Searches for new high-mass resonances in top-antitop and di-electron final states using the ATLAS detector

by

Elham E Khoda

BS.-MS., Indian Institute of Science Education and Research Kolkata, 2015

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

in

The Faculty of Graduate and Postdoctoral Studies

(Physics)

THE UNIVERSITY OF BRITISH COLUMBIA

(Vancouver)

March 2021

The following individuals certify that they have read, and recommend to the Faculty of Graduate and Postdoctoral Studies for acceptance, the dissertation entitled:

Searches for new high-mass resonances in top-antitop and di-electron final states using the ATLAS detector

submitted by	Elham E Khoda	in partial fulfillment of the requirements for
the degree of	Doctor of Philosophy	
in	Physics	

Examining Committee:

Alison Lister, Physics & Astronomy, UBC

Supervisor

Colin Gay, Physics & Astronomy, UBC

Supervisory Committee Member

Janis McKenna, Physics & Astronomy, UBC

University Examiner

Alexandre Bouchard, Statistics, UBC

University Examiner

Freya Blekman, Physics, Vrije Universiteit Brussel

External Examiner

Additional Supervisory Committee:

Gary Hinshaw, Physics & Astronomy, UBC

Supervisory Committee Member

Gordon Semenoff, Physics & Astronomy, UBC

Supervisory Committee Member

Abstract

The standard model of particle physics (SM) describes all the fundamental particles and their interactions. It is a very successful theory; however, many experimental observations - such as the origin of neutrino mass, particle origin of dark matter, e.t.c. - are either not consistent or not explained by the SM. So, it is inevitable that there has to be a (physics) model beyond the SM which will consistently explain all these observations, not covered by the SM. Several of such extensions predict new heavy particles that can interact with SM particles. This dissertation presents searches for the resonant production of such high-mass particles in dielectron and top-antitop final states. These searches use proton-proton collision data at the center-of-mass energy of 13 TeV collected by the ATLAS detector at the Large Hadron Collider (LHC) between 2015 and 2018. Electrons are stable and easy to reconstruct, but top-quarks decay instantaneously. Two dominant top-decay final states, all-hadronic and semi-leptonic, are studied in this dissertation. The combined mass distributions of all the final-state particles are used to perform model-dependent and model-independent statistical searches. No evidence for the existence of new particles is found in any of the explored final states. Hence, upper limits on production cross-section times branching ratio and lower limits on the mass of heavy Z' particles, predicted by the BSM models, are placed at a 95% confidence level. The dilepton resonance search excludes Z' boson below 3.6 TeV. The resonance search in the boosted all-hadronic top-antitop final state excludes Z' bosons with a mass lower than 4.1 TeV. Whereas in the semi-leptonic search, the same signal is expected to be excluded up to 3.6 TeV.

The dissertation also presents a new algorithm for splitting the merged charge clusters in the ATLAS pixel detector, based on a Mixture Density Network (MDN). The performance of this new algorithm is found to be better than the existing algorithm. As a result, the MDN-based algorithm is expected to be used as a default algorithm in ATLAS during the next data collection period, which will start in 2022.

Lay Summary

The Standard Model (SM) of particle physics describes the fundamental particles and their interactions. Although it is a very successful theory, there are many theoretical and experimental limitations to the SM. It only addresses ordinary matter particles and does not include dark matter and dark energy, the largest component of our universe. To solve the problems of the SM, several theories, which include new fundamental particles and interactions, have been proposed. Searches for such new particles decaying either into two electrons or two top-quarks, the heaviest known fundamental particles, are presented in this dissertation by analyzing the data collected during 2015-2018 by the ATLAS experiment at the Large Hadron Collider. Although no evidence for the existence of new particles is found we can put very high constraints on some of the new physics scenarios. Constraints derived in this dissertation improved significantly compared to the previous ones, and they will help to streamline the future searches of new physics.

Preface

This dissertation is based on research done with the ATLAS detector. The ATLAS collaboration has over 3000 research scientists and comprising of graduate students, postdocs, research scientists and professors. The operation of the experiment and collection of data is possible because thanks to the collective effort of all the members of the collaboration. For example, physics scattering processes are simulated by collaboration members, often with the help of experts outside the collaboration. The 'physics objects' such as electrons, muons, jets are reconstructed using algorithms and software developed by the collaboration members for everyone's use. The analyses included in this dissertation use software jointly developed by many members of the collaboration and rely on shared computing resources. This dissertation would not have possible without the effort of all the members.

Figures presented in this dissertation falls into different categories. Figures with "ATLAS", "ATLAS Simulation", "ATLAS Preliminary", "ATLAS Simulation Preliminary" or "ATLAS Online" labels are public results reviewed and approved by the collaborations. The figures with "ATLAS" or "ATLAS Simulation" are published in peer-reviewed journals. On the other hand, figures with the "ATLAS Work in Progress" label or without any label are not approved by the collaboration and are prepared by me. The sources of the figures which are not prepared by me or are directly not part of my research are mentioned in the figure captions.

My work is summarized in chapter 3, chapter 6, chapter 7, and chapter 8. These led to two journal publications, one conference proceeding, one conference note and one publication note. The work presented in chapter 8 is not public yet and the publication is in progress. My contribution work and relevant contributions of all the collaborators are listed below.

ATLAS pixel cluster splitting with Mixture Density Network is presented in chapter 3. It was done with the guidance of W. Fedorko and A. Lister. In the beginning, L. Gagnon provided lots of support by providing the existing neural network training code and other scripts. Several useful inputs were obtained from B. Nachman, G. Louppe, A. Kastanas, R. Owen, P. Butti, B. Murray, G. Facini, N. Styles, M. Danninger, N. Pettersson at different stages of the study. I was the only student working on this problem. I developed the training setup and strategy with the help of W. Fedorko. All the relevant code for developing the mixture density network was written by me. The results are public in my LHC Physics conference (2019) proceeding:

E.Khoda, *ATLAS pixel cluster splitting using Mixture Density Networks*
PoS LHCP2019 (2019) 009, LHCP Conference Proceedings (2019)

Di-lepton resonance search in the electron channel is presented in chapter 6 using 36 fb^{-1} proton-proton collision data collected with the ATLAS detector. This project was carried out by a team of around 20 active people. I was working with W. Fedorko from the UBC group. I was one

of the main analyzers in the electron channel along with C. Willis. My colleague from the UBC group S. Rettie was working in the muon final state, and it was helpful to discuss with him to get the full picture. I was involved in several aspects of the analysis.

- I developed the analysis code for electron channel with the help of D. Hayden.
- Together with C. Willis, I optimized the event selection.
- I estimated the background and make the distribution of the kinematic variables.
- The effects of the systematic uncertainties on the main discriminant was estimated by me.
- I performed the model-independent search with BumpHunter algorithm.
- With the help of W. Fedorko, I developed the strategy to include systematics uncertainties in the BumpHunter search algorithm.

Final and intermediate results were made public through one conference note and a journal publication.

ATLAS Collaboration, *Search for new high-mass resonances in the dilepton final state using proton-proton collisions at* $\sqrt{s} = 13$ *TeV with the ATLAS detector*
ATLAS-CONF-2016-045, Public Conference Note (2016) for the International Conference on High Energy Physics.

ATLAS Collaboration, *Search for new high-mass phenomena in the dilepton final state using* 36 fb^{-1} *proton-proton collision data at* $\sqrt{s} = 13$ *TeV with the ATLAS detector*
Journal of High Energy Physics, 10, 182 (2017).

Searches for top-antitop resonances in the boosted all-hadronic final state is presented in chapter 7. This search is performed using 139 fb^{-1} proton-proton collision data collected with the ATLAS detector. The analysis team consists of around 15 members in total. I was one of the main analyzers and was involved in almost every step of the analysis. This work was done together with my supervisor A. Lister. My collaborators who were helped me in this analysis are J. Cantero, Y-H. Chen, A. Dattagupta, S.S. Farahani, T. Ou, K. Terashi, T. Poulsen, and H. Wang.

- I was responsible for maintaining the analysis software framework with Y-H. Chen.
- I ran the full event selection on data and the simulations to select a small subset of data interesting for our search. The work was distributed between Y-H. Chen and me.
- I optimized the event selection and studies the b-tagging algorithms to find the one with the highest sensitivity. Some aspects of the event selections were studied by T. Poulsen.
- The impact of the Deep Neural Network based top-tagger on the $t\bar{t}$ mass reconstruction was studied by me.
- Together with T. Poulsen, I worked on the data-driven background estimation described in this dissertation and the spurious signal tests. The data-driven method is the main way of estimating background in this analysis. I also performed the signal injection test to demonstrate the ability to extract signals from data.

- I estimated the systematic uncertainties on the main discriminant, the mass of the $t\bar{t}$ system in collaboration with Y-H. Chen.
- I developed the strategy for the statistical analysis, to interpret the search results in the context of narrow width Z'_{TC2} resonance. I collaborated with T. Ou at the early stage of this analysis, and T. Poulsen in the later stage. We used the TRExFitter software developed by other members of the collaboration. All the necessary scripts to interface with this software based on our needs were written by me.
- Finally the resonance search results were interpreted in the context of the dark matter mediator model. This work was done in collaboration with T. Poulsen, S. Crepe-Renaudind, K. Pachal.

The search results are published. The dark-matter interpretations are made public in the form of a publication note.

ATLAS Collaboration, *Search for $t\bar{t}$ resonances in fully hadronic final states in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*
Journal of High Energy Physics, 10, 061, (2020)

ATLAS Collaboration, *Dark matter summary plots for s-channel mediators*
ATL-PHYS-PUB-2020-021, Publication Note (2020) for the International Conference on High Energy Physics.

Searches for top-antitop resonances in lepton+jets final state is presented in chapter 8. This search is also performed using 139 fb^{-1} proton-proton collision data collected with the ATLAS detector. This work is still ongoing. This work is being done with my supervisor A. Lister and the analysis team consists around 10 people. My collaborators who are actively working on this project are W. Barbe, S. Batlamous, S. Calvet, S. Crepe-Renaudin, F. Fassi, D. Hayden, K. Krowpman, B-E. Ngair . I am currently the analysis contact and coordinating the analysis activities in the analysis team. Being the analysis contact, I oversee the progress of different small segments of the project and help people if they are facing anhy problem. I am also one of the main analyzers and am involved in different aspects of the analysis.

- I am the only developer and maintainer of the analysis framework. My collaborators S. Calvet, Y-H. Chen, D. E. Ferreira de Lima helped me in the past.
- I optimized the event selection in collaboration with S. Batlamous and F. Fassi.
- I studied how to effectively use the b-tagging and top-tagging algorithms to reduce the dominant QCD and W+jets background.
- The lepton and the jet coming from boosted top decay are very close to each other. So, a non-standard lepton-jet overlap removal process is needed to maintain signal efficiency at high mass. The standard method used in the ATLAS experiment is not efficient for our search. I collaborated with K. Krowpman and M. Pinamonti to study alternative methods to remove the overlap between lepton and jet.
- I studied the isolation method for high momentum electrons and muon with helpful inputs from O. ducu.

- I developed a simplified method to reconstruct $t\bar{t}$ mass using the neural network-based regression method. During the summer of 2020, I supervised an undergraduate student, T. Zhang along with A. Lister to do a more detailed study of $t\bar{t}$ reconstruction.
- I am responsible for interpreting the search result and calculating the cross-section limits.
- I am responsible for writing parts of and coordinating the rest of the internal documentation.

