stored.

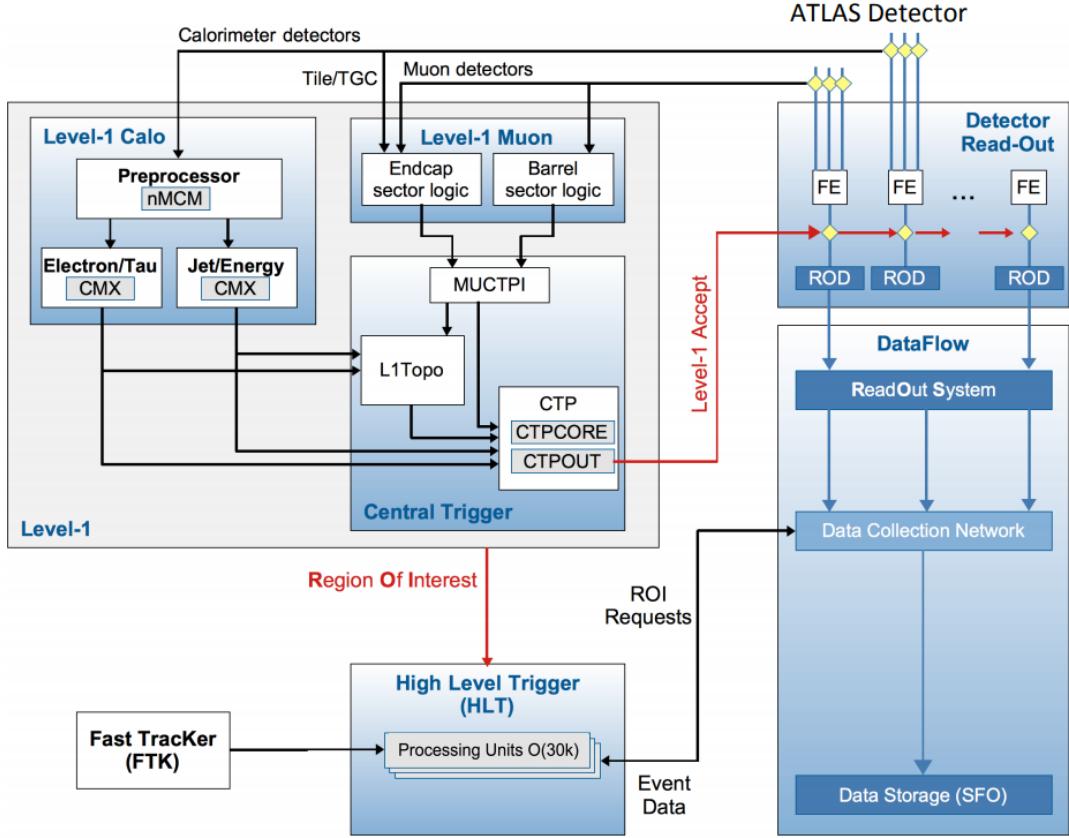


FIGURE 3.18. Flow diagram of the ATLAS trigger and data acquisition system used in Run 2.[46]

3.2.6.1. Level 1 Calorimeter

The Level 1 hardware trigger uses geometrically coarse information from some of the subdetectors. The data from the calorimeters is sent to the Level 1 Calorimeter (L1 Calo) system. L1 Calo uses low granularity information to identify Regions of Interest (RoIs) from objects that interact in the calorimeters (e.g., photons, electrons, jets, taus), events with high total energy, as well as events with an imbalance of energy coming from missing transverse energy. The information is fed into the L1 Calo system and through a preprocessor that allows L1 Calo to handle the effects within ATLAS from pileup events. Data from TileCal and the trigger portions of the muon systems

goes to the Level 1 Muon (L1 Muon) system which applies various logical processes to determine whether or not an event should be kept. Outputs from L1 Calo and L1 Muon are passed to the Central Trigger Processor (CTP) which provides a level 1 trigger accept and LHC timing information to the detector read out. At the same time the CTP gives RoIs to the HLT.

3.2.6.2. High Level Trigger

The HLT takes RoIs from the CTP as well as full detector granularity and makes a further decision whether that event should be saved. This is done using a computing farm of over 40,000 cores that run over 2,500 independent algorithms (trigger chains) on the RoIs. The HLT can provide partial and full event reconstruction depending on the event stream the event is decided to be within. The main event stream is the physics analysis stream which gets full event reconstruction, while the other streams typically only require partial event reconstruction. The other streams are used for a variety of things such as trigger level analysis, monitoring of the subsystems, and calibrating the detector.

CHAPTER IV

SIMULATION AND RECONSTRUCTION

This chapter presents details on the simulation of various physics processes and the reconstruction of physics objects for both simulated events and data events.

4.1. Simulation of pp Collisions

To draw conclusions from ATLAS experimental data it is necessary to make accurate theoretical predictions about the processes being searched for. Having accurate background models can help identify when a data signal is behaving in a way that might suggest new physics. Due to the stochastic nature of particle physics collisions and interactions, it is not practical to create exact predictions. Instead the ATLAS experiment uses Monte Carlo (MC) simulations to model physical behaviors. MC simulations are done by repeated random sampling of possible physical processes that can occur at any given time to a particle. The possibilities change based on factors such as particle energy and particle environment. A flow chart for the entire simulation chain is shown in Figure 4.1.

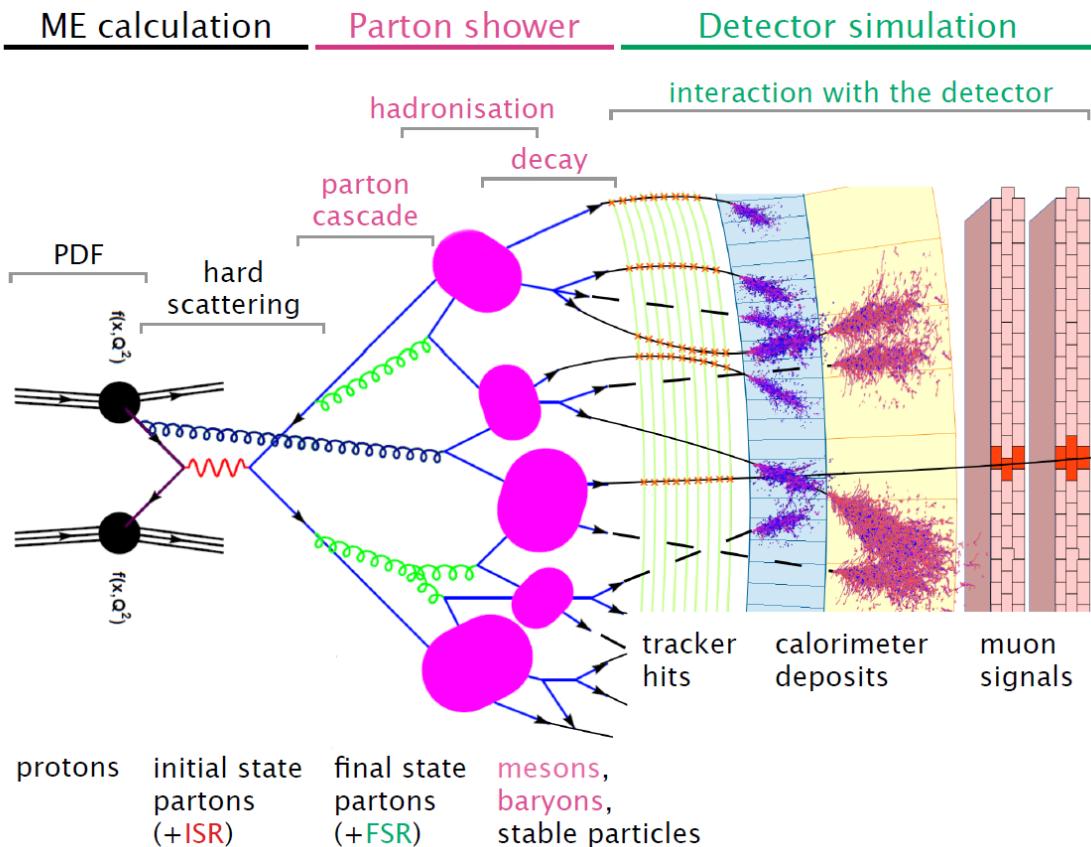


FIGURE 4.1. A pictoral view of the different steps for the creation of a MC event. [47]

4.1.1. Matrix Element Calculation and Parton Distribution Functions

Particle interactions at LHC energies do not involve the entire proton. The constituent partons that create the proton (the two up quarks, down quark, and the sea of gluons) are what interact in any given event. The gluons create many virtual quark-antiquark pairs which can interact as well. The valence quarks, the two up quarks and the down quark that make up the proton, are the major portion of interacting partons at low energies, mainly inelastic interactions. At LHC energies deep inelastic scattering is possible and the sea quarks play a more dominant role. Proton structure is described by a Parton Distribution Function (PDF) which gives

the probability of finding any parton with a particular momentum fraction, is shown in Figure 4.2.

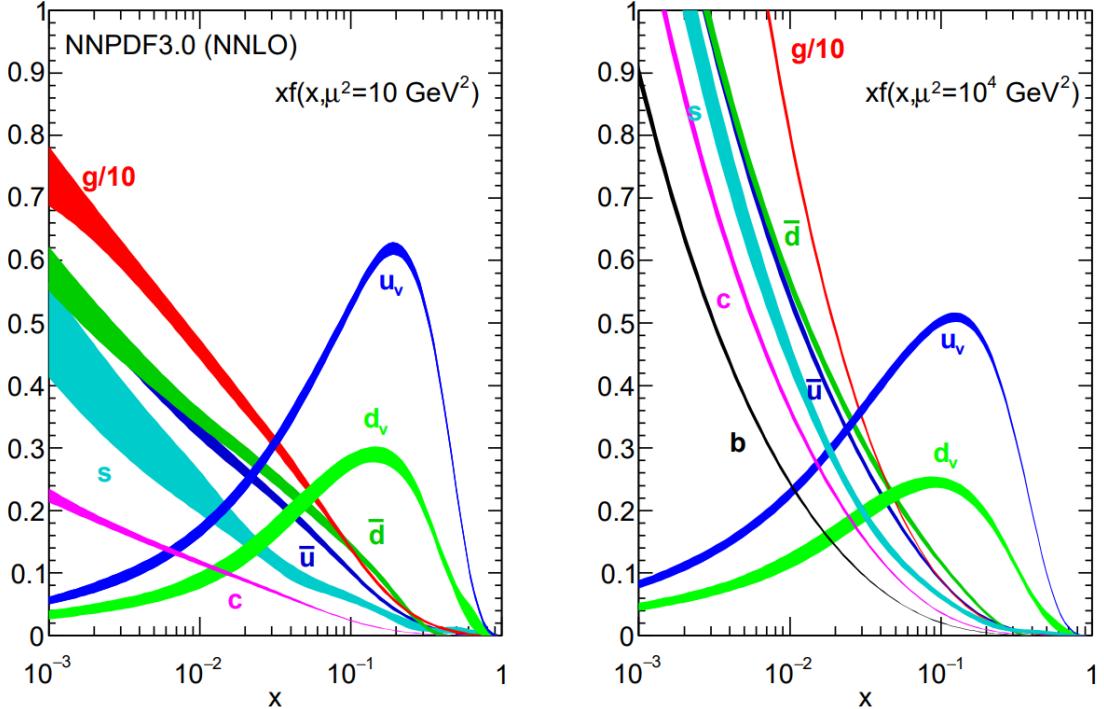


FIGURE 4.2. The bands are the momentum fraction, x , times the unpolarized parton distribution function obtained in NNLO NNPDF3.0 global analysis at scales $\mu^2 = 10 \text{ GeV}^2$ and $\mu^2 = 100 \text{ GeV}^2$ [10]

The PDFs and hard scattering processes are included in the calculation of the Matrix Elements (ME) of any interaction. Hard scattering processes can be described by Feynman diagrams, a representation of their amplitudes. Combining the PDFs and hard scattering amplitudes gives the probability of a particular interaction occurring. Calculation of the MEs is the first stage of simulation and is done to a specified order in perturbation theory: leading order (LO), next-to-leading order (NLO), etc. Higher order calculations lead to more accurate predictions but grow in complexity exponentially making them harder to calculate both theoretically and computationally, often restricting how accurate a process can be simulated.

4.1.2. Parton Shower Calculation

The next stage of simulating an event is the parton shower. These parton shower calculations deal with the quantum chromodynamic processes. In any interaction the particles that carry color can spontaneously emit gluons which can go on to create more gluons or quark-antiquark pairs. Depending on when this happens in the hard scattering process it is called initial state radiation (ISR) or final state radiation (FSR). The hard scattering partons as well as any additional radiated particles are used as inputs to parton shower calculations which determine how the quarks and gluons proceed through to the final state particles seen in the detector. This includes calculation of hadronization processes and further decay processes into the final state particles.

4.1.3. Detector Simulation

The final stage of creating an MC event is the detector simulation. The information from the event generators are processed using GEANT4 [48] and a detailed model of the ATLAS detector. GEANT4 simulates how various particles propagate through and interact with the material properties of the detector and where they leave energy which would then be measured by the ATLAS detector in an actual event. The result of this MC event construction flow is a collection of simulated data that is similar in structure to actual data collected using the ATLAS experiment. The energy deposits in both MC and real data are combined using the same software for physics objects are reconstruction. For MC events this allows for comparison between the physics object reconstruction and the truth record, or the types of particles fed into the detector simulation.

4.1.4. Monte Carlo Generators Used for LHC Physics

A variety of different MC generators are used in the creation of simulated events. Different generators specialize in simulating different physics processes to various levels of precision (eg., LO vs. NLO). The MC generators used in this search are summarized in this section.

MADGRAPH aMC@NLO [49]: An amplitude and event generator at LO and NLO for hard processes. Extendable to various models including effective field theory (EFT) models used in BSM searches. This generator is used to create the signal events searched for in this dissertation: discussed in Section 4.2.

POWHEG [50, 51]: **P**ositive **W**eight **H**ardest **E**mission **G**enerator is an NLO event generator that can be interfaced with other generators (i.e. PYTHIA) for showering.

PYTHIA [52]: A generator used most often for QCD final state hard processes and showering. It is commonly interfaced with other generators for showering within the ATLAS detector.

SHERPA [53, 54]: A multi-parton LO generator with an emphasis on merging ME and Parton Showering.

A common event file format developed at the Les Houches Accords [55] makes it possible for these generators to be interfaced in a straight forward way, typically with PYTHIA for showering. This allows a specialty generator to be created and used to generate hard processes and then simulate the rest of the event with common showering generators that might lack the ability to simulate the process in question.

4.2. Creation of Flavor Changing Neutral Current Signal Events

To create simulated signal events the typical Standard Model models must be extended to include higher order terms. A Universal FeynRules Output (UFO, [56])

model was created to include dimension 6 operators ([57, 58]). These individual operators are turned on for the specific final state being produced. The original operators can be reduced to a minimal set of coupling to anomalous final states (i.e. FCNC final states)[59], used for event production. This effective field theory method of signal production is beneficial as it allows for production of signal events that are not dependent on any particular BSM model. This method of including dimension-6 effective operators can be used to produce any of the top FCNC channels, for example in the $t \rightarrow qZ$ process [60] and $t \rightarrow qH$ [61].

4.2.1. FCNC Events Produced With MadGraph5 aMC@NLO

Signal events have been produced using MadGraph5 aMC@NLO following the work of Degrande et al.[62]. Before official ATLAS datasets can be produced and the entirety of the event reconstructed through the ATLAS detector, validation of the model must be performed. 10,000 events were produced locally at truth level for each decay channel to compare the kinematics of produced events in $t\bar{t} \rightarrow bWq\gamma$ to the kinematics of official production $t\bar{t}$ events.

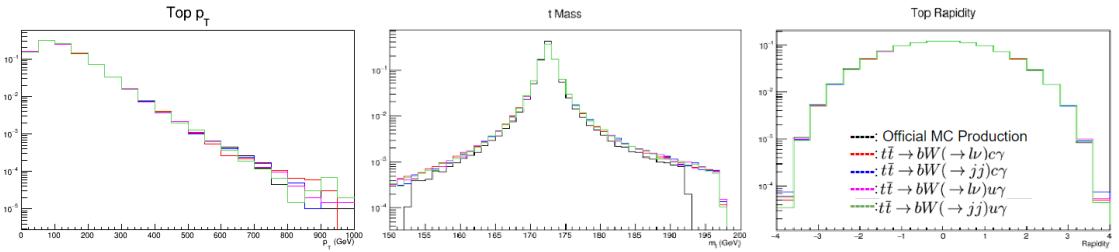


FIGURE 4.3. Normalized kinematics (p_T , m_t , and y_t) of individual top quarks produced by the model for each FCNC final state search and an official $t\bar{t}$ sample.

The minimal couplings mean there is one scalar coupling introduced for each decay mode, $t \rightarrow c\gamma$ and $t \rightarrow u\gamma$. All possible final states are shown in the figures

in this section: the leptonic and hadronic decays of the W boson from the top quark that decays through the typical standard model decay mode $t \rightarrow bW$.

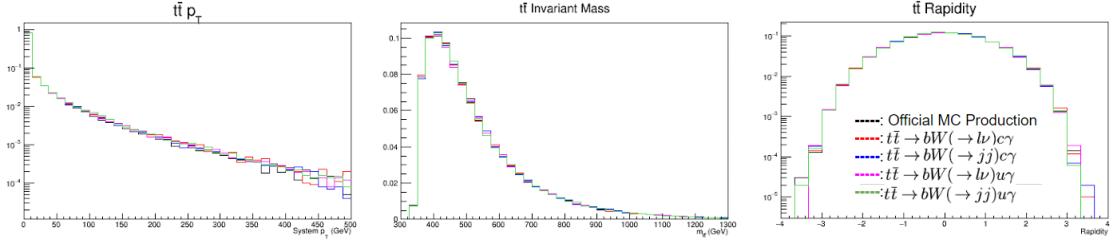


FIGURE 4.4. Normalized kinematics (p_T , $m_{t\bar{t}}$, and $y_{t\bar{t}}$) of the $t\bar{t}$ system produced by the model for each FCNC final state search and an official $t\bar{t}$ sample.

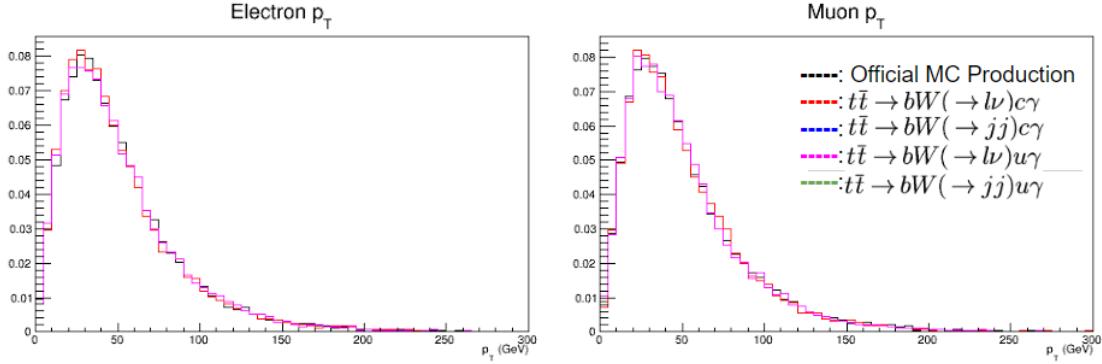


FIGURE 4.5. Normalized p_T of the electron and muons produced by the model for each FCNC final state search and an official $t\bar{t}$ sample.

The lepton validation plots in Figure 4.5 only show events where the W is forced to decay leptonically as well as the official sample which does not have a preference for the final state decay, i.e., the W bosons are allowed to decay leptonically or hadronically. No unexpected deviations from the Standard model produced $t\bar{t}$ samples are seen in any of the validation plots. The deviations seen in Figure 4.6 are misplaced quarks from NLO processes. The shifted mean values in the up and charm p_T spectrum are also expected. In those samples the up or charm quark is coming directly from the top quark as opposed to a W boson, which means it will have significantly boosted momentum.

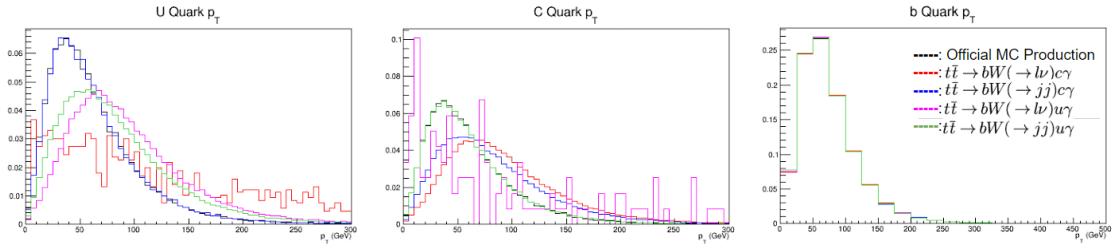


FIGURE 4.6. Normalized p_T of the up, charm, and bottom quarks produced by the model for each FCNC final state search and an official $t\bar{t}$ sample.

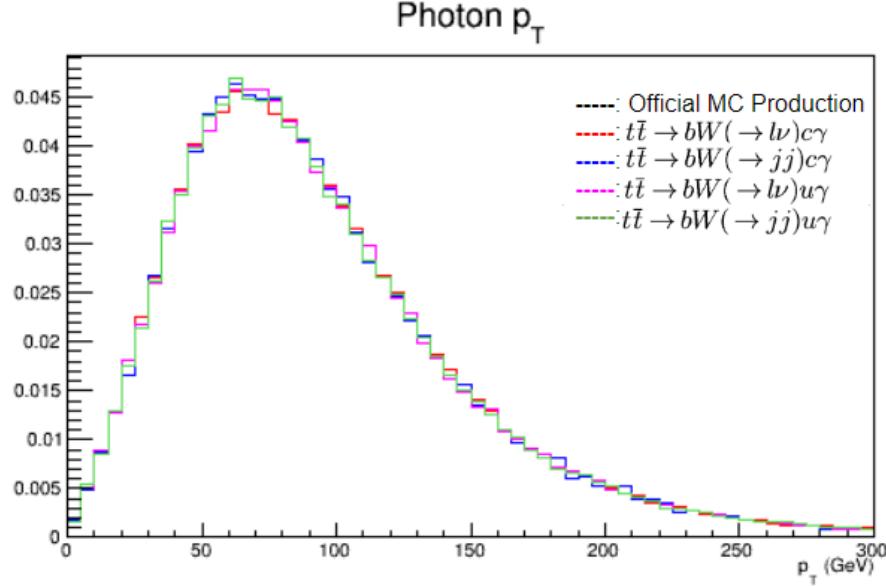


FIGURE 4.7. Normalized p_T of the photons produced by the model for each FCNC final state search and an official $t\bar{t}$ sample, there is 0 contribution from the official $t\bar{t}$ sample.

In the model there are left-and right-handed scalar couplings. Investigation into differences between the kinematics of the quarks produced using each of these couplings is shown in Figure 4.8.

$$\mathcal{L}_{\gamma tc} = -e\bar{c}\frac{i\sigma^{\mu\nu}q_\nu}{m_t}(\lambda_{ct}^L P_L + \lambda_{ct}^R P_R)tA_\mu + h.c.$$

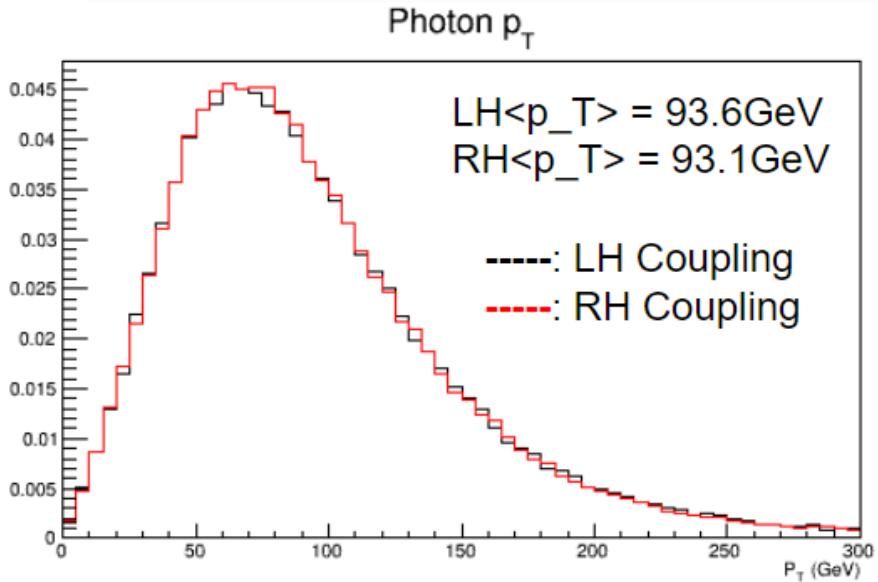


FIGURE 4.8. Normalized p_T of the photons produced by the model using the Left-handed (LH) and Right-handed (RH) couplings

No differences in final state kinematics were shown in the right-handed coupling compared to the left-handed coupling. Due to this, only one coupling value was used in the production. In the end only leptonic decays of the W were produced officially. The leptonic state is simpler to search for than a final state not involving leptons because of the much larger backgrounds from QCD processes. The lepton offers many handles for searching for these rare FCNC decays. Using the leptonic final state it is not necessary to use combinatorics for event reconstruction as each object is unique and comes from one particular object. In this analysis the final state involves a light jet and a photon (from the FCNC decay), and a lepton, photon, and b-jet (from the Standard Model top decay).

4.3. Object Reconstruction

After the events are simulated, or collected in case of real data, the collections of energy deposits within the detector systems must be transformed into meaningful physics objects through reconstruction. Reconstruction is typically done in two major parts using the specialized detectors covered in Chapter III. The Inner Detector and Muon System turn patterns of hits within the tracking detectors into tracks that have direction and momentum information. The calorimeter system transforms the energy deposits within the calorimeters into calibrated energy deposits with a particular position. These tracks and calorimeter deposits are used to create physics objects (electrons, muons, etc.) by using particle identification techniques to reconstruct the underlying physics event. For the analysis presented in this dissertation, the final state signal particles that need to be reconstructed are one lepton (an electron or a muon), one photon, two quarks (one light flavor and one b quark), and one neutrino (missing transverse energy as it is the only particle that does not interact with the detector). Each of these particles has a particular signature in the subdetectors of the ATLAS detector, shown in Figure 4.9.

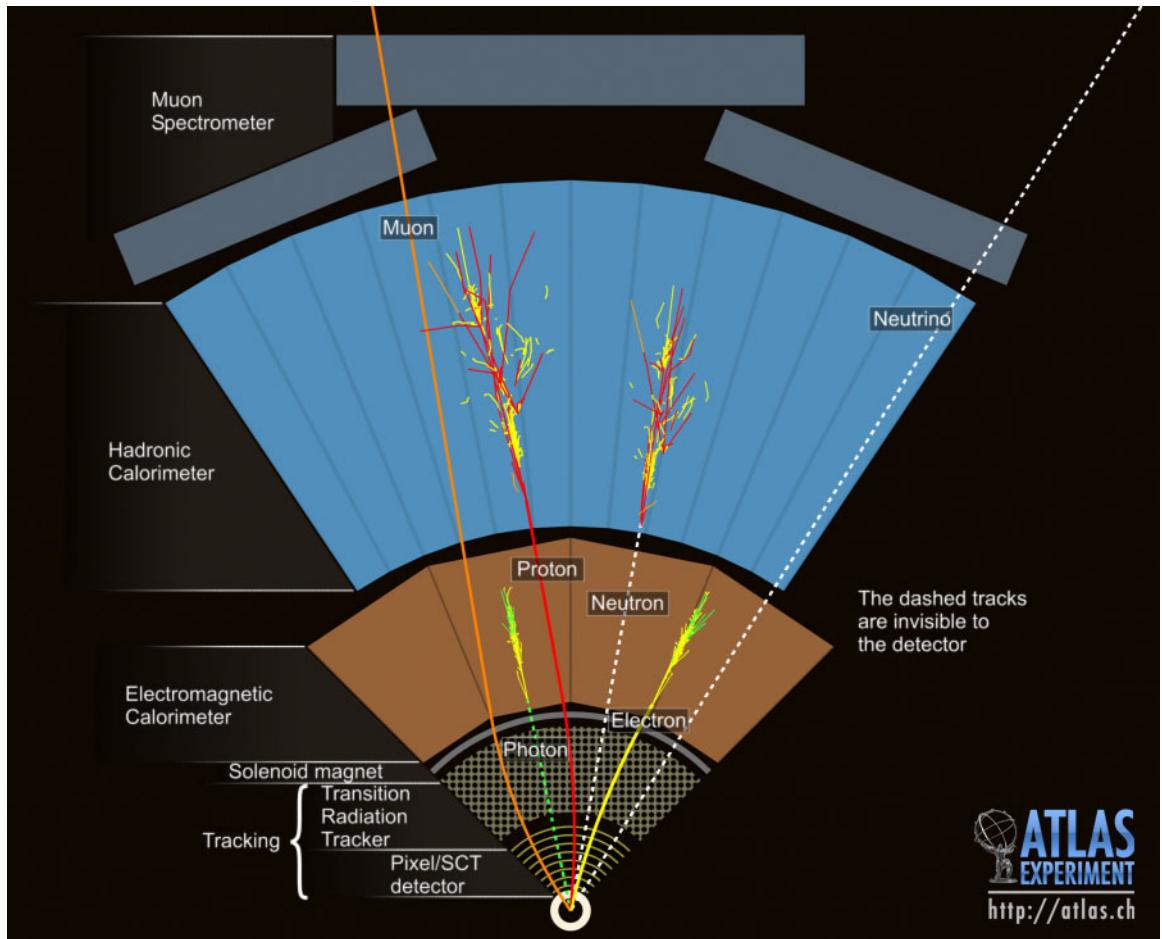


FIGURE 4.9. Cross section of a simulated ATLAS detector showing how various particles interact with ATLAS subsystems. Solid lines indicate interactions while dashed lines indicate that no interactions typically occur in that section of the detector. [63]

4.3.1. Electrons

Electrons interacting with within the ATLAS detector leave a track in the Inner Detector as well as a cluster of energy in the electromagnetic calorimeter. The track and cluster are required to be matched together to be identified as an electron candidate[64]. As electrons move through the detector they create electromagnetic showers through bremsstrahlung which can produce electron-positron pairs. The process continues as the particles continue to give energy to the detector. This

collection of electrons, positrons, and photons creates a signature energy cluster in the calorimeter.

Electron identification algorithms are applied to the electron candidates to separate prompt and isolated electron candidates from electrons that come from backgrounds such as converted photons and misidentified jets. The electron identification algorithms use a sliding window (3×5 in $\eta \times \phi$) in the high granularity section of the LAr electromagnetic calorimeter to search for electron cluster “seeds” greater than 2.5 GeV. Clusters are created around these seeds to form the electromagnetic shower and remove possible duplicate electron signals by containing them within the cluster. Further pattern recognition for the track fitting allows even larger amounts of energy into the shower to account for bremsstrahlung in the shower shape. Tracks and clusters are then matched to give electron candidates.

Electrons coming from background jets or photon conversion are called non-prompt as they do not originate from a signal object/the primary vertex. In order to reject these electrons, other discriminating variables are used in addition to the track-cluster matching. These variables include the amount of energy leakage into the hadronic calorimeter, the shower development throughout the electromagnetic calorimeter, and the amount of radiation measured in the TRT. Three electron identification working points are used: Loose, Medium, and Tight. Each of these operating points have their own level of background rejection and signal efficiency. Working points with higher background rejection are a subset of those with lower background rejection.

Isolation variables are another useful tool in the identification of signal electrons from converted photons produced in hadron decays and light hadron misidentification. These variables are defined by a cone size around the electron candidate and are the

sum of the transverse variable (momentum or energy) of all of the tracks within the cone, $p_T^{\text{cone}0.2}$ with a cone of $\Delta R = 0.2$ (or $10 \text{ GeV}/E_T$, for high energy electrons) and $E_{T,\text{Topo}}^{\text{varcone}0.4}$ with a cone defined in a similar manner.

Because the LAr calorimeter is a sampling calorimeter, the energy deposits must be calibrated and scaled such that the true electron energy is read out and not just the small amount of energy deposited into the active layers as discussed in Section 3.2.4. The energy scale is calibrated to be uniform throughout the detector. Any residual differences between data and simulation are corrected. The calibration strategy was developed for optimal performance in LHC Run 1[65] and updated for the conditions of LHC Run 2[66].

4.3.2. Muons

Muons behave differently from other particles as they traverse the detector. They act as minimum-ionizing-particles (MIPs) throughout the calorimeter. The Muon Spectrometer (MS), Section 3.2.5, specializes in precision measurements of muons. The Inner Detector (ID) plays a pivotal role in the identification of muons as it offers an independent measure of the muon characteristics. The muon reconstruction process uses a specific set of variables as well[67]. These variables include: q/p significance: the difference in the ratio of track charge and momentum measured with the ID and MS, ρ' : the difference between the transverse momenta measured with the ID and MS, and χ^2 of the combined track fit using tracks from both the ID and MS.

Muons are separated out into four separate types depending on their interactions with the various subdetectors. The best muon candidates are combined muons that use hits in the MS to trace back to a track in the ID in order to reconstruct the entire

muon track. Segment-tagged muons are muon candidates that leave a track in the ID but only a segment in the MS instead of a full track. Segment-tagged muons can occur because of the muon having low p_T or crossing through a region of the MS with reduced acceptance. Extrapolated muons require only tracks in the MS and are used in regions of η, ϕ phase space that the ID does not cover. Calorimeter-tagged muons are muons identified by MIPs in the calorimeters and are used to find muons that cross the ID and MS in regions where cabling might prevent particle detection.

Muons also have their own set of isolation criteria which is track-based $p_T^{\text{varcone}0.3}$, with a cone of $\Delta R = \min(0.3, 10 \text{ GeV}/p_T)$. Similar to electrons various working points are available at the analysis level for muons. These working points are named similarly: Loose, Medium, Tight, and High- p_T in order of background rejection.

High p_T jets that punch through the hadronic calorimeter can leave tracks in the MS which could be identified as muons. These would be identified as a bad or a fake muon because of the high-hit multiplicities they leave in the MS as opposed to a single track left by a muon as it is a MIP. Another source of bad muons is a mismeasured ID track that gets incorrectly matched to segments in the MS. Fake muons are a source of fake missing transverse energy, \cancel{E}_T

4.3.3. Photons

Photons behave very similarly to electrons in the calorimeter in that they also produce an electromagnetic shower in the calorimeter. However, they are neutrally charged particles meaning that they should not leave a track in the ID as they do not bend and produce bremsstrahlung photons traveling through the magnetic field. Prompt photons pair-produce electrons in the tracker, but this process can be identified as the associated cluster in the electromagnetic calorimeter

is matched to two tracks with opposite charge. This process produces what is called a converted photon. Unconverted photons have no matching tracks associated with an electromagnetic cluster.

Prompt photons produce narrower energy deposits in the electromagnetic calorimeter and have smaller leakage into the hadronic calorimeter compared to background photons. The energy contained within narrow structure in $\eta \times \phi$ strips compared to the energy contained in a larger section can help identify prompt from non-prompt photons [68]. Cuts on this and the other variables listed in Table 4.1 are tuned to reduce dependency of identification efficiency on the pileup conditions of Run 2.

Category	Description	Name	<i>loose</i>	<i>tight</i>
Acceptance	$ \eta < 2.37$, with $1.37 \leq \eta < 1.52$ excluded	-	✓	✓
Hadronic Leakage	Ratio of E_T in the first sampling layer of the hadronic calorimeter to E_T of the EM cluster (used over the range $0.8 < \eta $ or $ \eta > 1.52$)	R_{had_1}	✓	✓
	Ratio of E_T in the hadronic calorimeter to E_T of the EM cluster (used over the range $0.8 < \eta < 1.37$)	R_{had}	✓	✓
EM Middle Layer	Ratio of the energy in $3 \times 7 \eta \times \phi$ cells over the energy in 7×7 cells centered around the photon cluster position	R_η	✓	✓
	Lateral shower width, $\sqrt{(\sum E_i \eta_i^2)/(\sum E_i) - ((\sum E_i \eta_i)/(\sum E_i))^2}$, where E_i is the energy and η_i is the pseudorapidity of cell i and the sum is calculated within a window of 3×5 cells	ω_{η_2}	✓	✓
	Ratio of the energy in $3 \times 2 \eta \times \phi$ strips, over the energy of 3×6 cells centered around the photon cluster position	R_ϕ		✓
EM Strip Layer	Lateral shower width, $\sqrt{(\sum E_i(i - i_{\max})^2)/(\sum E_i)}$, where i runs over all strips in a window of $3 \times 2 \eta \times \phi$ strips, and i_{\max} is the index of the highest-energy strip calculated from three strips around the strip with maximum energy deposit	$\omega_s 3$		✓
	Total lateral shower width $\sqrt{(\sum E_i(i - i_{\max})^2)/(\sum E_i)}$, where i runs over all strips in a window of $20 \times 2 \eta \times \phi$ strips, and i_{\max} is the index of the highest-energy strip measured in the strip layer	$\omega_s \text{ tot}$		✓
	Energy outside the core of the three central strips but within seven strips divided by energy within the three central strips	f_{side}		✓
	Difference between the energy associated with the second maximum in the strip layer and the energy reconstructed in the strip with the minimum value found between the first and second maxima	ΔE_s		✓
	Ratio of the energy difference between the maximum energy deposit and the energy deposit in the secondary maximum in the cluster to the sum of these energies	E_{ratio}		✓
	Ratio of the energy in the first layer to the total energy of the EM cluster	f_1		✓

TABLE 4.1. Photon identification variables used for *loose* and *tight* photon identification, taken from [68]

4.3.4. Jets

Contrasting with electromagnetic showers produced by electrons and photons, hadronic showers form through QCD processs. Quarks very quickly undergo showering by emitting gluons which further produce quark-antiquark pairs, analogous to the photons and pair-produced electron-positron pairs of electromagnetic showers. When quarks have enough energy they hadronize by producing bound states of particles. These particles are typically pions or mesons that are measured by the ATLAS detector. The top quark is the only quark that decays before hadronization because it decays so fast (5×10^{-25} s). The spray of hadrons coming from a quark from the initial interaction is called a jet and is a collection of detector objects that are traced back and assigned to the quark(s) in the final state of the interaction. These algorithms are called jet-finding algorithms. Pictoral representations of the same event reconstructed with four various algorithms is shown in Figure 4.10.

The jets in this analysis use the anti- k_T algorithm [69] with a radius parameter $R = 0.4$. Jets are not physical objects but collections of clustered particles so how they are defined can change the physics objects that are eventually analyzed. The anti- k_T algorithm is preferred because it is infrared and collinear safe. Infrared safe jet algorithms do not merge two jets with a soft emission between them. Adding or removing a soft term between two jets should not change which objects are called jets. Collinear safe jet algorithms do not change the jet collection if the high transverse momentum particles is split or merged . Another added benefit of the anti- k_T jet finding algorithm is that it produces roughly circular jet objects, thereby simplifying the calculation of the energy density and simplifying the calibration of the jet object.

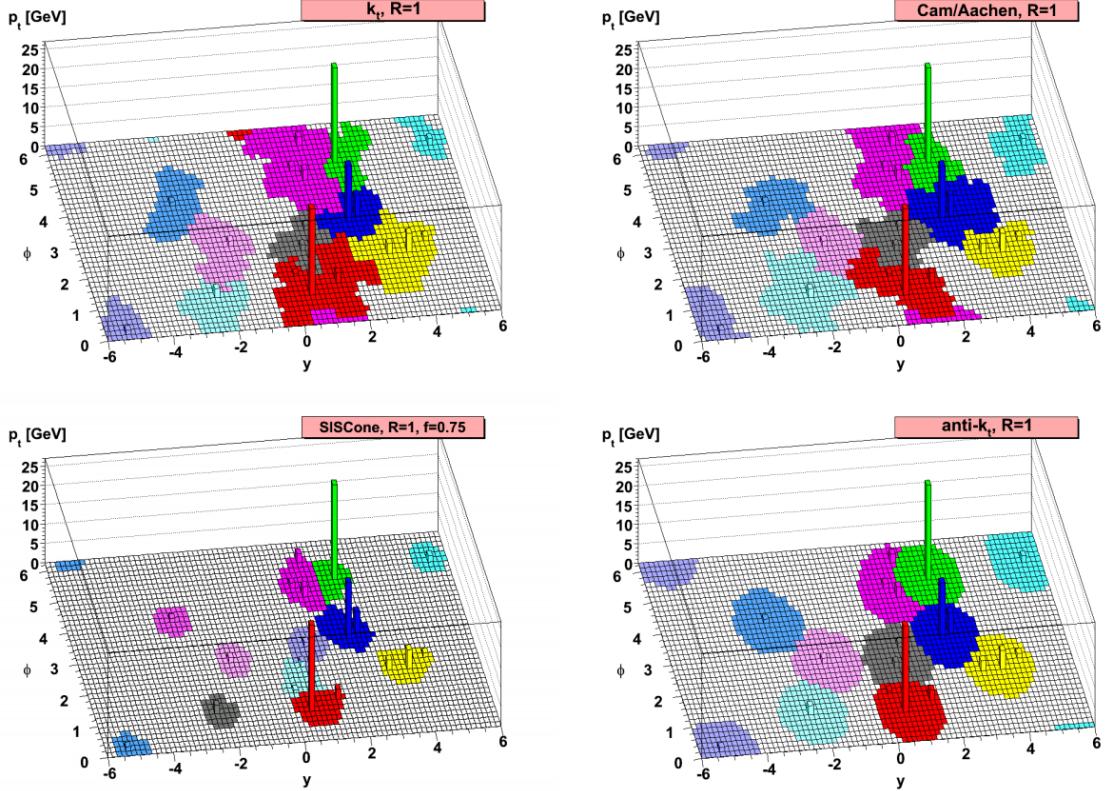


FIGURE 4.10. A sample parton-level event with many random soft jet objects, clustered with four different jets algorithms, illustrating the areas of the resulting hard jets. For k_T and Cambridge/Aachen the detailed shapes are in part determined by the specific set of ghosts used, and change when the ghosts are modified [69]

The $\text{anti-}k_T$ algorithm calculates the distance between an object i and all possible jet objects j (d_{ij}) and the beam (d_{iB})

$$d_{ij} = \min(k_{ti}^{2p}, k_{tj}^{2p}) \frac{\Delta_{ij}^2}{R}, \quad d_{iB} = k_{ti}^{2p}$$

where k_{ti} is the transverse momentum, Δ is the distance between the objects, and $p = -1$. This is a general form for the type of algorithm where the inclusive k_T algorithm has a p value of 1 and the inclusive Cambridge/Aachen algorithm has a p value of 0 [70]. The algorithm then follows that if d_{ij} is smaller than d_{iB} then objects

i and j are merged, otherwise i is labeled as a jet and removed from the list of entries of possible jet objects. This is repeated for all entries in the list of possible jet objects.

Jet cleaning is also applied to remove events with jets built from known noisy parts of the calorimeter due to particular calorimeter cells or non-collision background in those areas [71]. To reduce selecting jets that originate from pileup interactions, another requirement on the jet object is made on the jet vertex tagger [72, 73] as follows:

1. For jets with $20\text{GeV} < p_T < 60\text{GeV}$ and $|\eta| < 2.4$: if any jet is bad AND that jet is not marked as pileup by JVT, then reject the event
2. For jets with $20\text{GeV} < p_T < 60\text{GeV}$ and $|\eta| \geq 2.4$: if any jet is bad, then reject the event
3. For jets with $p_T \geq 60\text{GeV}$: if any jet is bad, then reject the event

4.3.4.1. B-Jets

While jets originate from any quark, jets coming from b quarks can be identified due to their decay products. B quarks hadronize into b-hadrons which have a relatively long lifetime compared to many other hadrons produced from light quarks. The longer lifetime and the relativistic speeds at which the hadrons travel mean the particle travels a measureable distance before it decays ($400 - 500\mu\text{m}$) [74]. Thus, the vertex reconstructed from the energy coming from a b hadron decay can be traced back to a point that does not correspond to the primary vertex of the event. A pictoral representation of a b quark decay is shown in Figure 4.11. The b-jet vertex is called the secondary vertex.

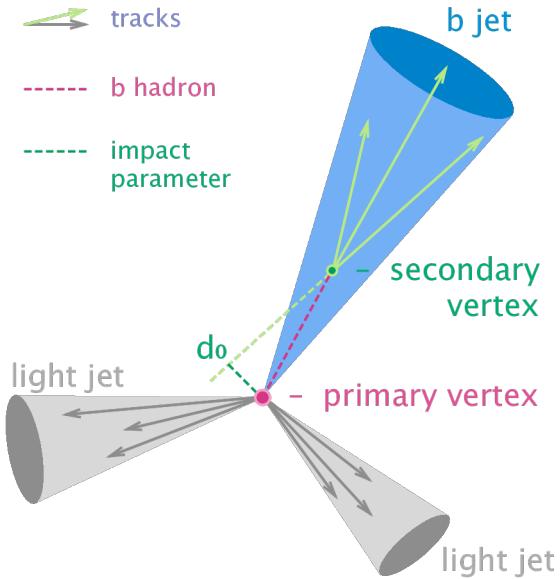


FIGURE 4.11. Pictoral representation of an event with a b-jet showing the secondary vertex and impact parameter [75]

In addition to the secondary vertex, other variables are helpful in identifying jets coming from b quarks. By back tracing the tracks within the displaced vertex the minimum distance between the track and the interaction point can be measured, known as the impact parameter. Reconstructing the decay chain of the jet is also used in determining the providence of the jet. This information is used in a multivariate analysis (MVA) to identify jets coming from b quarks and reject jets coming from light quarks.

The MVA used in this analysis is the MV2c10, the discriminant used for b-jet identification [76]. The output distributions for various flavors of jets as well as background rejection and signal efficiency plots are shown in Figure 4.12. The c10 in the algorithm name refers to the background training sample of the MVA consisting of a 10% fraction of c-jets. The 77% efficiency fixed-cut working point for b-jet identification was chosen for this analysis, discussed in Section 5.3. Differences in efficiency of b-tagging between data and simulation is taken into account with working

point specific scale factors provided by the ATLAS Flavour Tagging Combined Performance group.

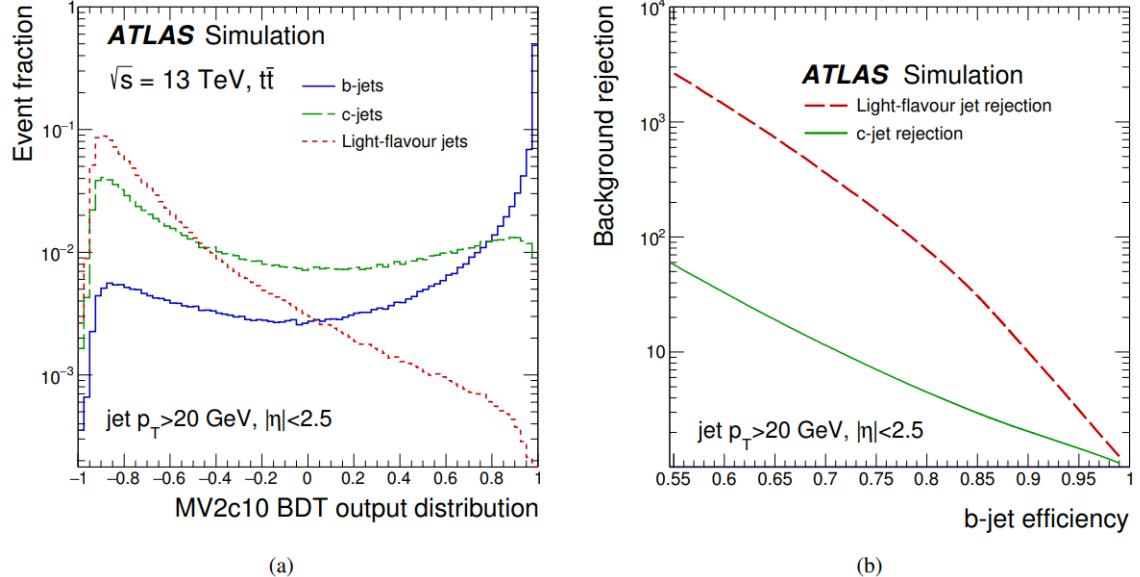


FIGURE 4.12. The MV2c10 output for b, c, and light flavored jets in simulated $t\bar{t}$ and the background rejection as a function of the b-jet efficiency [77]

4.3.5. Missing Transverse Energy

The remaining signal object that has yet to be discussed is the neutrino coming from the W boson decay. Neutrinos do not interact with the detectors as they pass through the ATLAS detector. The only way to measure any properties of the neutrino in ATLAS events is to use conservation of momentum. As previously mentioned the collision energy is unknown as partons do not carry a consistent fraction of the beam proton energy. However, in the transverse plane to the beamline the total momentum is known to be very small. Before the collision there is on the order of 1 GeV of momentum in the transverse plane. Therefore, the total transverse momentum of the collision products should be approximately zero.

Any imbalance in the momentum is referred to as Missing Transverse Momentum (\cancel{E}_T). The negative vector sum of all reconstructed objects plus an additional soft term are used to calculate the missing energy in the x-plane and the y-plane[78]. A magnitude and an azimuthal angle are calculated to give the \cancel{E}_T vector in the transverse plane but this does not directly correspond to a neutrino which also has a momentum in the z direction.

4.3.5.1. Neutrino Reconstruction

In this analysis the signal contains only one source of missing energy, therefore all of the missing energy can be used to reconstruct a neutrino object. There is an ambiguity in the choice of the neutrino z-momentum. To find the z-momentum a χ^2 minimization is done:

$$\begin{aligned}\chi^2 &= \chi_{\text{SMTop}}^2 + \chi_W^2 \\ \chi^2 &= \frac{(m_{\text{bjet},l,\nu} - m_t)^2}{\sigma_{\text{SMTop}}^2} + \frac{(m_{l,\nu} - m_W)^2}{\sigma_W^2}\end{aligned}$$

The widths σ_{SMTop} and σ_W^2 are determined from signal Monte Carlo. The event objects are combined to calculate the invariant mass of the top quark (the combination of the b-jet, lepton, and neutrino) and the W boson (combination of the lepton and neutrino). The χ^2 minimization is done while varying the z-momentum of the neutrino. The neutrino momentum that corresponds to the smallest χ^2 value is given assigned to the neutrino object for further use in the analysis. The χ^2 values are also used as a discriminating variable and fed into a neural network (Section 5.3).

CHAPTER V

SEARCH STRATEGY

This chapter will describe the major backgrounds and outline a search strategy used in the analysis. The kinematic regions for the signal will be defined as well as the introduction of neural networks to assist with the separation of signal and background like events.

5.1. Major Backgrounds

There are a large number of Standard Model processes that can end up in the signal region and share a similar final state topology as the studied signal process, $t\bar{t} \rightarrow bl\nu q\gamma$. All of these processes are modeled with Monte Carlo (MC) simulation, with the exception of particles that fake other particles (leptons and photons) which are done using data-driven techniques because they are poorly modeled with MC. A full list of the Monte Carlo samples used for each background can be found in Appendix B.

The dominant backgrounds in this search are Standard Model $t\bar{t}$, Standard Model $t\bar{t} + \gamma$, as well as W+jets and W+jets with an associated photon (W+jets+ γ). These along with minor backgrounds (Standard Model single top events, Z+jets (Z+jets+ γ), $t\bar{t}$ +Vector Boson, diboson) are all modeled with MC and some data-driven estimates and corrections are applied. The various backgrounds modeled in this search are summarized here:

SM Processes: SM $t\bar{t}$, W+jets, Z+jets, single top, diboson, $t\bar{t}+V$ are modeled with MC simulations. Control and validation regions are designed to test performance of the largest of these background processes, $t\bar{t}$ and W+jets. A

discussion of the modeling of the major SM processes without photons and a derivation of additional scale factors for these processes is shown in Section 5.6.1.

SM Processes with an associated photon: SM + associated photon processes: $t\bar{t} + \gamma$, W+jets+ γ , Z+jets+ γ are modeled with MC simulation and overlap removal is applied to the SM processes to remove events with similar phase spaces since these processes are a subset of the SM process. MC simulations are created with additional statistics for these processes. Further details are presented in Section 5.6.2.

Fake Leptons: Non-prompt or fake electrons and muons can arise from semi-leptonic decay of b and c quarks. For electrons additional contributions from photon conversions and jets in the electromagnetic calorimeter can lead to lepton fakes. Muons can be faked via energetic showers in the hadronic calorimeter or from hadrons that punch through the hadronic calorimeter. The matrix method is used to estimate the number of events with fake leptons as described in Section 5.7.2.

Fake Photons: The number of events with fake photons is estimated using a $Z \rightarrow e^+e^-$ tag-and-probe method discussed in Section 5.7.1. Further, the ABCD method is used in Section 5.7.3 to estimate the number of fake photon events that arise from a jet faking a photon.

5.1.1. Overlap Removal

Tracks and energy deposits within the detector can, in some cases, be used to reconstruct multiple objects. To prevent using these tracks and deposits multiple

times a standard overlap removal procedure is applied to objects. First, electrons that share tracks with any other electrons are removed. Any electron sharing a track with a muon is then also removed. Any jet that is found within $\Delta R < 0.2$ of an electron is removed. Then any jet with less than 3 tracks associated with it within $\Delta R < 0.2$ of a muon object is removed. After that any muon found within $\Delta R < 0.4$ of a jet is removed and any photon within $\Delta R < 0.4$ of an electron or muon object is removed.

5.1.2. Duplicate Event Removal

As specialized higher statistic samples are used for processes with prompt photons, a double counting of events could occur with the nominal MC samples. For example, in addition to the $t\bar{t}$ sample a sample of $t\bar{t} + \gamma$ events are used. This is true for the W+jets/Z+jets and special samples of W+jets+ γ and Z+jets+ γ . Therefore a truth based matching scheme is used to remove events in the nominal samples that match with the photon types produced in the specialized + γ samples i.e., they contain a truth photon that does not originate from a hadron or lepton.

5.2. Event Selection

This analysis is searching for $t\bar{t}$ candidate events where one of the top quarks decays through the most common decay path (a W boson that decays leptonically to an electron or muon and a bottom quark) and the other through the FCNC diagram to an up type light quark and a photon. The selected decay path is then $t\bar{t} \rightarrow Wbq\gamma \rightarrow l\nu b q\gamma$ such that the final state will contain at least two jets exactly one of which is b-tagged using the MV2c10 algorithm, exactly one massive lepton (an electron or muon), exactly one highly energetic photon, missing transverse energy from the

neutrino, and a large transverse W mass. The transverse W mass requirement selects events that have a lepton and \cancel{E}_T which are consistent with a leptonically decaying W boson, defined as:

$$m_T^W = \sqrt{2 \times p_{T_l} \times \cancel{E}_T \times (1 - \cos(\phi_l - \phi_{\text{MET}}))}$$

which is largest when the lepton and missing transverse energy are close in ϕ as would be expected from the decay products of a boosted W boson. Events are selected loosely to include all possible kinematic regions of interest and then skimmed down to the individual regions for various studies. This includes events that will not enter the signal region such as events with 0 photons or events with 0 b jets that are used to derive and test scale factors on the largest background samples and account for mismodeling of MC simulations.

For the events that have a chance of entering the final Signal Region i.e., events with exactly one b-tagged jet and exactly one photon, a neural network analysis is performed to help separate the signal from the background using a variety of high dimensional cuts. The neural network training and testing are described in Section 5.3. The neural network is then applied to all events with 1 b-tagged jet and 1 photon and greatly increases signal purity while separating out the most dominant backgrounds, in particular $t\bar{t}$ and $t\bar{t} + \gamma$.

5.3. Event Classification: Neural Network Optimization

To help distinguish signal events from the majority of background events, neural networks were employed for event classification. Neural networks are multivariate methods that take a variety of inputs and output a number between 0 and 1. The

output value is a discriminating variable that will be used to classify events and determine which events make it into the final Signal Region selection. Signal-like events accumulate towards 1 while background-like events cluster around 0. Two neural networks are trained, one for the electron+jets final state and one for the muon+jets final state. This section will discuss the neural network studies completed and their uses in the search for FCNC events.

5.3.1. Input Variables

A wide variety of input variables to the neural network were studied in detail. Studies were done using only low level variables such as the kinematic variables (p_T , η , ϕ , E) of the physics objects in the signal region. While a complex enough neural network should be able to figure out useful high level/event level variables (i.e., invariant masses, geometric separations), in practice a combination of some of these low level variables and high level variables used as inputs to the neural network proved to give the best separation and projected limits. Using physical intuition to guide the neural network proved to be a valuable tool.

Combinations of 29 input variables were tested at the outset. However variables such as η and ϕ tend to not have significant weights in the neural network and are left out in favor of the high level variables that include them (e.g., ΔR values). A measure of how different the variables are between signal and background is the separation. Table 5.1 shows the separation values for the variables that are inputs to the final neural network. Comparisons between the shapes of the input variables for the μ +jets channel are shown in Figures 5.1, 5.2, and 5.3

$$\text{Separation} = \sum_i^{bins} \frac{n_{si} - n_{bi}}{n_{si} + n_{bi}}$$

Variable	Separation e+jets	Separation μ +jets
$p_T(\gamma)$	22.97	24.01
$m_{q\gamma}$	22.65	28.31
γ_{iso}	18.62	41.32
m_{bW}	11.10	11.70
$m_{l\gamma}$	9.00	7.51
$\Delta R_{j\gamma}$	4.59	5.66
ΔR_{bl}	4.99	4.47
m_T^W	3.16	3.37
S_T	3.78	3.32
n_{jets}	1.70	2.03
χ_W^2	1.37	1.91
$p_T(q)$	2.46	2.82
$\Delta R_{l\gamma}$	1.40	1.19
E (lepton)	0.86	0.89
\cancel{E}_T	0.47	0.70
$p_T(b)$	0.51	0.53

TABLE 5.1. Separation of normalized variables between signal and background in the e+jets and μ +jets channels for the variables used as input to the final neural network.

Typically the kinematic variables with photon information have the biggest separation values. This is expected because the signal photon comes directly from the decay of a top quark and is much more energetic than background photons. Shape comparison plots for the e +jets channel and additional plots for other investigated variables are shown in Appendix C.2. The largest difference in separation between the e +jets and μ +jets channels is the photon isolation value. This is due to the fact that all backgrounds are included and fake photon contamination from a large Z+jets background is expected. Both networks perform similarly in their separation of signal and background events. The network is able to learn and compensate for this behavior with the help of other variables that include the lepton and photon: $\Delta R_{l\gamma}$ and $m_{l\gamma}$.

The neural networks are trained on MC events that have a chance of being in the signal region after basic event level cuts and are optimized for signal significance. Only events with 1 photon (> 15 GeV) and 1 bjet (MV2c10 77% working point) are classified by the neural network. The 77% working point was chosen by training the neural network on events with only 1 bjet at each working point: 70%, 77%, and 85%, and picking the network and working point with the best estimated significance. The b-tagging neural network study is shown in Section 5.3.4.

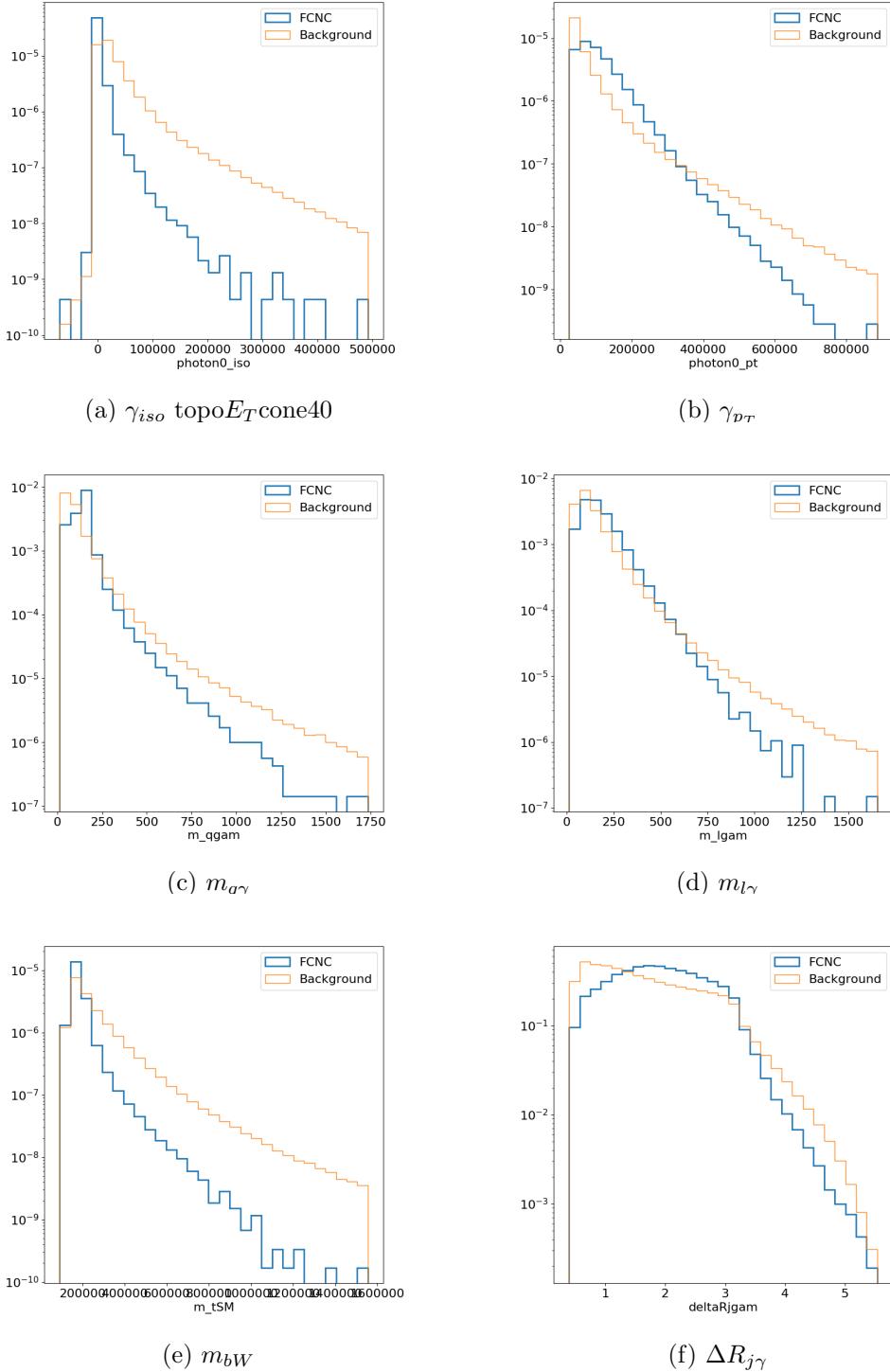


FIGURE 5.1. Normalized variables showing the shapes of neural network input variables for the $\mu+jets$ channel: γ_{iso} topo E_T cone40, γ_{p_T} , $m_{q\gamma}$, $m_{l\gamma}$, m_{bW} , and $\Delta R_{j\gamma}$

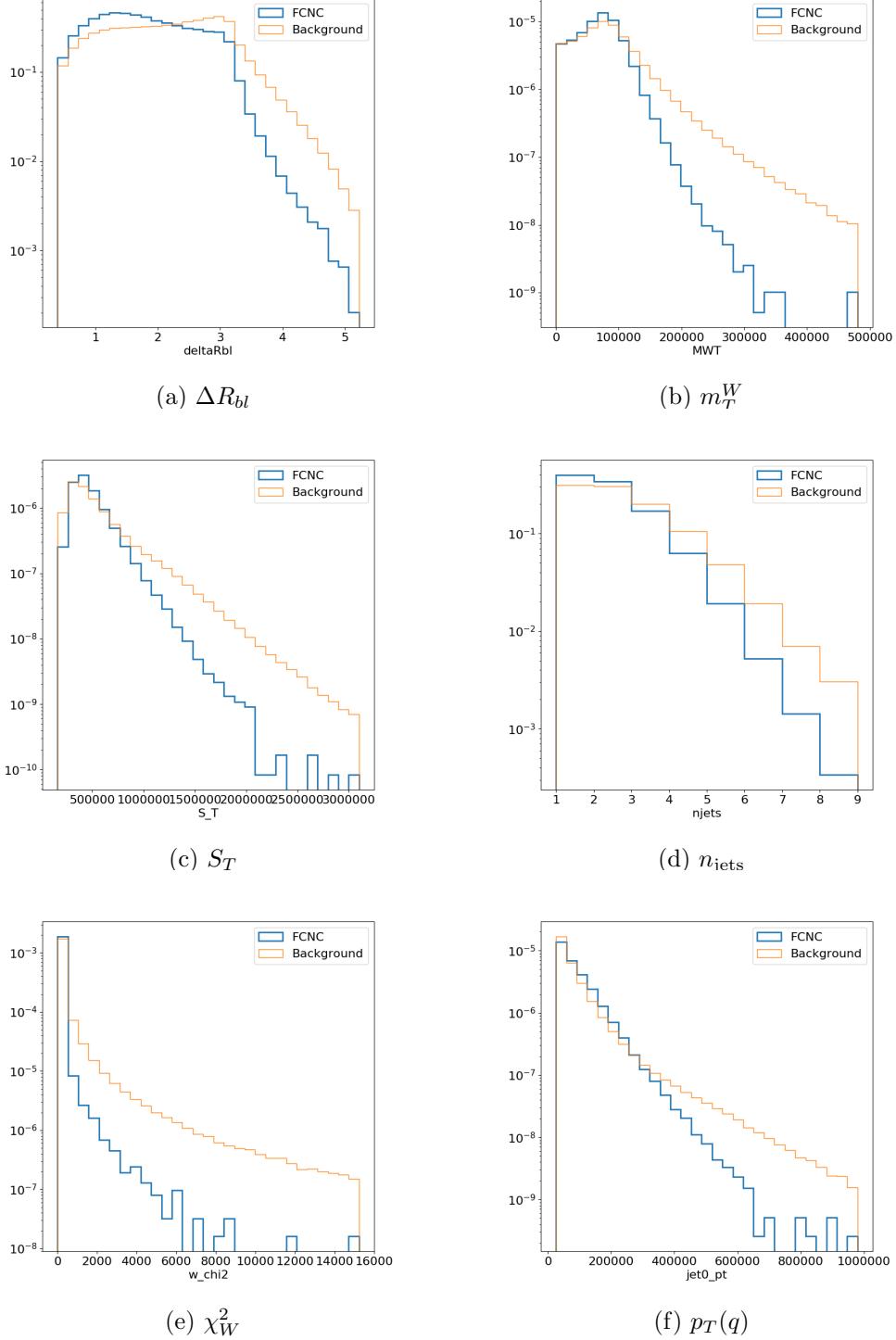


FIGURE 5.2. Normalized variables showing the shapes of neural network input variables for the $\mu+\text{jets}$ channel: ΔR_{bl} , m_T^W , S_T , n_{jets} , χ_W^2 , and $p_T(q)$

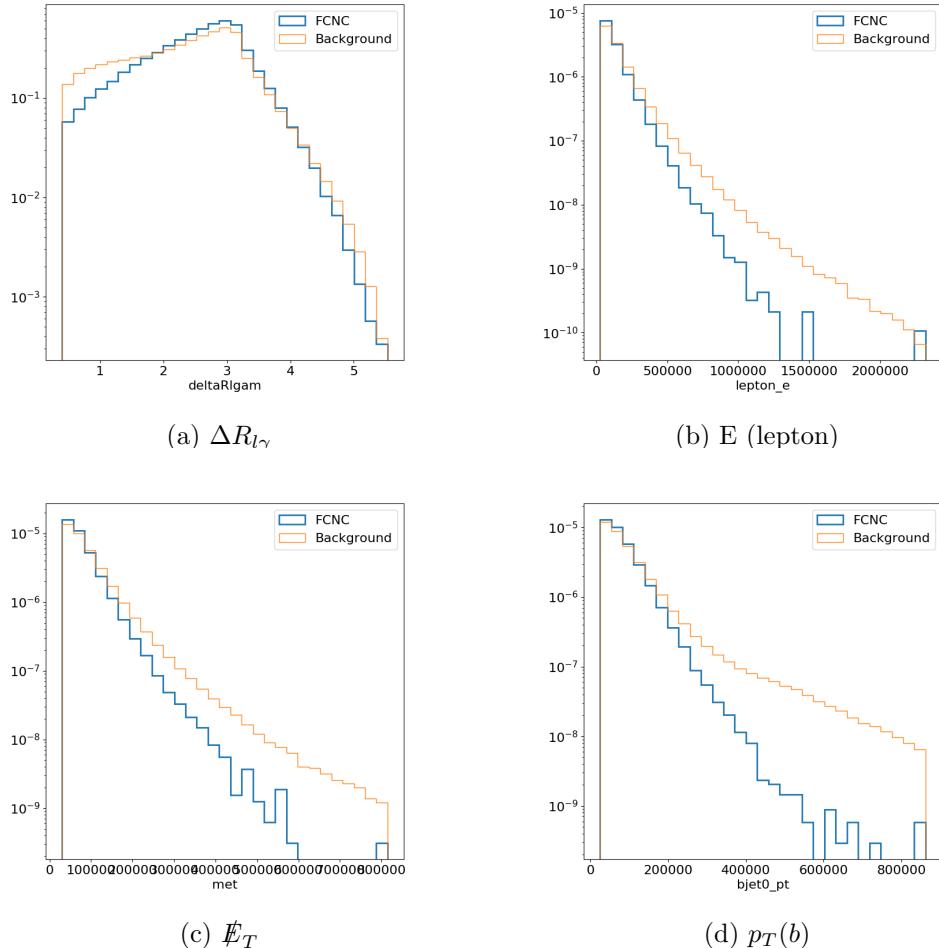


FIGURE 5.3. Normalized variables showing the shapes of neural network input variables for the $\mu+\text{jets}$ channel: $\Delta R_{l\gamma}$, E (lepton), \not{E}_T , and $p_T(b)$

5.3.2. Architecture

A variety of architectures of dense neural networks are studied using KERAS[79] on top of the TENSORFLOW backend [80]. Each network has a number of input nodes equal to the number of input variables. Networks with one, two, and three hidden layers are investigated each with 20 nodes. The output layer contains only a single node. Every node in one layer is connected to every node in the next layer and the previous layer. Every connection is assigned a weight that is optimized during the training of the network. For every node in the network a value is computed using the weights and input values of the previous nodes using an activation function. Nodes with the highest output of this function are more important to the fit. The activation function used on the internal nodes in this search is the Rectified Linear Unit activation function.

$$ReLU(x) = \begin{cases} x, & \text{if } x \geq 0 \\ 0, & \text{if } x < 0 \end{cases}$$

The output layer uses the sigmoid function, $\sigma(x)$, as an activation function. The sigmoid function maps the output smoothly to the range (0,1).

$$\sigma(x) = \frac{1}{1 + e^{-x}}$$

In every training step the weights of each node are updated following an optimization algorithm, in this case the ADAM optimizer[81]. This optimizer follows the steepest gradient to reach the minimum of the parameter of interest called the loss function. The loss function used for these classification neural networks is the binary cross

entropy:

$$\text{Loss} = -\frac{1}{N} \sum_{i=1}^N y_i \log(p(y_i)) + (1 - y_i) \log(1 - p(y_i))$$

where y is a binary indicator (0 or 1) if class label is the correct classification for observation and p is the predicted probability observation of the class label (0 or 1). The logarithmic nature of this loss function means it applies small values to correctly assigned events but more harshly punishes mismatching of events. Therefore having a similar number of signal and background events that get weighted similarly can improve the behavior of the network. In rare decay searches typically the amount of signal events is significantly smaller than the amount of background events in the training sample. By using the weight functionality in KERAS, the total number of signal events can be scaled to be similar to the number of background events.

Weighting the signal events this way allows the network to separate the signal and background events in a way that is significantly less harsh than without the weights by taking advantage of the loss function being used. This improves the estimated significance of the neural network cut after the signal events are rescaled to their proper normalization values.

Various hyperparameters are used as inputs into the neural network as well as the optimizer used. The ADAM optimizer has a default learning rate of 0.001 which remained constant throughout these studies. The learning rate corresponds to the amount that weights are updated during training. A learning rate that is too large can mean the network never settles into a local minima as it is always missing the minima or, at the very least, it can take much longer to converge into a minima. As the neural network training for this search always converged quickly and to a similar value after being tested multiple different times the learning rate was not adapted.

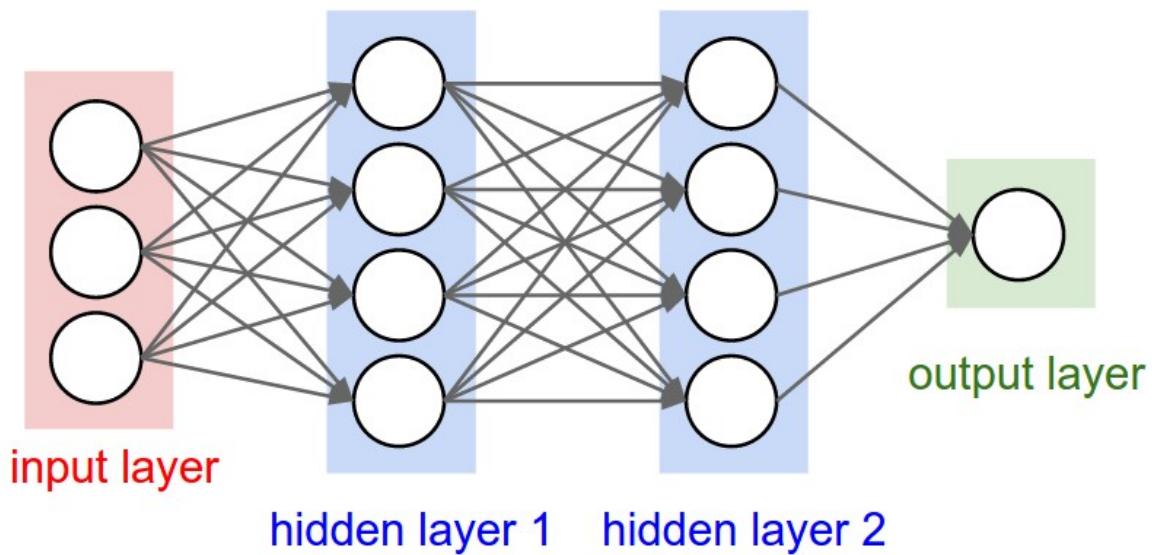


FIGURE 5.4. Pictoral representation of neural network architecture with 3 input variables, 2 hidden layers with 4 nodes each, and 1 output layer[82].

Another hyperparameter of note is the batch size that defines the number of samples that are propagated through the network at once. The batch size is crucial in determining how long the training of the network takes. A set of 1000 training samples with a batch size of 100 will propagate each set of 100 samples through the neural network every epoch, so 10 separate batches. A larger batch size means that each epoch of the training takes a shorter amount of time. However, as the weights are updated after each batch the network can take many more epochs to converge as the weights are being updated less frequently. A batch size of 100 was used while training the networks presented in this chapter. Larger batch sizes were tested with the only difference being the time each epoch took and the total time the network took to converge.

Epochs are the total number of times the network has been trained over the entire training set. All of the networks were allowed up to 200 epochs to converge with a KERAS patience value set to 50. The loss function minimization would be

done every batch and after each epoch the best possible value of the loss function is found. If this value is better than any previous epoch the network is allowed to train for 50 more epochs until 50 epochs have passed without finding a new minimum loss function value which then terminates the training. All models converge early and are terminated typically between epoch 80 and 120 meaning the loss function was minimized between epoch 30 and 70.

One method employed to avoid overtraining the network dropout regularization was used on each of the hidden layers. Dropout has the effect of simulating a large number of networks with very different network structures by removing nodes randomly throughout the training. A dropout rate of 20% was used, meaning that for every batch 20% of the weights of the hidden layer nodes were set to 0. This prevents the network from becoming overly dependent on any given node or learning the data ‘by heart’ as opposed to recognizing the trends in the sample.

5.3.2.1. Training and Validation of Neural Networks

The input variables into the neural network are preprocessed using the ROBUSTSCALAR method implemented in **scikit-learn**[83]. The preprocessing is done so that the input variables exist on a similar scale. As the network is tasked with learning how to combine these inputs through a series of linear combinations and nonlinear activation function values, a disparity in the scales of the input values can lead to awkward loss function topology that focuses on certain parameter gradients instead of treating them all similarly. Normalizing the values to a standard scale allows the network to learn the optimal parameters for each input node more quickly and efficiently. This means that less focus can be used on the optimization of the

hyperparameters for the network as the scales of the inputs do not need to be learned by the network itself.

Each input variable in the neural network, x , is scaled by the following equation:

$$z = \frac{x - m}{q_3 - q_1}$$

where m is the median of the distribution, and q_1 and q_3 are the first and third quartile. This changes the distribution of the input variable distributions to be centered around zero.

A second method to avoid overtraining the neural network is to make use of a train-test split to split the signal and background samples into 3 independent randomized sets before training the neural network. The samples are split into a training set of 64% of the samples, a test set containing 20% of the samples, and the remaining 16% are a validation set. The training and test sets are used during the training of the network while the validation set is used to compute performance of the trained neural network.

One measure of the performance of the network is the accuracy. The KERAS default accuracy measure is defined:

$$\text{accuracy} = \frac{N(\text{event}_{NN} \geq 0.5|\text{signal}) + N(\text{event}_{NN} < 0.5|\text{background})}{N(\text{signal}) + N(\text{background})}$$

where $N(\text{event}_{NN} \geq 0.5|\text{signal})$ ($N(\text{event}_{NN} \geq 0.5|\text{signal})$) is the number of signal (background) events with $P_{\text{signal}} \geq 0.5$ ($P_{\text{signal}} < 0.5$). Essentially, the accuracy is a measure of the mean of how often correct prediction values occur assuming a cut on the output of ≥ 0.5 .

5.3.3. Hidden Layer Studies

The general performance of the neural network was studied with a varying number of hidden layers (1, 2, and 3) in both the e +jets and μ +jets channels. All of the networks are trained on the same set of variables and with the same train-test split input data. For each of the channels the *Receiver Operating Characteristic* (ROC) curves are shown in Figure 5.5. The ROC curves show the value of $1 - \epsilon_{\text{bkg}}$ as a function of the true positive rate, ϵ_{signal} . A figure of merit is the Area Under the Curve (AUC) which is a measure of how close the resulting values are to the optimal value of unity.

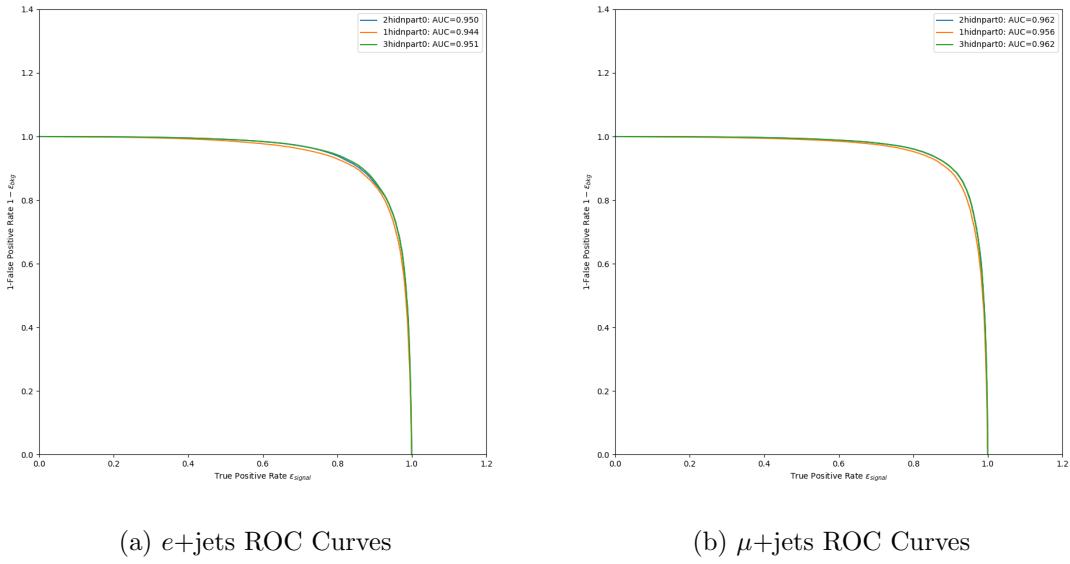


FIGURE 5.5. ROC Curves are shown for both search channels for a varying number of hidden layers. Orange lines correspond to one hidden layer, blue to 2 hidden layers and green to 3 hidden layers. The blue and green curves have near identical AUC values: 0.950 and 0.951 for the e +jets case and 0.962 for the μ +jets cases.