Table of Contents

Abstract . . . . . **iii**

Lay Summary . . . . . **iv**

Preface . . . . . **v**

Table of Contents . . . . . **ix**

List of Tables . . . . . **xv**

List of Figures . . . . . **xvii**

Acknowledgements . . . . . **xxii**

Dedication . . . . . **xxiv**

1 Theoretical background **1**
1.1 Historical introduction to the elementary particle physics 1
1.1.1 Electron, proton and neutron (1897 - 1932) 1
1.1.2 Photon (1900-1926) 2
1.1.3 Mesons (1934 - 1947) 2
1.1.4 Isospin and SU(2) 3
1.1.5 Antiparticles (1930-1956) 3
1.1.6 Neutrinos (1930-2000) 4
1.1.7 Strange particles (1947-1960) 4
1.1.8 The Eightfold Way (1961-1964) 5
1.1.9 The quark model (1964) 5
1.1.10 SU(3) flavor symmetry 6
1.1.11 SU(3) color 6
1.1.12 The fourth quark and more particles 6
1.1.13 Weak interaction and intermediate bosons 7
1.2 Standard model 8
1.2.1 Particles in the standard model 8
1.2.2 Quantum chromodynamics 8
1.2.3 Parton distribution function 9
1.2.4 Electroweak theory 10
1.2.5 Electroweak symmetry breaking 12
1.3 Top physics 16

1.3.1 Top-quark production . . . 16
1.3.2 Top-quark decay . . . 18
1.4 Success and limitations of standard model . . . 19
1.5 Beyond standard model . . . 21
1.5.1 GUT model based on E_6 algebra . . . 22
1.5.2 Sequential Standard Model . . . 22
1.5.3 Topcolor assisted Technicolor model . . . 23
1.5.4 Dark matter mediator models . . . 24

2 Experimental Setup . . . 25
2.1 The Large Hadron Collider . . . 25
2.1.1 The LHC acceleration chain . . . 25
2.1.2 Beam structure . . . 27
2.1.3 Luminosity and pileup . . . 27
2.1.4 LHC Run-2 performance . . . 28
2.2 The ATLAS detector . . . 28
2.2.1 Coordinate system . . . 28
2.2.2 Detector overview . . . 29
2.2.3 ATLAS magnets . . . 30
2.2.4 Inner detector . . . 31
2.2.5 Calorimeter . . . 32
2.2.6 Muon spectrometer . . . 34
2.2.7 Trigger and data acquisition . . . 36
2.2.8 Data quality . . . 37
2.2.9 ATLAS detector simulation . . . 38

3 Pixel Cluster splitting with Mixture Density Network . . . 39
3.1 Introduction . . . 39
3.2 Pixel cluster-splitting neural networks: current approach . . . 40
3.2.1 Network architectures . . . 40
3.2.2 Network training: least-squares formalism . . . 41
3.3 Mixture Density Networks . . . 41
3.3.1 MDN algorithm . . . 41
3.3.2 Toy Example: Inverse Problem . . . 42
3.3.3 Building MDN . . . 43
3.3.4 MDN training . . . 44
3.3.5 Overflow and Underflow problem . . . 44
3.3.6 The persistent NaN problem . . . 45
3.4 Results . . . 46
3.5 Conclusions . . . 48

4 Object Reconstruction . . . 49
4.1 Reconstruction overview . . . 49
4.2 Tracks, vertices and energy clusters . . . 50
4.2.1 Tracks . . . 50
4.2.2 Vertices . . . 51
4.2.3 Energy clusters . . . 52
4.3 Electrons . . . 53

4.3.1 Electron reconstruction . . . 53
4.3.2 Electron identification . . . 53
4.3.3 Electron isolation . . . 55
4.4 Muons . . . 56
4.4.1 Muon Reconstruction . . . 56
4.4.2 Muon Identification . . . 57
4.4.3 Muon Isolation . . . 57
4.5 Jets . . . 58
4.5.1 Jet reconstruction . . . 59
4.6 Truth jets . . . 60
4.7 Small-R jets . . . 60
4.7.1 EMTopo jets . . . 60
4.7.2 PFlow small-R jets . . . 61
4.8 Variable-R track jets . . . 62
4.9 Large-R jets . . . 63
4.9.1 Large-R jet calibration . . . 63
4.10 b-tagging . . . 65
4.10.1 Key ingredients for b-jet identification . . . 65
4.10.2 b-tagging algorithms . . . 66
4.10.3 b-tagging calibration . . . 69
4.11 Top tagging . . . 69
4.11.1 Jet substructure variables . . . 70
4.11.2 Jet truth labeling . . . 70
4.11.3 Top-tagging algorithms . . . 70
4.11.4 Top-tagging calibration . . . 71
4.12 Missing Transverse Energy . . . 71

5 Statistical methods . . . 73
5.1 Random variables and probability distributions . . . 73
5.1.1 Poisson distribution . . . 74
5.1.2 Gaussian distribution . . . 75
5.1.3 Chi-squared (χ^2) distribution . . . 75
5.1.4 Log-Normal distribution . . . 75
5.2 Likelihood function . . . 76
5.2.1 Systematic uncertainties . . . 77
5.3 Statistical inference . . . 77
5.3.1 Parameter estimation . . . 77
5.3.2 Hypothesis testing . . . 78
5.3.3 Profile likelihood ratio . . . 79
5.3.4 Test statistic . . . 80
5.3.5 Asymptotic approximation . . . 81
5.3.6 CL_s method . . . 83
5.3.7 Asimov data . . . 84
5.3.8 Toy example . . . 85
5.3.9 Wilks' test . . . 87
5.4 BumpHunter search . . . 88
5.4.1 Adding systematic uncertainties to the BumpHunter framework . . . 90

6 Search for high mass resonances in di-lepton final state . . . 93
6.1 Analysis overview . . . 93
6.2 Object definition . . . 94
6.3 Event selection . . . 95
6.4 Background modeling . . . 97
6.4.1 Background processes . . . 97
6.5 Signal modeling . . . 100
6.6 Dilepton invariant mass . . . 101
6.6.1 Differences between electron and muon channels . . . 101
6.6.2 Event yields . . . 101
6.7 Systematic uncertainties . . . 103
6.7.1 Theory uncertainties . . . 103
6.7.2 Experimental uncertainties . . . 106
6.7.3 Systematics summary . . . 109
6.8 BumpHunter search method . . . 110
6.9 Limit setting . . . 110
6.10 Conclusion . . . 111

7 Search for high mass resonances in boosted all-hadronic top-antitop final sate . . . 113
7.1 Analysis overview . . . 114
7.2 Top quark decay topology . . . 114
7.3 Background processes . . . 115
7.3.1 Standard model $t\bar{t}$ process . . . 115
7.3.2 Standard model multijet processes . . . 116
7.4 Signal processes . . . 116
7.4.1 Topcolor-assisted-Technicolor model . . . 116
7.4.2 Heavy Vector Triplet model . . . 117
7.5 Object definition . . . 117
7.6 Event selection . . . 121
7.7 Event categorization . . . 123
7.7.1 Control region (CR0b) . . . 123
7.7.2 Signal regions . . . 123
7.7.3 Validation regions . . . 124
7.7.4 Addition regions for multijet estimation . . . 125
7.7.5 Signal efficiencies . . . 127
7.7.6 Event yields . . . 128
7.8 Signal shape modeling using function fit . . . 128
7.9 Background modeling using function fit . . . 129
7.9.1 Why data-driven background estimation? . . . 130
7.9.2 Monte Carlo effective entries . . . 131
7.9.3 Fitting range and binning . . . 134
7.9.4 Background modeling function . . . 134
7.9.5 Wilks' test . . . 135
7.9.6 Background fitting with pseudo-experiments . . . 136
7.9.7 Wilks' test on pseudo-experiment . . . 137
7.9.8 Final choice of fit functions . . . 137
7.9.9 Spurious signal test . . . 138

7.9.10 Signal injection test . . . 139
7.10 $t\bar{t}$ invariant mass . . . 143
7.11 Systematic uncertainties . . . 143
7.11.1 Experimental signal uncertainties . . . 144
7.11.2 Background modeling . . . 145
7.11.3 Theory uncertainties . . . 147
7.11.4 Systematics summary . . . 147
7.12 Statistical analysis . . . 151
7.12.1 b-only profile likelihood fit . . . 152
7.12.2 s+b profile likelihood fit . . . 152
7.12.3 BumpHunter search results . . . 155
7.12.4 Model dependent search results . . . 155
7.13 Limit setting . . . 156
7.13.1 Comparison to previous results . . . 156
7.14 Inference with dark matter mediator model . . . 158
7.15 Conclusion . . . 159

8 Search for high mass resonances in the semi-leptonic top-antitop final sate . . . 160
8.1 Analysis overview . . . 161
8.2 Top quark decay topology . . . 162
8.3 Background processes . . . 162
8.4 Signal modeling . . . 163
8.5 Object definitions . . . 164
8.6 Event selection . . . 164
8.6.1 Boosted selection . . . 165
8.6.2 Resolved selection . . . 165
8.7 Overlap removal . . . 166
8.7.1 Electron-in-jet subtraction method . . . 168
8.8 Data-driven multijet background . . . 168
8.9 Event reconstruction . . . 170
8.9.1 Neutrino reconstruction . . . 170
8.9.2 Boosted region . . . 171
8.9.3 Resolved region . . . 171
8.9.4 $t\bar{t}$ invariant mass . . . 172
8.10 Kinematic distributions . . . 172
8.10.1 Resolved region . . . 172
8.10.2 Boosted region . . . 174
8.11 Statistical analysis . . . 176
8.11.1 b-only profile likelihood Asimov fit . . . 176
8.11.2 s+b profile likelihood Asimov fit . . . 176
8.11.3 Limit setting . . . 177
8.12 Conclusion . . . 179

9 Conclusions and outlook . . . 180

Bibliography . . . 183

Appendices . . . 199

Appendix A Jet Calibration **199**
A.1 EMTopo jet calibration 199
A.2 PFlow Energy scale correction 201
A.3 Large-R jet energy and mass scale calibration 201

Appendix B Additional studies in boosted all-hadronic $t\bar{t}$ resonance search **203**
B.1 Data-driven estimate of SM multijet process 203
B.2 Control plots 207
B.2.1 Signal region plots 207
B.3 All-hadronic stat 210
B.3.1 b-only profile likelihood fit: Asimov 210
B.3.2 s+b profile likelihood fit: Asimov 211
B.3.3 b-only profile likelihood fit: Data 215
B.3.4 s+b-only profile likelihood fit: Data 215
B.4 Model-dependent search 218

Appendix C Additional studies in semi-leptonic $t\bar{t}$ resonance search **219**
C.1 Statistical analysis 219
C.1.1 b-only profile likelihood Asimov fit 219
C.1.2 s+b profile likelihood Asimov fit 219

List of Tables

Table 1.1 he fermions of the Standard Model and their electric and color charges. . . . 9
Table 1.2 The force carries and the Higgs boson of the Standard Model and their charges. 9
Table 1.3 SM fermion fields and their electric charge (Q), third component of weak isospin (T^3) and weak hypercharge (Y). 11
Table 1.4 The best NNLO + NNLL theoretical prediction for $t\bar{t}$ pair production cross section at Tevatron and LHC. 17
Table 1.5 List of commonly used Z' models. 22

Table 3.1 The network structure and the hyperparameters used to train the MDNs. . . 44

Table 4.1 c-jet, τ and light-flavor jet rejections for the different b-jet tagging efficiency single-cut operating points and corresponding discriminant threshold values for the MV2 and the DL1 algorithms. 68

Table 6.1 Summary of the electron object definition. 95
Table 6.2 Summary of the jet object definition. 95
Table 6.3 Summary of the event selection 96
Table 6.4 Summary of the simulation information used to model the background processes in the dilepton resonance search. 98
Table 6.5 Trigger scheme used for the selection of a fake-enriched sample for 2015 data. 99
Table 6.6 Trigger scheme used for the selection of a fake-enriched sample for 2016 data. 99
Table 6.7 Summary of the simulation process used to generate the signal events used in the dilepton resonance search. 101
Table 6.8 Expected and observed event yields in the dielectron channel in different dilepton mass intervals. 103
Table 6.9 Summary of the pre-marginalised relative systematic uncertainties in the expected number of events at dilepton masses of 2 TeV and 4 TeV. 109
Table 6.10 Observed and expected upper limits on Z' signal production at 95% CL. . . . 112

Table 7.1 Summary of MC sample information for background processes used in the boosted all-hadronic $t\bar{t}$ resonance search. 116
Table 7.2 Leading-order theoretical cross-sections for the $Z'_{\mathrm{TC2}} \to t\bar{t}$ signal, before applying the k-factor of 1.3. 117
Table 7.3 Leading-order theoretical cross-sections for only the fully hadronic $t\bar{t}$ final state, $Z'_{\mathrm{HVT}} \to t\bar{t} \to q\bar{q}'b + q\bar{q}'b$ signal including all branching fractions. 118
Table 7.4 Summary of MC sample information for signal processes used in the boosted all-hadronic $t\bar{t}$ resonance search. 118
Table 7.5 Summary of the large-R jet object definition. 118

Table 7.6 Summary of the small-R jet object definition. . . . 119
Table 7.7 Summary of the variable-R track jet object definition. . . . 119
Table 7.8 Summary of the electron object definition. . . . 120
Table 7.9 Summary of the muon object definition. . . . 120
Table 7.10 Summary of the event selection for the $t\bar{t}$ resonance search in the boosted all-hadronic final state. . . . 123
Table 7.11 Summary of kinematic selections and categorization of signal, control and validation regions. . . . 125
Table 7.12 Event categorization used to model the multijet background from data according to whether the leading and subleading large-R jets are top-tagged or b-tagged. . . . 126
Table 7.13 Event yields in CR0b, SR1b and SR2b regions in 2015-2018 data. . . . 128
Table 7.14 Table showing the Wilks' p-values calculated by comparing Model 1 and Model 2. . . . 136
Table 7.15 Summary of the pre-fit relative systematic uncertainties in the expected number of events at Z'_{TC2} masses of 2 TeV and 4 TeV in 1 b-tag signal region (SR1b). 149
Table 7.16 Summary of the pre-fit relative systematic uncertainties in the expected number of events at Z'_{TC2} masses of 2 TeV and 4 TeV in 2 b-tag signal region (SR2b). 150
Table 7.17 Summary of the expected and observed mass exclusion limits of Z' signals with $\Gamma = 1.2\%, 1\%, 3\%$. . . . 157

Table 8.1 Simulation details of the signal processes used in the semi-leptonic $t\bar{t}$ resonance search. . . . 164
Table 8.2 List of the single-lepton high Level triggers (HLT) used in this analysis. All the triggers are used together with an 'OR' condition. . . . 165
Table 8.3 Summary of the event selection criteria used in the semi-leptonic $t\bar{t}$ resonance search. . . . 166
Table 8.4 Muon-jet and electron-muon overlap removal definition used in the semi-leptonic $t\bar{t}$ resonance analysis. . . . 168
Table 8.5 Definitions of the leptons used for the matrix method QCD estimation. . . . 169

Table B.1 Summary of the significance scan using Z'_{TC2} signals with relative width (Γ) of 1.2%. . . . 218

List of Figures

Figure 1.1 Octet of (a) spin-1/2 baryons and (b) mesons. 5
Figure 1.2 The scalar field potential, $V(\Phi)$ as a function of $|\Phi|$. 13
Figure 1.3 Reduced coupling-strength modifiers for fermions, κ_F ($F = t, b, \tau, \mu$) and weak gauge bosons, κ_V ($V = W, Z$) as a function of their masses. 15
Figure 1.4 LO Feynman diagrams of the top-quark pair production. 16
Figure 1.5 Summary of the LHC and Tevatron measurements of the top-pair production cross-section as a function of the center-of-mass energy. 17
Figure 1.6 The branching ratios (BR) of different $t\bar{t}$ decay channels. 18
Figure 1.7 Summary of several Standard Model total production cross section measurements, corrected for leptonic branching fractions, compared to the corresponding theoretical expectations. 20

Figure 2.1 Illustration of the various beam lines and experimental facilities of the CERN accelerator complex. 26
Figure 2.2 (a) The luminosity weighted distribution of mean number of interactions per bunch crossing (pp) for the 2015 to 2018 data taking periods. (b) The cumulative luminosity as a function of time. 29
Figure 2.3 Cut-away view of the full ATLAS detector. 30
Figure 2.4 Cut-away view of the ATLAS inner detector. 31
Figure 2.5 Cut-away view of the ATLAS calorimeter detector. 33
Figure 2.6 Cut-away view of the ATLAS muon spectrometer. 35
Figure 2.7 A schematic picture showing a quarter-section of the ATLAS muon system. . 36
Figure 2.8 A schematic picture ATLAS trigger and data acquisition system using in Run-2. 37

Figure 3.1 Schematic diagrams of charge diffusion in a pixel sensor (left), and merged clusters of two nearby tracks (right). 39
Figure 3.2 Performance of the MDN in modeling an inverse problem. 43
Figure 3.3 Schematic diagrams of (a) the pixel neural network work flow and (b) mixture density network. 44
Figure 3.4 Local y residuals and pulls for 1-particle IBL clusters. 47
Figure 3.5 Local x and y residuals and pulls for 2-particle barrel (without IBL) clusters. 47
Figure 3.6 Local x and y residuals and pulls for 3-particle endcap clusters. 48

Figure 4.1 Schematic diagram showing the particle interactions with the ATLAS detector. 50
Figure 4.2 A schematic diagram of the path of an electron through the detector. 54
Figure 4.3 Electron reconstruction efficiency as a function of electron E_T. 54
Figure 4.4 Electron identification efficiency as a function of electron E_T. 55

Figure 4.5 Muon reconstruction efficiency as a function of muon p_T. 58
Figure 4.6 A schematic diagram shows the calibration stages for EM-scale jets. 61
Figure 4.7 A flow chart of the particle flow algorithm. 62
Figure 4.8 Overview of the large-R jet reconstruction and calibration. 64
Figure 4.9 Schematic diagram comparing the b-jet topology with light jets. 65
Figure 4.10 DL1 b-tagger performance. 68
Figure 4.11 Schematic diagram of resolved and boosted top-decay topology. 69
Figure 4.12 DNN top-tagger performance. 71

Figure 5.1 p-value and Z-score. 80
Figure 5.2 Test statistic distribution of a 2-bin toy data. 82
Figure 5.3 Example test static distributions under the b and $s+b$ hypotheses. 84
Figure 5.4 Test statistic distribution along with the asymptotic formula for a toy Poisson counting model. 86
Figure 5.5 The CL_s scans for a top Poisson counting model. 87
Figure 5.6 Schematic diagram of BumpHunter search intervals. 88
Figure 5.7 BumpHunter tomography and scan plot showing the most significant interval. 89
Figure 5.8 BumpHunter statistic distribution. 90
Figure 5.9 BumpHunter scan on the dilepton mass distribution in presence of systematic uncertainties. 92

Figure 6.1 Feynman diagram of the Z' signal process decaying into two electrons or muons. 94
Figure 6.2 Acceptance times efficiency as a fucntion of dielectron truth mass. 96
Figure 6.3 Invariant mass distribution in the (a) di-electron and (b) di-muon channels. 102
Figure 6.4 Simulated signal templates for the Z'_χ signal in the (a) di-electron and (b) di-muon channels. 102
Figure 6.5 The PDF eigenvector up (down) variations are shown as a function of the di-electron invariant mass. 105
Figure 6.6 Theoretical uncertainties like PDF choice, PDF scales, top-quark, diboson background uncertainties and α_S variation, electroweak correction, and photon-induced correction on the Drell-Yan background. 106
Figure 6.7 Experimental systematic uncertainties as a function of dielectron invariant mass. 108
Figure 6.8 Dilepton BumpHunter search scan in the di-electron and di-muon channels. 110
Figure 6.9 Upper 95% CL limit on Z' signal cross-section × branching ratio as a function of Z' pole mass. 111

Figure 7.1 Feynman diagram of the Z' signal process where both the top quarks decays hadronically. 114
Figure 7.2 Schematic diagram showing the all-hadronic $t\bar{t}$ final state in resolved and boosted regime. 115
Figure 7.3 Trigger efficiency as a function of (a) trimmed large-R jet p_T and (b) Z'_{TC2} pole mass. 121
Figure 7.4 Schematic diagram illustrating the trigger efficiency turn-on. 122
Figure 7.5 Schematic diagram showing the signal and control region definitions with different event categories. 124

Figure 7.6 Pie charts showing background composition in control, validation and signal regions. . . . 125
Figure 7.7 Expected signal sensitivity for 3 TeV and 4 TeV Z'_{TC2} signal models. . . . 126
Figure 7.8 The $m_{t\bar{t}}^{\mathrm{reco}}$ distributions of the simulated Z' signal events. . . . 127
Figure 7.9 Acceptance and acceptance times selection efficiency as a function of $m_{t\bar{t}}^{\mathrm{gen}}$ in the signal regions. . . . 127
Figure 7.10 Simulated signal $m_{t\bar{t}}$ distribution fitted with the Crystal Ball and a Gaussian function. . . . 130
Figure 7.11 The effective entries in the $m_{t\bar{t}}$ distribution of the simulated $t\bar{t}$ events are compared with the nominal prediction. . . . 131
Figure 7.12 The effective entries in the data-driven multi-jet $m_{t\bar{t}}$ distributions are compared with the nominal estimates. . . . 132
Figure 7.13 The effective entries in the $m_{t\bar{t}}$ distribution fof the simulated multi-jet events are compared with the nominal prediction. . . . 132
Figure 7.14 The ratio of the MC multijet and data-driven multijet distributions in the (a) $1b$ signal region and (b) $2b$ signal region. The ratios are fit with 1^{st} and 2^{nd} order polynomial functions. The two fits look very similar, but for simplicity a 1^{st} order polynominal is used. . . . 133
Figure 7.15 The effective entries in the $m_{t\bar{t}}$ distribution of the combined multijet and MC $t\bar{t}$ distribution are compared with the combined multijet and MC $t\bar{t}$ background. . . . 133
Figure 7.16 The $m_{t\bar{t}}$ distributions fitted with 3-, 4- and 5-parameter function in the signal regions. . . . 135
Figure 7.17 4-parameter fits on the nominal spectrum and 1000 pseudo-experiments. . . . 136
Figure 7.18 Wilks' p-value for the pseudo-experiments comparing fits with 3-, 4- and 5-parameter fits. . . . 137
Figure 7.19 Schematic diagram showing the spurious signal test. . . . 138
Figure 7.20 Spurious signal pull for background functions with 3-, 4- and 5-parameters in (a) SR1b and (b) SR2b. The gray shaded area shows the ±50% around zero which is used for comparison. . . . 139
Figure 7.21 Extracted signal strength (μ_{ext}) as a function of injected signal strength (μ_{inj}) in (a) SR1b (b) SR2b. . . . 140
Figure 7.22 Ratios of extracted and injected signal events in the regions. . . . 141
Figure 7.23 b-only and $s+b$ fit on the signal injected spectrum and the expected limit calculated from the signal injected distributions. . . . 142
Figure 7.24 Observed $m_{t\bar{t}}^{\mathrm{reco}}$ distributions in data for (a) SR1b and (b) SR2b. . . . 143
Figure 7.25 Spurious signal uncertainty for the 4 TeV signal mass point in (a) SR1b and (b) SR2b. . . . 146
Figure 7.26 Fit parameter uncertainty as a function of $m_{t\bar{t}}$. . . . 147
Figure 7.27 Up and down variations of JES modeling uncertainty as a function of $m_{t\bar{t}}$ for the 4 TeV Z'_{TC2} signal. . . . 148
Figure 7.28 Up and down variations of JES baseline and tracking uncertainties as a function of $m_{t\bar{t}}$ for the 4 TeV Z'_{TC2} signal. . . . 148
Figure 7.29 Post-fit (b-only) nuisance parameter pulls observed in data. . . . 152
Figure 7.30 Post-fit ($s+b$) correlation matrix of the systematic NPs and signal strength for Z' with a mass of 4 TeV. . . . 153
Figure 7.31 Systematics ranking plot for Z' mass of 4 TeV. . . . 154

Figure 7.32 Local p_0-values as function of $m_{t\bar{t}}$. 155
Figure 7.33 Observed and expected 95% CL upper limits on the cross-section times branching fraction of the $Z' \to t\bar{t}$ as a function of the Z' mass. 157
Figure 7.34 Comparison of expected upper limits on the signal cross-section times branching fraction as a function of the Z' mass. 158
Figure 7.35 95% CL upper limit on the quark coupling, g_q, as a function mediator mass, $m_{Z'_A}$. 159

Figure 8.1 Feynman diagram of the Z' signal process where one of the top quarks decay hadronically and the other one decays leptonically. 161
Figure 8.2 Schematic diagram showing the single-lepton $t\bar{t}$ final state in the resolved and boosted regimes. 162
Figure 8.3 Distributions of (a) top and anti-top p_T and (b) $m^{\text{gen}}_{t\bar{t}}$ 163
Figure 8.4 Angular distance between the electron and b-quark at the parton level in the (a) SM $t\bar{t}$ and (b) 4 TeV Z' signal events. 167
Figure 8.5 $\Delta R(el_{\text{true}}, b_{\text{true}})$ as a function of truth electron p_T for (a) SM $t\bar{t}$ and (b) 4 TeV Z'.167
Figure 8.6 Schematic diagram demonstrating the electron-in-jets subtraction method. . 169
Figure 8.7 The p_T distributions of the lepton, close-to-lepton small-R jet, and E_T^{miss} distributions in the boosted electron and boosted muon channel. 173
Figure 8.8 The p_T distributions of the large-R jet , lepton, and close-to-lepton small-R jet in the boosted electron and muon channel. 175
Figure 8.9 The best-fit value of the normalization factors in the b-only fit. 177
Figure 8.10 The best-fit value of the signal strength and the normalization factors are shown in the $s+b$ fit. 177
Figure 8.11 Systematics ranking plot at Z'_{TC2} mass 4 TeV. 178
Figure 8.12 The expected cross-section 95% CL upper limits on Z'_{TC2} signals. 179

Figure A.1 EM topo jet energy response and η correction factor as a function of η_{det}. . . 200
Figure A.2 The response, $E^{\text{reco}}/E^{\text{truth}}$, of PFlow jet is shown as function of jet η for the MC dijet events. 202
Figure A.3 Large-R energy and mass response as a function of η_{det} 202

Figure B.1 $m_{t\bar{t}}$ distributions with the simulated multijet background in the two validation regions: (a) VR1b and (b) VR2b. 204
Figure B.2 $m_{t\bar{t}}$ distributions with simulated and data-driven multijet estimate are compared in the control region. 206
Figure B.3 $m_{t\bar{t}}$ distributions with data-driven multijet background in the two validation regions: (a) VR1b and (b) VR2b. 206
Figure B.4 $m_{t\bar{t}}$ distributions with the data-driven multijet background in the two signal regions: (a) SR1b and (b) SR2b. 207
Figure B.5 Leading and subleading large-R jet p_T distribution in SR1b (top row) and SR2b (bottom row). 208
Figure B.6 Distributions of the leading large-R jet mass and τ_{32} in the two signal regions.209
Figure B.7 The best-fit values of the NPs in the background-only likelihood fit on the Asimov data. 210
Figure B.8 Post-fit (ASIMOV) correlation matrix of the four background systematic NPs 210
Figure B.9 The best-fit signal strength value where statistical and all the systematics uncertainties are included in the likelihood. 211

Figure B.10 The best-fit signal strength value where the systematic uncertainties are not included in the likelihood. . . . 211
Figure B.11 Post-fit ($s+b$ fit) NP pulls where 4 TeV Z'_{TC2} signal model is used. . . . 212
Figure B.12 Post-fit ($s+b$ fit) correlation matrix of the systematic NPs and signal strength for Z' mass 4 TeV. . . . 213
Figure B.13 Systematics ranking plot based on there impact on the POI for the 4 TeV Z'_{TC2} signal model. . . . 214
Figure B.14 Post-fit (b-only) correlation matrix of the background systematic NPs observed in data. . . . 215
Figure B.15 The best-fit signal strength value observed on data in presence of statistical and all the systematic uncertainties. . . . 215
Figure B.16 The best-fit signal strength value observed on data in presence of only the statistical uncertainty. . . . 216
Figure B.17 Post-fit ($s+b$ fit) NP pulls where 4 TeV Z'_{TC2} signal model is used. . . . 217

Figure C.1 The best-fit values of the NPs in the background-only likelihood fit. . . . 220
Figure C.2 The best-fit values of the NPs in the signal+background likelihood fit where 4 TeV Z'_{TC2} signal model is used. . . . 221

Acknowledgements

It was my honor to get the opportunity to work with so many enthusiastic scientists from all over the world and be supported by a wonderful group of friends and my family during my PhD. Thank you for making my journey an amazing one.