The AUC for 2 hidden layers and 3 hidden layers are identical within rounding errors for both channels. As such the network with 2 hidden layers has been chosen as

it is computationally simpler. The normalized neural network output values are shown in Figure 5.8. Adding a second hidden layer significantly improves the performance of the network but a third layer does not. The output shapes change slightly when the third hidden layer is added due to the network learning differently from similar data. However, as the AUC shows, the performance of 2 and 3 hidden layers is identical. Figures 5.6 and 5.7 show the accuracy metric and the loss function as a function of the training epoch for the networks trained with 2 hidden layers. The accuracy plot behavior is expected as the validation data sets do not have dropout regularization applied to them. These networks are also trained without further reduction of Z+jets background meaning the e +jets sample has a larger background contamination that makes the validation testing more volatile. This is due to the increased number of similar events in that sample that can be more heavily dependent on specific weights across the network for identification.

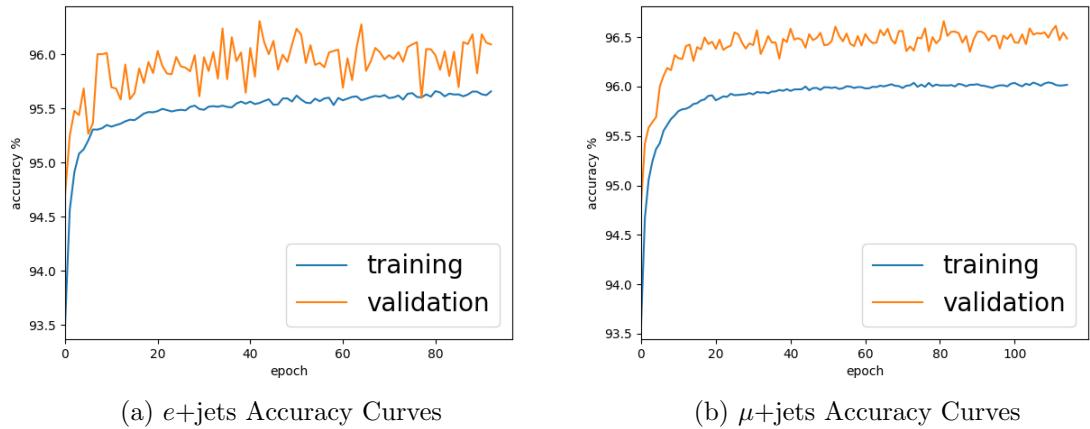


FIGURE 5.6. Accuracy plots for both channels for the 2 hidden layer neural network

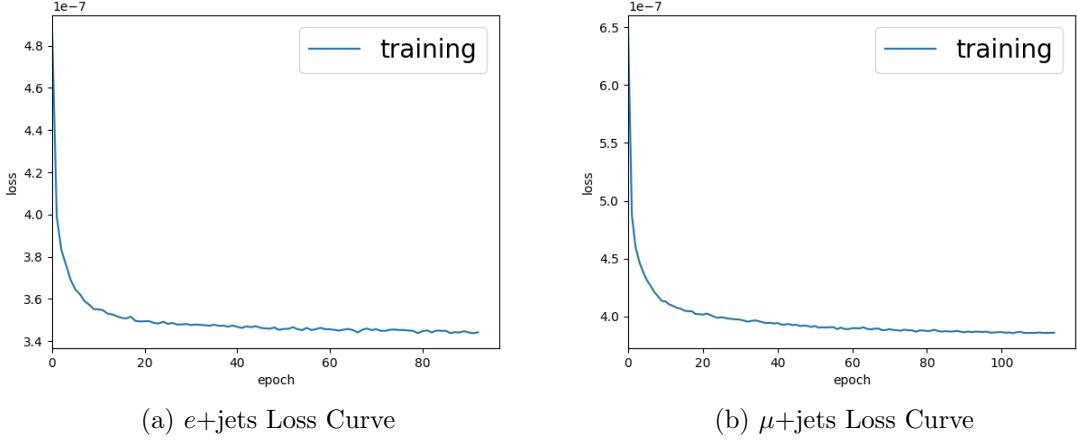


FIGURE 5.7. Loss plots for both channels for the 2 hidden layer neural network

The main metric used in choosing which network has the best physics reach is the significance:

$$\text{significance} = \frac{N_s}{\sqrt{N_s + N_b}}$$

where N_s is the number of signal events that pass the cut and N_b is the number of background events that pass the neural network cut. After the model has been fully trained, it is tested on all of the Monte Carlo for signal and background. The signal samples are normalized to various branching ratios (in the range $10^{-5} \rightarrow 3 \times 10^{-3}$) and full LHC Run-2 luminosity, and the significance is calculated as a function of the cut on the output of the neural network $P(\text{signal})$. The network with the output cut for the smallest branching ratio with a maximum significance of 2 is chosen, a rough estimate of where the expected limit could be set. The significance as a function of the neural network output cut is shown in Figure 5.9.

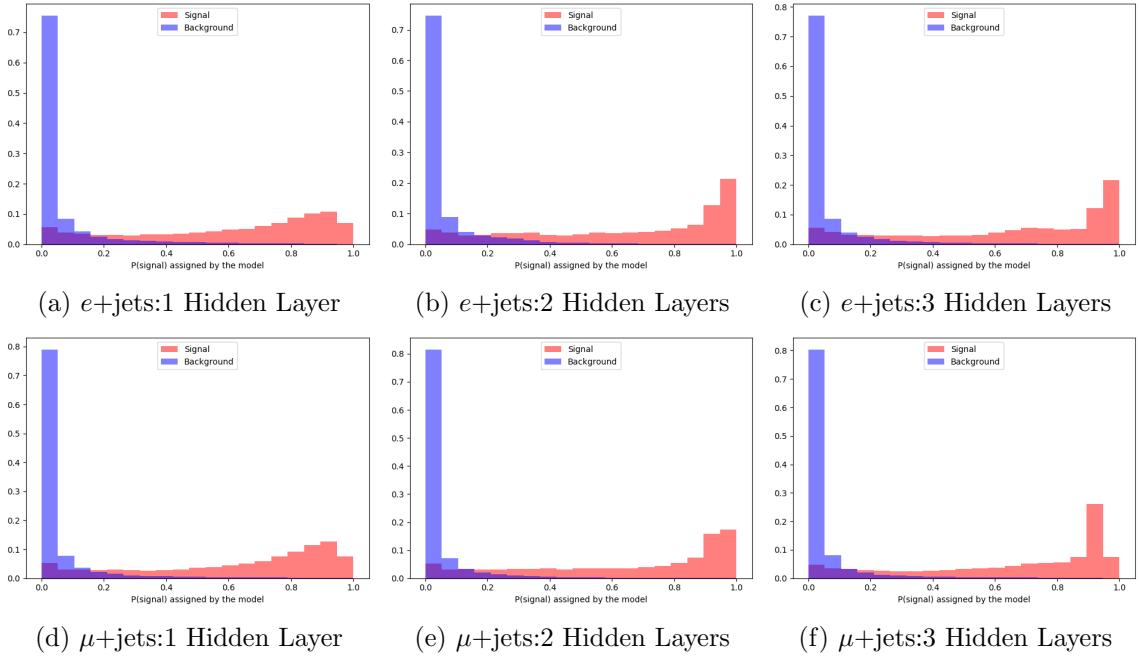


FIGURE 5.8. Normalized neural network output signal and background distribution plots are shown for both search channels for a varying number of hidden layers.

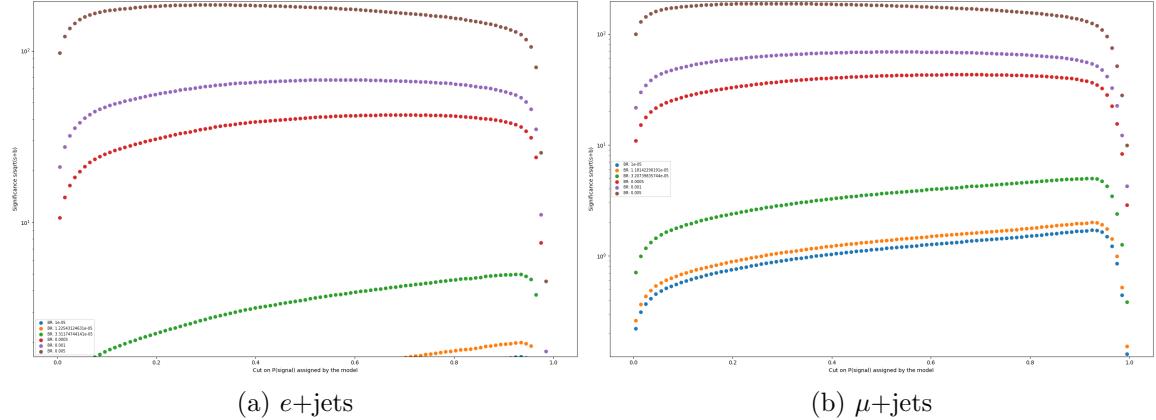


FIGURE 5.9. Significance plots for both channels for the 2 hidden layer neural network. The green points correspond to a branching ratio with a maximum significance of 5, and the orange to a maximum significance of 2. The $e+jets$ ($\mu+jets$) branching ratio with max significance of 2 is 1.22×10^{-5} (1.18×10^{-5}). The blue, red, purple, and brown points correspond to branching ratios of 1×10^{-5} , 5×10^{-4} , 1×10^{-3} , and 5×10^{-3} , respectively.

5.3.4. B-Tagging Working Point Studies

The b-tagging working point selection was performed with similar neural network studies. Three neural networks were trained with the datasets using the jet information and total scaled events for each of the major b-tagging working points: 70%, 77%, and 85%. Changing the working point alters a number of things about the signal and background data sets such as which jets are b tagged and therefore which jets are combined into the higher level variables (e.g., $m_{q\gamma}$ and m_{Wb}). The total number of events that pass the pre-selection to the neural network is also changed for all of the datasets since the neural networks are only trained on events with 1 b-tagged jet. Similar sets of plots to Section 5.3.3 will be presented in this section.

This selection of neural networks were trained in parallel with one, two, and three hidden layers. The only results shown are the 2 hidden layer outputs as they perform equally or better than the others as previously discussed. The accuracy and loss plots for these networks are shown in Figures 5.10 and 5.11. The neural network output and significance plots shown in Figures 5.12 and 5.13 follow.

The result of these studies is the choice of using the 77% working point for b-tagged jets. The branching ratio with significance of 2 is found for each network and reported in Table 5.2.

B-Tag Working Point	$e + \text{jets}$ Branching Ratio	$\mu + \text{jets}$ Branching Ratio
70%	1.25×10^{-5}	1.31×10^{-5}
77%	1.23×10^{-5}	1.18×10^{-5}
85%	1.27×10^{-5}	1.19×10^{-5}

TABLE 5.2. Branching ratio values with a significance of 2 after neural network optimization



FIGURE 5.10. Accuracy and loss plots for the $e+jets$ channel at 70%, 77%, and 85% b-tagging working points.

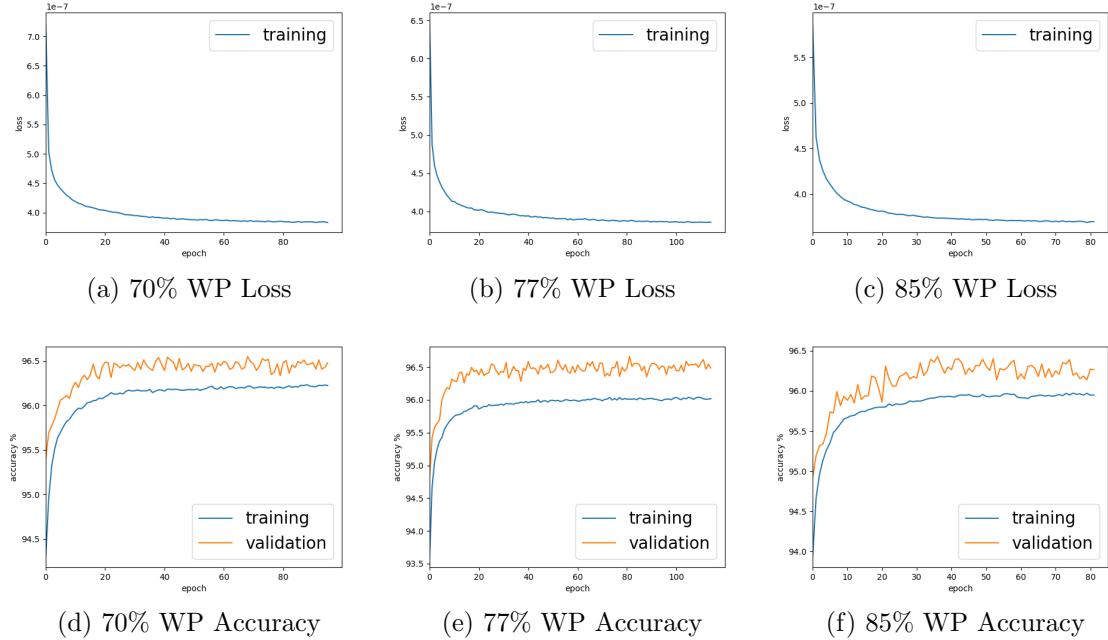


FIGURE 5.11. Accuracy and loss plots for the $\mu+jets$ channel at 70%, 77%, and 85% b-tagging working points.

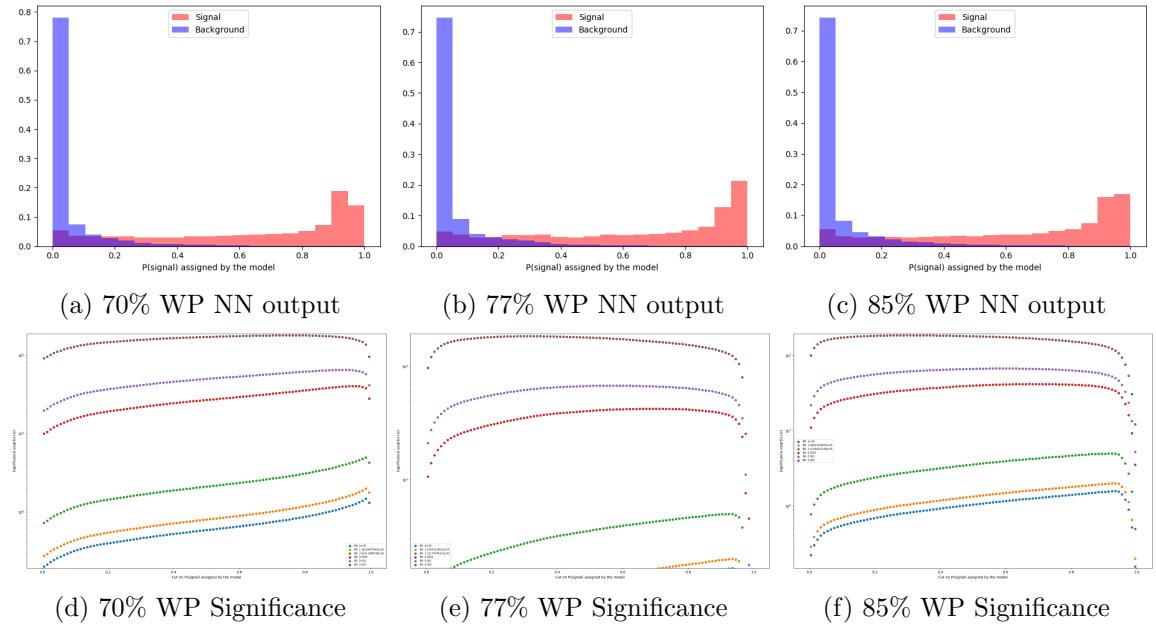


FIGURE 5.12. Neural network output and significance plots for the $e+jets$ channel at 70%, 77%, and 85% b-tagging working points.

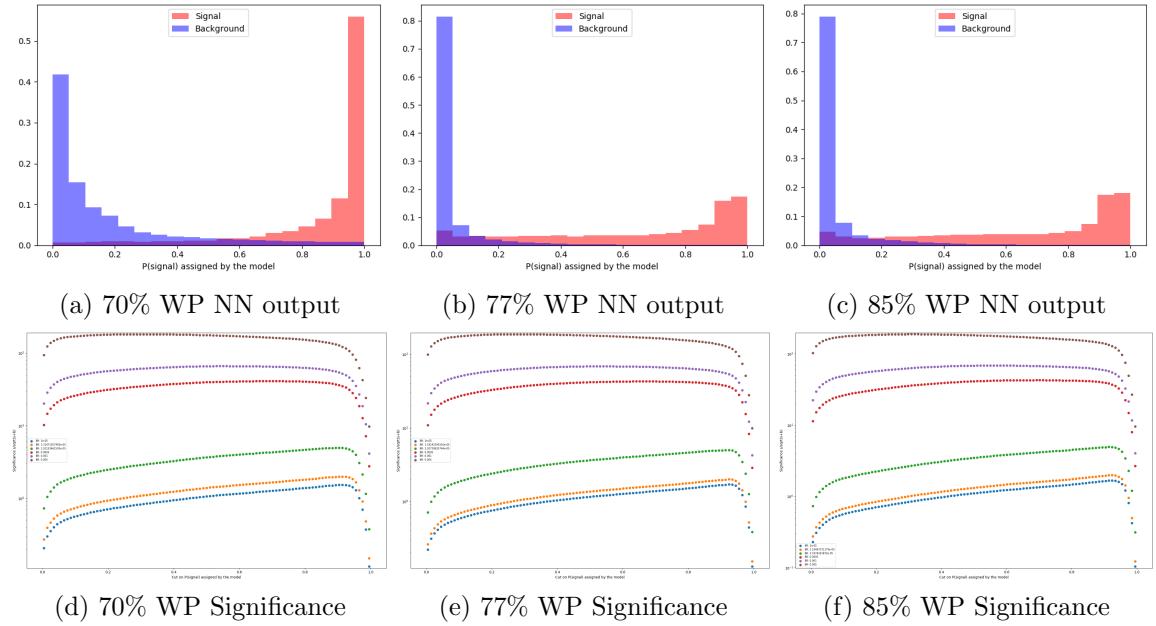


FIGURE 5.13. Neural network output and significance plots for the $\mu+jets$ channel at 70%, 77%, and 85% b-tagging working points.

5.4. Initial Event Selection

Initial event selection is done to ensure that the events accepted into the analysis are not contaminated by extremely noisy detector environments and that these events happened during times when the ATLAS detector was accepting events properly. All of the events have the same initial set of criteria for determining whether or not the event is looked at any further for this analysis, applying to both MC and Data. These initial checks are as follows:

- Only events occurring during runs good for physics
- Good Calorimeter status: ensures that the LAr and Tile calorimeters are not experiencing a noise burst at the time of the event
- Requires a primary vertex to be reconstructed for the event which ensures timing of further reconstructed objects are placed with the correct vertex
- Global Trigger Decision: selects events based on whether they passed one of the triggers including the trigger thresholds, further discussed in Section 5.4.1
- Trigger Match: select events where an electron or muon matches the trigger
- Overlap Removal as discussed in Section 5.1.1
- Ignore events that have a bad muon, which occur mostly in the transition region and the cathode strip chamber regions.
- Jet Cleaning: removes events with jets formed from calorimeter information from sources that are unrelated to the energy flow from the initial hard scatter interaction

These basic event selection values are applied to every event, in both MC and Data. Beyond these values, various kinematic cuts are included to form the additional analysis level objects and regions used in the analysis. These additional kinematic cuts are examined more closely in Section 5.5 and in the discussion of kinematic region creation throughout the rest of the analysis e.g., Section 5.6.

5.4.1. Triggers

Different HLT triggers are used for data taking periods for each year of Run 2. This analysis takes advantage of single lepton triggers for electrons and muons to dramatically reduce backgrounds due to QCD events without leptons.

Year	p_T threshold [GeV]	Identification Menu	Isolation Menu	L1 Seed
2015	≥ 24	Medium	None	L1EM20VH
	≥ 60	Medium	None	-
	≥ 120	Loose	None	-
2016-2018	≥ 26	Tight	Gradient (Loose)	-
	≥ 60	Medium	None	-
	≥ 140	Loose	None	-

TABLE 5.3. The electron trigger requirements in the event selections

Year	p_T threshold [GeV]	Identification Menu	Isolation Menu	L1 Seed
2015	≥ 20	None	Gradient (Loose)	L1MU15
	≥ 50	None	None	-
2016-2018	≥ 26	None	Gradient (Medium)	-
	≥ 50	None	None	-

TABLE 5.4. The muon trigger requirements in the event selections

5.5. Data and MC Pre-Selection Cuts

The Signal Region pre-selection is defined to select events that have an opportunity to enter the final search selection. This pre-selection selects events with exactly one massive lepton, at least two jets (at least one of which is b-tagged at the 77% working point), transverse momentum, and exactly one photon such that it resembles the expected final-state topology for the signal. All of the events have the same initial set of criteria for determining whether or not the event is further examined for this analysis, applying to both MC and Data. These initial checks are as follows:

- Exactly 1 lepton (electron or muon) $p_T > 25$ GeV
- At least two good jets ($p_T > 25$ GeV)
- At least one b-tag (MV2c10, 77% working point)
- $\cancel{E}_T > 30$ GeV and $m_T^W > 30$ GeV (for events with electrons)
- $\cancel{E}_T > 20$ GeV and $\cancel{E}_T + m_T^W > 60$ GeV (for events with muons)
- Exactly 1 photon, $p_T > 50$ GeV

These plots are also produced before additional scale factors are added to account for mismodelling of various processes. These include the fake rate scale factors for

processes where a truth electron or hadron is reconstructed to a photon and scaled to account for further mismodeling based on the order of the MC events produced (leading order, next-to-leading order, etc.). Only statistical uncertainties are shown.

Signal photons, which originate from a top quark decay, are very high p_T whereas background photons typically result from soft processes. A cut on the photon candidate p_T removes much of the backgrounds while keeping a majority of the signal. The photon p_T in the preselection region is shown in Figure 5.14.

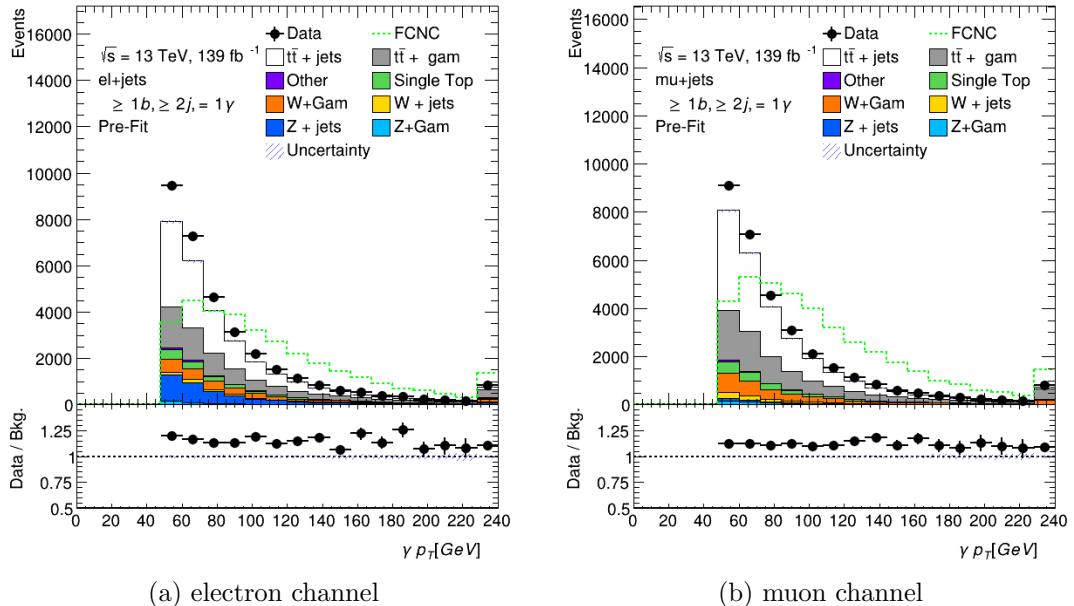


FIGURE 5.14. Photon p_T in the signal region pre-selection region. FCNC signal branching ratio is scaled to 1.5%.

Other variables of interest, i.e., those being used as inputs into the neural network are also showed in this section. Figure 5.15 shows the S_T and m_T^W distributions. Figure 5.16 shows the invariant mass distributions for both top quark candidates, m_{Wb} and $m_{q\gamma}$. The kinematic variables for the electron channel are shown in Figure 5.17 and for the muon channel in Figure 5.18. The neural network output of these events are shown in Figure 5.19.

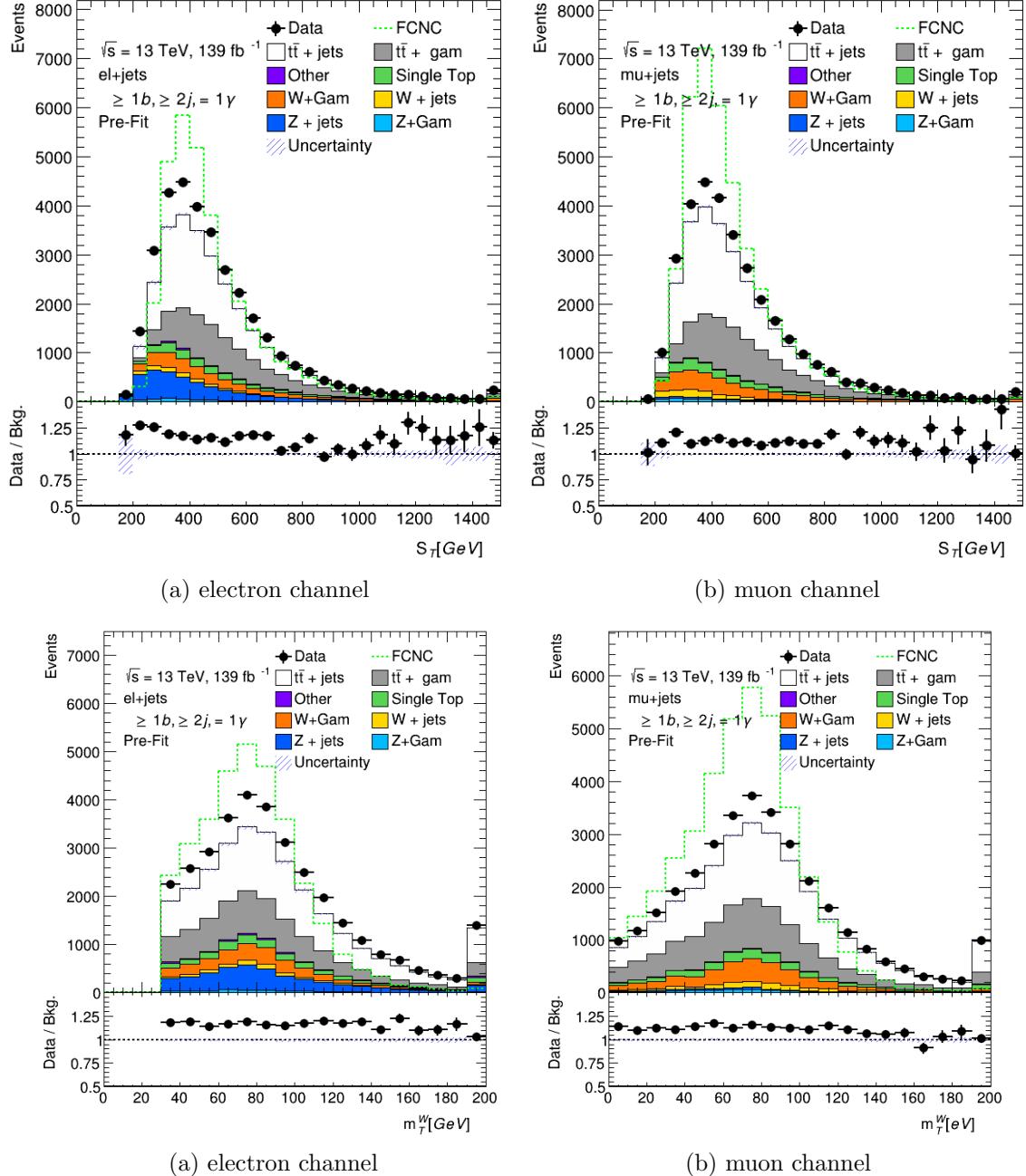


FIGURE 5.15. S_T and m_T^W in the signal region pre-selection region. FCNC signal branching ratio is scaled to 1.5%.

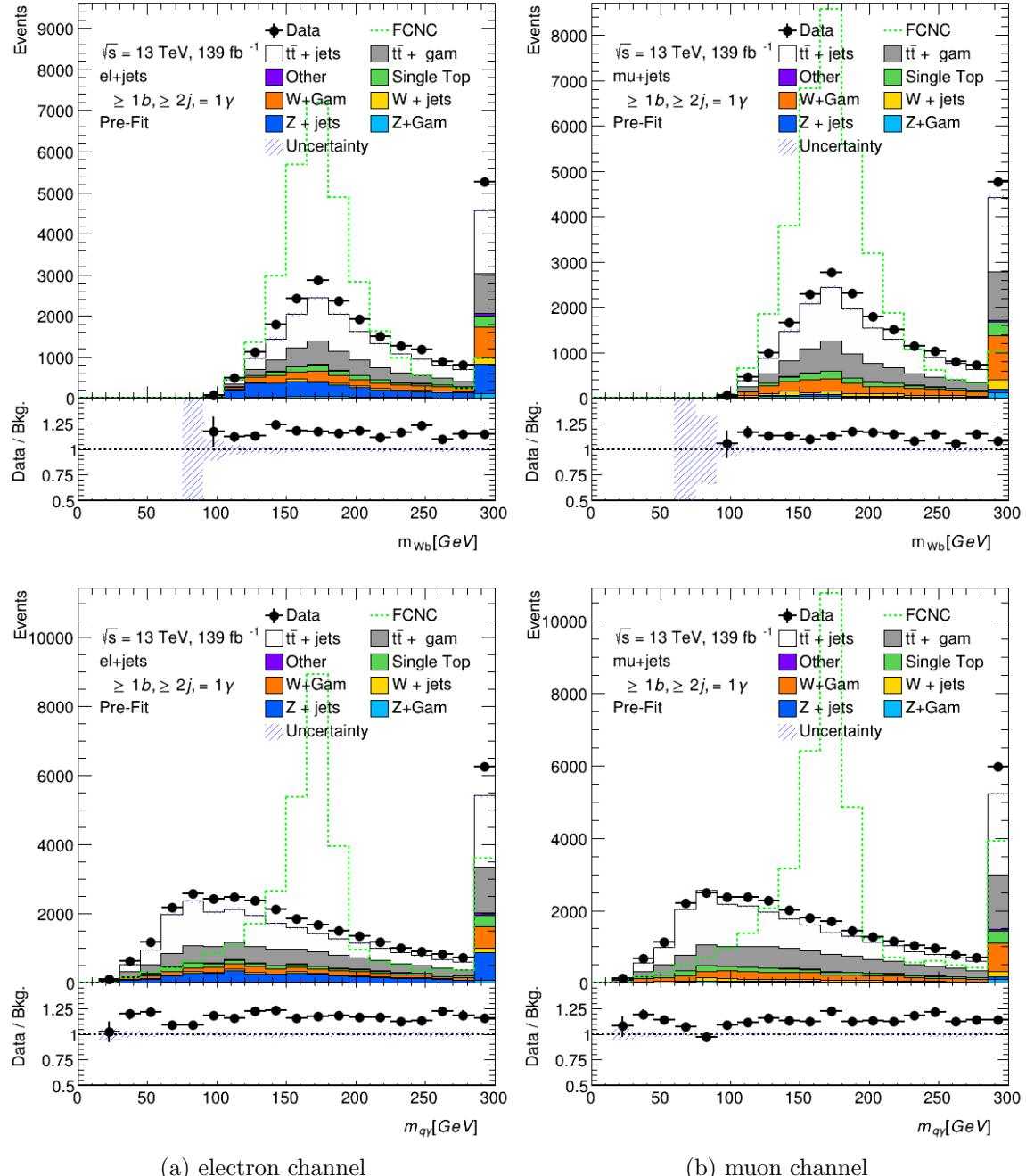


FIGURE 5.16. Top mass candidates in the signal region pre-selection: m_{Wb} and $m_{q\gamma}$. FCNC signal branching ratio is scaled to 1.5%.

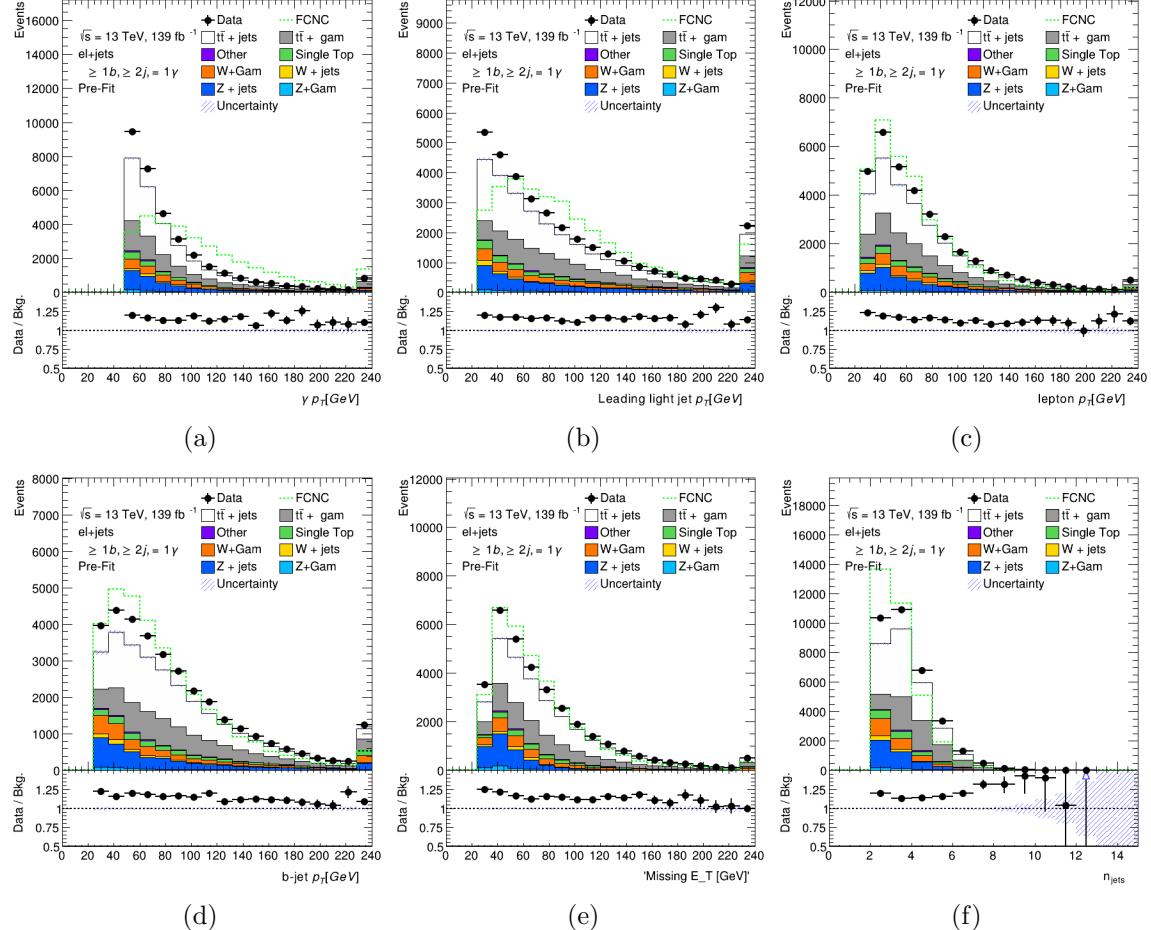


FIGURE 5.17. Photon p_T (a), leading light jet p_T (b), lepton p_T (c), b-jet p_T (d), \not{E}_T (e), and n_{jets} (f) plots in the signal region pre-selection for the electron+jets channel. FCNC signal branching ratio is scaled to 1%.

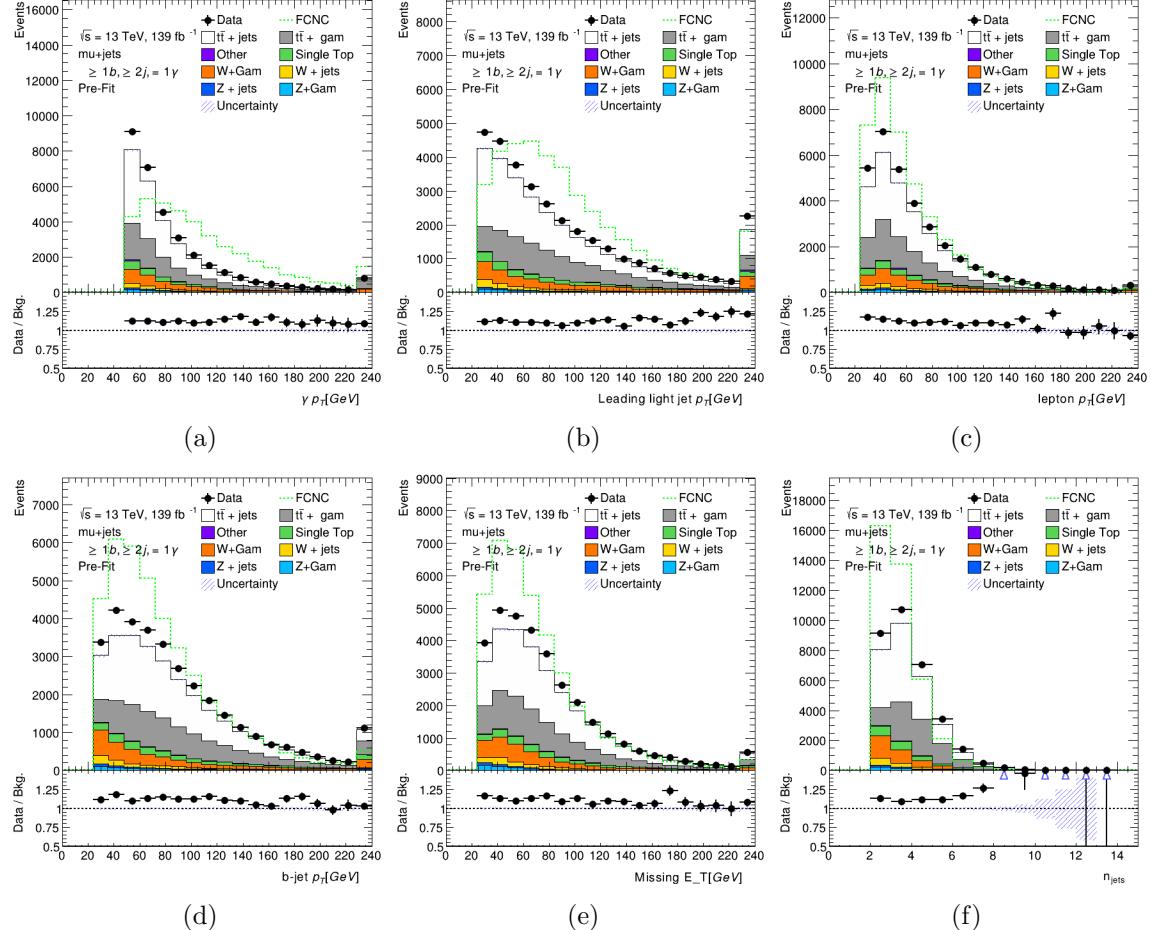


FIGURE 5.18. Photon p_T (a), leading light jet p_T (b), lepton p_T (c), b-jet p_T (d), \cancel{E}_T (e), and n_{jets} (f) plots in the signal region pre-selection for the muon+jets channel. FCNC signal branching ratio is scaled to 1.5%.