First and foremost, I thank my Supervisor, Alison Lister, for supporting me during the past five years. While you gave me the freedom to pursue various projects, you were there for me when I needed your guidance. Your guidance and support have been invaluable during my PhD, and I am fortunate to have you as my supervisor. Apart from being a wonderful thesis supervisor, you have been a great mentor and helped me figure out my post-PhD path.

When I started my Ph.D. in experimental particle physics, it was entirely a new area for me. I am grateful to Wojtek Fedorko for helping me find my footing. You have been my primary resource in getting my scientific questions answered, and you have helped me navigate my first ATLAS projects and authorship qualification task.

It has been a privilege to work under the guidance of my committee members. I thank Colin Gay, Gary Hinshaw, and Gordon Semenoff for their suggestions and advice throughout my PhD studies.

I would like to thank the ATLAS collaboration for providing me with the resources during my PhD studies. I would like to thank my colleagues from the dilepton resonance search. To Dan Hayden, your help and guide in the dilepton resonance search were valuable to me. Thank you for your support in the $t\bar{t}$ resonance search analysis and help me by proofreading this dissertation. Thanks to my colleagues, Christopher Willis, and Etienne Dreyer, for their assistance in this project.

I would like to thank my colleagues from the all-hadronic $t\bar{t}$ resonance search, Aparajita Dattagupta, Hulin Wang, Josu Cantero Garcia, Koji Terashi, Trine Poulsen, and Tong Ou. Thanks to Aparajita Dattagupta, Hulin Wang, Josu Cantero Garcia, Koji Terashi for leading the team. Special thanks to Yu-Heng for being the best collaborator I ever had. To Yu-Heng, it was my pleasure to work with you and I enjoyed every bit of it. I learned a lot from you. Thanks for always helping me when I needed it and keeping me motivated. To Trine, thanks for being an awesome collaborator. I enjoyed working with you. To Koji, your support and advice were very valuable. Thanks for your help.

I would like to thank my colleagues from the semi-leptonic $t\bar{t}$ resonance search. To Danilo Ferreira de Lima, your guidance and help were crucial to start with this analysis. Thanks for always answering my questions. To Samuel Calvet, I cannot lead the team without your help and support. Thanks for leading the team at the beginning and keeping me motivated with your new ideas. I am also very grateful to the whole analysis team for giving me the opportunity to lead the analysis team since late 2019.

I would like to thank Louis Guillaume Gagnon for his help with setting up the MDN project and his continuous support with technical stuff. To Sourav Sen, your help and support were very

valuable. Thanks for your help during thesis writing and preparation for the oral exam.

I am thankful to my colleagues at the UBC ATLAS group, Dominique, Robin H., Robin N., Matthias, Steffen, Sebastien, Alex, Vincent, Felix, and Chris, who have supported me all along and made this PhD experience a great one. It was great to share the same office with all of you and participate in the fun activities. I am very grateful to all of you for helping me with the proofread of this dissertation. I hope our paths cross in the future, and we work together. Special thanks to Felix and Wojtek for taking care of Flashy. I would also like to thank my friends Shreya and Anjishnu for their academic and non-academic help. To Shreya and Anjishnu, you made my stay at CERN entertaining and colorful.

I would like to thank Satyaki Bhattacharya for introducing me to experimental particle physics and supporting me always when I needed it. Your advice and guidance to figure out the post-PhD path were valuable. To Shilpi Jain, thanks for giving me company at CERN.

Moving to a foreign country, miles away from my family, would have been a challenging one. Special thanks to my friends, Soumalya, Avinash, Jibin, Tushar, Leya, Subrata, Santanil, Samyadeb, Prithviraj, for always being there for me and making this place feel like home. Although physically apart, I am thankful to my friends Dipanjan, Uditendu, Chitram, Shubradeep, Shibajyoti, Arindam, Shukla for always having my back through countless online chats for the past five years.

I would like to thank my parents, Golam Mortuza and Romeja Kathun, for all their love and the incredible sacrifices they made for me. Their constant support and encouragement are instrumental in my accomplishments. Without their hard work, I would not have reached where I am today. I am thankful to my lovely sister Mahfuja Kathun who could not witness this journey.

Last but not least, I am thankful to my partner Sreeparna Vappala who has been by my side throughout this PhD and loving unconditionally during my good and bad times. To Sreeparna, I am grateful to you for leaving your PhD and move to Vancouver and start a new PhD. Living with you was the best part of life, and I learned a lot from you in the past few years. Thank you for making me a better person. This Ph.D. absolutely would not be possible without you.

To my parents
with love and eternal appreciation

To my sister
in loving memory

CHAPTER

ONE

Theoretical background

1.1 Historical introduction to the elementary particle physics

A brief evolution of particle experiments is given in this section. It does not cover all the developments that happened over the last century, but highlights several important discoveries. This section is inspired by the particle physics book of David Griffiths [1], and the presentation style is adopted from the first chapter of that book.

1.1.1 Electron, proton and neutron (1897 - 1932)

The birth of elementary particle physics can be related to the discovery of the electron in 1897 by J. J. Thomson. Although there were speculations about atomic structure in earlier times, those were just philosophical concepts without any scientific evidence. Until J. J. Thomson's discovery, the particle nature of cathode rays was not known. Even physicist like Lord Kelvin wrote that "electricity is a continuous homogeneous liquid" [2]. In the same year, J. J. Thomson performed an experiment that showed cathode rays carry electric charge [3]. He showed that a hot filament coming from a cathode ray deflects in the presence of a magnetic field. Thus, it became clear that the cathode rays are not rays at all; instead, they are streams of particles. Thomson further measured the ratio of electric charge and mass, e/m, of the beam constituents. The e/m value measured by Thomson was almost several thousand times larger than that of any known ions. It indicated that either their charge was very large or the mass was very small. Thomson concluded that this particle had a very tiny mass with a negative charge based on other indirect evidence. Thomson initially called them *corpuscles* and their charge was called *electron*, the basic unit of electric charge [4]. Later the particle itself was called an electron. The term electron was introduced in 1891 by George Johnstone Stoney [5].

Thanks to the work of Thomson, atoms were then known to be full of these lightweight, negatively charged electrons. Since the atom as a whole is charge neutral, there has to be an equal amount of positive charges inside the atom. However, it was not known where this positive charge resided within the atom. Thomson came up with a theoretical model of the atom in 1904, where the positive charges were distributed homogeneously inside the atom, with the electrons embedded within it. This model is known as *plum pudding model* because electrons were like the plums inside the positively charged medium acting like the pudding. Later in 1911, it was proven wrong

by Rutherford's famous gold foil experiment [6]. He fired a beam of α-particles with charge $+2e$ through a thin sheet of gold foil. The gold foil experiment showed that the positive charges were concentrated in the core of the atom, *nucleus*. Later, Rutherford named the nucleus of the hydrogen atom a *proton*. His atomic model was quite simple, but it was disproved since negatively charged electrons do not collapse into the positively charged nucleus. In 1913, Niels Bohr proposed the energy levels in the atomic model, and he described the hydrogen atom model where a single electron circulates the proton [7]. But surprisingly, the heavier helium atom, carrying two electrons, weighs four times hydrogen, and lithium weighs seven times the weight of hydrogen. The puzzle was finally solved in 1932 with the discovery of *neutron*, the electrically neutral twin of the proton, by Chadwick [8]. Originally, Rutherford proposed the existence of one such neutral particle more than ten years prior, in 1920. The name 'neutron' was given by him.

1.1.2 Photon (1900-1926)

The discovery of another important particle happened also in the early nineteen hundreds. In the first two decades of the 20th century, lots of development happened in understanding quantum mechanics and which eventually lead to the discovery of the *photon*. Although there were some discussions in late 1800s, the first contribution came from Max Plank in 1900 while he was attempting to explain the blackbody spectrum. Max Plank found an empirical formula that fitted the experimental data. To explain the fit, Plank needed to assume electromagnetic radiation was quantized with a quantum that has energy $E = h\nu$, where ν is the frequency of the radiation and h is a constant adjusted by the fit. Later, this constant was named after him. The modern value of h is 6.626×10^{-34} J·s. Plank did not know why the radiation was quantized and assumed it was due to a pathological effect in the emission process. Then in 1905, Albert Einstein explained that quantization is a feature of the electromagnetic field itself [9]. Using Planck's idea, Einstein explained the *photoelectric effect* with the particle nature of light. But Einstein's concept was not taken very positively in the physics community at that time [1] and other physicists were not ready to accept the particle theory of light. The issue was finally settled down in 1923 by A. H. Compton with the discovery of *Compton effect* [10]. The experimental evidence indicates that light indeed behaves like a particle on a subatomic scale. The particle was called a *photon*.

1.1.3 Mesons (1934 - 1947)

By the mid-1930s, we had an atomic model, but we did not know what holds the nucleus together. The positively charged protons should strongly repel one another, so there must be some other stronger force that holds them together. The physicists called that force the *strong force*. In 1934, Yukawa proposed the first theory of the strong interaction [11]. Yukawa assumed that protons and neutrons were attached by some field, like electromagnetic or gravitational fields. He also thought there should be a proper quantization of that field. Then he tried to understand the properties of its quantum, the particle analogous to the photon. Yukawa predicted that the mass of such a particle should be around 200 times the electron mass or about one-sixth of the proton mass. He called these particles 'mesons'. It was later confirmed independently by Anderson and Neddermeyer by 1937 [12]. At that time, it was proposed that cosmic rays contain these middleweight mesons. But in 1946, it was shown that cosmic ray particles interact only very weakly with the atoms, whereas a much stronger interaction would be expected from Yukawa's meson. In 1947 this puzzle was solved by Powell's group that discovered that there are two middleweight parti-

[1] Though Einstein ended up getting Nobel prize for photoelectric effect.

cles in the cosmic rays [13]. They called them the *pion* (π) and *muon* (μ). Then it was understood that the true Yukawa meson is the pion. Two types of charged pions, π^+, and π^- were experimentally observed at that time. The neutral pion, π^0, was observed later. A muon, on the other hand, is just another charged particle that behaves very similar to electrons and has nothing to do with strong interaction.

1.1.4 Isospin and SU(2)

The concept of *isospin* was introduced by Heisenberg in the year 1932 after the discovery of the neutron [14]. It was observed that the newly discovered neutron had a very similar mass to a proton. A proton has a positive electric charge, and a neutron is neutral. Otherwise, they seemed identical in all other aspects. It was also found that the strong force that held the proton and the neutron together did not depend on the electron charge. So, in absence of the electric charge of the proton, it will not be possible to distinguish between the proton and the neutron. To explain this symmetry of the strong nuclear force, Heisenberg treated the proton and the neutron as two components of a *nucleon* doublet that is analogous to the two states of a spin-half particle. He introduced the concept of another conserved quantity, which allowed protons to convert into neutrons and vice-versa. It was named *isospin* by Eugene Wigner in 1937 to indicate that the new quantity is similar to spin in behavior, although it has nothing to do with actual spin [15]. This symmetry is expressed by a unitary transformation governed by the $SU(2)$ group in the abstract isospin space. The three pions represent a triplet state of $SU(2)$. In the past, it was believed that the isospin symmetry was an exact symmetry of the strong interaction, and it was violated by electromagnetic and weak interactions. The mass difference between proton and neutron was thought to be related to their charge content. If the mass difference was pure electrostatic, then the protons had to be more massive. On the other hand, it was known that the opposite was true, and proton decay was never observed. Isospin symmetry is an approximate symmetry, and neutrons have a slightly higher mass due to isospin breaking. After the discovery of the quark model in 1964, it was understood that mass difference is due to the mass difference of "up" and "down" quarks. Isospin symmetry is thus an "accidental" property shown by the up and down quarks.

1.1.5 Antiparticles (1930-1956)

Surprisingly, all the major developments in nonrelativistic quantum mechanics happened just within three years (1923-1926). The second development towards relativistic quantum mechanics, after the Klein-Gordon equation, was in 1927 by Dirac [16]. Dirac's equation gave a negative energy solution for every positive energy solution. Dirac proposed a sea of negative energy particles, known as hole theory, to explain it. Later in 1932, Anderson discovered the *positron*, a positively charged twin of the electron [17]. The discovery was confirmed soon after by Occhialini and Blacket, in 1933 [18]. In 1940, Feynman and Stuckelberg gave a much simpler interpretation of the negative energy states that appeared in Dirac's question [19]. They rewrote the negative energy states as positive-energy states of the positron. Later as quantum field theory started developing, it was clear that each particle should have its antiparticle. Anti-protons and anti-neutrons were observed first at the Berkeley Bevatron in 1955 [20] and 1956 [21], respectively. The photon, on the other hand, was it's own anti-particle.

1.1.6 Neutrinos (1930-2000)

The phenomena of radioactive beta decay was known since 1900, but it was observed that energy-momentum did not seem to be conserved in beta decays. In a two-body decay ($A \rightarrow B + e$) where a nucleus (A) converts into another nucleus (B) and an electron (e), the conservation of energy-momentum gives the electron energy to be

$$E_e = \left(\frac{m_A^2 - m_B^2 + m_e^2}{2m_A} \right) c^2 \, . \tag{1.1}$$

So, the electron energy should be a fixed number once we fix the mass of the three particles, A, B, e. However, the experimental data showed that Equation 1.1 only determines the maximum electron energy, and the electron energy varies a lot. Pauli solved this puzzle in 1932 and offered a theory of beta decay [22]. He proposed another neutral particle was emitted along with the electrons and initially was called the neutron [23]. Later in 1932 Fermi called them *neutrino*.

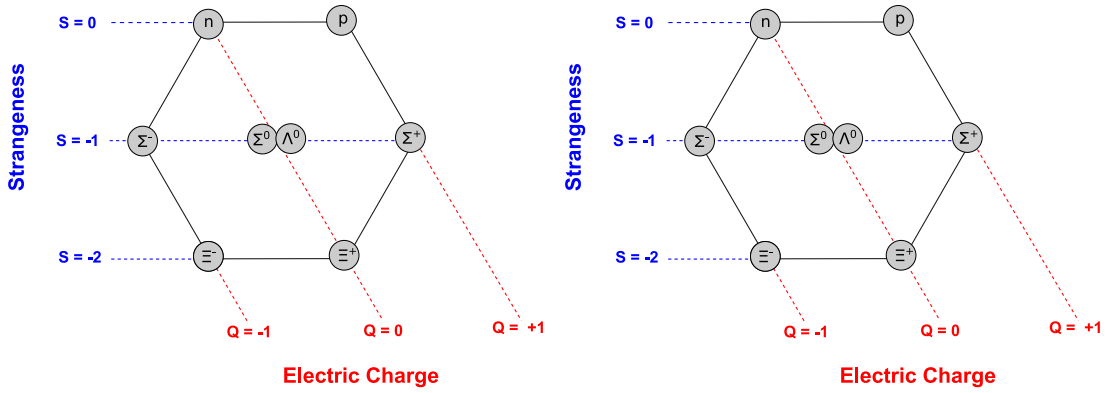
After the discovery of neutrons, more accurate measurements showed that neutrons are slightly heavier than protons. The neutron mass is greater than the combined mass of a proton and an electron. An isolated neutron decays into a proton and an electron within about 15 minutes in the same way as radioactive beta decay. It was understood that nuclear beta decay is a conversion of a neutron to a proton, an electron, and a neutrino. Neutrinos were first detected in 1956 by Clyde Cowan and Frederick Reines [24]. Later in 1962, Lederman, Schwartz, and Steinberger showed that there is more than one type of neutrino [25]. At that time, the neutrino mass was assumed to be zero. The third type of neutrino was proposed after the discovery of the third type of lepton, *tau* (τ) in 1975 [26]. The tau neutrino was detected 25 years later in 2000 by the DONUT experiment of Fermilab [27]. Over several years of experiments, it was found that neutrinos can convert from one type to the other. This phenomenon is generally known as neutrino oscillation and was first proposed by Bruno Pontecorvo in 1957 [28]. The oscillation probability is proportional to the neutrino mass difference square. Starting from 1998, several experimental pieces of evidence (Super-Kamiokande and Sudbury Neutrino Observatory) proved neutrino oscillations [29], and it became clear that the neutrinos have non-zero mass.

1.1.7 Strange particles (1947-1960)

In 1947, Rochester and Butler published their cloud chamber photograph showing a neutral particle decaying into two charged pions [30]. This particle showed a V-shaped decay pattern in cloud chamber photographs. It was later identified as a neutral kaon, K^0. In 1950 Anderson's group found another neutral 'V' particle, which decays into a proton and a pion [31]. It was later called the lambda particle, Λ. At this time, the concept of baryon number was already known, and these new particles were identified to be baryons. It was observed that these particles were produced copiously, but decayed very slowly. So, they were called strange particles. Later, we understood that these particles get produced by the strong interaction, but decay via the weak interaction. It was also observed that the strange particles were always produced in pairs. These productions follow a rule, e.g. Λ was produced always together with a K^+, but never with a K^-. To explain this behavior, Gell-Mann and Nishijihima assigned a new property to each particle called 'strangeness' [32–34]. Strangeness is conserved in any strong interaction, but is not conserved in the weak interaction. In the next few years more baryons were discovered, namely the Σ, Ξ particles.

1.1.8 The Eightfold Way (1961-1964)

With the discovery of strange particles, the number of particles started exploding. To find an underlying symmetry of these particles, first Sakata gave a model extending the isospin formalism by strangeness [35]. Thus p, n, and Λ form the Sakata $SU(2)$ triplet. Ikeda, Ogawa, and Ohnuki introduced the $SU(3)$ symmetry for the Sakata triplet [36]. However, this theory did do work well for all the hadrons. Recognizing the success of the $SU(3)$ symmetry, Murray Gell-Mann and Yuval Ne'eman proposed the *Eightfold Way* in 1961 [37, 38]. The name comes from Gell-Mann's 1961 paper. The Eightfold Way organizes all the mesons and baryons into geometric patterns according to their charge and strangeness. It was observed, that the relation between the hadrons is similar to that of the representation theory of $SU(3)$. Spin-1/2 baryons are organized into an octet, a hexagonal array, and two at the center as shown in Figure 1.1a. It consists of neutron, proton, $\Sigma^-, \Sigma^0, \Sigma^+, \Lambda^0, \Xi^-$ and Ξ^0. Similarly, meson octet was also build with eight lightest mesons, K^0, K^+, K^-, and $\bar{K}^0$, π^+, π^0, π^- and η as shown in Figure 1.1b. The three horizontal lines crossing the hexagonal shape determine the strangeness and corresponds to different values for mesons and baryons. For mesons, the top line has $S = 1$, the middle line $S = 0$, and the bottom line $S = -1$. The three lines in the baryon octet have $S = 0$, $S = -1$, and $S = -2$ respectively. The Eightfold Way also arranged the spin-3/2 baryons using the same principle, and they formed a decuplet. Nine out of the ten particles were already observed experimentally. Gell-Mann named the tenth particle Ω^-, which was expected to have strangeness -3 and electric charge -1. Gell-Mann also predicted the mass of this new particle. The particle was discovered in 1964 by the particle accelerator group at Brookhaven [39].



(a) Baryon octet (b) Meson octet

Figure 1.1: Octet of (a) spin-1/2 baryons and (b) mesons.

1.1.9 The quark model (1964)

It was still not understood why the hadrons follow the pattern predicted by the Eightfold Way. In 1964 Gell-Mann and Zweig independently proposed that all hadrons are composite particles [40–42]. The elementary constituents were called *quarks* by Gell-Mann. They proposed that three types (known as 'flavors') of quarks form an Eightfold Way triangle. The quarks have a fractional electric charge. They also proposed the anti-particles to each quark with opposite electric charge and strangeness. The anti-quarks form an upside-down triangle. Two composition rules were proposed for the quark bound states, baryons and mesons. Baryons are compositions of three

quarks[2]. Mesons are quark-antiquark bound states[3]. While enumerating the mesons with quark-antiquark combinations, a total of nine mesons were found. But in the Eightfold Way, there were only eight. The ninth meson should be a $s\bar{s}$ state with electric charge (Q) = 0 and $S = 0$. It turned out that such particle had already been found experimentally, the η' meson. η' was a singlet in the Eightfold Way. Finding the combinations for the baryons was a little bit more involved, but the quark model explained the previously predicted octets and decuplets nicely. Although the quark model predicts everything very nicely, isolated quarks were never found experimentally. This phenomenon was explained with the introduction of quark confinement [47–49]. The exact mechanism is still not completely known. The first experimental evidence was found in the deep inelastic scattering experiment at the Stanford Linear Accelerator Center in 1968 [50, 51].

1.1.10 SU(3) flavor symmetry

The idea of isospin symmetry that originated from proton and neutron could be extended to quarks. The strong interaction treats different quark types or flavors equally. The initial underlying symmetry is $SU(2)$ symmetry. The up/down flavor symmetry can be extended to include the strange quark. The three quark states can be expressed as a triplet, and the underlying symmetry will be $SU(3)$. In the limit of $m_u = m_d = m_s$ where all three quark masses are the same, this is an exact symmetry. But in reality, all three quark masses are different, and this is an approximate symmetry. It is know how to correct results obtained using this approximate symmetry, so it still makes this approximation useful.

1.1.11 SU(3) color

The spin-1/2 and spin-3/2 baryons with multiple quarks are expected to follows Pauli's exclusion principle. In 1964 after the discovery of the quark model, Oscar W Greenberg introduced a color charge to explain how quarks states can be represented by anti-symmetric wave functions [52]. Greenberg proposed the quarks possess an additional $SU(3)$ symmetry with three color charges, red, blue, and green. Eventually, it was predicted that quarks might interact via gluons, which correspond to the adjoint representation of the $SU(3)$ color symmetry. Unlike $SU(3)$ flavor, the $SU(3)$ color symmetry is an exact symmetry. The discovery of color later lead to the theory of the strong interaction known as Quantum Chromodynamics (QCD). In the next few years, the development of the QCD gauge theory started. The development of gauge theories started in 1954, when Yang and Mills formulated a non-abelian gauge theory with $SU(2)$ isospin to describe the strong interaction [53]. The quantization of non-Abelian gauge theories was studied by a number of other theorists in the 1960s [54–56]. Further developments happened in 1971 when 't Hooft showed that non-abelian gauge theories are renormalizable [57, 58]. The concept of color as the source of a strong field was developed into the theory of QCD by Harald Fritzsch and Heinrich Leutwyler, along with Gell-Mann in 1973 [59]. The first direct evidence of gluons were seen at the PETRA electron-positron collider of DESY, Germany, in 1979 [60–63].

1.1.12 The fourth quark and more particles

The fourth type of quark, the charm quark, was predicted before the discovery of the gluon. Bjorken and Glashow predicted it in 1964, but there was very little evidence to support its ex-

[2] In July 2015 LHCb experiment observed resonances involving exotic baryons with pentaquark states [43].

[3] Tetraquark bound states were discovered by Belle experiment in 2007 [44]. The observed by BESIII experiment in 2013 [45] and LHCb experiment in 2014 [46].

istence [64]. The credit is usually given to Glashow, Iliopoulos, and Maiani. In 1970, they proposed the GIM mechanism to explain the observed suppression of flavor changing neutral currents (FCNC) [65]. The GIM mechanism required a fourth flavor of quark. The discovery of the charm quark could be connected to the discovery of J/ψ meson. This J meson was first observed at Brookhaven by C. C. Ting's group in 1974 [66]. Ting named it J, but at the same time, a similar observation was made at SLAC by Burton Richter's group [67]. Richter called it ψ particle. So this particle is now called J/ψ. The true nature of this new particle was debated for a few months. Finally, the quark model explained it as a bound state of the charm quark, c. It was a great success of the quark model. With the addition of this new elementary particle, the quark model predicted charmed baryons and meson, which started to appear in experiments starting from 1975.