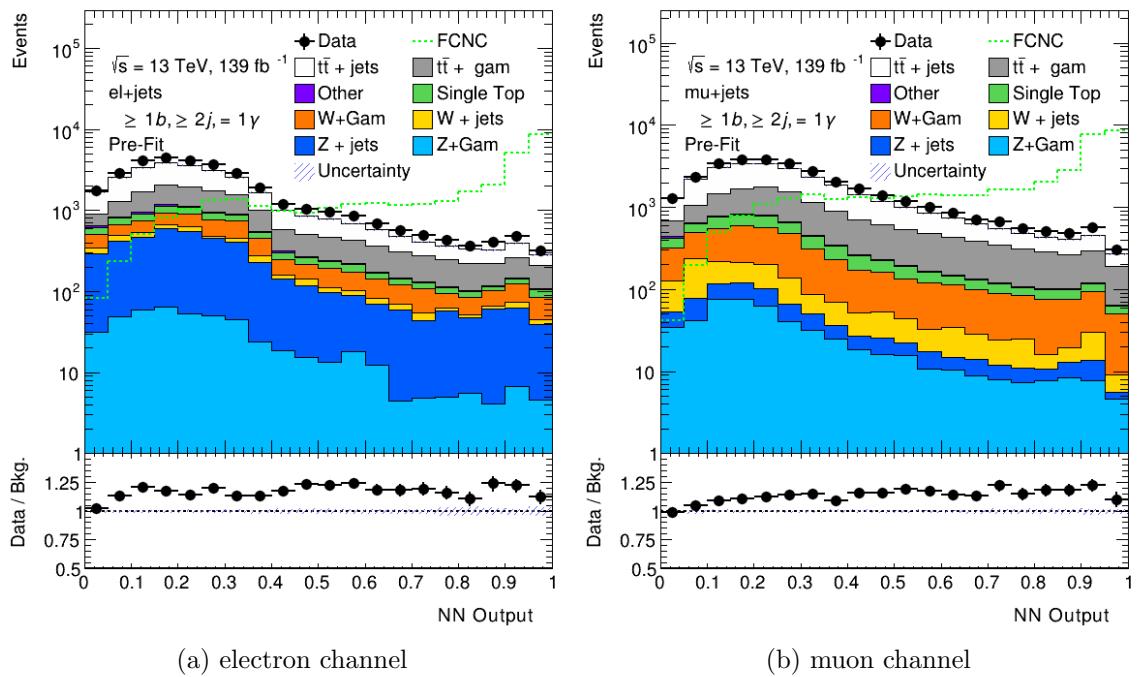


FIGURE 5.19. Output of the Neural Network in the signal region pre-selection region. FCNC signal branching ratio is scaled to 1.5%.

5.6. Background Evaluation: Control and Validation Regions

Orthogonal regions to the signal region have been created to test the performance of Monte Carlo samples. Control and validation regions are designed to isolate specific physics processes to determine and test the efficacy of scale factors that will be applied to the final signal region Monte Carlo events. These control and validation regions need to be kinematically similar to the signal region such that derived scale factors can be translated directly into the signal region and orthogonal to make sure that there is little signal contamination in the regions. Regions have been created to test the major backgrounds expected in the signal region: $t\bar{t}$, W+jets, as well as similar events produced with an associated photon: $t\bar{t} + \gamma$ and W+Jets+ γ . Events without real photons are described in Section 5.6.1 and regions with a real photon are described in Section 5.6.2.

5.6.1. Backgrounds Without Photons

Various background processes that do not have a real photon produced in the events can still enter the signal region if an electron or jet is mis-reconstructed as a photon. Of these processes the largest contributors in the signal region are Standard Model $t\bar{t}$ and W+jets. As the LHC attains higher and higher energies the QCD multijet backgrounds become increasingly hard to model due to the non-perturbative nature of the interactions. A data-driven technique to study these backgrounds was developed by scaling the major backgrounds without photons to account for the QCD backgrounds that contribute extra jets to the major backgrounds. Designing a single control region satisfactorily close to the signal region is impossible. Thus, two control regions are designed, one which is W+jets rich and the other $t\bar{t}$ rich. Scale factors for these backgrounds are derived simultaneously and tested in a third similar region for

validation before being applied to other regions. These control and validation regions are defined as follows:

- All of the Initial Event Selection as outlined in Section 5.4
- Exactly 1 lepton (electron or muon) $p_T > 25$ GeV
- Number of Jets ($p_T > 25$ GeV) to define the regions
 - * Control Region 1 (W+Jets enriched): $n_{\text{jets}} = 3$
 - * Validation Region: $n_{\text{jets}} = 4$
 - * Control Region 2 ($t\bar{t}$ enriched): $n_{\text{jets}} \geq 5$
- $\cancel{E}_T > 30$ GeV and $m_T^W > 30$ GeV (for events with electrons)
- $\cancel{E}_T > 20$ GeV and $\cancel{E}_T + m_T^W > 60$ GeV (for events with muons)
- Exactly 1 b-tagged jet (MV2c10, 77% working point)
- 0 photons, $p_T > 15$ GeV

The efficiency of scale factors derived using control regions 1 ($n_{\text{jets}}=3$) and 2 ($n_{\text{jets}} \geq 5$) are then tested in the validation region ($n_{\text{jets}}=4$). The scale factors for the $t\bar{t}$ and W+jets MC are derived using:

$$\begin{bmatrix} N(W)_{3j} & N(t\bar{t})_{3j} \\ N(W)_{5+j} & N(t\bar{t})_{5+j} \end{bmatrix} \begin{bmatrix} W_{SF} \\ t\bar{t}_{SF} \end{bmatrix} = \begin{bmatrix} N(\text{data-bkg})_{3j} \\ N(\text{data-bkg})_{5j} \end{bmatrix}$$

Figure 5.20 shows the S_T distribution in both electron and muon channels before scale factors are calculated for all three kinematically separate regions. The large mismodelling occurs at low S_T values as expected as QCD processes will typically

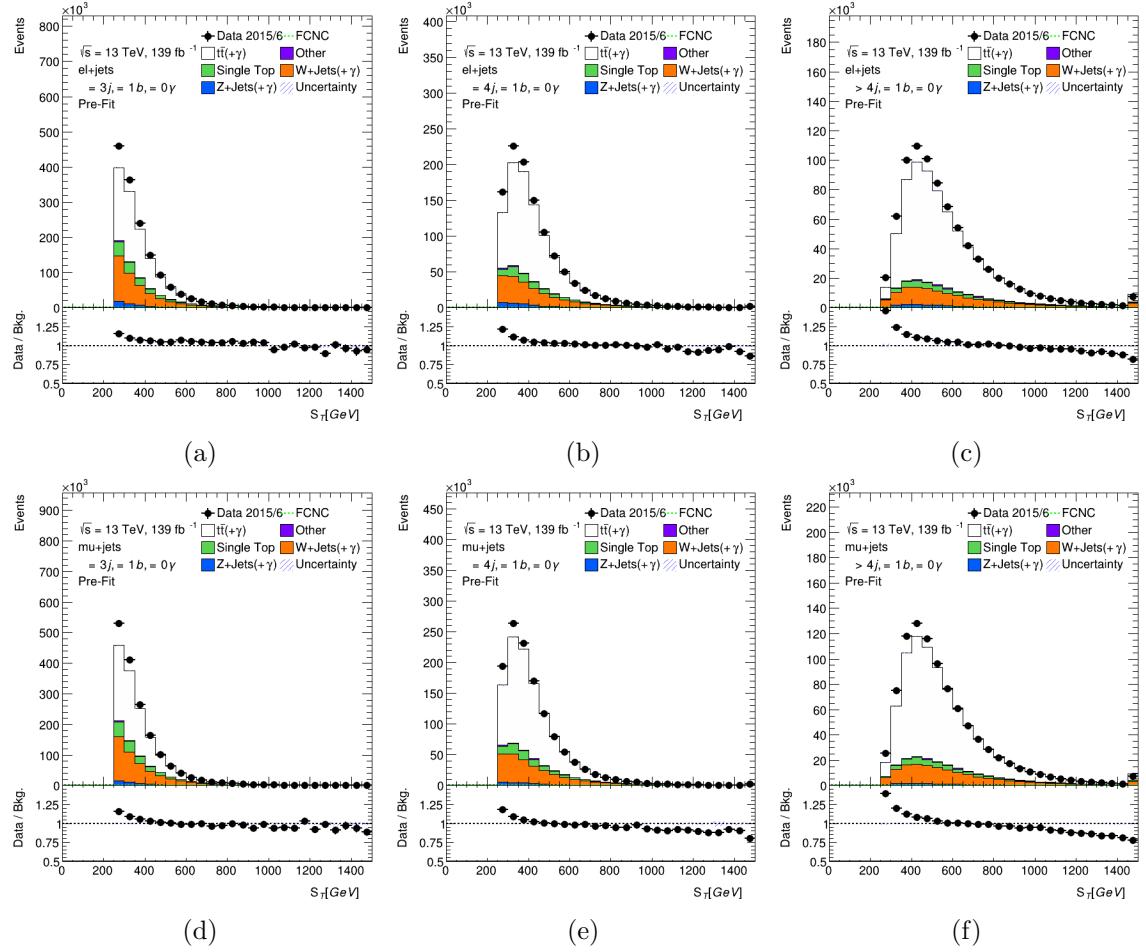


FIGURE 5.20. S_T distributions in the 3(a,d), 4(b,e), and 5+(c,f) jets control and validation regions. The electron channel is shown on the top and the muon channel on the bottom, before scale factors are determined.

add low energy jets to the events. The Figures 5.21(electron channel) and 5.22(muon channel) show various event-level variable plots for the validation region after the scale factors have been applied. The problem areas in the S_T distributions are not present in regions containing a photon. The scale factors do an excellent job scaling all of the kinematic regimes within regions enriched with signal like events, shown in Figures 5.33(f) and 5.35(f).

The derived scale factors using these regions are shown in Table 5.5.

Sample	e+jets SF	μ +jets SF
W+jets	1.22	1.25
$t\bar{t}$	1.06	1.01

TABLE 5.5. Derived $t\bar{t}$ and W+jets scale factors for QCD multijet backgrounds

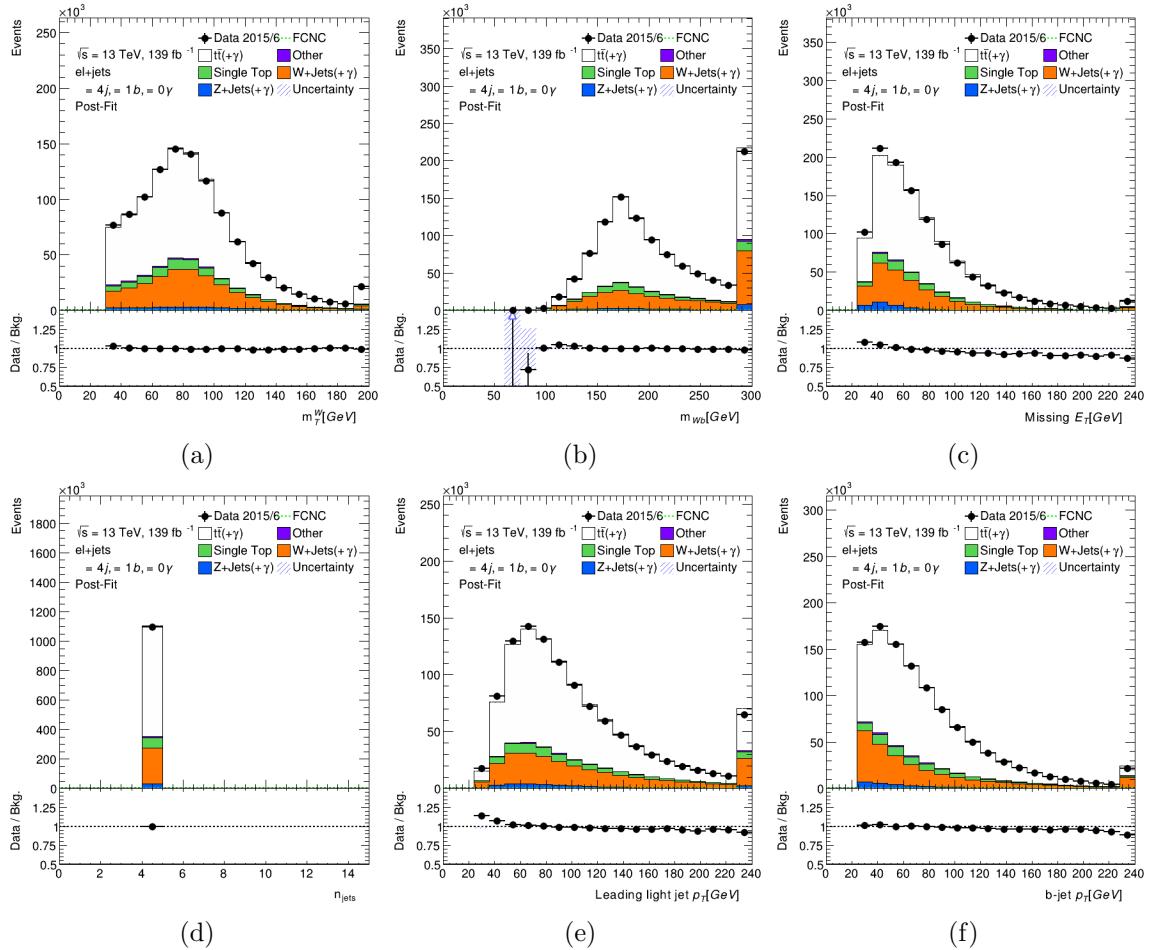


FIGURE 5.21. Event-level plots for the =4 jet validation region after scale factors have been applied in the electron channel. FCNC signal branching ratio is scaled to 1.5%.

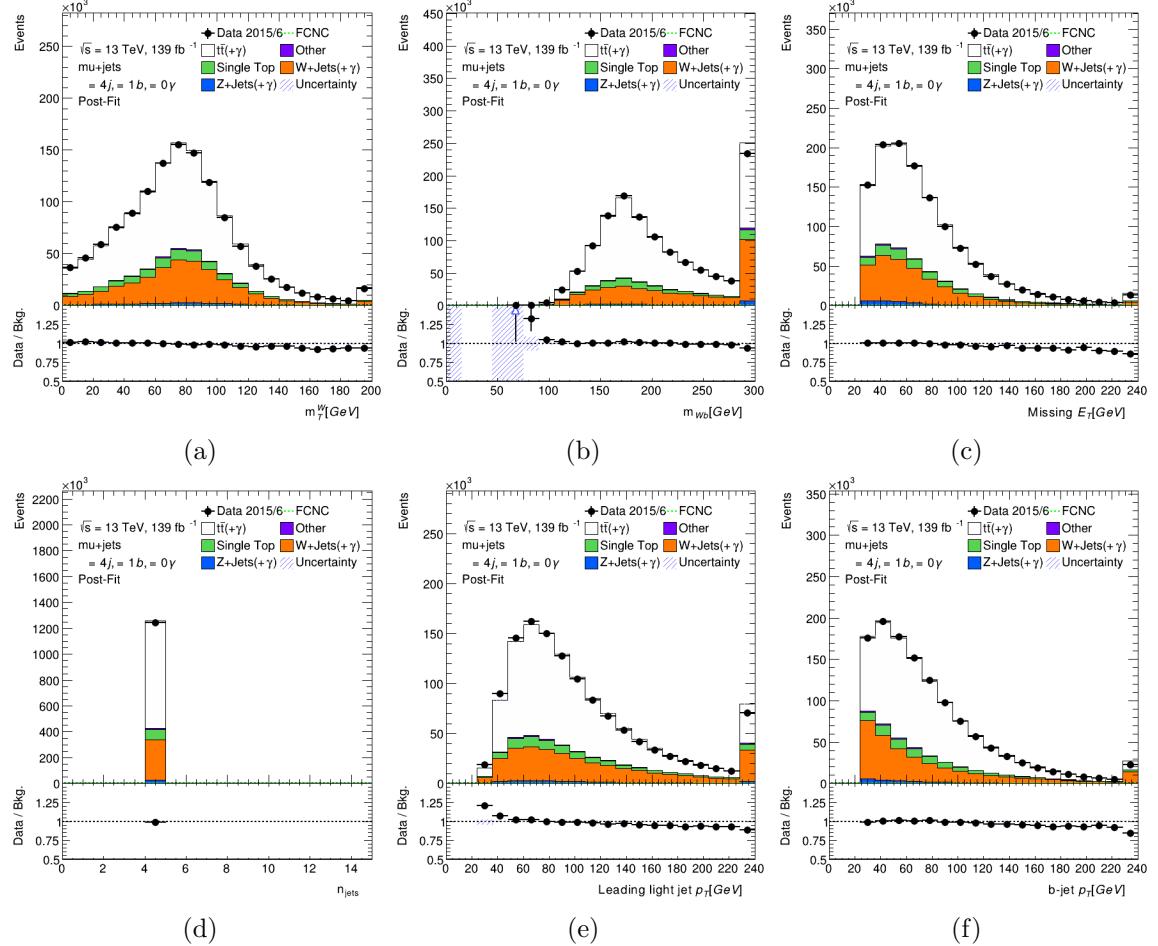


FIGURE 5.22. Event-level plots for the $=4$ jet validation region after scale factors have been applied in the muon channel. FCNC signal branching ratio is scaled to 1.5%.

5.6.2. Background With Photons

Standard Model processes that are produced with an extra real photon are an irreducible background for this search as they can share the same final state as the signal events. The largest contributors of these irreducible backgrounds are the major background samples discussed in the previous section with an associated photon ($t\bar{t}+\gamma$ and $W+jets+\gamma$). Special Monte Carlo samples are produced for these samples (along with $Z+jets+\gamma$) that have higher statistics of these photon enriched events than the nominal samples. However, as these samples ($X+jets+\gamma$) are subsets of the nominal sample($X+jets$), duplicate events must be removed from the nominal sample. This is done using the **MCTruthClassifier** tool that is detailed further in Section 5.7. All events with a photon from the hard scattering are removed from the $X+jets$ samples as they are contained within the $X+jets+\gamma$ samples.

5.6.2.1. $W+\gamma$ Control Region

A validation region for $W+jets+\gamma$ was created as it is one of the more dominant backgrounds other than $t\bar{t}$ and $t\bar{t} + \gamma$ events. The normalization for the $W+jets+\gamma$ validation region enters as a free parameter into the final fit. The region selection for the $W+jets+\gamma$ is as follows:

- All of the Initial Event Selection as outlined in Section 5.4
- Exactly 1 lepton (electron or muon) $p_T > 25$ GeV
- At least 2 Jets ($p_T > 25$ GeV)
- $\cancel{E}_T > 30$ GeV and $m_T^W > 30$ GeV (for events with electrons)
- $\cancel{E}_T > 20$ GeV and $\cancel{E}_T + m_T^W > 60$ GeV (for events with muons)

- Exactly 0 b-tagged jet (MV2c10, 77% working point)
- Exactly 1 photon, $p_T > 50$ GeV
- Photon isolation cuts: $\text{topo}E_T\text{cone}40<4$ GeV
- Z mass cut $|m_{l\gamma} - m_Z|V > 5$ GeV

Distributions of kinematic variables in the electron (muon) channels are shown in Figure 5.23 (5.24).

Channel:	e+jets	μ +jets
W+jets+ γ SF	PUT IN VALUE	PUT IN VALUE

TABLE 5.6. Fit result W+jets+ γ normalization scale factors including both statistical and systematic uncertainties

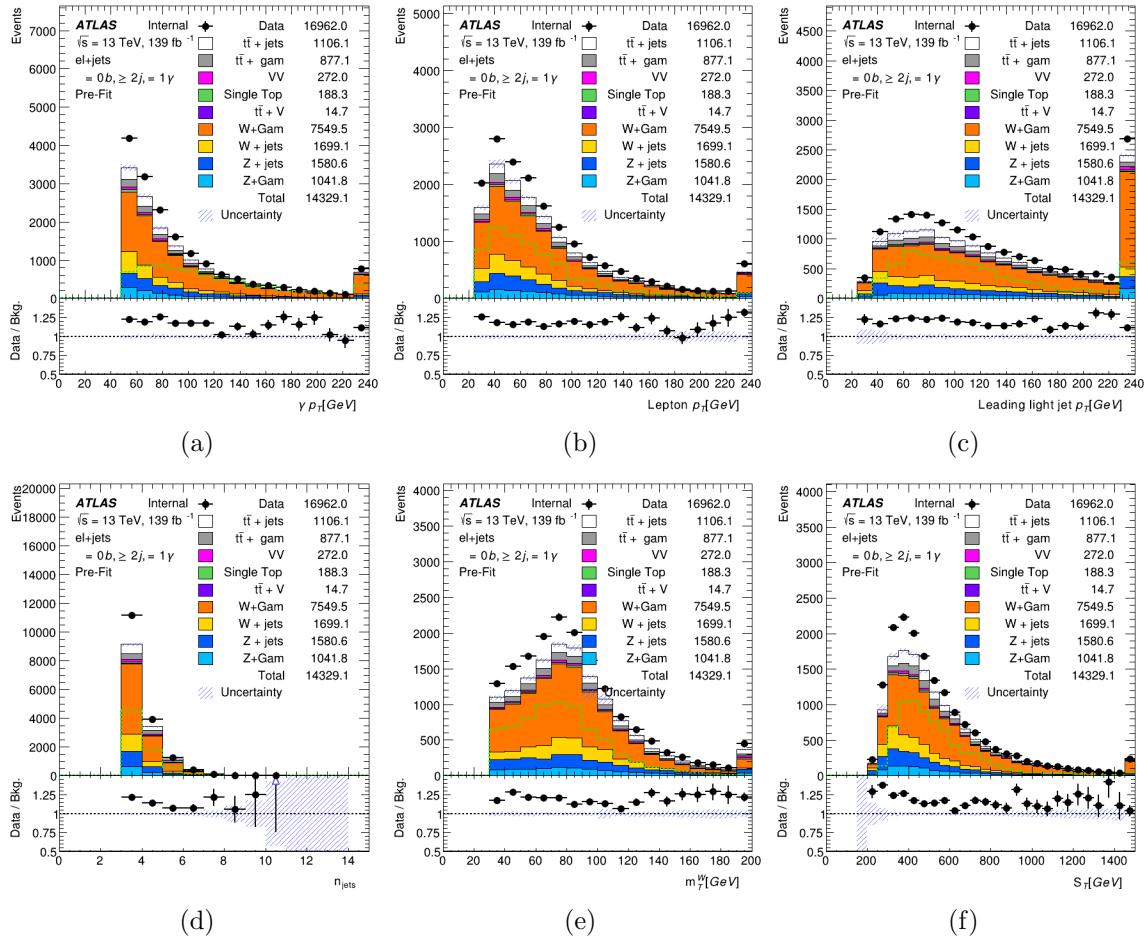


FIGURE 5.23. $W+jets+\gamma$ validation region plots for the electron channel. The FCNC signal sample is scaled to 1.5%.

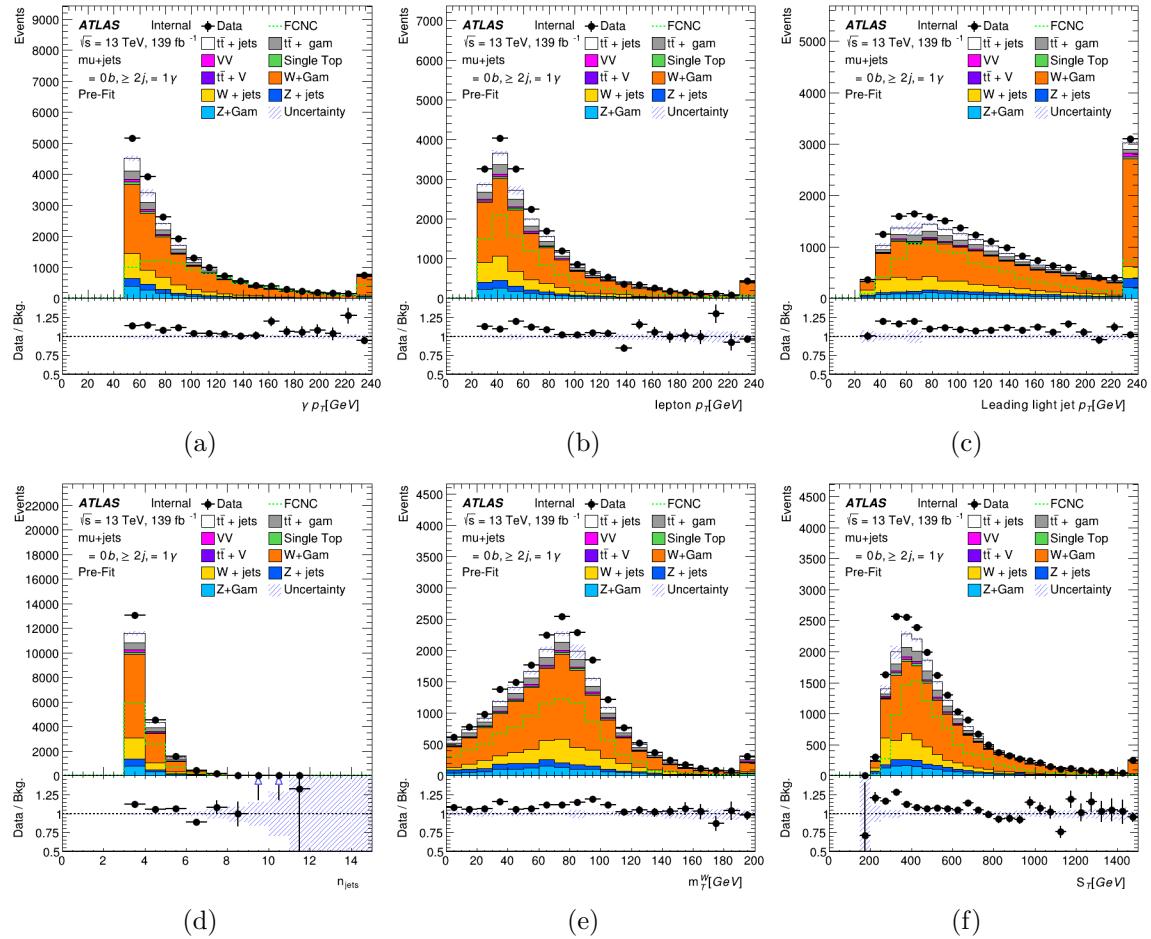


FIGURE 5.24. $W + jets + \gamma$ validation region plots for the muon channel. The FCNC signal sample is scaled to 1.5%.

5.6.2.2. $t\bar{t} + \gamma$ Control Region

Another validation region was created for the other largest photon enriched samples, $t\bar{t} + \gamma$. The normalization for the $t\bar{t} + \gamma$ validation region enters as a free parameter into the final fit, and the region selection is as follows:

- All of the Initial Event Selection as outlined in Section 5.4
- Exactly 1 lepton (electron or muon) $p_T > 25$ GeV
- At least 4 Jets ($p_T > 25$ GeV)
- $\cancel{E}_T > 30$ GeV and $m_T^W > 30$ GeV (for events with electrons)
- $\cancel{E}_T > 20$ GeV and $\cancel{E}_T + m_T^W > 60$ GeV (for events with muons)
- At least 1 b-tagged jet (MV2c10, 77% working point)
- Exactly 1 photon, $p_T > 50$ GeV
- Photon isolation cuts: $\text{topo}E_T\text{cone}40<4$ GeV
- Reverse Neural Network Cut: $\text{NNOutput}<0.93$ electron channel, $\text{NNOutput}<0.92$ muon channel

Distributions of kinematic variables in the electron (muon) channels are shown in Figure 5.25 (5.26).

Channel:	e+jets	μ +jets
$t\bar{t} + \gamma$ SF	PUT IN VALUE	PUT IN VALUE

TABLE 5.7. Fit result $t\bar{t} + \gamma$ normalization scale factors including both statistical and systematic uncertainties

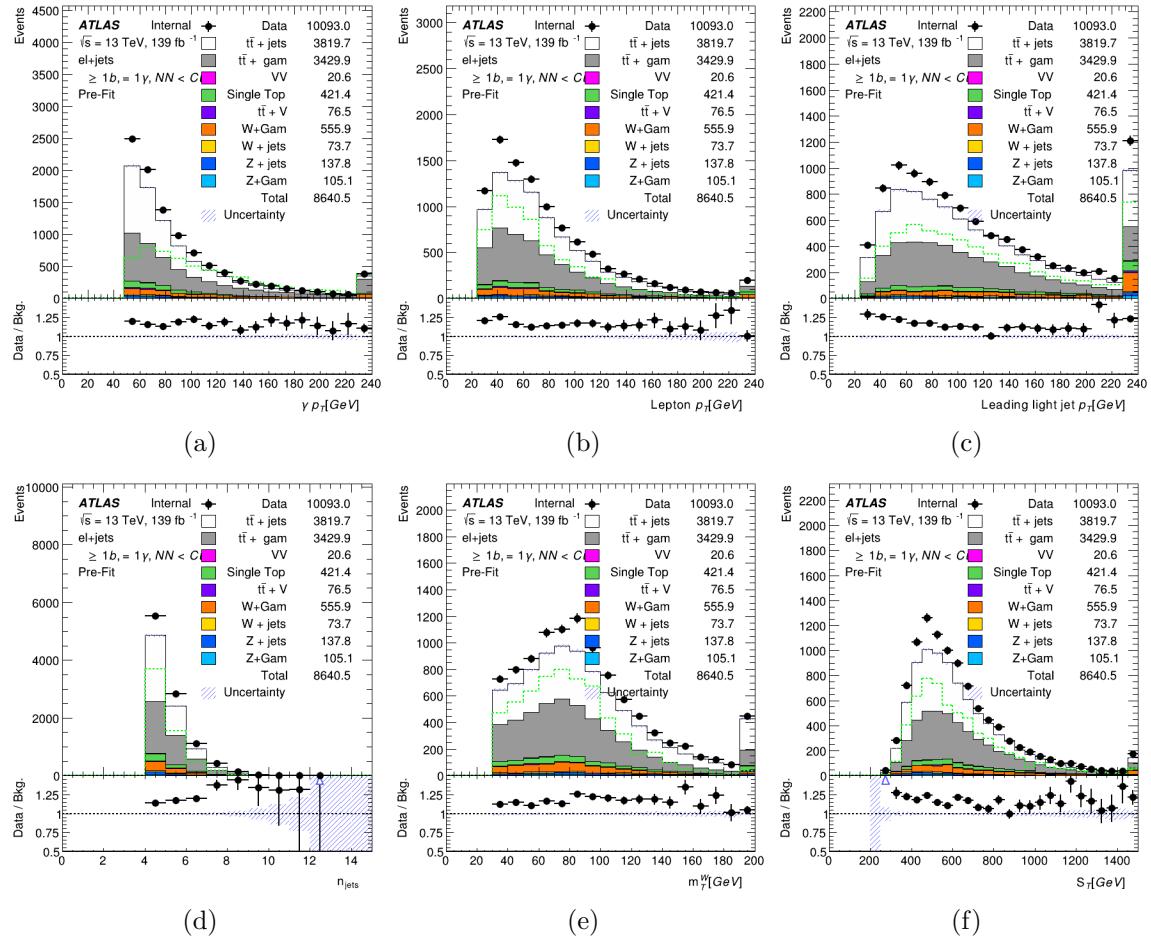


FIGURE 5.25. $t\bar{t}$ +jets+ γ validation region plots for the electron channel. The FCNC signal sample is scaled to 1.5%.

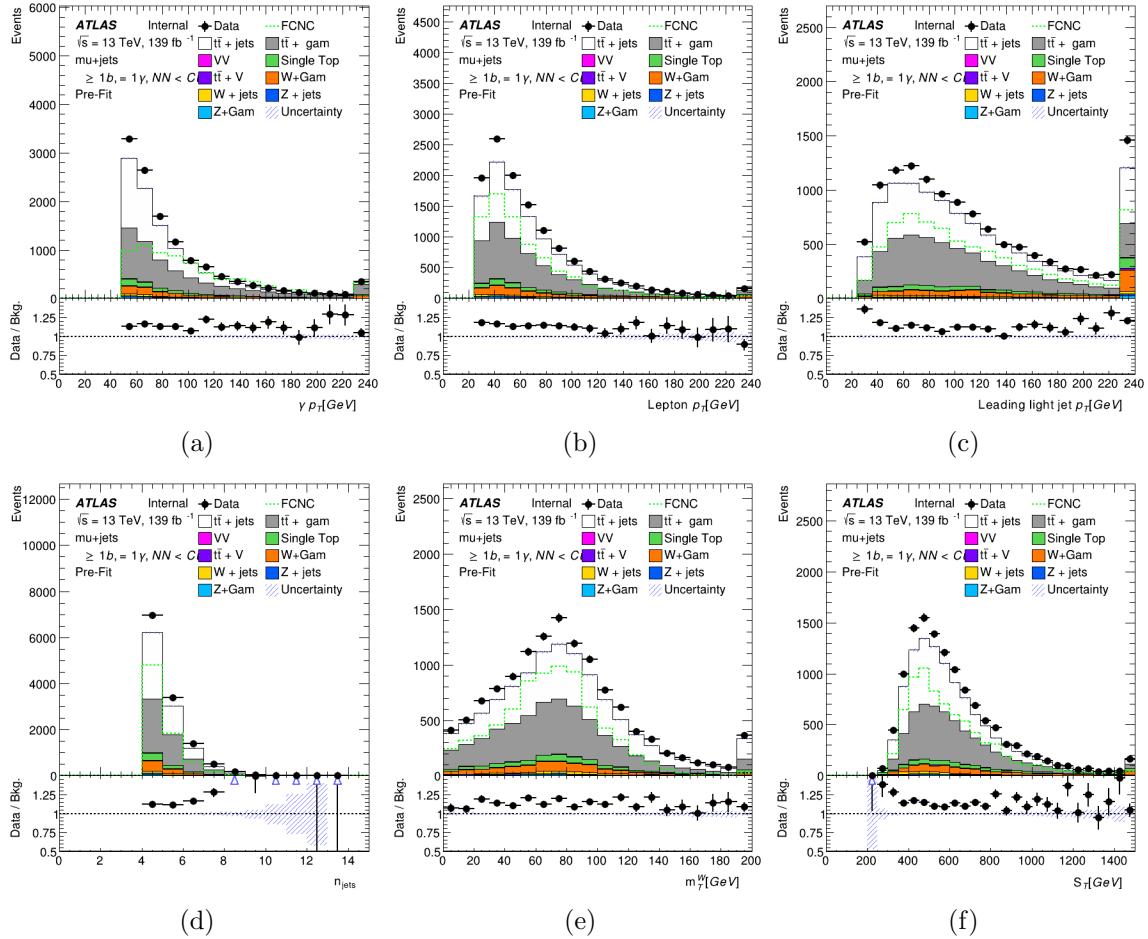


FIGURE 5.26. $t\bar{t}+\text{jets}+\gamma$ validation region plots for the muon channel. The FCNC signal sample is scaled to 1.5%.

5.7. Fake Rates

Photons and leptons can be faked by various other particles depending on how they interact within the detector. Section 5.7.2 discusses how jets faking leptons are accounted for, and Sections 5.7.1 and 5.7.3 discuss how electrons and jets can appear as photons and enter the signal region. Photon fake rates and scale factors are determined using data-driven techniques and truth information in the MC samples. Information in the truth record of the reconstructed photons is found using the **MCTruthClassifier** tool which performs a *Truth to Cluster* matching algorithm based on geometric separations, ΔR , between various truth-level physics objects.

A photon is considered an electron fake if the truth particle ID is equal to the PDG ID of an electron or if the truth particle ID is equal to the PDG ID of a photon but a truth electron is within a distance of $\Delta R < 0.05$. For the second case the photon is assumed to have been originating from the electron.

A photon is considered a hadronic fake if the truth photon originates from any hadron or when the truth particle is a hadron. These hadrons can be any meson or baryon within the initial hard interaction.

5.7.1. Electron → Photon Fakes

In multiple scenarios it is possible to reconstruct an electron incorrectly within the ATLAS detector i.e., if the track is unable to be associated to the shower in the electromagnetic calorimeter the object can be reconstructed as a photon instead of an electron. Additionally if an electron radiates all of its energy as a photon the object will be correctly reconstructed as a photon but it is not a prompt photon from the hard interaction. The second type of faked photon does not correspond to a genuine signal like photon as it originates from an electron. As other background events can

enter the signal region through these fake processes, understanding how often this happens is imperative. Modeling of the detector in MC is known to be inaccurate when regarding fakes. As such a data driven method has been used to calculate scale factors for events with MC photons that are not matched to truth photons in order to calculate appropriate scale factors for these MC events.

A tag-and-probe method is employed to determine the fake rate for $e \rightarrow \gamma$ in data and MC in a sample consisting mostly of $Z \rightarrow e^+e^-$ events. Two separate regions are created, one with two opposite sign electrons and the other with a single electron and a single photon. These events have similar cuts to the preselection cuts discussed in Section 5.5.

- All of the Initial Event Selection as outlined in Section 5.4
- At least 2 Jets ($p_T > 25$ GeV)
- At least 1 b-tagged jet (MV2c10, 77% working point)
- $\cancel{E}_T > 25$ GeV
- At least 1 electron $p_T > 25$ GeV
- Further, a $Z \rightarrow ee$ region is created with:
 - * Exactly 2 opposite site electrons > 25 GeV
 - * No good photons > 20 GeV
- In addition, a $Z \rightarrow e\gamma$ region is created with:
 - * Exactly 1 electron > 25 GeV
 - * Exactly 1 photon > 20 GeV

Distributions of the p_T spectra of tag electrons and the probe electrons and photons (both converted and unconverted) in data and MC are shown in Figures 5.27 and 5.28.