In 1973, Kobayashi and Maskawa extended the GIM mechanism to explain the CP violation in Kaon decays [68] and predicted the existence of the third generation quarks [69]. Two years later, in 1975, there was the first observation of the third generation elementary particle, τ lepton. Following that, in 1977, a new meson, *upsilon*, was discovered by the E288 experiment at Fermilab [70]. It was recognized as a bound state of a fifth quark, b (beauty or bottom). Prediction of the sixth quark, the top quark (t), was also made by Kobayashi and Maskawa. It took almost 20 years to confirm the existence of the top quark. It was discovered in 1995 by the Tevatron CDF [71] and D0 [72] experiments at Fermilab.

1.1.13 Weak interaction and intermediate bosons

Quantum Electrodynamics (QED) was introduced by Dirac in the late 1920s, and it laid the foundation for Quantum Field Theory (QFT). Major developments in QED calculations and the introduction of Feynman diagrams happened a bit later in 1947. Fermi proposed the theory of beta decay, $n \to pe^-\bar{\nu}_e$, in 1934, which matched loosely with QED. Fermi teated this process as a contact interaction and did not introduce any mediator particle. The strength of the weak interaction was given by the Fermi constant, G_F. Since the weak force has an extremely short range, Fermi's theory gave an excellent approximation at low energy. In 1956 T. D. Lee and C. N. Yang suggested that the weak interaction violates parity [73, 74]. Parity violation in weak interaction was experimentally shown by C. S. Wu and the collaborators in 1957 [75]. Fermi's theory still survived with some modifications. Fermi's original theory had a vector current of the form $\bar{\psi}_p\gamma^\mu\psi_n$. To include parity violation, it was replaced by a vector-minus-axial-vector (V-A) current, $\bar{\psi}_p\gamma^\mu(1-\gamma^5)\psi_n$. V-A theory was introduced in 1958 by Feynman, Gell-Mann, Sudarshan, and Marshak [76, 77]. Fermi's contact interaction theory was not renormalizable and was expected to fail at high energy. So a theory with a mediator, *intermediate vector boson*, was developed. Theorists started predicting the properties of the mediator particle. It gradually became clear that the new particle had to be a massive boson with electric charge. Initially, two charged bosons $W^\pm$ were proposed. At the same time, physicists tried to understand the underlying symmetry of the weak interaction. In analogy to the isospin in the strong interaction, a weak isospin doublet structure was proposed with electron and neutrino. The anticipation of a possible $SU(2)$ structure of weak interaction leads to the prediction of a neutral boson (W^0). Later in 1968, Glasgow, Salam, and Weinberg predicted the electroweak (EW) theory by combining weak and electromagnetic interaction [78–80]. They predicted three massive intermediate vector mediators charged $W^\pm$ and neutral Z boson. These particles were discovered by UA1 and UA2 experiments of CERN in the year 1983 [81–84].

1.2 Standard model

The term "standard model" was first used by Abraham Pais and Sam Treiman in 1975 while referring to the electroweak theory with four quarks [85]. The standard model (SM) of particle physics is a gauge field theory, invariant under the non-abelian $SU(3)_C \times SU(2)_L \times U(1)_Y$ gauge group. The subscripts do not have any meaning related to group theory, but refer to the physical application. The subscript C stands for the color of the strong interaction, L denotes the weak isospin, which refers to the left-chiral nature of the fermion coupling, and Y is the hypercharge. The SM Lagrangian density can be written as a combination of strong QCD and electroweak (EW) terms,

$$\mathcal{L}_{\text{SM}} = \mathcal{L}_{\text{QCD}} + \mathcal{L}_{\text{EW}}. \tag{1.2}$$

The two components will be described in subsection 1.2.2 and subsection 1.2.4, respectively. This summary is based on reference [86], and [87], and some inputs from [88]. Most of the experimental measurements are taken from the review of the Particle Data Group [29].

1.2.1 Particles in the standard model

Our current understanding is that ordinary matter is made out of three kinds of elementary particles: leptons, quarks and the force mediators. There are six leptons and they are classified in three generations. Each generation contains a charged lepton and the corresponding neutrino, which is electrically neutral. The quarks come in six flavors and are also divided into three generations. Each generation contains an up-type and down-type quark. The quarks also come in three colors, so there are 18 different kinds of quarks based on their color and flavor. The name of the quarks and leptons and their electric charge (Q) and color charges are summarized in Table 1.1. All these fermions participate in the weak interaction, so they also have weak isospin charge. The quarks and leptons have associated anti-particles.

There are four spin-1 bosons in the SM and they are the mediators of the electromagnetic, weak and strong force. The only spin-0 boson in the SM is the Higgs boson. The SM bosons are shown in Table 1.2. The gravity force carrier is hypothesized to be a spin-2 boson and is not discovered yet, nor it is included in our current standard model.

1.2.2 Quantum chromodynamics

Quantum chromodynamics is the gauge field theory that describes the strong interaction. The gauge group is denoted as $SU(3)_C$, where the suffix C denotes color. The eight gluons are represented by eight gluon fields, G^a, where the index a runs from 1 to 8. The quark field kinetic term is written using a covariant derivative, which leaves the kinetic term invariant under $SU(3)$ transformation.

$$D_\mu = \partial_\mu + i g_s t_C^a G_\mu^a \tag{1.3}$$

The matrices t_C^a are the eight Gell-Mann matrices and are a three-dimensional generalization of the Pauli matrices. They form an irreducible representation of $SU(3)$ and satisfy $[t_C^a, t_C^b] = 2if^{abc}t_C^c$, where f^{abc} are the antisymmetric structure constants of $SU(3)$. The corresponding gauge charge is color. The Gluon field strength can be written as

$$G_{\mu\nu}^a = \partial_\mu G_\nu^a - \partial_\nu G_\mu^a - g_s f^{abc} G_\mu^b G_\nu^c \,, \tag{1.4}$$

where g_s is the coupling strength of the strong interaction. The quark fields are denoted by q_r, where the flavor index takes value $r = u, d, s, c, t, b$. If we include the kinetic term along with the quark mass term the full Lagrangian can be written as

Table 1.1: The fermions of the Standard Model and their electric and color charges. Data collected from the Review of Particle Physics [29].

Generation	Particle Name	Symbol	EM Charge (Q)	Strong Charge (Color)	Mass (MeV)
			Leptons		
1st	electron	e	-1	0	$0.511 \pm 0.31 \times 10^{-8}$
	electron neutrino	ν_e	0	0	$< 1.1 \times 10^{-6}$
2nd	muon	μ	-1	0	$105.66 \pm 0.24 \times 10^{-5}$
	muon neutrino	ν_μ	0	0	< 0.19
3rd	tau	τ	-1	0	1776.86 ± 0.12
	tau neutrino	ν_τ	0	0	< 18.2
			Quarks		
1st	up	u	2/3	R/G/B	$2.16^{+0.49}_{-0.26}$
	down	d	-1/3	R/G/B	$4.67^{+0.48}_{-0.17}$
2nd	charm	c	2/3	R/G/B	$(1.27 \pm 0.02) \times 10^3$
	strange	s	-1/3	R/G/B	93^{+11}_{-5}
3rd	top	t	2/3	R/G/B	$(172.76 \pm 0.30) \times 10^3$
	bottom	b	-1/3	R/G/B	$(4.18^{+0.03}_{-0.02}) \times 10^3$

Table 1.2: The force carries and the Higgs boson of the Standard Model and their charges. Mass values are collected from the Review of Particle Physics [29].

Particle Name	Symbol	EM Charge (e)	Weak Charge (Isospin)	Strong Charge (Color)	Mass (GeV)
photon	γ	0	no	no	$< 1 \times 10^{-24}$
Z-boson	Z^0	0	no	no	91.188 ± 0.002
W-boson	$W^\pm$	± 1	yes	no	80.379 ± 0.012
gluon	g	0	no	yes	0
Higgs	h	0	yes	no	125.10 ± 0.14

$$\mathcal{L}_{\text{QCD}} = -\frac{1}{4} G^a_{\mu\nu} G^{a\mu\nu} + \sum_r \bar{q}_r \left(i\gamma^\mu D_\mu - m_r \right) q_r \,, \tag{1.5}$$

where γ^μ are the Dirac matrices.

1.2.3 Parton distribution function

The parton model was proposed by Feynman in 1969 to explain high energy hadron collisions [89]. This model proposed that any hadron can be considered as a composition of point-like constituents called 'partons'. In the same year, Bjorken and Paschos made a similar assumption to understand electron-proton deep inelastic scattering [90]. The parton model avoided introducing

any hypothesis about the nature of the constituents. Once the quark model was widely accepted, we understood that the partons are quarks and gluons.

The quark model describes the proton as a static three quark bound state. It consists of two up (u) quarks and one down (d) quarks. But a proton is much more complex. The quarks inside the proton interact strongly with each other via gluon exchange. The gluons can also split into a quark-antiquark pair, which essentially forms a sea of virtual gluons, quarks and antiquarks. The three initial quarks, uud, usually are called *valance quarks*, and the other quarks and antiquarks produced via gluon splitting are called *sea quarks*. Since the quarks inside the proton are always interacting, the quark momentum follows a distribution. These distributions are described via the parton distribution functions (PDF). A PDF gives us the probability of a parton having a momentum fraction between x and $x + \delta x$ of a proton. PDFs are probability density functions and are parameterized in terms of momentum fraction, x, and momentum transfer in the scattering, Q^2. The quark-parton model, explains the underlying process of a deep inelastic scattering as elastic scattering between an electron and a parton with momentum fraction, x. Thus the cross-sections can be described in terms of the PDFs. This description is experimentally verified. PDFs cannot be calculated from first principles of QCD, therefore, they have to be measured by experiments. Nonetheless their energy dependence (Q^2) is calculable using Dokshitzer-Gribov-Lipatov-Altarelli-Parisi (DGLAP) [91–93] equations. This ability enables us to calculate the PDFs relevant for LHC experiments extrapolating from lower energy collisions. It is crucial to have a very good knowledge of these PDFs to measure the cross-sections of all SM processes at LHC.

1.2.4 Electroweak theory

Electroweak theory is a unified gauge field theory of the weak and electromagnetic interactions. It is a non-abelian gauge theory described by the $SU(2)_L \times U(1)_Y$ gauge group. The weak interaction is chiral, and the subscript L denotes the left chiral nature of $SU(2)$ coupling. In the electroweak model of Glashow, Salam, and Weinberg, the $U(1)$ gauge symmetry of electromagnetism is replaced by a new local gauge symmetry $U(1)_Y$, where the subscript Y denotes weak hypercharge. Weak hypercharge relates the electric charge (Q) and the third component of weak isospin (T^3) as

$$Y = Q - T^3 \tag{1.6}$$

The Lagrangian density of the electroweak physics can be broken down into several terms:

$$\mathcal{L}_{\text{EW}} = \mathcal{L}_{\text{gauge}} + \mathcal{L}_{\text{fermion}} + \mathcal{L}_{\text{Higgs}} + \mathcal{L}_{\text{Yukawa}} \tag{1.7}$$

The gauge fields related to $SU(2)_L$ are denoted as W^a, where $a = 1, 2, 3$. The coupling strength of these gauge bosons is g. The gauge term looks very similar to the QCD one (Equation 1.4), but now $f^{abc} = \epsilon^{abc}$, the totally antisymmetric three-index tensor. Therefore, the $SU(2)_L$ field strength tensor can be written as

$$W^a_{\mu\nu} = \partial_\mu W^a_\nu - \partial_\nu W^a_\mu - g\epsilon^{abc} W^b_\mu W^c_\nu \,. \tag{1.8}$$

The $U(1)_Y$ part describes a gauge boson B, which has a coupling strength g'. The field strength tensor looks very similar to electromagnetism,

$$B_{\mu\nu} = \partial_\mu B_\nu - \partial_\nu B_\mu \,. \tag{1.9}$$

So the electroweak gauge term can be written as

$$\mathcal{L}_{\text{gauge}} = -\frac{1}{4} W^a_{\mu\nu} W^{a\mu\nu} - \frac{1}{4} B_{\mu\nu} B^{\mu\nu} \tag{1.10}$$

The gauge term gives the kinematics as well as the interactions of the gauge fields. Explicit mass terms of the gauge bosons are not gauge invariant, hence cannot be added to the Lagrangian. So, in presence of unbroken gauge invariance, the gauge bosons are massless.

The $\mathcal{L}_{\text{fermion}}$ component of the EW Lagrangian describes the fermion kinetic terms. Fermions are spin 1/2 particles and can be categorized into three different types, namely, Dirac (massive), Weyl (massless), and Majorana (particles that are their own anti-particles) fermions. All SM fermions are massive. Based on our current understanding, all fermions except the neutrinos are Dirac fermions. It is yet to be determined whether neutrinos are Dirac or Majorana fermions. The kinetic term can be written using the Dirac equation with the following covariant derivative

$$D_\mu = \partial_\mu \; + \; ig' B_\mu Y \; + \; ig t_L^a W_\mu^a \, . \tag{1.11}$$

Because of the chiral nature of the weak interaction, it only couples to left-chiral (left-handed) fermions and right-chiral (right-handed) anti-fermions[4]. Thus the left-handed leptons (e_L, μ_L, τ_L) and corresponding neutrinos $(\nu_e, \nu_\mu, \nu_\tau)$ each form a doublet under $SU(2)_L$, whereas the right handed leptons (e_R, μ_R, τ_R) forms a singlet under $SU(2)_L$ group. In the quark sector also the left-handed up-and down-type quarks form a doublet in each generation, whereas the right-handed quarks form singlets. Unlike the lepton sector, the quark sector has two singlets of the form u_R and d_R corresponding to the up-and down-type quarks, respectively. The fermions and their electric charge (Q), third component of the weak isospin (T^3) and weak hypercharge (Y) are summarized in Table 1.3.