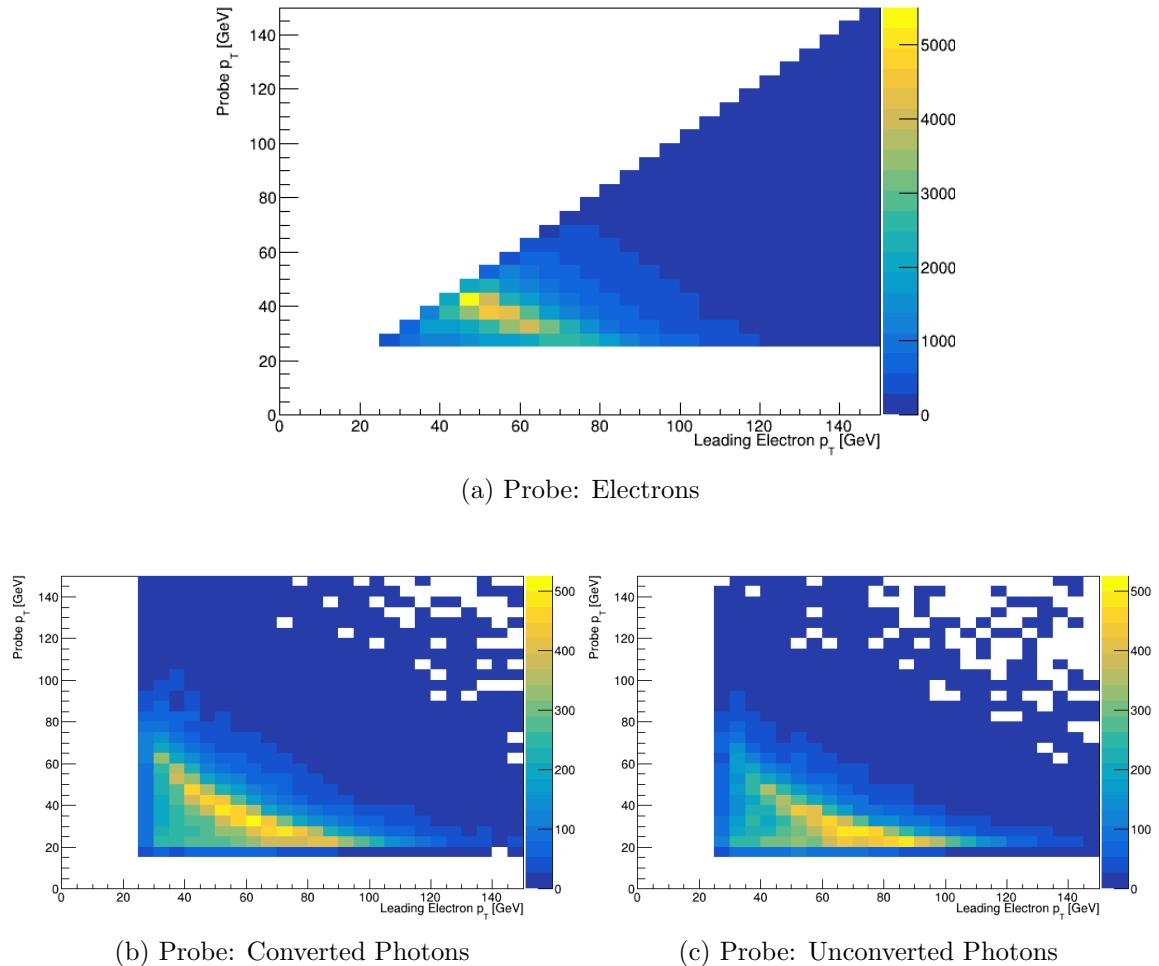
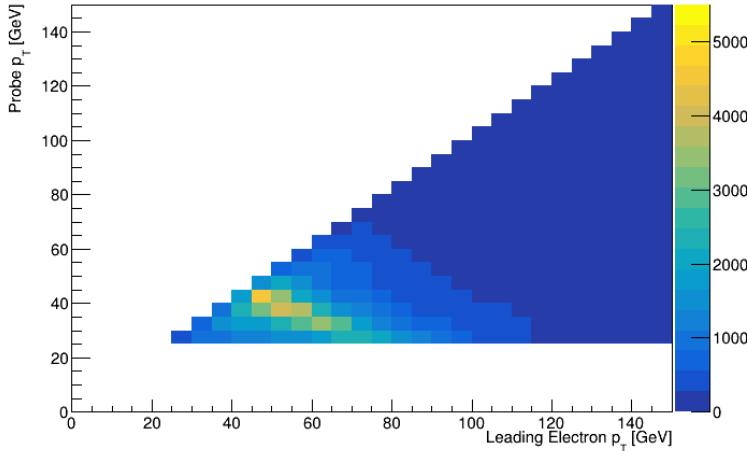


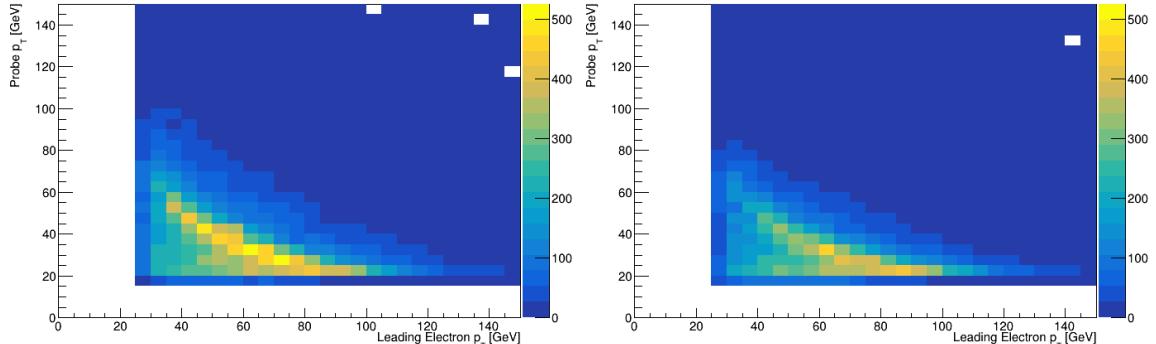
FIGURE 5.27. Probe p_T vs leading electron p_T in $Z \rightarrow e^+e^-$ Data Events

A fake rate can be inferred from MC defined as:

$$\text{FR}_{\text{MC}}^{\text{e-fake}} = \frac{N_{e,\gamma}}{N_{e,e}}$$



(a) Probe: Electrons



(b) Probe: Converted Photons

(c) Probe: Unconverted Photons

FIGURE 5.28. Probe p_T vs leading electron p_T in $Z \rightarrow e^+e^-$ MC Events

where $N_{e,\gamma}$ ($N_{e,e}$) is the number of events observed in the $Z \rightarrow e\gamma$ ($Z \rightarrow ee$) regions. The MC subindex means that this is the fake rate derived using MC events. A data driven fake rate is also defined in each region by subtracting the backgrounds that do not contribute to the Z-boson invariant mass peak using a sideband fit to the $m(l, \gamma)$ distribution defined as:

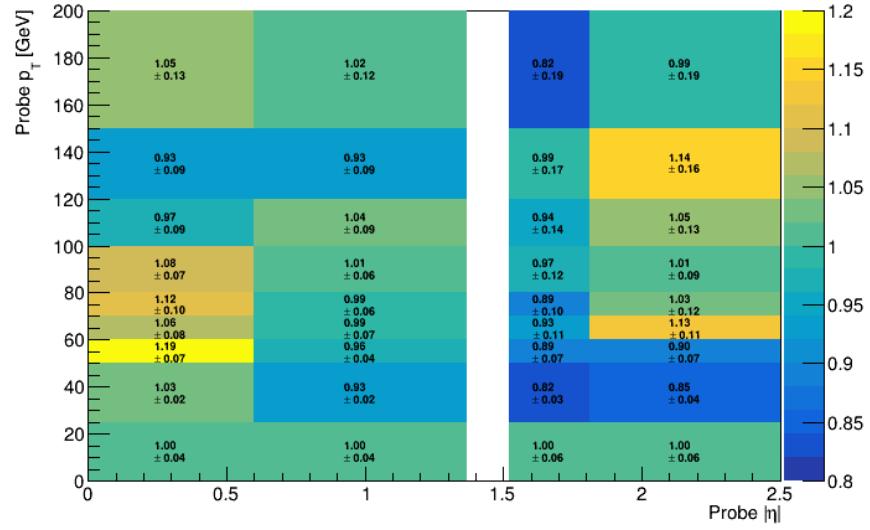
$$\text{FR}_{\text{d.d.}}^{\text{e-fake}} = \frac{N_{e,\gamma}^{\text{data}} - N_{e,\gamma}^{\text{non-Z}}}{N_{e,e}^{\text{data}} - N_{e,e}^{\text{non-Z}}}$$

where the tails of the Z-boson peak are included as $N_{e, \gamma/e}^{\text{non-Z}}$. Combining the fake rates from the data driven and MC method, a scale factor can be applied as a correction to the samples where the truth MC photon comes from an electron is calculated by

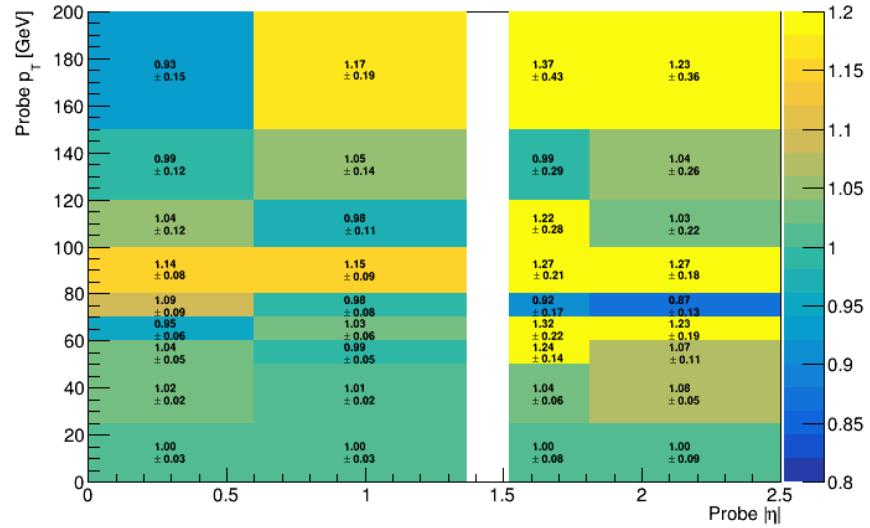
$$\text{SF}_{\text{FR}}^{\text{e-fake}} = \frac{\text{FR}_{\text{d.d.}}^{\text{e-fake}}}{\text{FR}_{\text{MC}}^{\text{e-fake}}}.$$

To give a sense of size the overall scale factor is derived to be $0.97 \pm 0.01(\text{stat})$. In practice the scale factor is derived in bins of probe p_T and η as well as converted and unconverted photon type.

As shown in Figure 5.29, the 2D scale factors generally agree with the overall scale factor derived for all photons and $\eta - \phi$ bins but additional correction factors are calculated based on the conversion type and photon kinematic information. Systematic variations are taken into account for this scale factor by considering larger regions around the Z invariant mass peak. The nominal sample value is calculated with a window of width 10 GeV (5 GeV on either side of the Z mass), varying this to 5 GeV, 15 GeV, and 20 GeV and recalculating the values in each bin a systematic variation is calculated based on the deviations using these larger more background enriched windows. Appendix D details the scale factors calculated with these invariant mass windows. A conservative systematic uncertainty estimate of 5% is set on events with this scale factor based on the median deviation from nominal for each bin using all of these invariant mass windows.



(a) Converted Photons



(b) Unconverted Photons

FIGURE 5.29. 2-Dimensional scale factors derived using the $Z \rightarrow e^+e^-$ events for unconverted and converted photon types shown with statistical uncertainties

5.7.2. Jet \rightarrow Lepton Fakes

The final state is determined by the decay of the W boson throughout this search, resulting in a charged lepton and a neutrino. However, this can be faked by

the QCD multi-jet background if a jet is misidentified as a lepton. Non-prompt and fake leptons can pass the event selection and come from predominantly QCD multi-jet processes with an associated photon. These leptons can arise from semi-leptonic decay of bottom and charm quarks. The electron final state has additional contributions from photon conversions in the electromagnetic calorimeter while muons can be faked from highly energetic hadrons that punch-through to the hadronic calorimeter before beginning their shower leaving only tracks similar to muons throughout the electromagnetic calorimeter.

The estimation for the lepton fake background in the signal region channels follows the fully data-driven Matrix Method[84] approach which compares the number of leptons with looser lepton identification and isolation requirements than those used in the search. The baseline preselection cuts are also influenced by the use of the Matrix Method in that the calculations done in [84] in the single lepton channel have the following cuts, as seen in the selection of various regions:

- $\cancel{E}_T > 30 \text{ GeV}$ and $m_T^W > 30 \text{ GeV}$ (for events with electrons)
- $\cancel{E}_T > 20 \text{ GeV}$ and $\cancel{E}_T + m_T^W > 60 \text{ GeV}$ (for events with muons)

Electrons in the loose sample are required to have *Medium* identification criteria but no isolation requirement is applied. Similarly for muons the identification criteria is *Medium* and the requirement on the isolation is dropped. The Matrix Method works under the assumption that the tight sample will contain mostly real leptons whereas the loose sample will be enriched with fake lepton events. Therefore it follows that the number of leptons in the loose (tight) sample is simply a combination of the number of real and fake leptons in those samples:

$$N^{\text{loose}} = N_{\text{real}}^{\text{loose}} + N_{\text{fake}}^{\text{loose}}$$

$$N^{\text{tight}} = N_{\text{real}}^{\text{tight}} + N_{\text{fake}}^{\text{tight}}$$

Then ϵ_{real} is defined as the probability of a real lepton in the loose sample to pass the tight selection and ϵ_{fake} as the probability of a fake lepton in the loose sample to pass the tight selection. Then the number of events with fake leptons in the tight sample can be estimated by applying the following equation:

$$N_{\text{fake}}^{\text{tight}} = \frac{\epsilon_{\text{fake}}}{\epsilon_{\text{real}} - \epsilon_{\text{fake}}} (\epsilon_{\text{real}} N^{\text{loose}} - N^{\text{tight}})$$

The real efficiencies, ϵ_{real} , are estimated using the tag-and-probe method in $Z \rightarrow e^+e^-$ and $Z \rightarrow \mu^+\mu^-$, regions and the fake efficiencies, ϵ_{fake} , are estimated in data samples that are dominated by non-prompt and fake leptons. The efficiency measurements are provided centrally in ATLAS and are explained in much more detail in [84]. Individual weights are applied to the event identification calculated as a function of these efficiencies which follow:

$$w_i = \frac{\epsilon_{\text{fake}}}{\epsilon_{\text{real}} - \epsilon_{\text{fake}}} (\epsilon_{\text{real}} - \delta_i)$$

where δ_i is 1 if the event passes the tight selection and 0 otherwise. These event weights are applied to the loose data sample to estimate the fake contributions from jets faking leptons.

5.7.3. Jet \rightarrow Photon Fakes

The ABCD method is used to determine the number of hadrons that fake photons in the analysis. As was shown in Section 5.6.1 the multi-jet background is poorly modeled, particularly at the high energies of Run-2. An accurate estimation of this

background that does not rely on the inaccurate MC modelling is done created a data driven technique. Scale factors to match MC predictions to data are derived in control regions with enriched hadronic fake contributions and applied in the signal region. The majority of hadronic fake photon events come from $t\bar{t}$ events where a final state jet radiates a non-prompt photon. Additional small contributions come from similar topologies from W+jets and single top events.

The ABCD method requires four orthogonal and non-correlated regions of phase space. Events that pass the preselection cuts and the additional requirements for the final signal region cuts (photon p_T , b-jet multiplicity, and photon multiplicity) but not the final neural network cut for the signal region can be used to craft these additional regions. The neural network cut is not used here to preserve statistics in the regions and should not effect the scale factor derivation as the regions are expected to be independent. The photon isolation, topo E_T cone40, and the transverse W mass, m_T^W are chosen. The electron (muon) region requirement of $m_T^W > 30$ GeV ($\cancel{E}_T + m_T^W > 60$ GeV) from the Matrix Method can be reversed to create the ABCD regions used here. A pictoral representation of the regions is shown in Figure 5.30. If these variables are uncorrelated then it follows that the ratio of isolated vs. non-isolated photons should not change depending on the kinematic variables unrelated to the photon. The hadronic fake rate is enhanced when the photons are classified as converted photons in the detector as seen in Figure 5.31.

$$\frac{N_D^{\text{h-fake}}}{N_C^{\text{h-fake}}} = \frac{N_A^{\text{h-fake}}}{N_B^{\text{h-fake}}} \text{ and } \frac{N_D^{\text{h-fake}}}{N_A^{\text{h-fake}}} = \frac{N_C^{\text{h-fake}}}{N_B^{\text{h-fake}}}$$

To quantify the correlation between these variables, a double ratio θ_{MC} is defined as a correction factor to ensure closure in accounting for imperfect simulation of fake

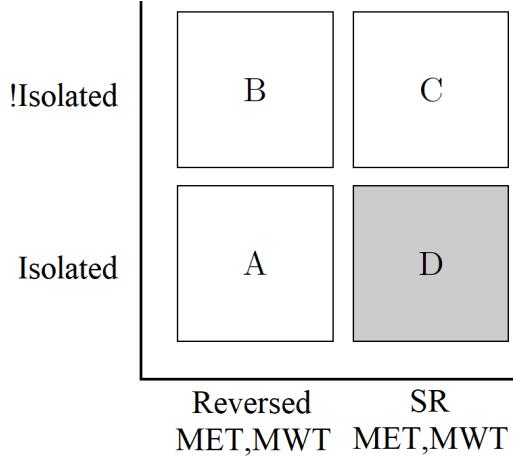


FIGURE 5.30. Pictoral representation of the regions used for the ABCD method in this search

photons in MC events. This correction factor is defined as:

$$\theta_{\text{MC}} = \frac{N_{\text{D},\text{MC}}^{\text{h-fake}} / N_{\text{C},\text{MC}}^{\text{h-fake}}}{N_{\text{A},\text{MC}}^{\text{h-fake}} / N_{\text{B},\text{MC}}^{\text{h-fake}}}$$

The number of events in MC samples where the truth photon comes from a hadron are listed in Table 5.8. A closure test is performed on these yields for the correction factor θ_{MC} as shown in Table 5.9. A systematic uncertainty of 50% is set to this correction factor to account for further imperfect MC modelling.

Using these the hadron fake contribution in the signal region (D) is estimated with the following:

$$N_{\text{D},\text{est.}}^{\text{h-fake}} = \frac{N_{\text{A},\text{data}}^{\text{h-fake}} \times N_{\text{C},\text{data}}^{\text{h-fake}}}{N_{\text{B},\text{data}}^{\text{h-fake}}} \times \theta_{\text{MC}}$$

Where the number of estimated hadronic fake events $N_{\text{A/B/C,data}}^{\text{h-fake}}$ is defined by subtracting the sum of all background events from the total number of data events in the respective regions. A scale factor for hadronic fakes can then be calculated with the esimates for the number of fakes in MC and data in the signal region using:

Channel:	e+jets	μ +jets
$N(A)_{\text{MC}}^{\text{h-fake}}$ Converted	481 ± 11	124 ± 7
$N(B)_{\text{MC}}^{\text{h-fake}}$ Converted	4921 ± 37	1126 ± 19
$N(C)_{\text{MC}}^{\text{h-fake}}$ Converted	24606 ± 99	26575 ± 108
$N(D)_{\text{MC}}^{\text{h-fake}}$ Converted	2631 ± 31	2805 ± 35
$N(A)_{\text{MC}}^{\text{h-fake}}$ Unconverted	414 ± 12	112 ± 6
$N(B)_{\text{MC}}^{\text{h-fake}}$ Unconverted	2236 ± 23	551 ± 14
$N(C)_{\text{MC}}^{\text{h-fake}}$ Unconverted	11661 ± 136	12338 ± 61
$N(D)_{\text{MC}}^{\text{h-fake}}$ Unconverted	2284 ± 38	2419 ± 41

TABLE 5.8. Monte Carlo hadronic photon origin event yields in each region for converted and unconverted photons, including statistical errors

Channel:	$\frac{N_{\text{MC}}^{\text{h-fake}}(A)}{N_{\text{MC}}^{\text{h-fake}}(B)} [\%]$	$\frac{N_{\text{MC}}^{\text{h-fake}}(D)}{N_{\text{MC}}^{\text{h-fake}}(C)} [\%]$	θ_{MC}
Electron Channel, Converted	9.8 ± 0.2	10.7 ± 0.1	1.09 ± 0.02
Electron Channel, Unconverted	18.5 ± 0.6	19.6 ± 0.4	1.06 ± 0.04
Muon Channel, Converted	11.0 ± 0.6	10.6 ± 0.1	0.96 ± 0.06
Muon Channel, Unconverted	20.3 ± 1.2	19.5 ± 0.3	0.96 ± 0.06

TABLE 5.9. Closure test results and the resulting correction factor values, θ_{MC} for hadronic fake estimates including statistical errors

$$\text{SF}^{\text{h-fake}} = \frac{N_{\text{D,est.}}^{\text{h-fake}}}{N_{\text{D,MC}}^{\text{h-fake}}}$$

This scale factor is calculated in terms of converted and unconverted photon types and is then applied to events where the MC photon comes from a truth hadron.

The results shown in Table 5.10 are consistent with Figure 5.31. Shown in Figure 5.31 the separation of photons leads to smaller scale factors as the hadronic fakes are pulled out. However, for unconverted photons the scale factors are similar as the bulk of the photons are in region D. In addition to the systematic uncertainty propagated from θ_{MC} another estimation was done using a direct comparison of the scale factors

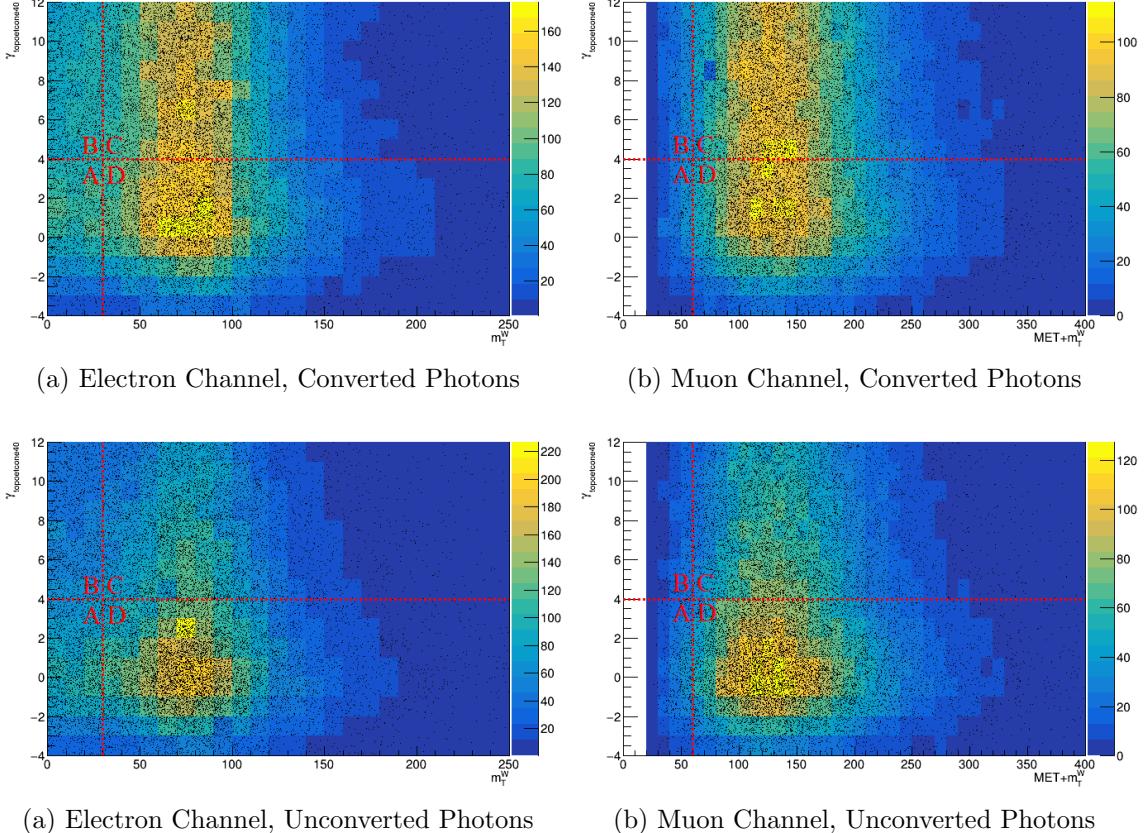


FIGURE 5.31. Regions used for the ABCD method shown for both converted and unconverted photon types. MC events are shown in the gradient histogram and data events are shown with the black points. All final signal region cuts are included except the neural network cut to give enough statistics in the regions for the ABCD method to be a reasonable estimate of the fake rate scale factor.

in region D for events by reversing the final neural network cut. This allows the estimation to be completely blinded while still directly computing the scale factors.

Given that the amount of preselection events expected to pass the final event selection is very small (on the order of 1%), removing the few events that could end up in the signal region should result in sufficiently similar scale factors. These scale factors are calculated and an additional systematic effect is estimated by comparing

Channel:	Converted	Unconverted
Electron Channel	$1.28 \pm 0.16 \pm 0.30$	$1.99 \pm 0.21 \pm 0.69$
Muon Channel	$1.23 \pm 0.30 \pm 0.41$	$2.15 \pm 0.46 \pm 0.80$

TABLE 5.10. Hadron scale factors determined in data for both channels and both photon types using the ABCD method. Statistical and systematic uncertainties are shown.

the two sets of scale factors. The maximal value of the deviations is taken to be a further source of systematic uncertainty.

5.8. Signal Region

After the calculation and application of various scale factors kinematic and event-level distributions with the final selection cuts are shown in this section. The signal MC is scaled to $\text{BR}(t \rightarrow q\gamma) = 0.01\%$. Figures 5.32 and 5.33 show the distributions in the electron channel while Figures 5.34 and 5.35 show the muon channel.

Sample	Events e+jets	Events μ +jets
FCNC, signal	VALUE ± err	VALUE ± err
$t\bar{t}$	VALUE ± err	VALUE ± err
$t\bar{t} + \gamma$	VALUE ± err	VALUE ± err
W+jets	VALUE ± err	VALUE ± err
W+jets+ γ	VALUE ± err	VALUE ± err
Z+jets	VALUE ± err	VALUE ± err
Z+jets+ γ	VALUE ± err	VALUE ± err
Single top	VALUE ± err	VALUE ± err
Diboson (VV)	VALUE ± err	VALUE ± err
$t\bar{t} + V$	VALUE ± err	VALUE ± err
Total MC	VALUE ± err	VALUE ± err
Data	VALUE ± err	VALUE ± err

TABLE 5.11. Signal region event yields after all scale factors and cuts are applied

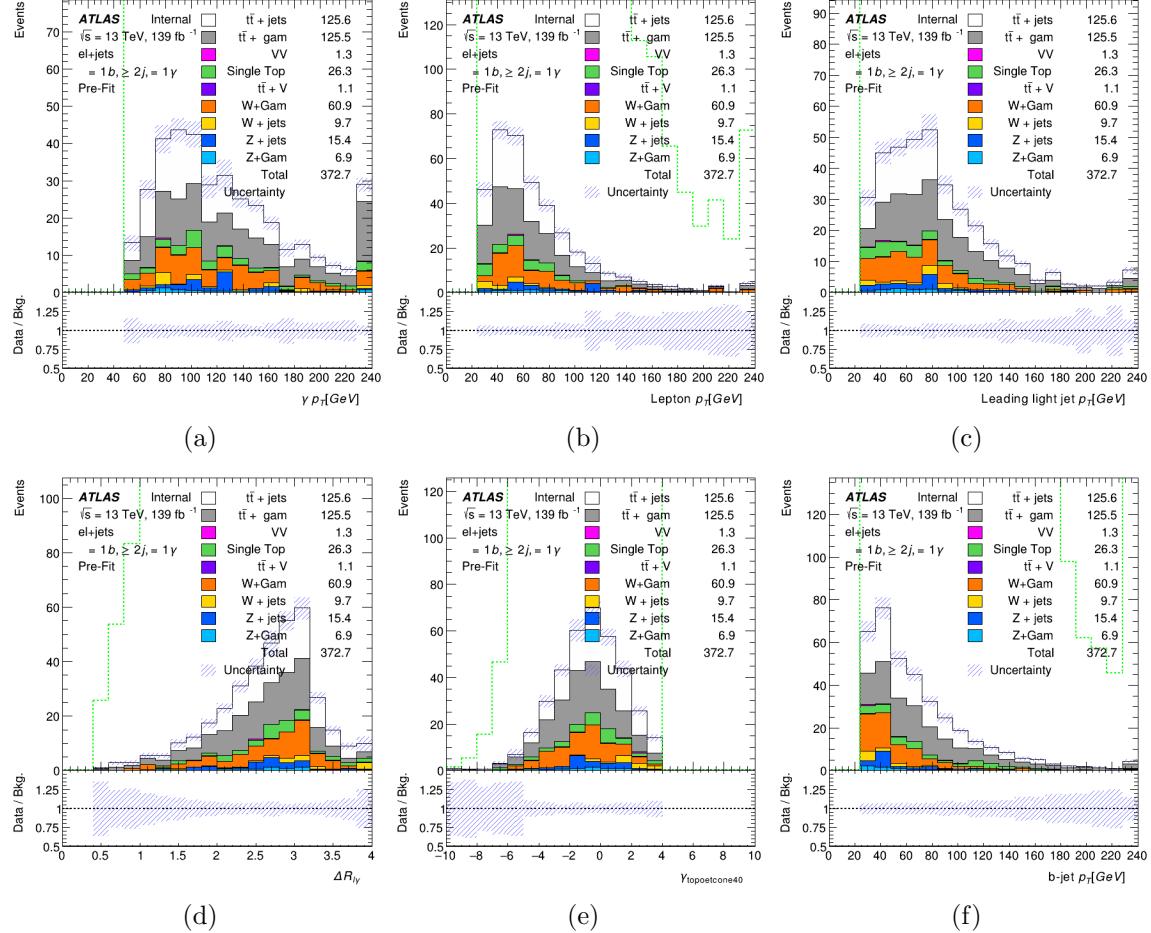


FIGURE 5.32. Photon p_T (a), lepton p_T (b), leading light jet p_T (c), $\Delta R_{l\gamma}$ (d), $\gamma_{topoetcone40}$ (e), and b-jet p_T (f) distributions in the final signal region with scale factors applied for the electron channel

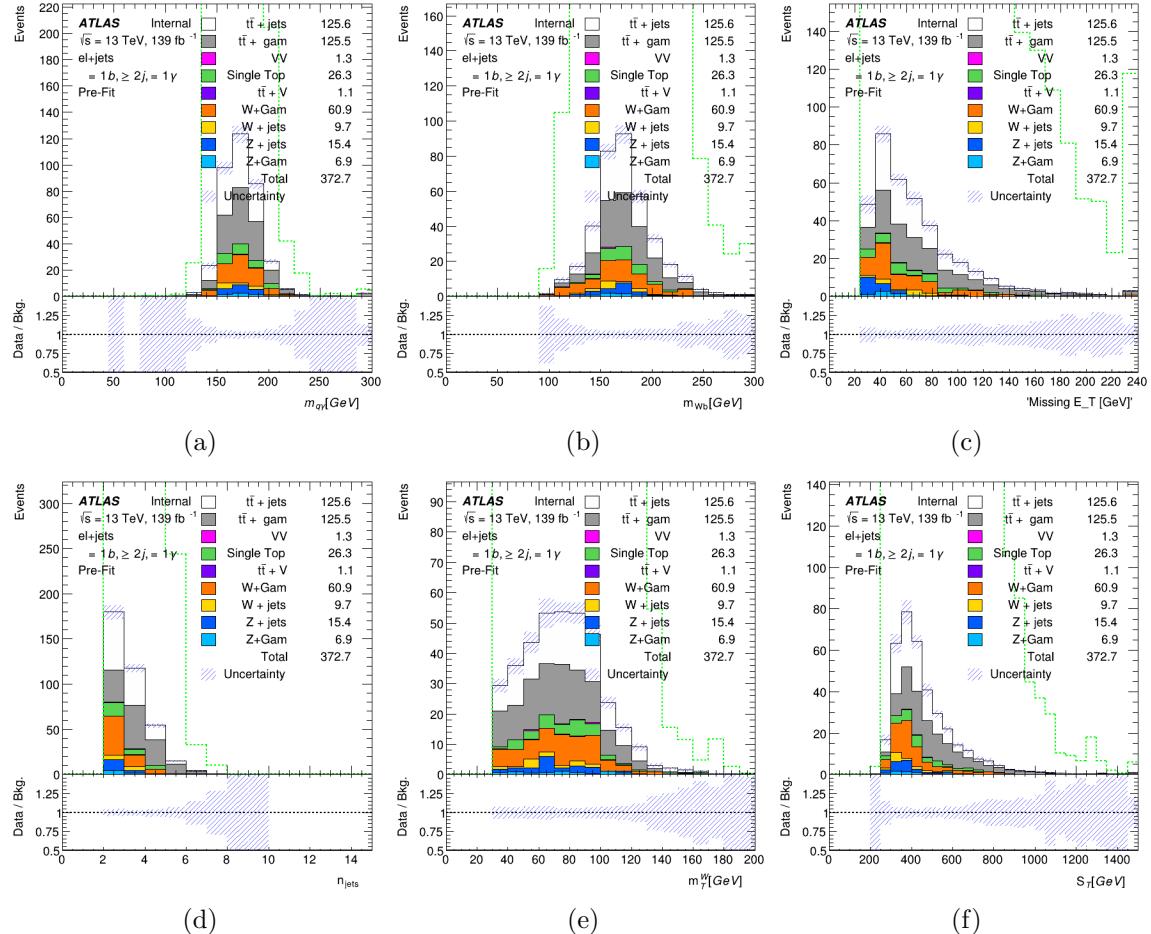


FIGURE 5.33. FCNC top candidate mass (a), Standard Model top candidate mass (b), \cancel{E}_T (c), N_{jets} (d), m_T^W (e), and S_T (f) distributions in the final signal region with scale factors applied for the electron channel

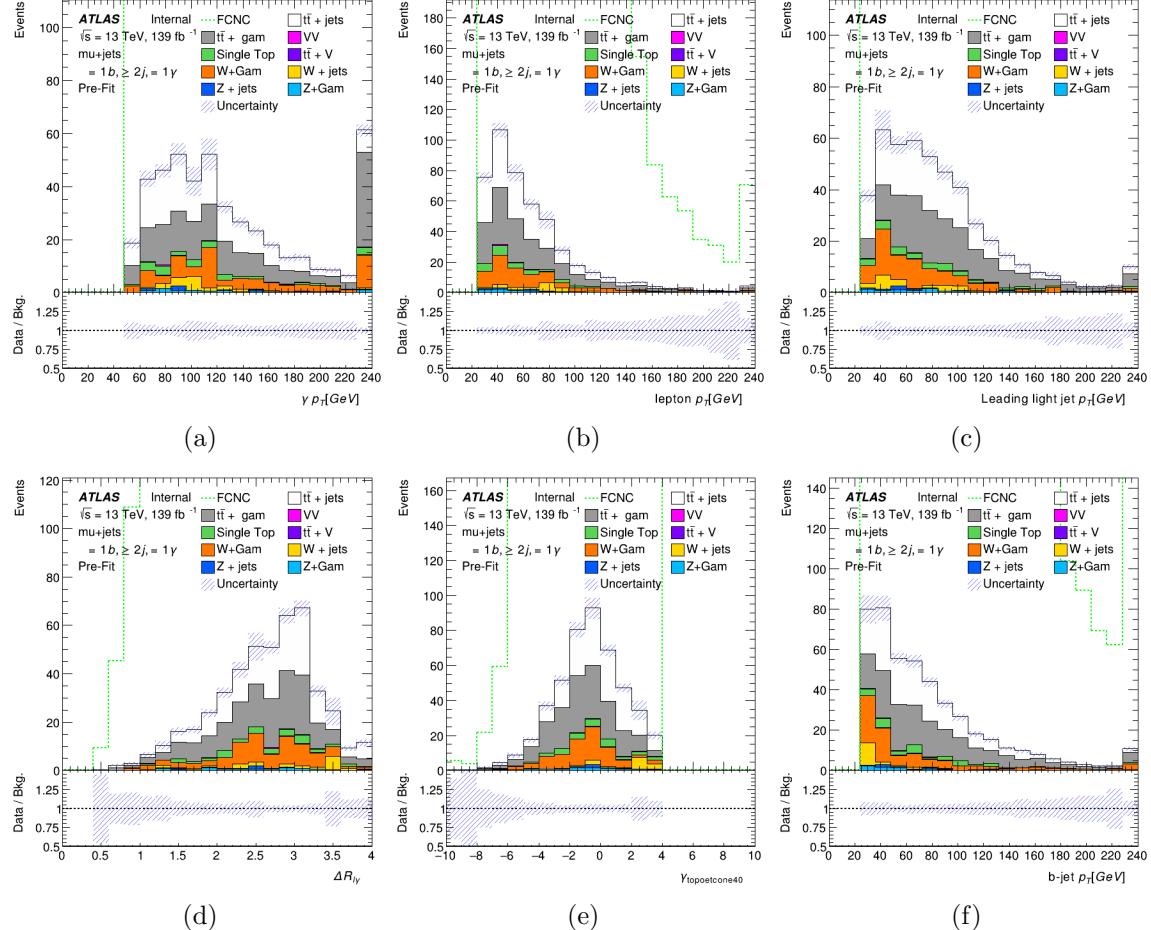


FIGURE 5.34. Photon p_T (a), lepton p_T (b), leading light jet p_T (c), $\Delta R_{l\gamma}$ (d), $\gamma_{\text{topoetcone40}}$ (e), and b-jet p_T (f) distributions in the final signal region with scale factors applied for the muon channel

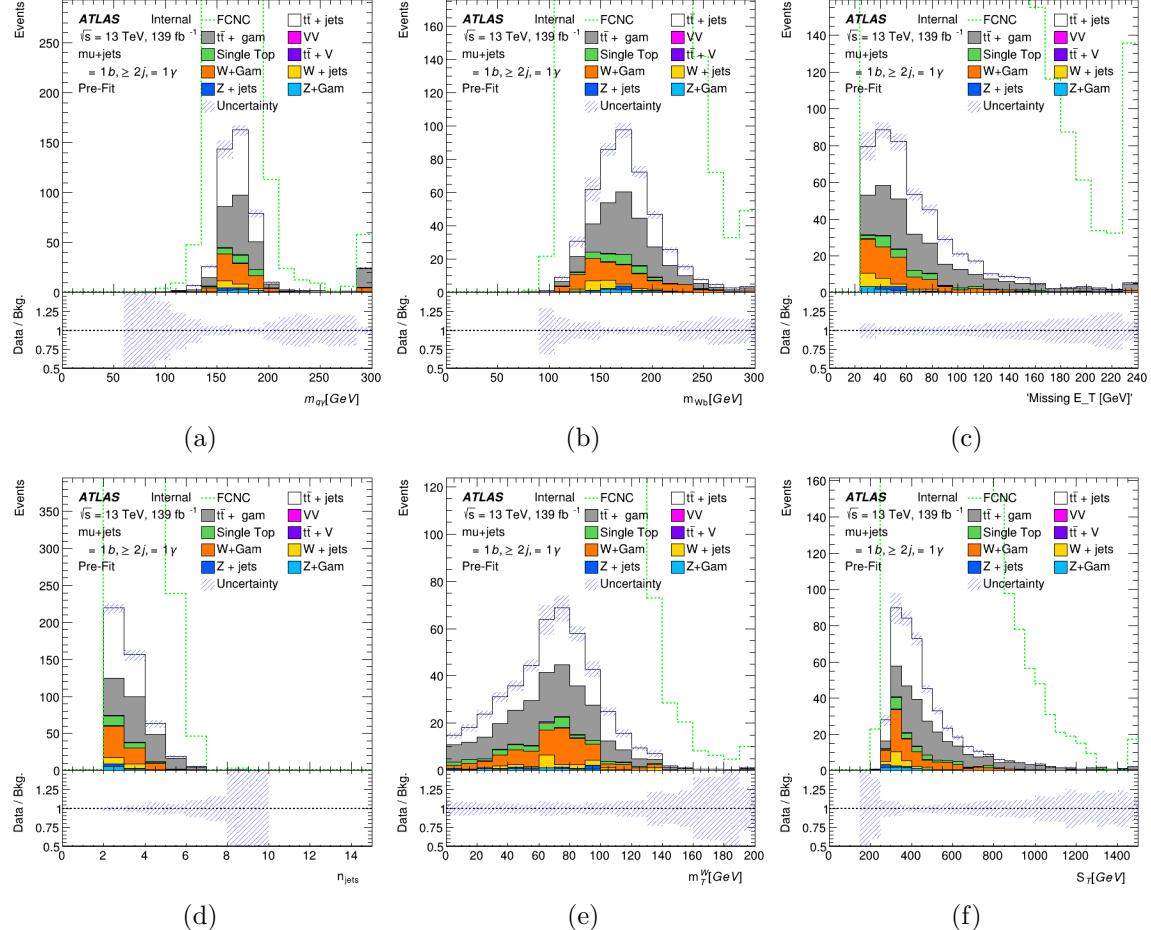


FIGURE 5.35. FCNC top candidate mass (a), Standard Model top candidate mass (b), \cancel{E}_T (c), N_{jets} (d), m_T^W (e), and S_T (f) distributions in the final signal region with scale factors applied for the muon channel

CHAPTER VI

ANALYSIS AND RESULTS

This search is to attempt to observe the FCNC decay of a top quark $t \rightarrow q\gamma$ or to set an upper limit on the branching ratio of this decay process if no observation occurs. As there was no significant excess of events in the collected data an upper limit on the branching ratio is set. This chapter will discuss the methods used to set the upper limit on the branching ratio as well as discussion of the systematic uncertainties within the experiment.

6.1. Systematic Uncertainties

Various sources of uncertainty are considered for any analysis in high energy physics. Statistical and systematic uncertainties are studied and propagated through to the final upper limit set on the branching ratio. The statistical errors are related to the amount of data collected or MC events that are available and created for analysis. More data events or more MC simulated events lead to a smaller statistical error due to random fluctuations of stochastic processes. In this section the systematic errors will be discussed. Systematic errors are derived from limitations from detector construction or a lack of complete understanding of physics objects or algorithms used in reconstruction. Some sources of systematic uncertainties are provided centrally within ATLAS from studies carried out by other analysis teams such as errors on the jet energy resolution (JER) or luminosity amongst others. Other sources are particular to the analysis such as errors propagating from deriving data-driven backgrounds as discussed in the previous chapter. This section will briefly introduce and discuss the theoretical systematic uncertainties considered in this

analysis including the modeling and experimental uncertainties as well as a discussion on systematic smoothing, symmetrization, and pruning used in the final fit.

6.1.1. Theoretical uncertainties

Cross Section: Various cross sections for MC samples are separately varied up and down by one standard deviation for each process: $t\bar{t}$ ($\pm 5.6\%$ [85]), $t\bar{t}+\gamma$ ($\pm 8.0\%$ [86]), single top(t-channel top $^{+4.0\%}_{-3.4\%}$, t-channel anti-top $^{+5.0\%}_{-4.5\%}$, s-channel top $^{+3.6\%}_{-3.1\%}$, s-channel anti-top $^{+4.8\%}_{-4.3\%}$, tW-channel $\pm 5.3\%$ [87]), V+jets ($\pm 5\%$ [88]), and diboson VV ($\pm 6\%$ [88]). Uncertainties on the prompt photon contribution from $W/t\bar{t}+\text{jets}+\gamma$ enter the fit as other nuisance parameters as the normalization is allowed to vary in the fit.

Renormalization and Factorization Scale: The effect of the choice of renormalization and factorization scales (μ_r and μ_f) is estimated varying them independently or simultaneously up and down by a factor of 2 compared to the nominal value. This is done using event weights and the maximum and deviations from the nominal are taken to be the up and down variations.

PDF Uncertainty: The PDF uncertainty is estimated using the PDF4LHC15 error set which contains 30 eigen variations which enter the fit as individual nuisance parameters.

Initial and Final State Radiation: The effects of ISR are estimated by decreasing the total QCD radiation activity in the events. These variations are available for single top and $t\bar{t}$ processes. For single top processes the A14 tune is varied up (down) using the event weight Var3cUp (Var3cDown) and dividing (multiplying) μ_r and μ_f by a factor of 2. For $t\bar{t}$ processes a sample with

parameter $h_{\text{damp}} = 3m_{\text{top}}$ is used in the same manner varying Var3cUp while the down variation uses event weights in the nominal sample. FSR uncertainties are estimated varying parameters using the A14 tune by varying event weight (Var2) up and down.

Normalization of Prompt Photon Contribution: The normalization of the prompt photon contributions that stem from the $W + \text{jets} + \gamma$ and $t\bar{t} + \text{jets} + \gamma$, the major backgrounds, enter the fit as free floating nuisance parameters.

$t\bar{t}$ Matrix Element and Shower Generator: To estimate the uncertainty based on our choice of MC generator and showering algorithm used to create the nominal sample (POWHEG-Box + PYTHIA) is replaced by samples generated using MADGRAPH5_aMC@NLO + PYTHIA and POWHEG-Box + HERWIG.

6.1.2. Experimental Uncertainties

Luminosity: The uncertainty in the combined 2015–2018 integrated luminosity is 1.7 [89], obtained using the LUCID-2 detector [90] for the primary luminosity measurements using x-y beam separation scans.

Pile-up: Events are re-weighted in the MC samples to match the number of interactions per bunch crossing in data. Systematic uncertainties for pile-up are evaluated by scaling these distributions up and down.

Lepton Identification and Trigger: Lepton efficiencies contain the trigger, reconstruction, identification, and isolation efficiencies. Scale factors are used to correct any deviation from data in the MC simulation. Correction factors are derived from $Z \rightarrow ll$ and $J/\psi \rightarrow ll$ decays for electrons [64] and muons [67].

Lepton Energy Scale and Resolution: The lepton energy (momentum) is calibrated using MC -based techniques. Correction factors derived from dileptonic Z boson decay channels are applied to account for detector calibration mismodeling. For electrons, the energy scale and resolution are calculated together with photons as the EGamma energy scale and resolution. Various efficiencies are derived for the EGamma scale and resolution, e.g., varying the amount of material in front of the calorimeters, using various MC generators, and considering varying background fits [64, 68].

Photon Efficiency: Scale factors for isolation are measured as described in Ref. [91] while additional scale factors for the photon ID efficiency are derived for photons using enriched $Z \rightarrow ee$ events where the similarity between electrons and photons in our detector is exploited using the matrix method as described in Section 5.7.1. These sets of scale factors are combined into a single set that is applied to MC simulation based on photon information.

Photon Energy Scale and Resolution: Photon energy scale and resolution are calculated together with electron energy scale and resolution.

Jet Energy Scale: The jet energy scale (JES) and its uncertainty are derived using combinations of measurements in both simulation and data. The *CategoryReduction* reduction parameter set (30 nuisance parameters) are varied up and down and used for categories such as *in-situ* jet energy corrections, flavor composition and response, η inter-calibration, b-jet energy scale, and pile-up.

Jet Energy Resolution: The jet energy resolution (JER) is measured independently for data and MC using two techniques [92] which results in 8 separate nuisance

parameters. These nuisance parameters are one sided, as such the uncertainty is symmetrized.

Jet Vertex Tagging: The cut on the jet vertex tagging (JVT) discriminant is varied up and down [72] and the certainty on the JVT scale factor is calculated which is then propagated through this analysis.

b-tagging: Efficiencies for the various b-tagging working points are measured in data and scale factors are derived for simulation depending on the jet flavor. The uncertainties on these scale factors are provided by the Flavor Tagging group for b, c, and light jets.

\cancel{E} *Uncertainties:* Uncertainties on the lepton, photon, jet scale and resolution are propagated through to the missing transverse energy. As such the impact on the \cancel{E} uncertainty is estimated when evaluating the shift on the other variables.

Further Background Estimation: Uncertainties on the data-driven scale factors calculated for this analysis as described in Section 5.7 are applied and the scale factors are varied up and down by one standard deviation.

6.1.3. Symmetrization, Smoothing, and Pruning of Systematic Uncertainties

Symmetrization and smoothing are methods used to minimize statistical fluctuations in various systematic sources. Symmetrization centers the systematic uncertainty around a mean value and smoothing averages the expected number of events across bins in order to remove statistical fluctuations.

6.1.3.1. Symmetrization

Two-sided symmetrization is preformed when up and down variations are provided for any given systematic. The difference between these variations is calculated and the half sum of the absolute deviations from the nominal is taken as a symmetric variation:

$$\text{Symmetric Variation} = \frac{|\text{up} - \text{nominal}| + |\text{down} - \text{nominal}|}{2}$$

The nominal value is then varied up and down by this symmetrized value. However, if only an up or a down variation is provided for a systematic one-sided symmetrization is used by mirroring the absolute deviation about the nominal value. Experimental systematic sources are generally symmetrized while signal and background modelling contributions are not.

6.1.3.2. Smoothing

Smoothing is a technique used to average statistics across bins. This prevents large statistical spikes in many systematic uncertainties that are expected to provide small contributions. The smoothing algorithm depends on two parameters, the tolerance and maximal number of slope changes taking advantage of bin and neighboring bin information. Distributions are rebinned until the statistical uncertainty of each bin is below the tolerance and then the number of slope changes in the distribution is checked. If the number of slope changes is smaller than the threshold of four the distribution is kept, else the first step is performed again with the tolerance value halved. The smoothing algorithm 353QH [93] is run to avoid artificially flat uncertainties being introduced in the first two steps. All uncertainties

are smoothed unless stated otherwise. Smoothing does not change the overall normalizations of the uncertainty.

6.1.3.3. Pruning

Pruning uncertainties is done in order to reduce the number of nuisance parameters (NP) and stabilize the fit. Uncertainties that would only have a small impact on the end result are removed. An initial fit is calculated for each NP using the $\pm 1\sigma$ variation, if the effect on the uncertainty is less than the given threshold of 1% the contribution is removed from further fits.

6.2. Statistical Treatment of Results

A profile likelihood fit is performed simultaneously on the SR and $W + \gamma + \text{jets}$ and $t\bar{t} + \gamma + \text{jets}$ CRs. The **TRExFitter** framework [94] was used for this analysis which provides a framework built upon existing code such as the **RooStats** project [95]. Following [96] as well as the fit preformed in [33] we can define a likelihood function \mathcal{L} that is generally defined in the following manner:

$$\mathcal{L} = \mathcal{L}(\mu, \theta | \vec{x})$$

where the parameter μ is defined as the signal strength and θ as the set of nuisance parameters for the systematic uncertainties. The parameter of interest in the fit is proxied as \vec{x} . The likelihood is computed using the following:

$$\mathcal{L} = \prod_r^{N_{\text{regions}}} \prod_i^{N_{\text{r,bins}}} P(N_{r,i} | N_{r,i}^{\text{s}} + \sum_b N_{r,i}^{\text{b}}) \prod_j^{N_{\text{NP}}} G(x|1, \sigma_j)$$

using the following parameters:

- N_{regions} is the number of regions considered in the fit
- $N_{r,\text{bins}}$ is the number of bins in region r
- r is the index that runs over each of the different regions
- i is the index that runs over the number of bins in the region being considered
- $P(x|\lambda)$ is the Poisson function with a mean λ
- $N_{r,i}$ is the observed number of events in bin i of region r
- $N_{r,i}^s$ is the expected number of signal events in bin i of region r
- N_{bkg} is the number of background processes that are considered in the fit
- b is the index that runs over all categories of backgrounds considered
- $N_{r,i}^b$ is the expected number of events of background b in bin i of region r
- N_{NP} is the total number of nuisance parameters (NPs) considered in the fit
- j is the index that runs over the number of NPs
- $G(x|\mu, \sigma_j)$ is a Gaussian function with mean μ and width σ_j for the source of systematic uncertainty j (this Gaussian is replaced with a Poisson function for statistical uncertainties due to MC statistics)

The signal strength, μ , enters the likelihood as

$$N_{r,i}^s = \mu \cdot N_{\text{input},r}^s \cdot \rho_{r,i}$$

with the number of signal events in region r , $N_{\text{input},r}^s$ is scaled using the effective cross section $\sigma_{\text{eff}}^{\text{coup}}$ and $\rho_{r,i}$, the fraction of signal events in the respective bin i and region r .

The decay mode effective cross section is given by:

$$\sigma_{\text{eff, input}}^{\text{decay}} = 2 \times \sigma(pp \rightarrow t\bar{t}) \times \mathcal{B}(t \rightarrow Wb) \times \mathcal{B}(W \rightarrow l\nu) \times \mathcal{B}_{\text{input}}(t \rightarrow q\gamma)$$

where $\sigma(pp \rightarrow t\bar{t}) = 831.76$ pb, $\mathcal{B}(t \rightarrow bW) \approx 1$, and $\mathcal{B}(W \rightarrow l\nu) = 32.58\%$. The effective cross sections and couplings in the signal region are shown assuming the branching ratio to the FCNC decay mode $\mathcal{B}(t \rightarrow q\gamma) = 10^{-3}$ for $q=u,c$. This implies that the total effective cross section used for the FCNC signal samples, $\sigma_{\text{eff, input}}^{\text{decay}}$, is 542 fb. The final states of this analysis include both up type quark decay modes and the assumption is made that $\mathcal{B}(t \rightarrow u\gamma) = \mathcal{B}(t \rightarrow c\gamma)$. As can be seen in figures of the Signal Region distributions in Section 5.8 the single top event rate is around 7% of the total MC background. Due to this the effect of the production mode ($u \rightarrow t\gamma$) is not further considered as the final state for single top production with an association photon has significant kinematic differences to the regions in this analysis. These events would appear as a further subset of the single top background. Reconstructing an invariant mass around the top quark mass from the photon and the leading light jet is a key input to the neural network discriminator. This is unlikely to occur in the production mode.

A profile likelihood approach for treatment of nuisance parameters[97] (systematic uncertainties presented in Section 6.1) is used.

Physically we can interpret this signal strength as the ratio of the signal cross section to the predicted signal cross section. However if both the decay mode and

production (prod) mode are considered then we can interpret μ as

$$\mu = \frac{N_{\text{fit}}^{\text{decay}} + N_{\text{fit}}^{\text{prod.}}}{N_{\text{input}}^{\text{decay}} + N_{\text{input}}^{\text{prod.}}}$$

then it follows that

$$\mu = \frac{\epsilon^{\text{decay}} \times \mathcal{A}^{\text{decay}} \times \sigma_{\text{eff, fit}}^{\text{decay}} \times L + \epsilon^{\text{prod.}} \times \mathcal{A}^{\text{prod.}} \times \sigma_{\text{eff, fit}}^{\text{prod.}} \times L}{\epsilon^{\text{decay}} \times \mathcal{A}^{\text{decay}} \times \sigma_{\text{eff, input}}^{\text{decay}} \times L + \epsilon^{\text{prod.}} \times \mathcal{A}^{\text{prod.}} \times \sigma_{\text{eff, input}}^{\text{prod.}} \times L}$$

where the fitted (input) number of signal events per coupling (production/decay) $N_{\text{fit}}^{\text{coupling}}$, the signal efficiencies σ^{coupling} , the detector acceptances $\mathcal{A}^{\text{coupling}}$, the fitted (input) effective cross sections $\sigma_{\text{eff, fit(input)}}^{\text{coupling}}$, and the luminosity L . If it is then assumed that the detector acceptances and signal efficiencies are independent from the coupling strength and the fact that the signal samples have the same dependence on the coupling strength ($\sigma_{\text{eff}}^{\text{decay}}/\sigma_{\text{eff}}^{\text{prod.}} = \text{constant}$) we can simplify the expression for the signal strength to the following form:

$$\mu = \frac{\sigma_{\text{eff, fit}}^{\text{decay}}}{\sigma_{\text{eff, input}}^{\text{decay}}} = \frac{\sigma_{\text{eff, fit}}^{\text{prod.}}}{\sigma_{\text{eff, input}}^{\text{prod.}}}$$

The background-only hypothesis is fulfilled when $\mu = 0$. For this search the signal strength can be classified in terms of the branching ratio (BR)

$$\mu = \frac{\mathcal{B}(t \rightarrow q\gamma) \times \mathcal{B}(t \rightarrow Wb)}{\mathcal{B}_{\text{input}}(t \rightarrow q\gamma) \times \mathcal{B}_{\text{input}}(t \rightarrow Wb)}$$

Under the assumption that the top quark nominally decays only to a b quark and a W boson and the negligible decay mode being searched for this can be rewritten as

$$\mu = \frac{\mathcal{B}(t \rightarrow q\gamma) \times |1 - \mathcal{B}(t \rightarrow q\gamma)|}{\mathcal{B}_{\text{input}}(t \rightarrow q\gamma) \times |1 - \mathcal{B}_{\text{input}}(t \rightarrow q\gamma)|}$$

$$\mu \approx \frac{\mathcal{B}(t \rightarrow q\gamma)}{\mathcal{B}_{\text{input}}(t \rightarrow q\gamma)}$$

Values of μ between 0 (background-only hypothesis) and 2 (2x the nominal branching ratio) are tested in order to calculate expected and observed limits using the CL_s technique[98] and presented in Section 6.4. The CL_s technique makes use of a test statistic based on the negative log likelihood motivated by Wilks' theorem that allows approximate asymptotic behavior $-2\ln\mathcal{L}(\mu)$ as a χ^2 goodness of fit [99].

6.3. Nuisance Parameters

The statistical analysis allows the uncertainties discussed in Section 6.1 to enter into the fits as nuisance parameters (NPs). Each of these NPs is considered in all of the fit regions independently and dropped (pruned away) if its effect is less than 1%. Many of the parameters are dropped and for some only the normalization or the shape impact is dropped. A large number of NPs are considered for each region and include theory, experimental, and modeling uncertainties.

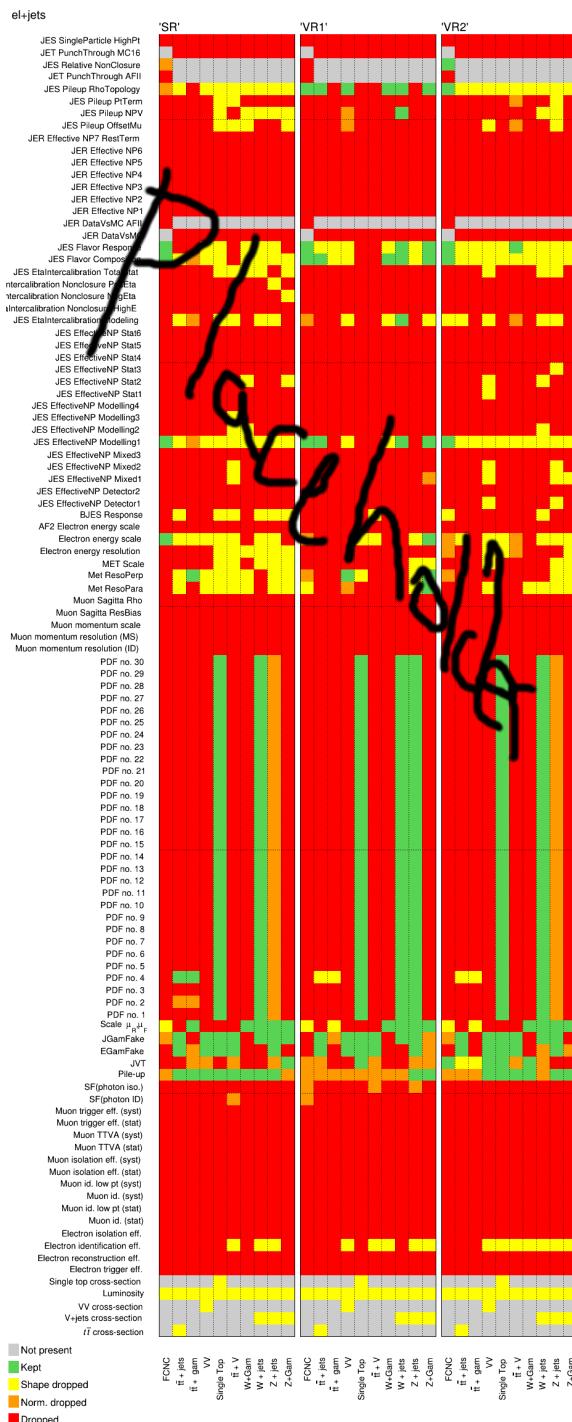


FIGURE 6.1. Overview of nuisance parameters after pruning for the $e+jets$ channel. If shape or normalization impacts are smaller than 1% that part of the nuisance parameter will be dropped in the final fit.

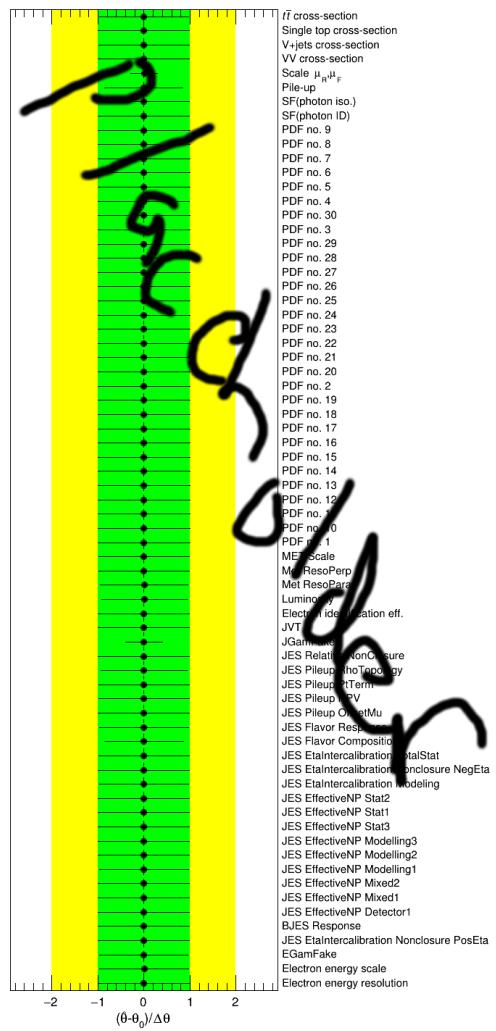


FIGURE 6.2. Pull values for the various nuisance parameters considered in the fit for the $e+jets$ channel.

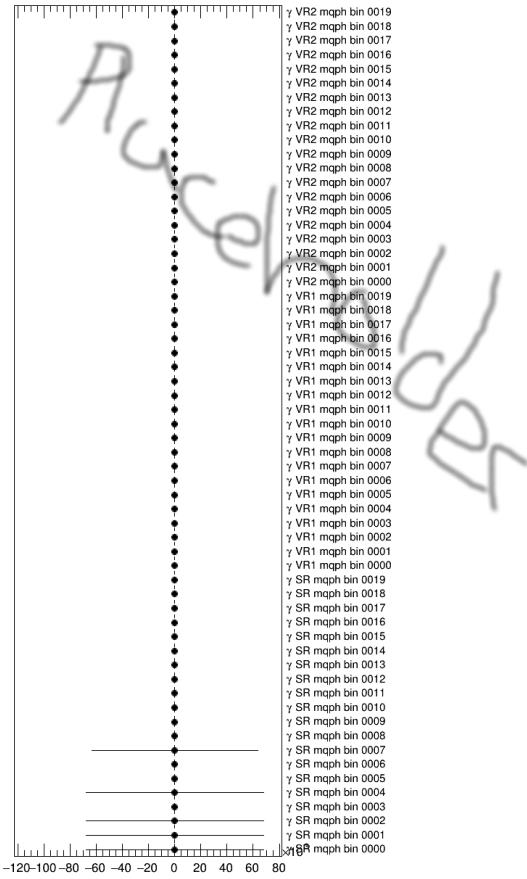


FIGURE 6.3. Bin by bin normalization γ factors used in each region for the $e+jets$ channel.

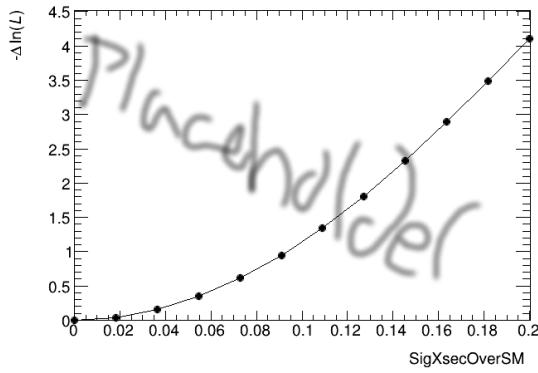


FIGURE 6.4. Negative-log likelihood (goodness of fit) as a function of signal strength using data in all regions for the $e+jets$ channel.

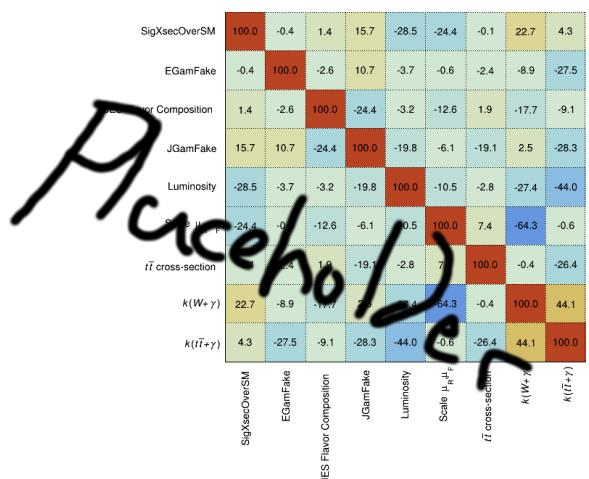


FIGURE 6.5. Correlation matrix with at least one coefficient above 20% for $e+jets$ channel

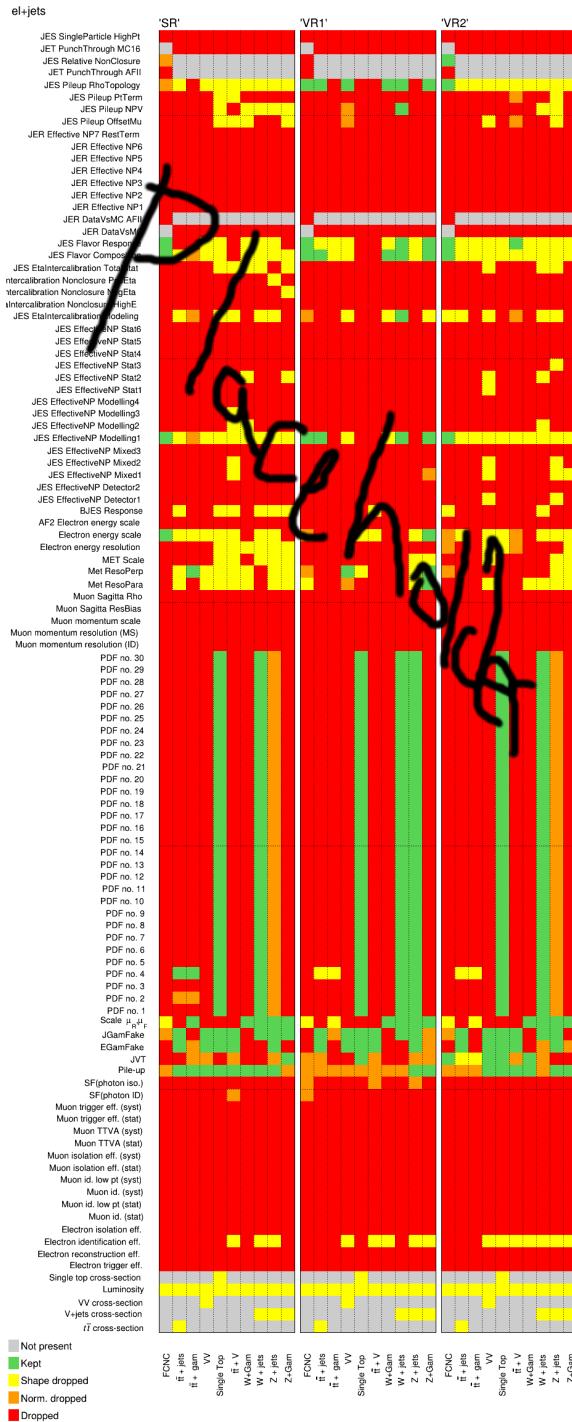


FIGURE 6.6. Overview of nuisance parameters after pruning for the $\mu+jets$ channel. If shape or normalization impacts are smaller than 1% that part of the nuisance parameter will be dropped in the final fit.

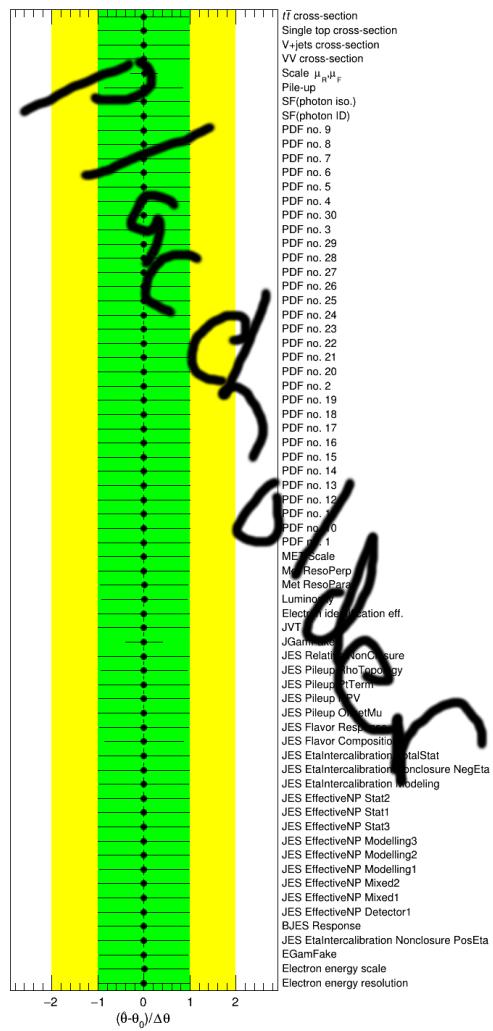


FIGURE 6.7. Pull values for the various nuisance parameters considered in the fit for the $\mu + \text{jets}$ channel.

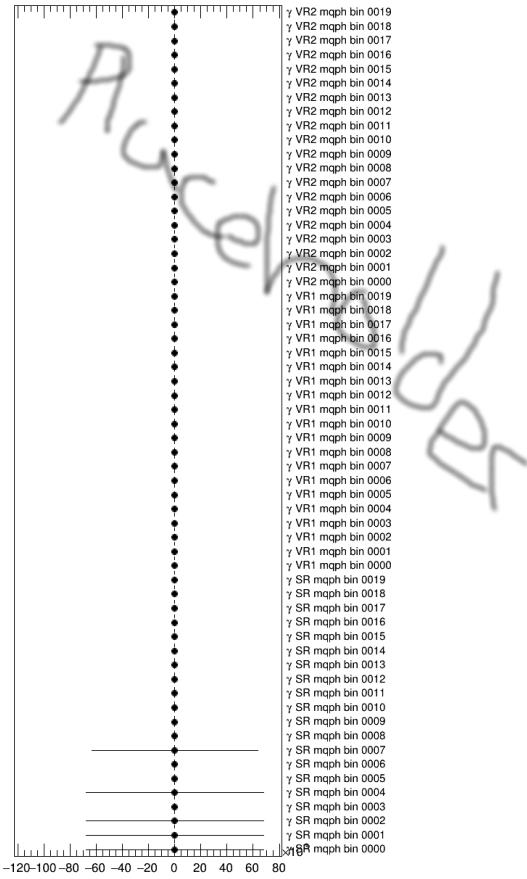


FIGURE 6.8. Bin by bin normalization γ factors used in each region for the $\mu + \text{jets}$ channel.

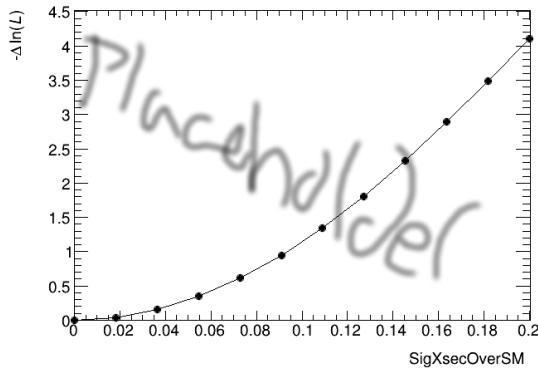


FIGURE 6.9. Negative-log likelihood (goodness of fit) as a function of signal strength using data in all regions for the $\mu + \text{jets}$ channel.

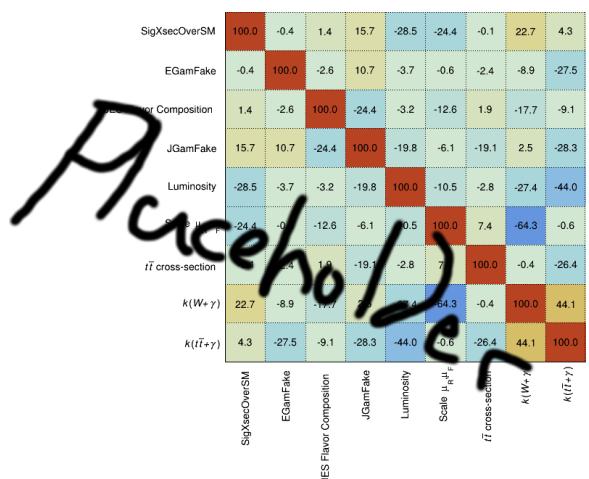


FIGURE 6.10. Correlation matrix with at least one coefficient above 20% for μ +jets channel

6.4. Limit on Branching Ratio $t \rightarrow q\gamma$

Combining searches with small stats from Lizas: [100]

CHAPTER VII

COMPLEMENTARY SEARCHES AND OUTLOOK

7.1. Comparison with Complementary Searches

7.2. Future Directions

HL-LHC and Beyond Future prospectives at Linear Colliders? - <https://www.sciencedirect.com>

7.3. Conclusion

APPENDIX A

DERIVATION INFORMATION (TOPQ1)

Preselection is applied to both data and MC samples using the derivation framework in order to reduce the xAOD sample size. TOPQ1 derivations were used and further skimmed for the specific n-tuples used in this analysis. TOPQ1 derivations are used for standard top analyses with at least one lepton and have the following definitions:

Particle	Definition
Electron	$ \eta < 2.5$ and Electrons.DFCommonElectronsLHLoose
Muon	$ \eta < 2.5$ and Muons.muonType=0 and Muons.DFCommonGoodMuon
Small-R Jet	$ \eta < 2.5$ and AntiKt4EMTopoJets.DFCommonJets_Calib
Large-R Jet	$ \eta < 2.5$ and AntiKt10LCTopoTrimmedPtFrac5SmallR20Jets.DFCommonJets_Calib

TABLE A.1. TOPQ1 Object Definitions

In addition to these object definitions TOPQ1 Derivations require trigger skimming of ≥ 1 lepton with $p_T > 20\text{GeV}$.

APPENDIX B

COMPLETE LIST OF MONTE CARLO SAMPLES USED

The following MC samples have been used to simulate the signal and various Standard Model backgrounds at center of mass energy of 13 TeV. MC16a, MC16d, and MC16e samples correspond to run conditions for years 2015/16, 2017, and 2018 respectively. The newest unprescaled p-tag samples at the time of writing were used for each sample.

Process	MC DSID	Generator	Generator Tune
Diboson processes, Full Simulation			
$ZZ \rightarrow llll$	364250	SHERPA2.2.2	NNPDF3.0 NNLO
$ZW \rightarrow ll\nu\nu$	364253	SHERPA2.2.2	NNPDF3.0 NNLO
$WW/ZZ \rightarrow ll\nu\nu$	364254	SHERPA2.2.2	NNPDF3.0 NNLO
$WZ \rightarrow l\nu\nu\nu$	364255	SHERPA2.2.2	NNPDF3.0 NNLO
$ZZ \rightarrow qq\nu\nu$	364355	SHERPA2.2.1	NNPDF3.0 NNLO
$ZZ \rightarrow qql\bar{l}$	364356	SHERPA2.2.1	NNPDF3.0 NNLO
$WZ \rightarrow qq\nu\nu$	363357	SHERPA2.2.1	NNPDF3.0 NNLO
$WZ \rightarrow qql\bar{l}$	363358	SHERPA2.2.1	NNPDF3.0 NNLO
$WW \rightarrow qql^-\nu$	363359	SHERPA2.2.1	NNPDF3.0 NNLO
$WW \rightarrow qq\ell^+\nu$	363360	SHERPA2.2.1	NNPDF3.0 NNLO
$WZ \rightarrow qql\nu$	363489	SHERPA2.2.1	NNPDF3.0 NNLO
$t\bar{t} + V$ processes			
$t\bar{t}W$	410155	MADGRAPH5_AMC@NLO + PYTHIA 8 + EVTGEN	A14NNPDF23LO
$t\bar{t}Z \rightarrow t\bar{t}\nu\nu$	410156	MADGRAPH5_AMC@NLO + PYTHIA 8 + EVTGEN	A14NNPDF23LO
$t\bar{t}Z \rightarrow t\bar{t}qq$	410157	MADGRAPH5_AMC@NLO + PYTHIA 8 + EVTGEN	A14NNPDF23LO
$t\bar{t}Z \rightarrow t\bar{t}ee$	410218	MADGRAPH5_AMC@NLO + PYTHIA 8 + EVTGEN	A14NNPDF23LO
$t\bar{t}Z \rightarrow t\bar{t}\mu\mu$	410219	MADGRAPH5_AMC@NLO + PYTHIA 8 + EVTGEN	A14NNPDF23LO
$t\bar{t}Z \rightarrow t\bar{t}\tau\tau$	410220	MADGRAPH5_AMC@NLO + PYTHIA 8 + EVTGEN	A14NNPDF23LO

TABLE B.1. List of MC16a (r-tag 9364) samples used in this search for the $t\bar{t} + V$ and diboson processes. All samples are available and used for MC16d (r-tag 10201) and MC16e (r-tag 10724) unless otherwise stated.

Process	MC DSID	Generator	Generator Tune
Signal samples for FCNC in decay mode: $pp \rightarrow t\bar{t} \rightarrow b\bar{b}l\nu q\gamma$, NLO, AFII			
$t\gamma$, anti-top	410980	MADGRAPH5_AMC@NLO + PYTHIA 8 + EVTGEN	A14 NNPDF23LO
$t\gamma$, top	410981	MADGRAPH5_AMC@NLO + PYTHIA 8 + EVTGEN	A14 NNPDF23LO
$t\gamma$, anti-top	410984	MADGRAPH5_AMC@NLO + PYTHIA 8 + EVTGEN	A14 NNPDF23LO
$t\gamma$, top	410985	MADGRAPH5_AMC@NLO + PYTHIA 8 + EVTGEN	A14 NNPDF23LO
$t\bar{t}$ processes, Full Simulation			
$t\bar{t}$ (NLO)	410470	POWHEG-BOX + PYTHIA 8 + EVTGEN	A14 NNPDF23LO
$t\bar{t} + \gamma$ (LO)	410389	POWHEG-BOX + PYTHIA 8 + EVTGEN	A14 NNPDF23LO
$t\bar{t}$ processes, Variation Samples, NLO, AFII			
$t\bar{t}$ (single lepton)	410557	POWHEG-BOX + HERWIG 7 + EVTGEN	MMHT2017lo68cl
$t\bar{t}$ (di-lepton)	410558	POWHEG-BOX + HERWIG 7 + EVTGEN	MMHT2017lo68cl
$t\bar{t}$ (single lepton)	410464	MADGRAPH5_AMC@NLO + PYTHIA 8 + EVTGEN	A14NNPDF23LO
$t\bar{t}$ (di-lepton)	410465	MADGRAPH5_AMC@NLO + PYTHIA 8 + EVTGEN	A14NNPDF23LO
$t\bar{t}$ (single lepton, $h_{\text{damp}} = 3m_{\text{top}}$)	410480	POWHEG-BOX + PYTHIA 8 + EVTGEN	A14NNPDF23LO
$t\bar{t}$ (di-lepton, $h_{\text{damp}} = 3m_{\text{top}}$)	410482	POWHEG-BOX + PYTHIA 8 + EVTGEN	A14NNPDF23LO
Single-top processes, NLO, Full Simulation			
t-channel, top	410658	POWHEG-BOX + PYTHIA 8 + EVTGEN	A14 NNPDF23LO
t-channel, anti-top	410659	POWHEG-BOX + PYTHIA 8 + EVTGEN	A14 NNPDF23LO
s-channel, top	410644	POWHEG-BOX + PYTHIA 8 + EVTGEN	A14 NNPDF23LO
s-channel, anti-top	410645	POWHEG-BOX + PYTHIA 8 + EVTGEN	A14 NNPDF23LO
tW-channel, top	410646	POWHEG-BOX + PYTHIA 8 + EVTGEN	A14 NNPDF23LO
tW-channel, anti-top	410647	POWHEG-BOX + PYTHIA 8 + EVTGEN	A14 NNPDF23LO
Single-top processes, Variation Samples, NLO			
t-channel, anti-top, AFII	411032	POWHEG-BOX + HERWIG 7 + EVTGEN	MMHT2017lo68cl
t-channel, top, AFII	411033	POWHEG-BOX + HERWIG 7 + EVTGEN	MMHT2017lo68cl
t-channel, AFII	410560	MADGRAPH5_AMC@NLO + PYTHIA 8 + EVTGEN	A14NNPDF23LO
s-channel, anti-top, AFII	411034	POWHEG-BOX + HERWIG 7 + EVTGEN	MMHT2017lo68cl
s-channel, top, AFII	411035	POWHEG-BOX + HERWIG 7 + EVTGEN	MMHT2017lo68cl
s-channel, AFII	410561	MADGRAPH5_AMC@NLO + PYTHIA 8 + EVTGEN	A14NNPDF23LO
tW-channel, anti-top, AFII	411036	POWHEG-BOX + HERWIG 7 + EVTGEN	MMHT2017lo68cl
tW-channel, top, AFII	411037	POWHEG-BOX + HERWIG 7 + EVTGEN	MMHT2017lo68cl
tW-channel, anti-top, AFII	410654	POWHEG-BOX + HERWIG 7 + EVTGEN	MMHT2017lo68cl
tW-channel, top, AFII	410655	POWHEG-BOX + HERWIG 7 + EVTGEN	MMHT2017lo68cl
tW-channel, AFII	412002	MADGRAPH5_AMC@NLO + PYTHIA 8 + EVTGEN	A14NNPDF23LO

TABLE B.2. List of MC16a (r-tag 9364) samples used in this search for the signal sample, $t\bar{t}$, and single-top processes. All samples are available and used for MC16d (r-tag 10201) and MC16e (r-tag 10724) unless otherwise stated.