Table 1.3: SM fermion fields and their electric charge (Q), third component of weak isospin (T^3) and weak hypercharge (Y) are summarized for first generation fermions. The same is true for the second and third generation leptons and quarks.

Fermion Type	Field	Q	T_3	Y
Quarks	$Q_L = \begin{pmatrix} u_L \\ d_L \end{pmatrix}$	+2/3 -1/3	1/2 -1/2	$\frac{1}{6}$
	u_R	+2/3	0	+2/3
	d_R	-1/3	0	-1/3
Leptons	$L_L = \begin{pmatrix} \nu_L \\ e_L \end{pmatrix}$	0 -1	+1/2 -1/2	$-\frac{1}{2}$
	e_R	-1	0	-1

Note that the left-handed and right-handed fermions carry different charges under $SU(2)_L \times U(1)_Y$ gauge group, hence it is not possible to mix them. But the Dirac mass term mixes the left and right-handed components and takes the following form $-m\bar{\psi}\psi = -m\bar{\psi}_R\psi_L - m\bar{\psi}_L\psi_R$. So, it is not possible to insert an explicit mass term for fermion fields. As a result, the fermion part of the EW Lagrangian only contains the kinematic term, and can be written as

$$\mathcal{L}_{\text{fermion}} = \sum_{i=1}^{3} \bar{Q}_L^i \gamma^\mu D_\mu Q_L^i \; + \; \bar{u}_R^i \gamma^\mu D_\mu u_R^i \; + \; \bar{d}_R^i \gamma^\mu D_\mu d_R^i \; + \; \bar{L}_L^i \gamma^\mu D_\mu L_L^i \; + \; \bar{e}_R^i \gamma^\mu D_\mu e_R^i \tag{1.12}$$

[4]The left and right components of the fermion fields can be obtained by projection operators: $\psi_L = \frac{1}{2}(1-\gamma^5)\psi$ and $\psi_R = \frac{1}{2}(1+\gamma^5)\psi$

The Higgs field is an $SU(2)_L$ doublet of the complex scaler field and can be written as

$$\Phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} \phi_1 + i\phi_2 \\ \phi_3 + i\phi_4 \end{pmatrix} . \tag{1.13}$$

The Higgs field transforms as a singlet under $SU(3)_C$, like the lepton fields, and it has a hypercharge $Y = 1/2$. The covariant derivative takes the same form as Equation 1.11, only the hypercharge will be replaced by that of the Higgs field. Thus the Lagrangian density of the Higgs sector can be given by

$$\mathcal{L}_{\text{Higgs}} = (D\mu\Phi)^\dagger (D_\mu \Phi) - V(\Phi) , \tag{1.14}$$

where the first term is the kinetic and gauge interaction term and the second term is the potential term of the form

$$V(\Phi) = \mu^2 \Phi^\dagger \Phi + \lambda \left(\Phi^\dagger \Phi\right)^2 . \tag{1.15}$$

It is necessary to have $\lambda > 0$, to make sure $V(\Phi)$ is bound, and allows a stable vacuum state. We will discuss more about the Higgs potential in the next section.

The Yukawa terms describes the coupling between the Higgs doublet and the fermions, which is needed to generate fermion masses. The Yukawa term can be written as

$$\mathcal{L}_{\text{Yukawa}} = \sum_{i=1}^{3} \sum_{j=1}^{3} \left[y^l_{ij} \bar{e}_{Ri} \Phi^\dagger L_{Lj} + y^u_{ij} \bar{u}_{Ri} \tilde{\Phi}^\dagger Q_{Lj} + y^d_{ij} \bar{d}_{Ri} \Phi^\dagger Q_{Lj} \right] + \text{h.c.} \tag{1.16}$$

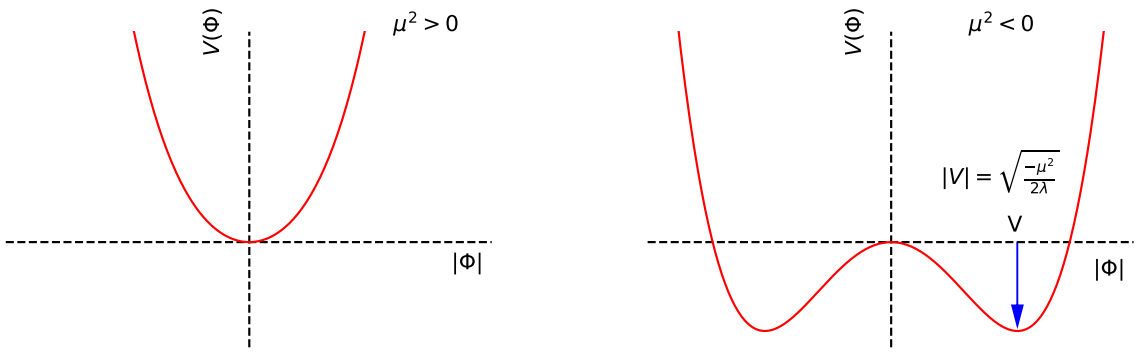
Here the indices i and j denote the three fermion generation and h.c. means the Hermitian conjugate term of the previous terms. The field $\tilde{\Phi}$ is the conjugate Higgs doublet, $\tilde{\Phi} = i\sigma^2 \Phi^*$. The $y^l_{ij}, y^u_{ij}, y^d_{ij}$ are the Yukawa coupling terms.

1.2.5 Electroweak symmetry breaking

As discussed in the last sections, in the presence of gauge invariance, all SM particles shall be massless. However, experimental evidence does not support it, and we know that all the SM particles have non-zero mass. Therefore, to solve this problem, the concept of *spontaneous symmetry breaking* is used to break the $SU(2)_L \times U(1)_Y$ symmetry spontaneously to $U(1)_Q$. It is done via the Brout-Englert-Higgs mechanism, or the Higgs mechanism, for short [94–96]. The $SU(2)_L$ doublet Higgs field was introduced in the SM to break the symmetry.

For $\mu^2 > 0$, the minimum of the potential $V(\Phi)$ is located at $\Phi = 0$. Electroweak symmetry is unbroken in the vacuum for this case, since a gauge transformation acting on the vacuum state $\Phi = 0$ does not change the vacuum state. When $\mu^2 < 0$, the potential minima are located away from zero, as shown in the schematic diagram Figure 1.2. In this case, the vacuum is not invariant under $SU(2)_L \times U(1)_Y$. As soon as we choose a specific ground state, the symmetry gets spontaneously broken. Without loss of generality we can choose the ground state such that $\langle 0|\phi_1|0\rangle = \langle 0|\phi_2|0\rangle = \langle 0|\phi_4|0\rangle$ and $\langle 0|\phi_3|0\rangle = v = \sqrt{-\mu^2/\lambda}$, where $|0\rangle$ represents the vacuum state. v is called the vacuum expectation value (VEV) of the Higgs field. Hence after the symmetry breaking, we are left with $U(1)_Q$. So, the vacuum state can be written as

$$\langle 0|\Phi|0\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v \end{pmatrix} \tag{1.17}$$



(a) $\mu^2 > 0$ (b) $\mu^2 < 0$

Figure 1.2: The scalar field potential, $V(\Phi)$ as a function of $|\Phi|$ for (a) $\mu^2 > 0$ and (b) $\mu^2 < 0$

Based on perturbative calculations, the Higgs field can be expanded around the minimum. Hence can be written as

$$\Phi(x) = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + h(x) \end{pmatrix} . \tag{1.18}$$

$h(x)$ here denotes a scalar field, and corresponds to the Higgs boson. This particular gauge choice is the basis of the Higgs mechanism.

Higgs mass

Once we choose the VEV the Lagrangian of the Higgs field can be written as

$$\mathcal{L}_{\text{Higgs}} = (D_\mu h)^\dagger (D^\mu h) - \frac{1}{2}(-2\mu^2)h^2 - \lambda v h^3 - \frac{1}{4}\lambda h^4 + \dots \tag{1.19}$$

where the covariant derivative is given by Equation 1.11. The second term in Equation 1.19 can be parameterized in terms of tree-level mass of the Higgs boson, m_H

$$m_H = \sqrt{-2\mu^2} = \sqrt{2\lambda} v . \tag{1.20}$$

Thus the Higgs mass is related to the VEV and the quartic coupling and needs to be determined experimentally.

Gauge boson mass and their interaction

The gauge boson interactions and their masses are obtained from the gauge kinetic term. We break the $SU(2)_L \times U(1)_Y$ symmetry to $U(1)_Q$, so it is convenient to write the gague bosons in a different basis where W^1_μ and W^2_μ get mixed

$$W^\pm_\mu = \frac{W^1_\mu \mp i W^2_\mu}{\sqrt{2}} . \tag{1.21}$$

The other two gauge bosons, W^3_μ, B_μ respect the $U(1)_Q$ symmetry, but they are not physical since the two terms mix together. The physical bosons, which can independently propagate, can be written by diagonalizing the mass matrix. The mass eigenstates can be written as

$$Z_\mu = \cos\theta_W W^3_\mu - \sin\theta_W B_\mu \,, \tag{1.22}$$

$$A_\mu = \sin\theta_W W^3_\mu + \cos\theta_W B_\mu \tag{1.23}$$

Here the weak mixing angle is defined as

$$\theta_W = \arctan\left(\frac{g'}{g}\right) \quad \text{,and} \quad \cos\theta_W = \frac{g}{\sqrt{g^2+g'^2}}, \quad \sin\theta_W = \frac{g'}{\sqrt{g^2+g'^2}} \,. \tag{1.24}$$

The mass terms of the gauge boson in this basis can be written as

$$\mathcal{L} \supset \frac{(g^2+g'^2)v^2}{8} Z_\mu Z^\mu + \frac{g^2 v^2}{4} W^+_\mu W^{-\mu} \,. \tag{1.25}$$

So, the mass of the Z and W bosons can be written as

$$m_Z = \frac{1}{2} v \sqrt{g^2+g'^2} \quad \text{and} \quad m_W = \frac{1}{2} g v \,. \tag{1.26}$$

Therefore the mass of the W boson is directly given by the Higgs VEV. m_W and g are directly measured, so we can determine $v \simeq 246$ GeV. The mass of the photon, m_A is zero.

Gauge boson interaction terms with the Higgs field are also generated from the gauge kinetic term and can be written as

$$\mathcal{L} \supset \frac{g^2 v}{2} h W^+_\mu W^-_\mu + \frac{g^2}{4} h h W^+_\mu W^-_\mu + \frac{(g^2+g'^2)v}{4} h Z_\mu Z^\mu + \frac{(g^2+g'^2)}{8} h h Z_\mu Z^\mu \tag{1.27}$$

All these couplings can be uniquely determined if we know the Higgs VEV, v, W and Z boson masses.

Fermion mass

Fermion masses come through the Yukawa term, where the fermion states are written in the weak-eigenstate basis. y^l, y^u and y^d shown in Equation 1.16 are 3×3 matrices corresponding to lepton, up-and down-type quarks. Lepton masses are directly obtained from the Yukawa coupling and can be expressed as

$$m_l = \frac{y^l v}{\sqrt{2}} \,, \tag{1.28}$$

where $l = e, \mu, \tau$. The fermion masses are given by the Higgs VEV. Neither the fermion masses nor the Yukawa couplings are predicted by the SM. They are determined from measurements. The experimental data of measured fermion masses validates the nature of Higgs coupling. It is observed that the couplings to the Higgs field are proportional to their masses as shown in Figure 1.3. Neutrinos are assumed to be massless since the right-handed neutrinos are not defined in the SM. After the experimental confirmation of neutrino oscillations, we know that the neutrinos are massive, but the exact mechanism of generating neutrino mass is still not determined.