Process	MC DSID	Generator	Generator Tune
W+Jets, Full Simulation			
$W \rightarrow e\nu$, $\max(H_T, p_T^V) < 70\text{GeV}$, c veto, b veto	364170	SHERPA2.2.1	NNPDF3.0 NNLO
$W \rightarrow e\nu$, $\max(H_T, p_T^V) < 70\text{GeV}$, c filter, b veto	364171	SHERPA2.2.1	NNPDF3.0 NNLO
$W \rightarrow e\nu$, $\max(H_T, p_T^V) < 70\text{GeV}$, b filter	364172	SHERPA2.2.1	NNPDF3.0 NNLO
$W \rightarrow e\nu$, $70\text{GeV} < \max(H_T, p_T^V) < 140\text{GeV}$, c veto, b veto	364173	SHERPA2.2.1	NNPDF3.0 NNLO
$W \rightarrow e\nu$, $70\text{GeV} < \max(H_T, p_T^V) < 140\text{GeV}$, c filter, b veto	364174	SHERPA2.2.1	NNPDF3.0 NNLO
$W \rightarrow e\nu$, $70\text{GeV} < \max(H_T, p_T^V) < 140\text{GeV}$, b filter	364175	SHERPA2.2.1	NNPDF3.0 NNLO
$W \rightarrow e\nu$, $140\text{GeV} < \max(H_T, p_T^V) < 280\text{GeV}$, c veto, b veto	364176	SHERPA2.2.1	NNPDF3.0 NNLO
$W \rightarrow e\nu$, $140\text{GeV} < \max(H_T, p_T^V) < 280\text{GeV}$, c filter, b veto	364177	SHERPA2.2.1	NNPDF3.0 NNLO
$W \rightarrow e\nu$, $140\text{GeV} < \max(H_T, p_T^V) < 280\text{GeV}$, b filter	364178	SHERPA2.2.1	NNPDF3.0 NNLO
$W \rightarrow e\nu$, $280\text{GeV} < \max(H_T, p_T^V) < 500\text{GeV}$, c veto, b veto	364179	SHERPA2.2.1	NNPDF3.0 NNLO
$W \rightarrow e\nu$, $280\text{GeV} < \max(H_T, p_T^V) < 500\text{GeV}$, c filter, b veto	364180	SHERPA2.2.1	NNPDF3.0 NNLO
$W \rightarrow e\nu$, $280\text{GeV} < \max(H_T, p_T^V) < 500\text{GeV}$, b filter	364181	SHERPA2.2.1	NNPDF3.0 NNLO
$W \rightarrow e\nu$, $500\text{GeV} < \max(H_T, p_T^V) < 1000\text{GeV}$	364182	SHERPA2.2.1	NNPDF3.0 NNLO
$W \rightarrow e\nu$, $\max(H_T, p_T^V) > 1000\text{GeV}$	364183	SHERPA2.2.1	NNPDF3.0 NNLO
$W \rightarrow \mu\nu$, $\max(H_T, p_T^V) < 70\text{GeV}$, c veto, b veto	364156	SHERPA2.2.1	NNPDF3.0 NNLO
$W \rightarrow \mu\nu$, $\max(H_T, p_T^V) < 70\text{GeV}$, c filter, b veto	364157	SHERPA2.2.1	NNPDF3.0 NNLO
$W \rightarrow \mu\nu$, $\max(H_T, p_T^V) < 70\text{GeV}$, b filter	364158	SHERPA2.2.1	NNPDF3.0 NNLO
$W \rightarrow \mu\nu$, $70\text{GeV} < \max(H_T, p_T^V) < 140\text{GeV}$, c veto, b veto	364159	SHERPA2.2.1	NNPDF3.0 NNLO
$W \rightarrow \mu\nu$, $70\text{GeV} < \max(H_T, p_T^V) < 140\text{GeV}$, c filter, b veto	364160	SHERPA2.2.1	NNPDF3.0 NNLO
$W \rightarrow \mu\nu$, $70\text{GeV} < \max(H_T, p_T^V) < 140\text{GeV}$, b filter	364161	SHERPA2.2.1	NNPDF3.0 NNLO
$W \rightarrow \mu\nu$, $140\text{GeV} < \max(H_T, p_T^V) < 280\text{GeV}$, c veto, b veto	364162	SHERPA2.2.1	NNPDF3.0 NNLO
$W \rightarrow \mu\nu$, $140\text{GeV} < \max(H_T, p_T^V) < 280\text{GeV}$, c filter, b veto	364163	SHERPA2.2.1	NNPDF3.0 NNLO
$W \rightarrow \mu\nu$, $140\text{GeV} < \max(H_T, p_T^V) < 280\text{GeV}$, b filter	364164	SHERPA2.2.1	NNPDF3.0 NNLO
$W \rightarrow \mu\nu$, $280\text{GeV} < \max(H_T, p_T^V) < 500\text{GeV}$, c veto, b veto	364165	SHERPA2.2.1	NNPDF3.0 NNLO
$W \rightarrow \mu\nu$, $280\text{GeV} < \max(H_T, p_T^V) < 500\text{GeV}$, c filter, b veto	364166	SHERPA2.2.1	NNPDF3.0 NNLO
$W \rightarrow \mu\nu$, $280\text{GeV} < \max(H_T, p_T^V) < 500\text{GeV}$, b filter	364167	SHERPA2.2.1	NNPDF3.0 NNLO
$W \rightarrow \mu\nu$, $500\text{GeV} < \max(H_T, p_T^V) < 1000\text{GeV}$	364168	SHERPA2.2.1	NNPDF3.0 NNLO
$W \rightarrow \mu\nu$, $\max(H_T, p_T^V) > 1000\text{GeV}$	364169	SHERPA2.2.1	NNPDF3.0 NNLO
$W \rightarrow \tau\nu$, $\max(H_T, p_T^V) < 70\text{GeV}$, c veto, b veto	364184	SHERPA2.2.1	NNPDF3.0 NNLO
$W \rightarrow \tau\nu$, $\max(H_T, p_T^V) < 70\text{GeV}$, c filter, b veto	364185	SHERPA2.2.1	NNPDF3.0 NNLO
$W \rightarrow \tau\nu$, $\max(H_T, p_T^V) < 70\text{GeV}$, b filter	364186	SHERPA2.2.1	NNPDF3.0 NNLO
$W \rightarrow \tau\nu$, $70\text{GeV} < \max(H_T, p_T^V) < 140\text{GeV}$, c veto, b veto	364187	SHERPA2.2.1	NNPDF3.0 NNLO
$W \rightarrow \tau\nu$, $70\text{GeV} < \max(H_T, p_T^V) < 140\text{GeV}$, c filter, b veto	364188	SHERPA2.2.1	NNPDF3.0 NNLO
$W \rightarrow \tau\nu$, $70\text{GeV} < \max(H_T, p_T^V) < 140\text{GeV}$, b filter	364189	SHERPA2.2.1	NNPDF3.0 NNLO
$W \rightarrow \tau\nu$, $140\text{GeV} < \max(H_T, p_T^V) < 280\text{GeV}$, c veto, b veto	364190	SHERPA2.2.1	NNPDF3.0 NNLO
$W \rightarrow \tau\nu$, $140\text{GeV} < \max(H_T, p_T^V) < 280\text{GeV}$, c filter, b veto	364191	SHERPA2.2.1	NNPDF3.0 NNLO
$W \rightarrow \tau\nu$, $140\text{GeV} < \max(H_T, p_T^V) < 280\text{GeV}$, b filter	364192	SHERPA2.2.1	NNPDF3.0 NNLO
$W \rightarrow \tau\nu$, $280\text{GeV} < \max(H_T, p_T^V) < 500\text{GeV}$, c veto, b veto	364193	SHERPA2.2.1	NNPDF3.0 NNLO
$W \rightarrow \tau\nu$, $280\text{GeV} < \max(H_T, p_T^V) < 500\text{GeV}$, c filter, b veto	364194	SHERPA2.2.1	NNPDF3.0 NNLO
$W \rightarrow \tau\nu$, $280\text{GeV} < \max(H_T, p_T^V) < 500\text{GeV}$, b filter	364195	SHERPA2.2.1	NNPDF3.0 NNLO
$W \rightarrow \tau\nu$, $500\text{GeV} < \max(H_T, p_T^V) < 1000\text{GeV}$	364196	SHERPA2.2.1	NNPDF3.0 NNLO
$W \rightarrow \tau\nu$, $\max(H_T, p_T^V) > 1000\text{GeV}$	364197	SHERPA2.2.1	NNPDF3.0 NNLO

TABLE B.3. List of MC16a (r-tag 9364) samples used in this search for the W+jets background. All samples are available and used for MC16d (r-tag 10201) and MC16e (r-tag 10724). All processes are simulated for up to 2 partons at NLO and 4 partons at LO.

Process	MC DSID	Generator	Generator Tune
Z+Jets, Full Simulation			
$Z \rightarrow ee, \max(H_T, p_T^V) < 70\text{GeV}, c \text{ veto}, b \text{ veto}$	364114	SHERPA2.2.1	NNPDF3.0 NNLO
$Z \rightarrow ee, \max(H_T, p_T^V) < 70\text{GeV}, c \text{ filter}, b \text{ veto}$	364115	SHERPA2.2.1	NNPDF3.0 NNLO
$Z \rightarrow ee, \max(H_T, p_T^V) < 70\text{GeV}, b \text{ filter}$	364116	SHERPA2.2.1	NNPDF3.0 NNLO
$Z \rightarrow ee, 70\text{GeV} < \max(H_T, p_T^V) < 140\text{GeV}, c \text{ veto}, b \text{ veto}$	364117	SHERPA2.2.1	NNPDF3.0 NNLO
$Z \rightarrow ee, 70\text{GeV} < \max(H_T, p_T^V) < 140\text{GeV}, c \text{ filter}, b \text{ veto}$	364118	SHERPA2.2.1	NNPDF3.0 NNLO
$Z \rightarrow ee, 70\text{GeV} < \max(H_T, p_T^V) < 140\text{GeV}, b \text{ filter}$	364119	SHERPA2.2.1	NNPDF3.0 NNLO
$Z \rightarrow ee, 140\text{GeV} < \max(H_T, p_T^V) < 280\text{GeV}, c \text{ veto}, b \text{ veto}$	364120	SHERPA2.2.1	NNPDF3.0 NNLO
$Z \rightarrow ee, 140\text{GeV} < \max(H_T, p_T^V) < 280\text{GeV}, c \text{ filter}, b \text{ veto}$	364121	SHERPA2.2.1	NNPDF3.0 NNLO
$Z \rightarrow ee, 140\text{GeV} < \max(H_T, p_T^V) < 280\text{GeV}, b \text{ filter}$	364122	SHERPA2.2.1	NNPDF3.0 NNLO
$Z \rightarrow ee, 280\text{GeV} < \max(H_T, p_T^V) < 500\text{GeV}, c \text{ veto}, b \text{ veto}$	364123	SHERPA2.2.1	NNPDF3.0 NNLO
$Z \rightarrow ee, 280\text{GeV} < \max(H_T, p_T^V) < 500\text{GeV}, c \text{ filter}, b \text{ veto}$	364124	SHERPA2.2.1	NNPDF3.0 NNLO
$Z \rightarrow ee, 280\text{GeV} < \max(H_T, p_T^V) < 500\text{GeV}, b \text{ filter}$	364125	SHERPA2.2.1	NNPDF3.0 NNLO
$Z \rightarrow ee, 500\text{GeV} < \max(H_T, p_T^V) < 1000\text{GeV}$	364126	SHERPA2.2.1	NNPDF3.0 NNLO
$Z \rightarrow ee, \max(H_T, p_T^V) > 1000\text{GeV}$	364127	SHERPA2.2.1	NNPDF3.0 NNLO
$Z \rightarrow \mu\mu, \max(H_T, p_T^V) < 70\text{GeV}, c \text{ veto}, b \text{ veto}$	364100	SHERPA2.2.1	NNPDF3.0 NNLO
$Z \rightarrow \mu\mu, \max(H_T, p_T^V) < 70\text{GeV}, c \text{ filter}, b \text{ veto}$	364101	SHERPA2.2.1	NNPDF3.0 NNLO
$Z \rightarrow \mu\mu, \max(H_T, p_T^V) < 70\text{GeV}, b \text{ filter}$	364102	SHERPA2.2.1	NNPDF3.0 NNLO
$Z \rightarrow \mu\mu, 70\text{GeV} < \max(H_T, p_T^V) < 140\text{GeV}, c \text{ veto}, b \text{ veto}$	364103	SHERPA2.2.1	NNPDF3.0 NNLO
$Z \rightarrow \mu\mu, 70\text{GeV} < \max(H_T, p_T^V) < 140\text{GeV}, c \text{ filter}, b \text{ veto}$	364104	SHERPA2.2.1	NNPDF3.0 NNLO
$Z \rightarrow \mu\mu, 70\text{GeV} < \max(H_T, p_T^V) < 140\text{GeV}, b \text{ filter}$	364105	SHERPA2.2.1	NNPDF3.0 NNLO
$Z \rightarrow \mu\mu, 140\text{GeV} < \max(H_T, p_T^V) < 280\text{GeV}, c \text{ veto}, b \text{ veto}$	364106	SHERPA2.2.1	NNPDF3.0 NNLO
$Z \rightarrow \mu\mu, 140\text{GeV} < \max(H_T, p_T^V) < 280\text{GeV}, c \text{ filter}, b \text{ veto}$	364107	SHERPA2.2.1	NNPDF3.0 NNLO
$Z \rightarrow \mu\mu, 140\text{GeV} < \max(H_T, p_T^V) < 280\text{GeV}, b \text{ filter}$	364108	SHERPA2.2.1	NNPDF3.0 NNLO
$Z \rightarrow \mu\mu, 280\text{GeV} < \max(H_T, p_T^V) < 500\text{GeV}, c \text{ veto}, b \text{ veto}$	364109	SHERPA2.2.1	NNPDF3.0 NNLO
$Z \rightarrow \mu\mu, 280\text{GeV} < \max(H_T, p_T^V) < 500\text{GeV}, c \text{ filter}, b \text{ veto}$	364110	SHERPA2.2.1	NNPDF3.0 NNLO
$Z \rightarrow \mu\mu, 280\text{GeV} < \max(H_T, p_T^V) < 500\text{GeV}, b \text{ filter}$	364111	SHERPA2.2.1	NNPDF3.0 NNLO
$Z \rightarrow \mu\mu, 500\text{GeV} < \max(H_T, p_T^V) < 1000\text{GeV}$	364112	SHERPA2.2.1	NNPDF3.0 NNLO
$Z \rightarrow \mu\mu, \max(H_T, p_T^V) > 1000\text{GeV}$	364113	SHERPA2.2.1	NNPDF3.0 NNLO
$Z \rightarrow \tau\tau, \max(H_T, p_T^V) < 70\text{GeV}, c \text{ veto}, b \text{ veto}$	364128	SHERPA2.2.1	NNPDF3.0 NNLO
$Z \rightarrow \tau\tau, \max(H_T, p_T^V) < 70\text{GeV}, c \text{ filter}, b \text{ veto}$	364129	SHERPA2.2.1	NNPDF3.0 NNLO
$Z \rightarrow \tau\tau, \max(H_T, p_T^V) < 70\text{GeV}, b \text{ filter}$	364130	SHERPA2.2.1	NNPDF3.0 NNLO
$Z \rightarrow \tau\tau, 70\text{GeV} < \max(H_T, p_T^V) < 140\text{GeV}, c \text{ veto}, b \text{ veto}$	364131	SHERPA2.2.1	NNPDF3.0 NNLO
$Z \rightarrow \tau\tau, 70\text{GeV} < \max(H_T, p_T^V) < 140\text{GeV}, c \text{ filter}, b \text{ veto}$	364132	SHERPA2.2.1	NNPDF3.0 NNLO
$Z \rightarrow \tau\tau, 70\text{GeV} < \max(H_T, p_T^V) < 140\text{GeV}, b \text{ filter}$	364133	SHERPA2.2.1	NNPDF3.0 NNLO
$Z \rightarrow \tau\tau, 140\text{GeV} < \max(H_T, p_T^V) < 280\text{GeV}, c \text{ veto}, b \text{ veto}$	364134	SHERPA2.2.1	NNPDF3.0 NNLO
$Z \rightarrow \tau\tau, 140\text{GeV} < \max(H_T, p_T^V) < 280\text{GeV}, c \text{ filter}, b \text{ veto}$	364135	SHERPA2.2.1	NNPDF3.0 NNLO
$Z \rightarrow \tau\tau, 140\text{GeV} < \max(H_T, p_T^V) < 280\text{GeV}, b \text{ filter}$	364136	SHERPA2.2.1	NNPDF3.0 NNLO
$Z \rightarrow \tau\tau, 280\text{GeV} < \max(H_T, p_T^V) < 500\text{GeV}, c \text{ veto}, b \text{ veto}$	364137	SHERPA2.2.1	NNPDF3.0 NNLO
$Z \rightarrow \tau\tau, 280\text{GeV} < \max(H_T, p_T^V) < 500\text{GeV}, c \text{ filter}, b \text{ veto}$	364138	SHERPA2.2.1	NNPDF3.0 NNLO
$Z \rightarrow \tau\tau, 280\text{GeV} < \max(H_T, p_T^V) < 500\text{GeV}, b \text{ filter}$	364139	SHERPA2.2.1	NNPDF3.0 NNLO
$Z \rightarrow \tau\tau, 500\text{GeV} < \max(H_T, p_T^V) < 1000\text{GeV}$	364140	SHERPA2.2.1	NNPDF3.0 NNLO
$Z \rightarrow \tau\tau, \max(H_T, p_T^V) > 1000\text{GeV}$	364141	SHERPA2.2.1	NNPDF3.0 NNLO

TABLE B.4. List of MC16a (r-tag 9364) samples used in this search for the Z+jets background. All samples are available and used for MC16d (r-tag 10201) and MC16e (r-tag 10724). All processes are simulated for up to 2 partons at NLO and 4 partons at LO.

Process	MC DSID	Generator	Generator Tune
Z+Jets+γ, Full Simulation			
$Z \rightarrow ee, 7\text{GeV} < p_T^\gamma < 15\text{GeV}$	366140	SHERPA2.2.4	NNPDF3.0 NNLO
$Z \rightarrow ee, 15\text{GeV} < p_T^\gamma < 35\text{GeV}$	366141	SHERPA2.2.4	NNPDF3.0 NNLO
$Z \rightarrow ee, 35\text{GeV} < p_T^\gamma < 70\text{GeV}$	366142	SHERPA2.2.4	NNPDF3.0 NNLO
$Z \rightarrow ee, 70\text{GeV} < p_T^\gamma < 140\text{GeV}$	366143	SHERPA2.2.4	NNPDF3.0 NNLO
$Z \rightarrow ee, p_T^\gamma > 140\text{GeV}$	366144	SHERPA2.2.4	NNPDF3.0 NNLO
$Z \rightarrow \mu\mu, 7\text{GeV} < p_T^\gamma < 15\text{GeV}$	366145	SHERPA2.2.4	NNPDF3.0 NNLO
$Z \rightarrow \mu\mu, 15\text{GeV} < p_T^\gamma < 35\text{GeV}$	366146	SHERPA2.2.4	NNPDF3.0 NNLO
$Z \rightarrow \mu\mu, 35\text{GeV} < p_T^\gamma < 70\text{GeV}$	366147	SHERPA2.2.4	NNPDF3.0 NNLO
$Z \rightarrow \mu\mu, 70\text{GeV} < p_T^\gamma < 140\text{GeV}$	366148	SHERPA2.2.4	NNPDF3.0 NNLO
$Z \rightarrow \mu\mu, p_T^\gamma > 140\text{GeV}$	366149	SHERPA2.2.4	NNPDF3.0 NNLO
$Z \rightarrow \tau\tau, 7\text{GeV} < p_T^\gamma < 15\text{GeV}$	366150	SHERPA2.2.4	NNPDF3.0 NNLO
$Z \rightarrow \tau\tau, 15\text{GeV} < p_T^\gamma < 35\text{GeV}$	366151	SHERPA2.2.4	NNPDF3.0 NNLO
$Z \rightarrow \tau\tau, 35\text{GeV} < p_T^\gamma < 70\text{GeV}$	366152	SHERPA2.2.4	NNPDF3.0 NNLO
$Z \rightarrow \tau\tau, 70\text{GeV} < p_T^\gamma < 140\text{GeV}$	366153	SHERPA2.2.4	NNPDF3.0 NNLO
$Z \rightarrow \tau\tau, p_T^\gamma > 140\text{GeV}$	366154	SHERPA2.2.4	NNPDF3.0 NNLO
W+Jets+γ, Full Simulation			
$W \rightarrow e\nu, 7\text{GeV} < p_T^\gamma < 15\text{GeV}$	364521	SHERPA2.2.2	NNPDF3.0 NNLO
$W \rightarrow e\nu, 15\text{GeV} < p_T^\gamma < 35\text{GeV}$	364522	SHERPA2.2.2	NNPDF3.0 NNLO
$W \rightarrow e\nu, 35\text{GeV} < p_T^\gamma < 70\text{GeV}$	364523	SHERPA2.2.2	NNPDF3.0 NNLO
$W \rightarrow e\nu, 70\text{GeV} < p_T^\gamma < 140\text{GeV}$	364524	SHERPA2.2.2	NNPDF3.0 NNLO
$W \rightarrow e\nu, p_T^\gamma > 140\text{GeV}$	364525	SHERPA2.2.2	NNPDF3.0 NNLO
$W \rightarrow \mu\nu, 7\text{GeV} < p_T^\gamma < 15\text{GeV}$	364526	SHERPA2.2.2	NNPDF3.0 NNLO
$W \rightarrow \mu\nu, 15\text{GeV} < p_T^\gamma < 35\text{GeV}$	364527	SHERPA2.2.2	NNPDF3.0 NNLO
$W \rightarrow \mu\nu, 35\text{GeV} < p_T^\gamma < 70\text{GeV}$	364528	SHERPA2.2.2	NNPDF3.0 NNLO
$W \rightarrow \mu\nu, 70\text{GeV} < p_T^\gamma < 140\text{GeV}$	364529	SHERPA2.2.2	NNPDF3.0 NNLO
$W \rightarrow \mu\nu, p_T^\gamma > 140\text{GeV}$	364530	SHERPA2.2.2	NNPDF3.0 NNLO
$W \rightarrow \tau\nu, 7\text{GeV} < p_T^\gamma < 15\text{GeV}$	364531	SHERPA2.2.2	NNPDF3.0 NNLO
$W \rightarrow \tau\nu, 15\text{GeV} < p_T^\gamma < 35\text{GeV}$	364532	SHERPA2.2.2	NNPDF3.0 NNLO
$W \rightarrow \tau\nu, 35\text{GeV} < p_T^\gamma < 70\text{GeV}$	364533	SHERPA2.2.2	NNPDF3.0 NNLO
$W \rightarrow \tau\nu, 70\text{GeV} < p_T^\gamma < 140\text{GeV}$	364534	SHERPA2.2.2	NNPDF3.0 NNLO
$W \rightarrow \tau\nu, p_T^\gamma > 140\text{GeV}$	364535	SHERPA2.2.2	NNPDF3.0 NNLO

TABLE B.5. List of MC16a (r-tag 9364) samples used in this search for the W+jets+ γ and Z+jets+ γ backgrounds. All samples are available and used for MC16d (r-tag 10201) and MC16e (r-tag 10724). All processes are simulated for up to 1 parton at NLO and 3 partons at LO.

APPENDIX C

ADDITIONAL PLOTS FROM NN STUDIES

C.1. Additional Shape Comparison Plots: $\mu + \text{jets}$ channel

Various additional plots are shown in this appendix from the neural network creation and studies. Figure C.1 and C.2 show additional shape comparisons in variables which are not included in the final neural network model as they do not significantly change the fit values. In the cases of p_T or E variables with the higher separation value were used as there is a large correlation between the two values and the other is shown in this appendix. ΔR_{jb} was not included as the other 3 ΔR values had higher separation values and they are all related to each other as they are the geometrically related.

The neutrino reconstruction is done using a minimization of

$$\chi_\nu^2 = \chi_{bW}^2 + \chi_W^2$$

. All three were investigated for their separation values and the χ_W^2 value had the largest separation.

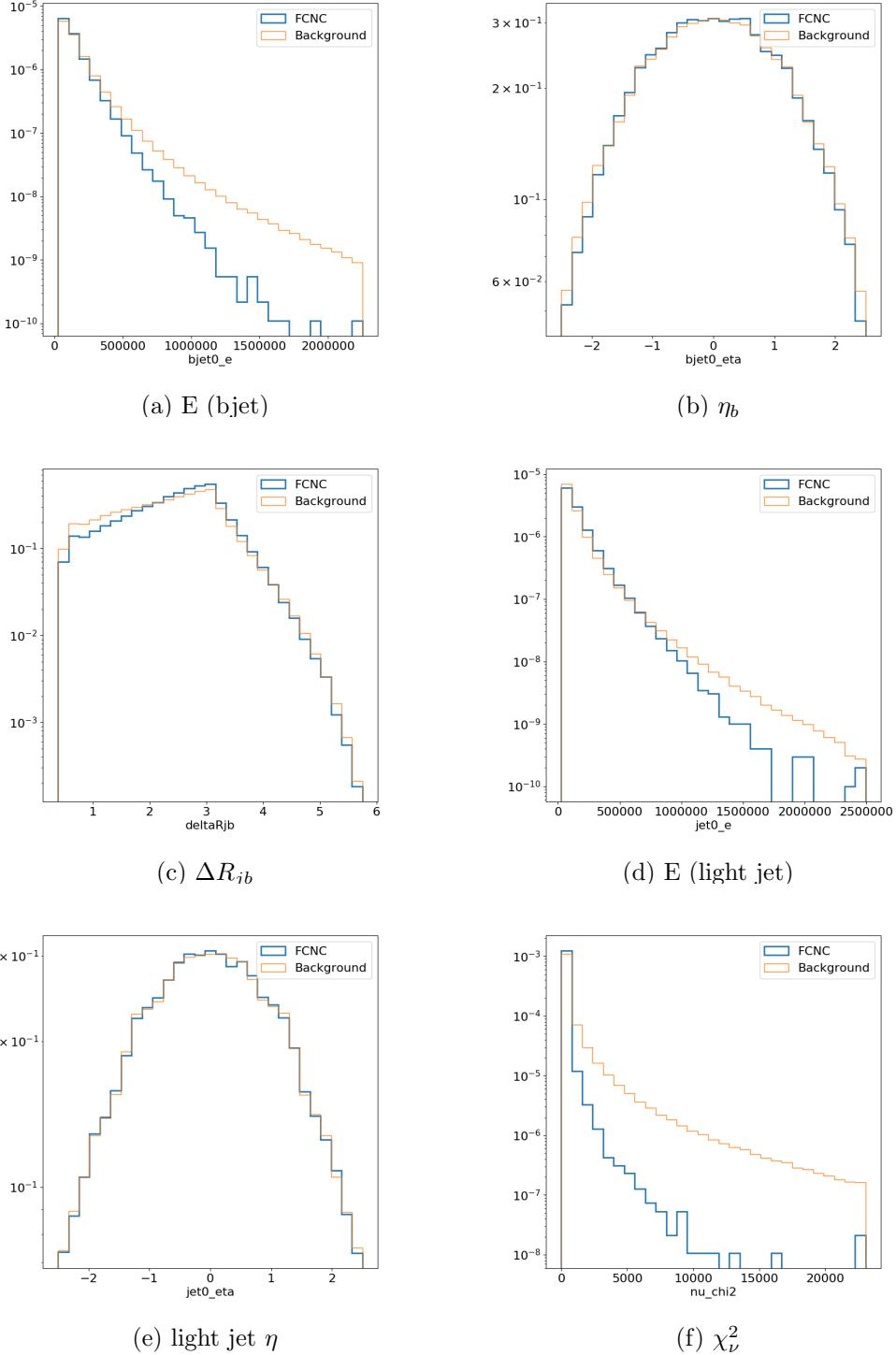


FIGURE C.1. Normalized variables showing the shapes of neural network input variables for the μ +jets channel: [E (bjet), η_b , ΔR_{jb} , E (light jet), light jet η , and χ^2_ν the total χ^2 fit value

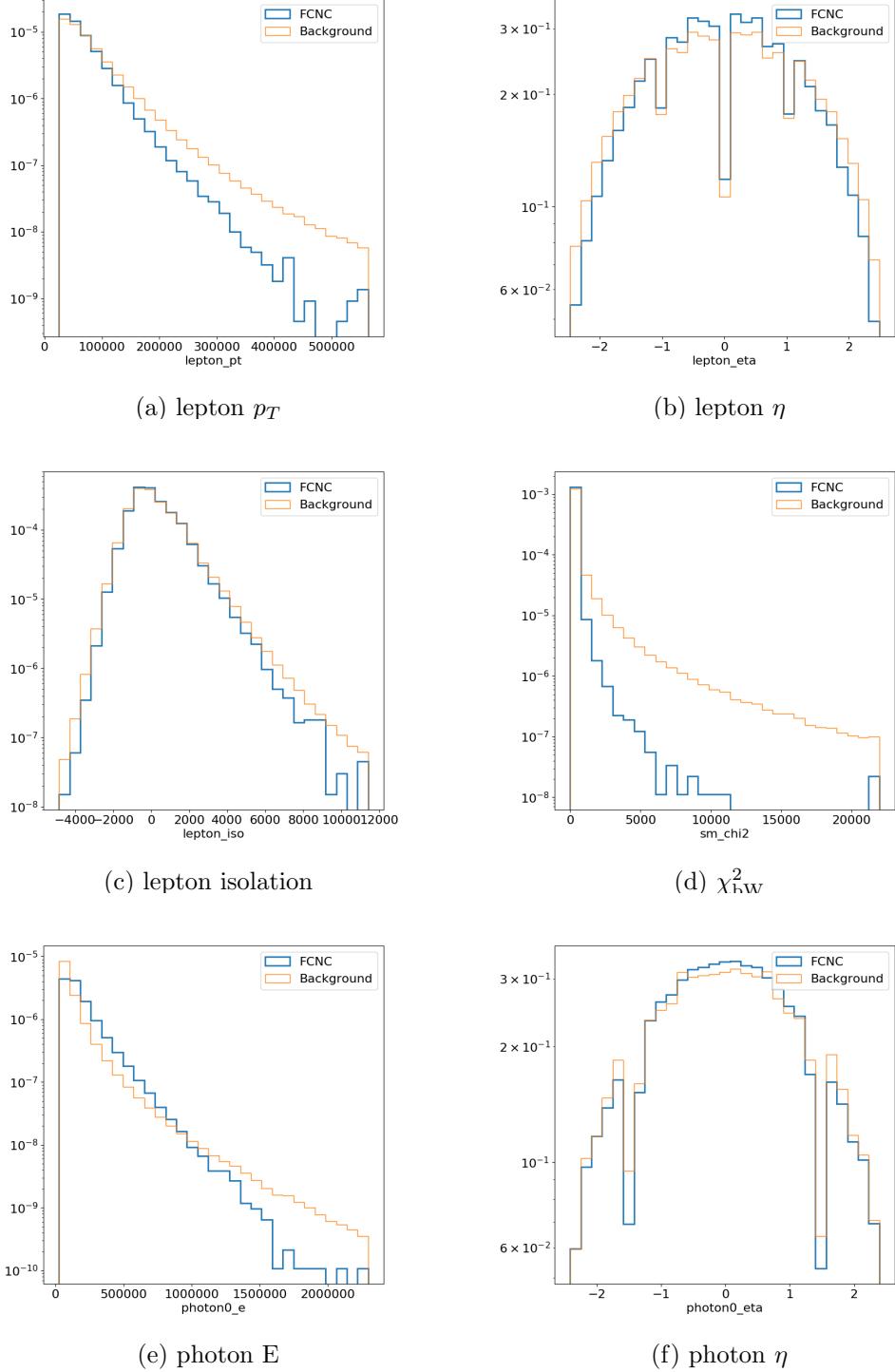


FIGURE C.2. Normalized variables showing the shapes of neural network input variables for the $\mu+jets$ channel: [lepton p_T , lepton η , lepton isolation , χ^2_{bW} the bW χ^2 value from neutrino reconstruction ,photon E, and photon η .

C.2. Shape Comparison Plots: $e+jets$ channel

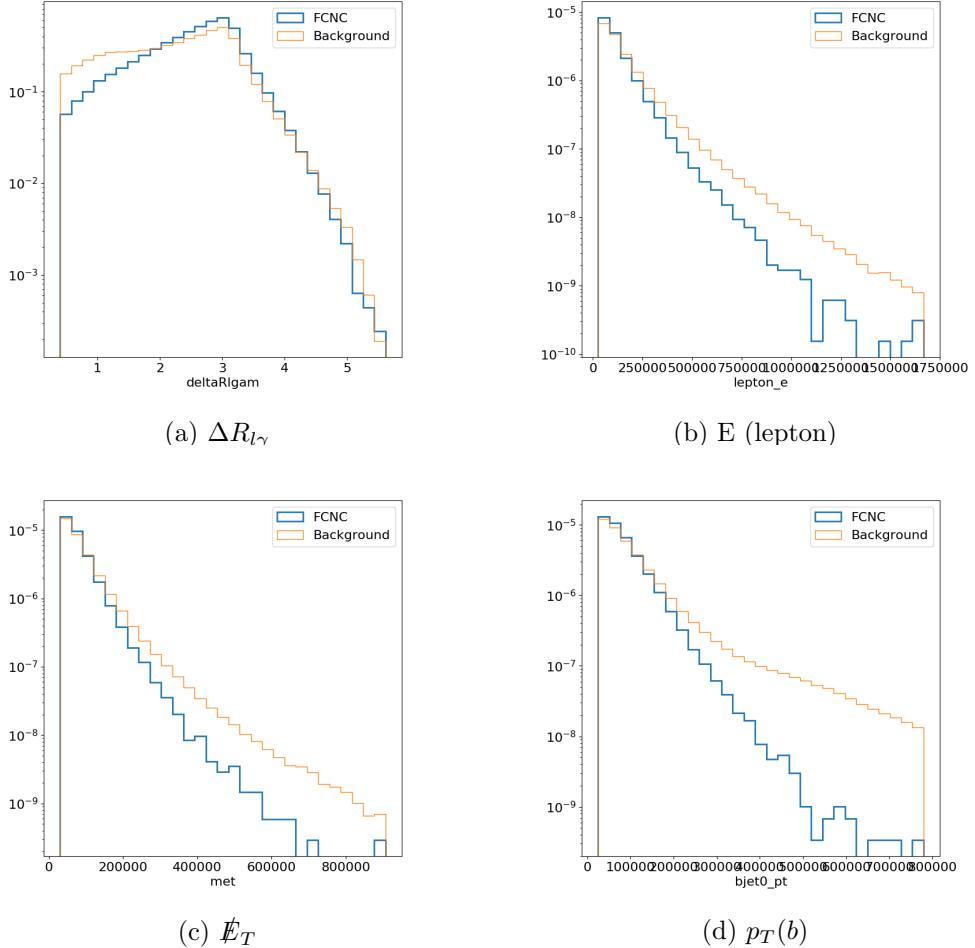


FIGURE C.3. Normalized variables showing the shapes of neural network input variables for the $e+jets$ channel: $\Delta R_{l\gamma}$, E (lepton), \cancel{E}_T , and $p_T(b)$

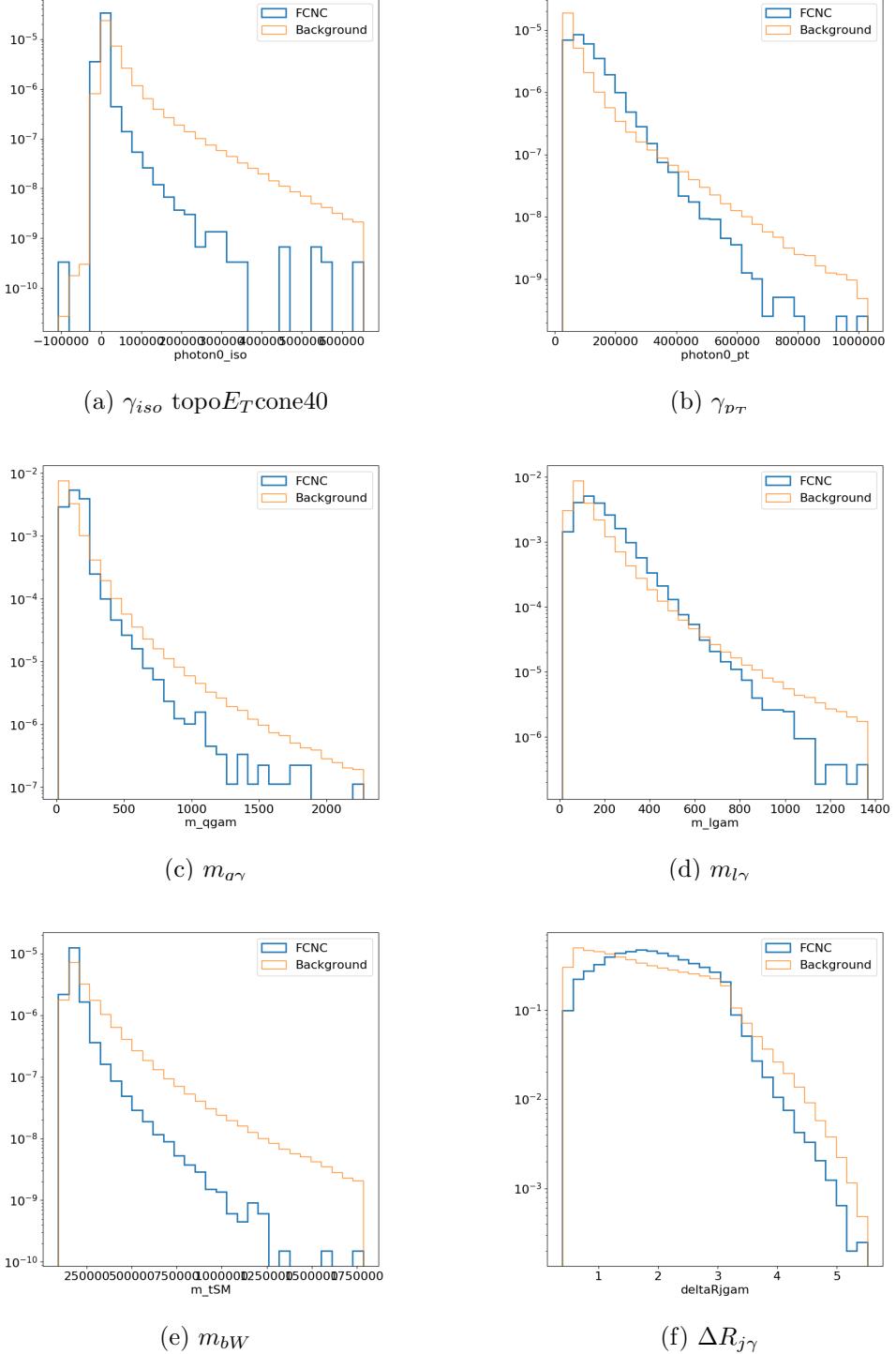


FIGURE C.4. Normalized variables showing the shapes of neural network input variables for the $e+jets$ channel: γ_{iso} topo E_T cone40, γ_{p_T} , $m_{q\gamma}$, $m_{l\gamma}$, m_{bW} , and $\Delta R_{j\gamma}$

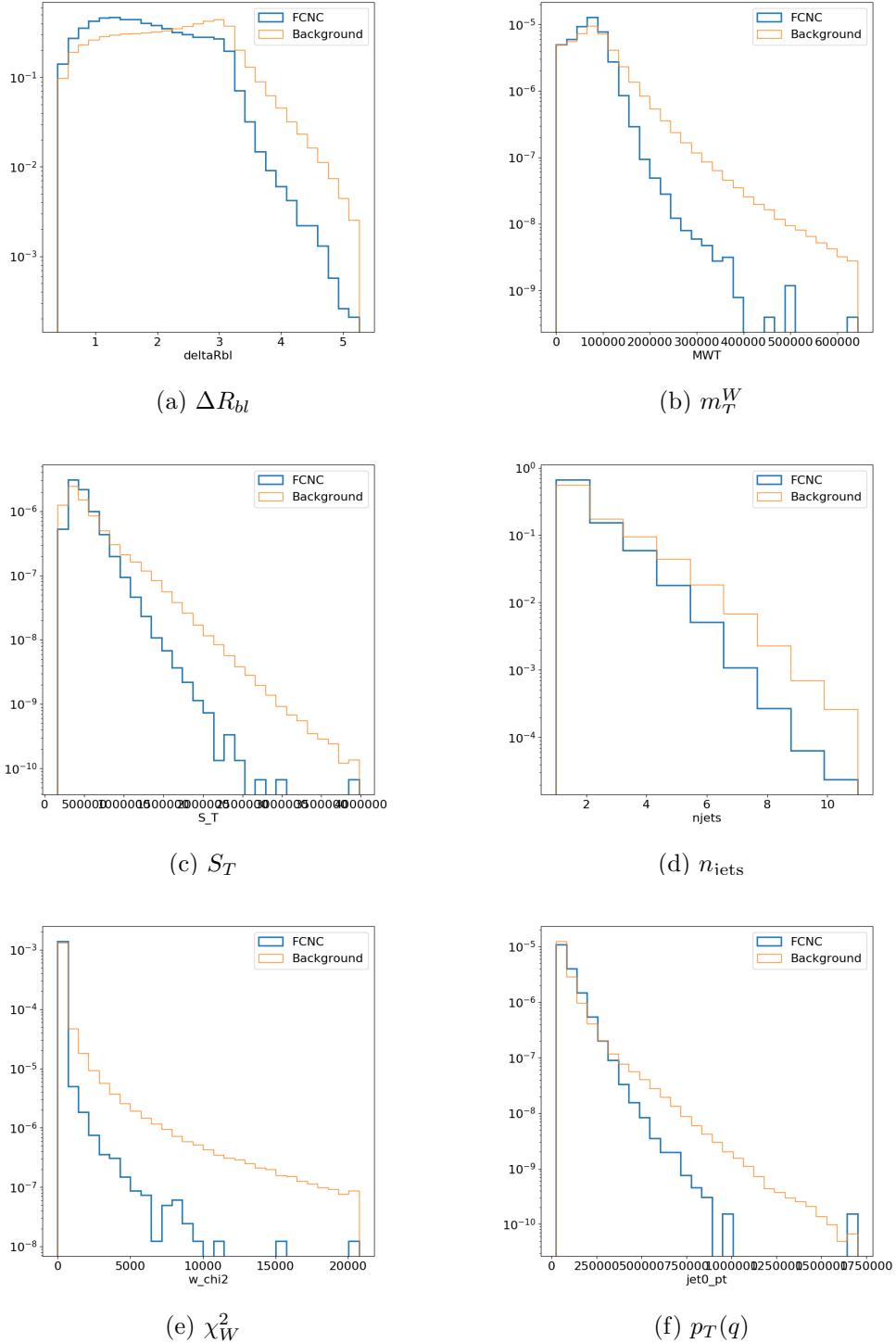


FIGURE C.5. Normalized variables showing the shapes of neural network input variables for the $e+jets$ channel: ΔR_{bl} , m_T^W , S_T , n_{jets} , χ_W^2 , and $p_T(q)$

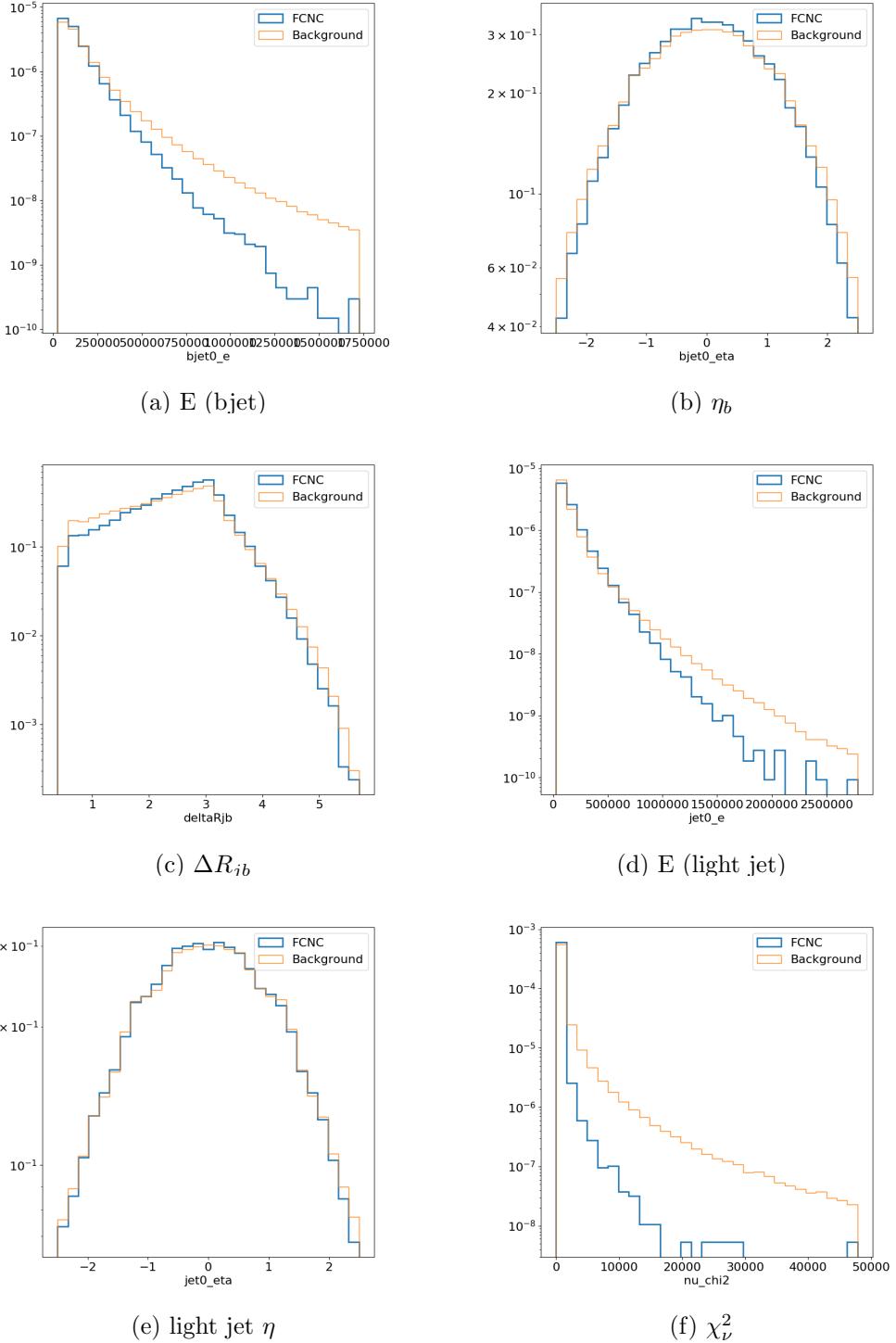
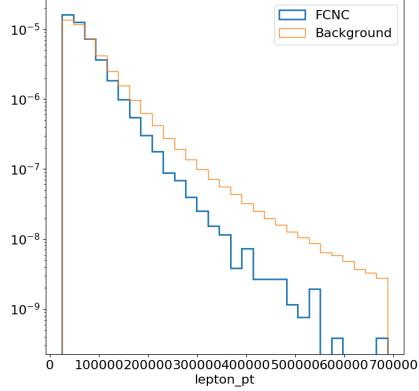
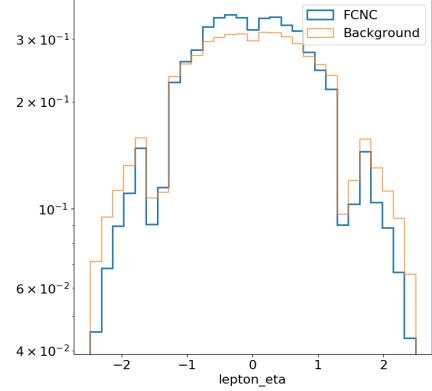


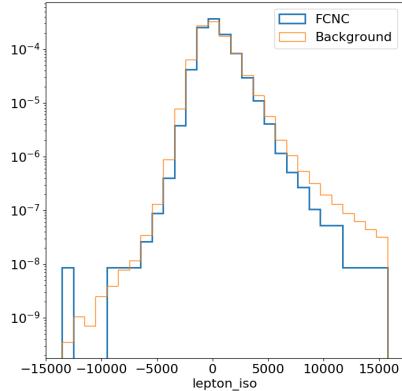
FIGURE C.6. Normalized variables showing the shapes of neural network input variables for the $e+jets$ channel: [E (bjet), η_b , ΔR_{jb} , E (light jet), light jet η , and χ^2 the total χ^2 fit value



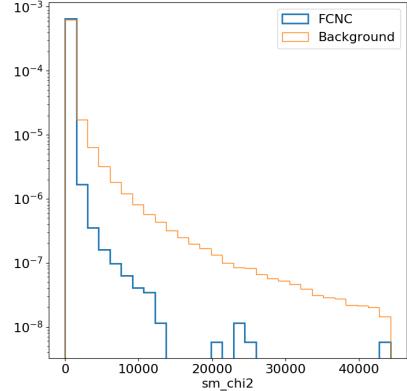
(a) lepton p_T



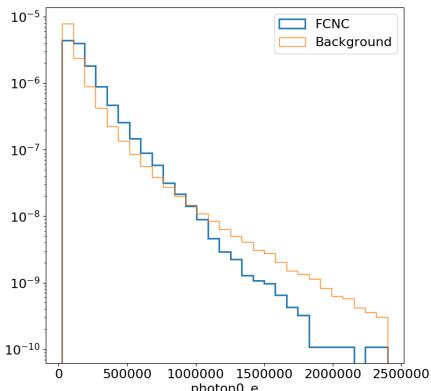
(b) lepton η



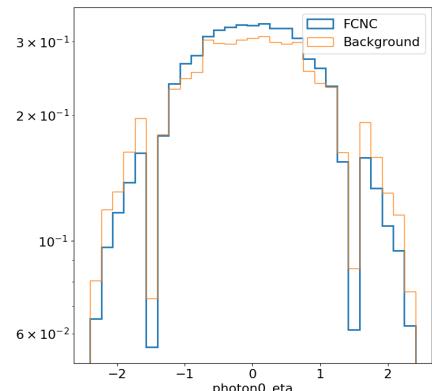
(c) lepton isolation



(d) χ^2_{bW}



(e) photon E



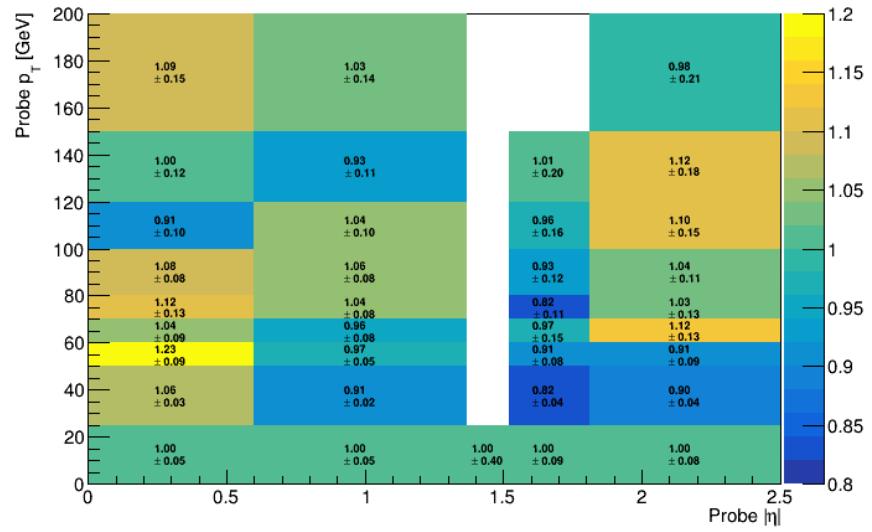
(f) photon η

FIGURE C.7. Normalized variables showing the shapes of neural network input variables for the $e+jets$ channel: [lepton p_T , lepton η , lepton isolation, χ^2_{bW} the $bW\chi^2$ value from neutrino reconstruction, photon E, and photon η .

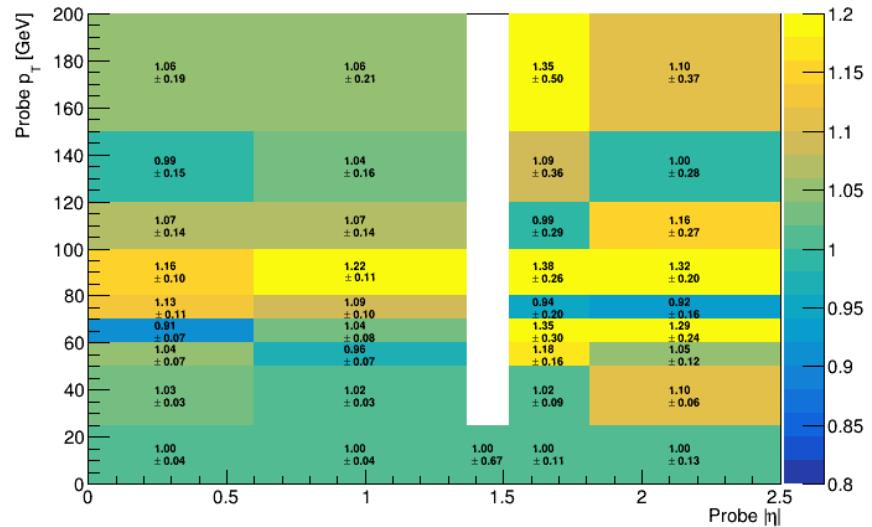
APPENDIX D

SYSTEMATIC VARIATIONS FOR ELECTRON TO PHOTON FAKE RATE CALCULATION

This appendix shows variation plots for the systematic uncertainty estimation on the fake rates for photons faked by electrons using various m_{ll} invariant mass windows around the Z boson mass (5 GeV, 15 GeV, and 20 GeV window sizes). From values and deviations from the nominal a systematic uncertainty was estimated which is then applied to events with a photon which is actually a truth electron.

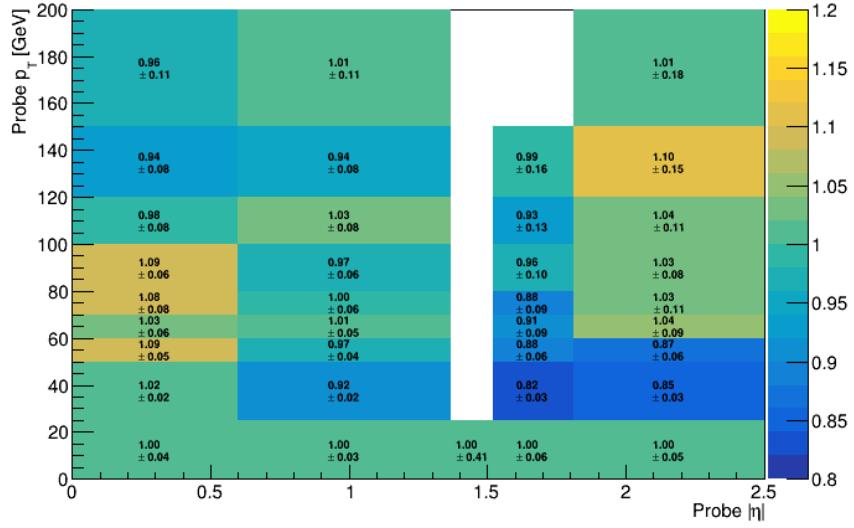


(a) Converted Photons

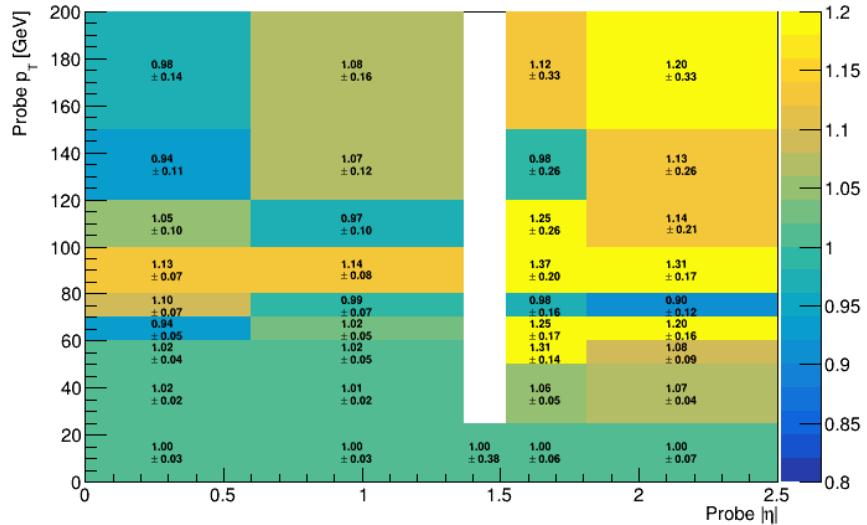


(b) Unconverted Photons

FIGURE D.1. 2-Dimensional scale factors derived using the $Z \rightarrow e^+e^-$ events within a 5 GeV mass window around the Z boson (2.5 GeV on each side) for unconverted and converted photon types shown with statistical uncertainties.

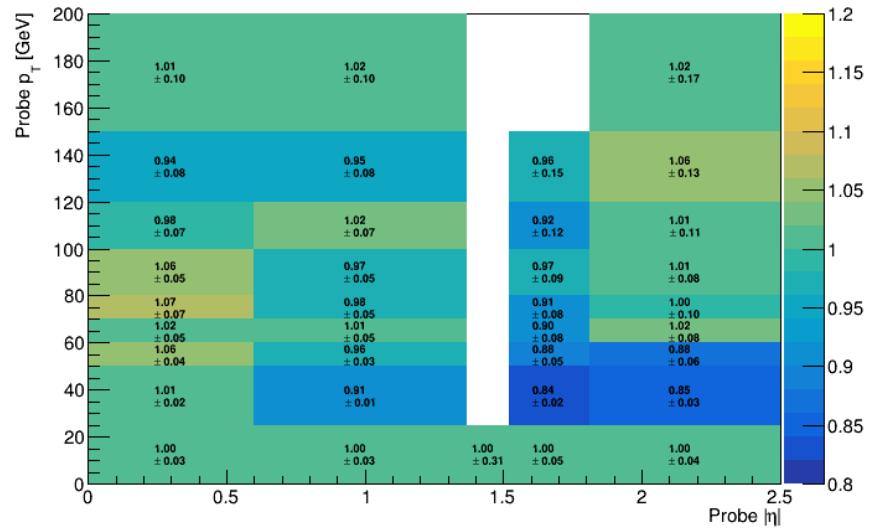


(a) Converted Photons

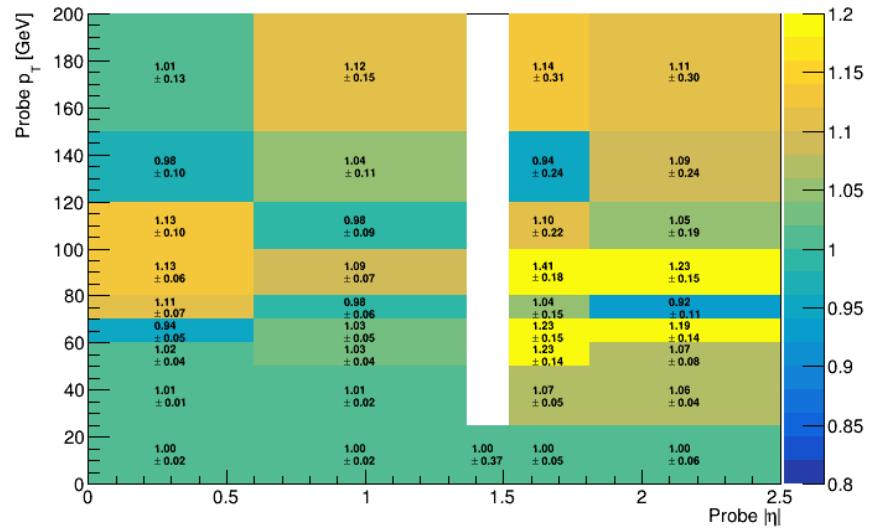


(b) Unconverted Photons

FIGURE D.2. 2-Dimensional scale factors derived using the $Z \rightarrow e^+e^-$ events within a 15 GeV mass window around the Z boson (7.5 GeV on each side) for unconverted and converted photon types shown with statistical uncertainties.



(a) Converted Photons



(b) Unconverted Photons

FIGURE D.3. 2-Dimensional scale factors derived using the $Z \rightarrow e^+e^-$ events within a 20 GeV mass window around the Z boson (10 GeV on each side) for unconverted and converted photon types shown with statistical uncertainties.

APPENDIX E
SYSTEMATIC PULL PLOTS

Test

REFERENCES CITED

- [1] S. L. Glashow. Partial Symmetries of Weak Interactions. *Nucl. Phys.*, 22:579–588, 1961.
- [2] Peter W. Higgs. Broken Symmetries and the Masses of Gauge Bosons. *Phys. Rev. Lett.*, 13:508–509, 1964. [,160(1964)].
- [3] F. Englert and R. Brout. Broken Symmetry and the Mass of Gauge Vector Mesons. *Phys. Rev. Lett.*, 13:321–323, 1964. [,157(1964)].
- [4] Steven Weinberg. A Model of Leptons. *Phys. Rev. Lett.*, 19:1264–1266, 1967.
- [5] Abdus Salam. Weak and Electromagnetic Interactions. *Conf. Proc.*, C680519:367–377, 1968.
- [6] Georges Aad et al. Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC. *Phys. Lett.*, B716:1–29, 2012.
- [7] Serguei Chatrchyan et al. Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC. *Phys. Lett.*, B716:30–61, 2012.
- [8] F. Abe et al. Observation of top quark production in $\bar{p}p$ collisions. *Phys. Rev. Lett.*, 74:2626–2631, 1995.
- [9] The ATLAS collaboration. Measurement of the top quark mass in the $t\bar{t} \rightarrow \text{lepton+jets}$ channel from $\sqrt{s}=8$ TeV ATLAS data. 2017.
- [10] M. Tanabashi et al. Review of Particle Physics. *Phys. Rev.*, D98(3):030001, 2018.
- [11] J. A. Aguilar-Saavedra. Top flavor-changing neutral interactions: Theoretical expectations and experimental detection. *Acta Phys. Polon.*, B35:2695–2710, 2004.
- [12] J. J. Cao, G. Eilam, M. Frank, K. Hikasa, G. L. Liu, I. Turan, and J. M. Yang. SUSY-induced FCNC top-quark processes at the large hadron collider. *Phys. Rev.*, D75:075021, 2007.
- [13] Junjie Cao, Zhaoxia Heng, Lei Wu, and Jin Min Yang. R-parity violating effects in top quark FCNC productions at LHC. *Phys. Rev.*, D79:054003, 2009.
- [14] G. Eilam, J. L. Hewett, and A. Soni. Rare decays of the top quark in the standard and two Higgs doublet models. *Phys. Rev.*, D44:1473–1484, 1991. [Erratum: Phys. Rev.D59,039901(1999)].

- [15] Yuval Grossman. Introduction to flavor physics. In *Flavianet School on Flavour Physics Karlsruhe, Germany, September 7-18, 2009*, pages 111–144, 2014. [,73(2014)].
- [16] Nicola Cabibbo. Unitary Symmetry and Leptonic Decays. *Phys. Rev. Lett.*, 10:531–533, 1963. [,648(1963)].
- [17] Makoto Kobayashi and Toshihide Maskawa. CP Violation in the Renormalizable Theory of Weak Interaction. *Prog. Theor. Phys.*, 49:652–657, 1973.
- [18] Lincoln Wolfenstein. Parametrization of the Kobayashi-Maskawa Matrix. *Phys. Rev. Lett.*, 51:1945, 1983.
- [19] S. L. Glashow, J. Iliopoulos, and L. Maiani. Weak Interactions with Lepton-Hadron Symmetry. *Phys. Rev.*, D2:1285–1292, 1970.
- [20] S. W. Herb et al. Observation of a Dimuon Resonance at 9.5-GeV in 400-GeV Proton-Nucleus Collisions. *Phys. Rev. Lett.*, 39:252–255, 1977.
- [21] S. Abachi et al. Search for high mass top quark production in $p\bar{p}$ collisions at $\sqrt{s} = 1.8$ TeV. *Phys. Rev. Lett.*, 74:2422–2426, 1995.
- [22] LHC Top Working Group. Lhctopwg summary plots, 2019.
- [23] V. M. Abazov et al. A precision measurement of the mass of the top quark. *Nature*, 429:638–642, 2004.
- [24] Markus Cristinziani and Martijn Mulders. Top-quark physics at the Large Hadron Collider. *J. Phys.*, G44(6):063001, 2017.
- [25] David Atwood, Laura Reina, and Amarjit Soni. Phenomenology of two Higgs doublet models with flavor changing neutral currents. *Phys. Rev.*, D55:3156–3176, 1997.
- [26] J. A. Aguilar-Saavedra and B. M. Nobre. Rare top decays $t \rightarrow c$ gamma, $t \rightarrow cg$ and CKM unitarity. *Phys. Lett.*, B553:251–260, 2003.
- [27] J. A. Aguilar-Saavedra. Effects of mixing with quark singlets. *Phys. Rev.*, D67:035003, 2003. [Erratum: Phys. Rev.D69,099901(2004)].
- [28] Jin Min Yang, Bing-Lin Young, and X. Zhang. Flavor changing top quark decays in r parity violating SUSY. *Phys. Rev.*, D58:055001, 1998.
- [29] Kaustubh Agashe, Gilad Perez, and Amarjit Soni. Collider Signals of Top Quark Flavor Violation from a Warped Extra Dimension. *Phys. Rev.*, D75:015002, 2007.
- [30] Mark Trodden. Electroweak baryogenesis. *Rev. Mod. Phys.*, 71:1463–1500, 1999.

- [31] G. C. Branco, P. M. Ferreira, L. Lavoura, M. N. Rebelo, Marc Sher, and Joao P. Silva. Theory and phenomenology of two-Higgs-doublet models. *Phys. Rept.*, 516:1–102, 2012.
- [32] Andreas Crivellin and Ulrich Nierste. Chirally enhanced corrections to FCNC processes in the generic MSSM. *Phys. Rev.*, D81:095007, 2010.
- [33] Georges Aad et al. Search for flavour-changing neutral currents in processes with one top quark and a photon using 81 fb^{-1} of pp collisions at $\sqrt{s} = 13 \text{ TeV}$ with the ATLAS experiment. 2019.
- [34] Fabian Kuger. Signal formation processes in micromegas detectors and quality control for large size detector construction for the atlas new small wheel. 08 2017.
- [35] Esma Mobs. The CERN accelerator complex. Complex des accélérateurs du CERN. Jul 2016. General Photo.
- [36] The ATLAS Collaboration. Luminosity public results run 2, 2019.
- [37] The ATLAS Experiment at the CERN Large Hadron Collider. *J. Instrum.*, 3:S08003. 437 p, 2008.
- [38] M Capeans, G Darbo, K Einsweiller, M Elsing, T Flick, M Garcia-Sciveres, C Gemme, H Pernegger, O Rohne, and R Vuillermet. ATLAS Insertable B-Layer Technical Design Report. Technical Report CERN-LHCC-2010-013. ATLAS-TDR-19, Sep 2010.
- [39] *ATLAS inner detector: Technical Design Report*, 1. Technical Design Report ATLAS. CERN, Geneva, 1997.
- [40] Yosuke Takubo. ATLAS IBL operational experience. *PoS*, Vertex2016:004, 2017.
- [41] Karolos Potamianos. The upgraded Pixel detector and the commissioning of the Inner Detector tracking of the ATLAS experiment for Run-2 at the Large Hadron Collider. *PoS*, EPS-HEP2015:261, 2015.
- [42] Bartosz Mindur. ATLAS Transition Radiation Tracker (TRT): Straw tubes for tracking and particle identification at the Large Hadron Collider. *Nucl. Instrum. Meth.*, A845:257–261, 2017.
- [43] *ATLAS liquid-argon calorimeter: Technical Design Report*. Technical Design Report ATLAS. CERN, Geneva, 1996.
- [44] *ATLAS tile calorimeter: Technical Design Report*. Technical Design Report ATLAS. CERN, Geneva, 1996.

- [45] ATLAS Outreach. ATLAS Fact Sheet : To raise awareness of the ATLAS detector and collaboration on the LHC. 2010.
- [46] Moritz Backes. The ATLAS Trigger System: Ready for Run-2. *PoS*, LeptonPhoton2015:045, 2016.
- [47] Chaowaroj Wanotayaroj and Jim Brau. Search for a Scalar Partner of the Top Quark in the Jets+MET Final State with the ATLAS detector, Nov 2016. Presented 25 Oct 2016.
- [48] S. Agostinelli et al. GEANT4: A Simulation toolkit. *Nucl. Instrum. Meth.*, A506:250–303, 2003.
- [49] J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, O. Mattelaer, H. S. Shao, T. Stelzer, P. Torrielli, and M. Zaro. The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations. *JHEP*, 07:079, 2014.
- [50] Stefano Frixione, Paolo Nason, and Carlo Oleari. Matching NLO QCD computations with Parton Shower simulations: the POWHEG method. *JHEP*, 11:070, 2007.
- [51] Simone Alioli, Paolo Nason, Carlo Oleari, and Emanuele Re. A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX. *JHEP*, 06:043, 2010.
- [52] Torbjorn Sjostrand, Stephen Mrenna, and Peter Z. Skands. A Brief Introduction to PYTHIA 8.1. *Comput. Phys. Commun.*, 178:852–867, 2008.
- [53] T. Gleisberg, Stefan. Hoeche, F. Krauss, M. Schonherr, S. Schumann, F. Siegert, and J. Winter. Event generation with SHERPA 1.1. *JHEP*, 02:007, 2009.
- [54] Enrico Bothmann et al. Event Generation with Sherpa 2.2. *SciPost Phys.*, 7:034, 2019.
- [55] Johan Alwall et al. A Standard format for Les Houches event files. *Comput. Phys. Commun.*, 176:300–304, 2007.
- [56] Celine Degrande, Claude Duhr, Benjamin Fuks, David Grellscheid, Olivier Mattelaer, and Thomas Reiter. UFO - The Universal FeynRules Output. *Comput. Phys. Commun.*, 183:1201–1214, 2012.
- [57] R. A. Coimbra, P. M. Ferreira, R. B. Guedes, O. Oliveira, A. Onofre, R. Santos, and Miguel Won. Dimension six FCNC operators and top production at the LHC. *Phys. Rev.*, D79:014006, 2009.

- [58] B. Grzadkowski, M. Iskrzynski, M. Misiak, and J. Rosiek. Dimension-Six Terms in the Standard Model Lagrangian. *JHEP*, 10:085, 2010.
- [59] J. A. Aguilar-Saavedra. A Minimal set of top anomalous couplings. *Nucl. Phys.*, B812:181–204, 2009.
- [60] Zenro Hioki, Kazumasa Ohkuma, and Akira Uejima. Effective Lagrangian Approach to Top Decay via Flavor Changing Neutral Current. In *29th International Conference on Lepton and Photon Interactions (LP2019) Toronto, Ontario, Canada, August 5-10, 2019*, 2019.
- [61] Artur Amorim, Juan Pedro Araque, Nuno Filipe Castro, Jose Santiago, and Rui Santos. Production of $t\gamma$, tZ and tH via Flavour Changing Neutral Currents. *PoS*, TOP2015:067, 2016.
- [62] Celine Degrande, Fabio Maltoni, Jian Wang, and Cen Zhang. Automatic computations at next-to-leading order in QCD for top-quark flavor-changing neutral processes. *Phys. Rev.*, D91:034024, 2015.
- [63] Joao Pequenao. Event Cross Section in a computer generated image of the ATLAS detector. Mar 2008.
- [64] Morad Aaboud et al. Electron reconstruction and identification in the ATLAS experiment using the 2015 and 2016 LHC proton-proton collision data at $\sqrt{s} = 13$ TeV. *Eur. Phys. J.*, C79(8):639, 2019.
- [65] Electron and photon energy calibration with the ATLAS detector using LHC Run 1 data. *Eur. Phys. J.*, C74(10):3071, 2014.
- [66] ATLAS Collaboration. Electron and photon energy calibration with the ATLAS detector using data collected in 2015 at $\sqrt{s} = 13$ TeV, Aug 2016.
- [67] Sébastien Rettie. Muon identification and performance in the ATLAS experiment. *PoS*, DIS2018:097, 2018.
- [68] Morad Aaboud et al. Measurement of the photon identification efficiencies with the ATLAS detector using LHC Run 2 data collected in 2015 and 2016. *Eur. Phys. J.*, C79(3):205, 2019.
- [69] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. The anti- k_t jet clustering algorithm. *JHEP*, 04:063, 2008.
- [70] M. Wobisch and T. Wengler. Hadronization corrections to jet cross-sections in deep inelastic scattering. In *Monte Carlo generators for HERA physics. Proceedings, Workshop, Hamburg, Germany, 1998-1999*, pages 270–279, 1998.

- [71] Selection of jets produced in 13TeV proton-proton collisions with the ATLAS detector. Technical Report ATLAS-CONF-2015-029, CERN, Geneva, Jul 2015.
- [72] The ATLAS collaboration. Tagging and suppression of pileup jets. 2014.
- [73] S. Schramm. Howtocleanjets2016 : Jet-based event cleaning, 2016.
- [74] Albert M Sirunyan et al. Measurement of b hadron lifetimes in pp collisions at $\sqrt{s} = 8$ TeV. *Eur. Phys. J.*, C78(6):457, 2018. [Erratum: Eur. Phys. J.C78,no.7,561(2018)].
- [75] N. Bartosik. File:b-tagging diagram.png, 2016.
- [76] Francesco Armando Di Bello. Optimisation of the ATLAS *b*-tagging algorithms for the 2017-2018 LHC data-taking. *PoS*, EPS-HEP2017:733, 2017.
- [77] Morad Aaboud et al. Measurements of b-jet tagging efficiency with the ATLAS detector using $t\bar{t}$ events at $\sqrt{s} = 13$ TeV. *JHEP*, 08:089, 2018.
- [78] Morad Aaboud et al. Performance of missing transverse momentum reconstruction with the ATLAS detector using proton-proton collisions at $\sqrt{s} = 13$ TeV. *Eur. Phys. J.*, C78(11):903, 2018.
- [79] F. Chollet et al. Keras: The python deep learning library, 2019.
- [80] Martín Abadi et al. TensorFlow: Large-Scale Machine Learning on Heterogeneous Distributed Systems. 2016.
- [81] Diederik P. Kingma and Jimmy Ba. Adam: A Method for Stochastic Optimization. 2014.
- [82] A. Karpathy. Cs231n convolutional neural networks for visual recognition, 2015.
- [83] Fabian Pedregosa et al. Scikit-learn: Machine Learning in Python. *J. Machine Learning Res.*, 12:2825–2830, 2011.
- [84] The ATLAS collaboration. Estimation of non-prompt and fake lepton backgrounds in final states with top quarks produced in proton-proton collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector. 2014.
- [85] ATLAS Collaboration. Nnlo+nnll top-quark-pair cross sections, 2015.
- [86] The ATLAS collaboration. Measurements of inclusive and differential cross-sections of $t\bar{t}\gamma$ production in leptonic final states in a fiducial volume at $\sqrt{s} = 13$ TeV in ATLAS. 2018.
- [87] ATLAS Collaboration. Nlo single-top channel cross sections, 2017.

- [88] ATLAS Collaboration. Pmgweakbosonprocesses, 2019.
- [89] Luminosity determination in pp collisions at $\sqrt{s} = 13$ TeV using the ATLAS detector at the LHC. Technical Report ATLAS-CONF-2019-021, CERN, Geneva, Jun 2019.
- [90] G. Avoni et al. The new lucid-2 detector for luminosity measurement and monitoring in atlas. *JINST*, 13(07):P07017, 2018.
- [91] Arthur Alexis Jules Lesage. Lepton and photon performance at ATLAS and CMS. In *5th Large Hadron Collider Physics Conference (LHCP 2017) Shanghai, China, May 15-20, 2017*, 2017.
- [92] Jet Calibration and Systematic Uncertainties for Jets Reconstructed in the ATLAS Detector at $\sqrt{s} = 13$ TeV. Technical Report ATL-PHYS-PUB-2015-015, CERN, Geneva, Jul 2015.
- [93] J H Friedman. Data analysis techniques for high energy particle physics. (SLAC-176):96 p, Oct 1974.
- [94] ATLAS Collaboration. Trexfitter (previously known as tthfitter), 2019.
- [95] Lorenzo Moneta, Kevin Belasco, Kyle S. Cranmer, S. Kreiss, Alfio Lazzaro, Danilo Piparo, Gregory Schott, Wouter Verkerke, and Matthias Wolf. The RooStats Project. *PoS*, ACAT2010:057, 2010.
- [96] Luca Lista. Practical Statistics for Particle Physicists. In *Proceedings, 2016 European School of High-Energy Physics (ESHEP2016): Skeikampen, Norway, June 15-28 2016*, pages 213–258, 2017.
- [97] Glen Cowan, Kyle Cranmer, Eilam Gross, and Ofer Vitells. Asymptotic formulae for likelihood-based tests of new physics. *Eur. Phys. J.*, C71:1554, 2011. [Erratum: Eur. Phys. J.C73,2501(2013)].
- [98] Alexander L. Read. Presentation of search results: The CL(s) technique. *J. Phys.*, G28:2693–2704, 2002. [,11(2002)].
- [99] S. S. Wilks. The Large-Sample Distribution of the Likelihood Ratio for Testing Composite Hypotheses. *Annals Math. Statist.*, 9(1):60–62, 1938.
- [100] Thomas Junk. Confidence level computation for combining searches with small statistics. *Nucl. Instrum. Meth.*, A434:435–443, 1999.