The neutrino weak or flavor eigenstates are related to the mass eigenstates by the 3×3 unitary Pontecorvo-Maki-Nakagawa-Sakata (PMNS) matrix.

$$\begin{pmatrix} \nu_e \\ \nu_\mu \\ \nu_\tau \end{pmatrix} = \begin{pmatrix} U_{e1} & U_{e2} & U_{e3} \\ U_{\mu 1} & U_{\mu 2} & U_{\mu 3} \\ U_{\tau 1} & U_{\tau 2} & U_{\tau 3} \end{pmatrix} \begin{pmatrix} \nu_1 \\ \nu_2 \\ \nu_3 \end{pmatrix} \tag{1.29}$$

The $W^{\pm}$ couples to the leptons belonging in the same generation. On the other hand, cross-generational couplings of the W boson exist in the quark sector. In 1963 Cabibbo proposed that the weak interactions of the quarks have the same strength as the leptons, but the weak eigenstates in the quark sector are different from the mass eigenstate. The mass eigenstate of the quarks can be determined from the weak eigenstate using the Cabibbo-Kobayashi-Maskawa (CKM) matrix,

$$\begin{pmatrix} d_1 \\ s_1 \\ b_1 \end{pmatrix} = V_{\mathrm{CKM}} \begin{pmatrix} d \\ s \\ b \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix}. \tag{1.30}$$

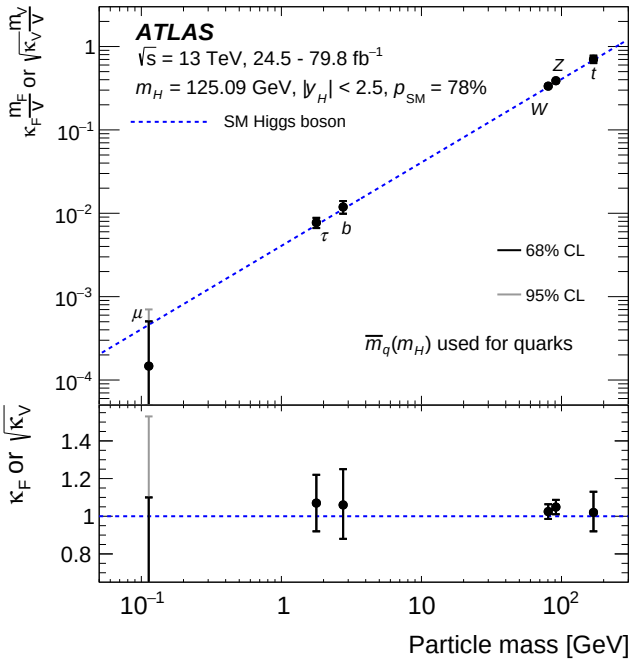


Figure 1.3: Reduced coupling-strength modifiers for fermions, κ_F ($F = t, b, \tau, \mu$) and weak gauge bosons, κ_V ($V = W, Z$) as a function of their masses m_F, m_V, and the vacuum expectation value of the Higgs field $v = 246$ GeV. The SM prediction for both cases is also shown (dotted line). The black error bars represent 68% CL intervals for the measured parameters. For κ_μ the light error bars indicate the 95% CL interval. The coupling modifiers are measured assuming no BSM contributions to the Higgs boson decays, and the SM structure of loop processes. The lower panel shows the ratios of the measured values to their SM predictions. Figure taken from [97]

There are nine elements in the CMK matrix, but they are not all independent. So they can be expressed in term of three 'generalized Cabibbo angles'. The elements of the CKM matrix are

measured in different flavor physics experiments, and the values can be found in the particle data group review [29]. Flavor changing charged current in the weak interaction is observed due to the non-zero off-diagonal terms of the CKM matrix. The observed CP violation in the SM can be accommodated by a single complex phase in the CKM matrix.

1.3 Top physics

The top quark is the heaviest known elementary particle with a mass of around 172.5 GeV. Being so much heavier than the W boson, the top quark is the only quark that can decay into an on-shell W-boson. Top-quark lifetime is very short (5×10^{-25} s). Moreover the top quarks decay before forming hadron bound states as the characteristic hadronization time of QCD is around 28×10^{-25} s, almost $\mathcal{O}(10)$ times longer than the top-quark lifetime. The top quark is also the only quark whose Yukawa coupling is close to unity. Therefore top quarks play a very important role in the standard model of particle physics and its extensions. This dissertation summarizes two searches for new physics involving top quark. This section gives a brief overview of top-quark production and its decay modes, relevant to the studies presented here. A more comprehensive summary of the top-quark phenomenology can be found in [29].

1.3.1 Top-quark production

In hadron colliders like LHC and Tevatron, the top quarks are produced predominantly in pairs through quark anti-quark annihilation ($q\bar{q} \rightarrow t\bar{t}$) and gluon fusion ($gg \rightarrow t\bar{t}$) at leading order (LO) of QCD. At the Tevatron $p\bar{p}$ collider ($\sqrt{s} = 1.96$ TeV), the dominant production mode was $q\bar{q}$ annihilation, which produced almost 85% of $t\bar{t}$ events. Whereas at the LHC almost 80 – 90% $t\bar{t}$ pairs are produced via gluon fusion [29]. Figure 1.4 shows the four LO Feynman diagrams for top-quark pair production via $q\bar{q}$ annihilation and gluon fusion.

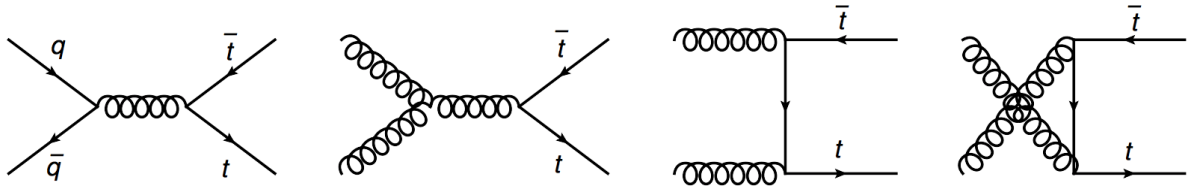


Figure 1.4: LO Feynman diagrams of the top-quark pair production.

$t\bar{t}$ pair production total cross sections are predicted at next-to-next-to-leading order (NNLO) including next-to-next-to-leading-log (NNLL) soft-gluon resummation [98]. In general $m_t = 172.5$ GeV is used while calculating the cross-sections for LHC. The cross-sections are summarized in Table 1.4, where the first uncertainty is from the scale dependence and the second one comes from the variations of parton distribution function. A summary of LHC and Tevatron measurements of the top-pair production cross-section as a function of the center-of-mass energy is shown in Figure 1.5. It can be seen that the measured cross-section values are very consistent with the theoretical predictions. This figure also shows that the production cross-section is almost the same in the proton-proton and proton-antiproton collision processes above the center-of-mass 10 TeV. Furthermore, the $t\bar{t}$ production cross-section is expected to increase only by a factor of 1.8 when the LHC center-of-mass energy changes to 14 TeV.

The secondary production mode of top quark at the LHC is via electroweak single-top production. Single-top quark production in the SM is via charged current, which involves tWb vertex. Single-top can be produced via s-channel ($q\bar{q}' \to t\bar{b}$), t-channel ($gq \to q't$) and Wt associate production ($bg \to Wt$). At the LHC, the contribution from s-channel is very low.

Table 1.4: The best NNLO + NNLL theoretical prediction for $t\bar{t}$ pair production cross section at Tevatron and LHC. The top quark mass (m_t) is assumed to be 173.3 GeV, close to Tevatron + LHC average. The values are taken from LHC Top Working group calculations [99].

Collider, $\sqrt{s}$	$\sigma_{t\bar{t}}$ [pb]
Tevatron 1.96 TeV	$7.16^{+0.11+0.17}_{-0.20-0.12}$
LHC 7 TeV	$177.3^{+4.6+9.0}_{-6.0-9.0}$
LHC 8 TeV	$252.9^{+6.4+11.5}_{-8.6-11.5}$
LHC 13 TeV	$831.8^{+19.8+35.1}_{-29.2-35.1}$
LHC 14 TeV	$984.5^{+23.2+41.3}_{-34.7-41.3}$

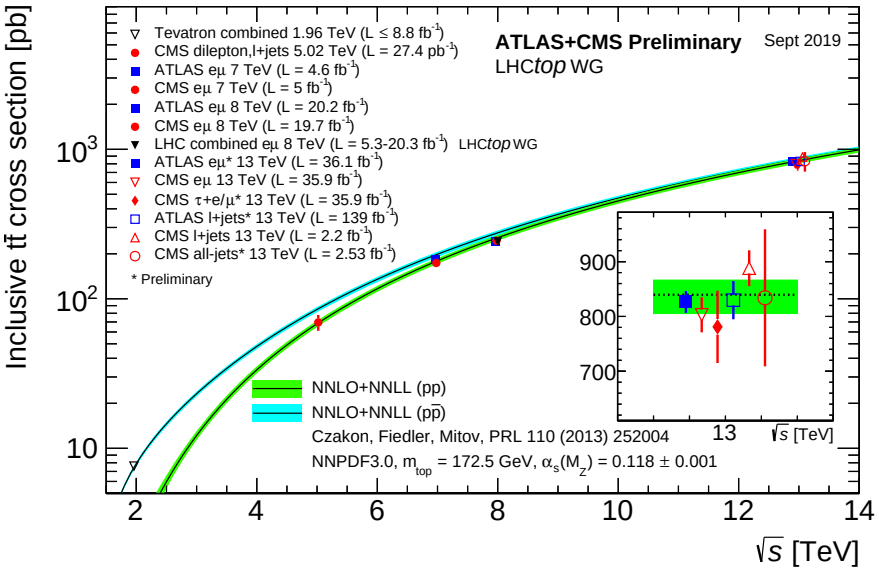


Figure 1.5: Summary of the LHC and Tevatron measurements of the top-pair production cross-section as a function of the center-of-mass energy compared to the NNLO QCD calculation complemented with NNLL resummation (Top++2.0). The theory band represents uncertainties due to renormalization and factorization scale, parton density functions and the strong coupling. The theoretical calculations and the measurements are quoted at m_t = 172.5 GeV. There are two different theory lines, the pp line is used to compare LHC measurements and the $p\bar{p}$ line is used to compare the Tevatron measurement. Figure taken from [100]

1.3.2 Top-quark decay

The top-quark almost always (99.9%) decays into a b-quark and a W boson. This is a consequence of the V_{tb} term of the CKM matrix being very close to unity. The b-quarks hadronize and form b-jets. 68% of time the W boson decays hadronically into two quarks ($u\bar{d}$ or $c\bar{s}$). And it decays leptonically into a charged lepton and corresponding neutrino ($l\nu_l$) roughly 32% of time [29]. We generally call the corresponding top-quarks containing hadronically or leptonically decaying W bosons as 'hadronic' or 'leptonic' top, respectively. $t\bar{t}$ processes are typically classified according to the final-state signatures, determined by the decay channels of the two W-bosons. The three different $t\bar{t}$ final-state signatures are

- **Fully hadronic channel:** Both W bosons decay hadronically, $t\bar{t} \rightarrow (bW^+)(\bar{b}W^-) \rightarrow (bq\bar{q}_2)(\bar{b}q_3\bar{q}_4)$. The corresponding branching fraction is $\approx 46\%$. We expect six jets in the final state at LO.
- **Single lepton (semi-leptonic, lepton+jets) channel:** Here one of the W boson decays hadonically and the other one leptonically, $t\bar{t} \rightarrow (bW^+)(\bar{b}W^-) \rightarrow (bq\bar{q}_2)(\bar{b}l^-\bar{\nu}_l)$ or $(bl^+\nu_l)(\bar{b}q\bar{q}_2)$. If we consider three generations of leptons, the total branching fraction of the single lepton channel is around 45%. Four jets, one charged lepton and missing transverse energy corresponding to the neutrino is expected at LO. Generally, only electron and muon final states are considered as single lepton final state and tau final states are categorized separately.
- **Dilepton (fully lepton) channel:** In this decay channel both the W boson decays leptonically giving rise to final state like this, $t\bar{t} \rightarrow (bW^+)(\bar{b}W^-) \rightarrow (bl^+\nu_l)(\bar{b}l_2^-\bar{\nu}_{l_2})$. This channel has 9% branching ratio. Two jets (b-jets), two opposite charged leptons and missing transverse energy are expected at LO.

All different $t\bar{t}$ decay branching fractions are shown in Figure 1.6. Tau final states are not considered in the lepton+jets final state.

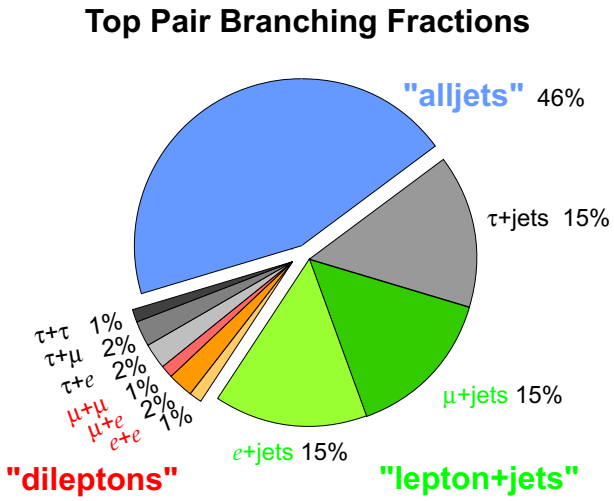


Figure 1.6: The branching ratios (BR) of different $t\bar{t}$ decay channels are shown in this pie chart. The blue color shows the fully hadronic (alljets) BR of 46%, green sections are showing the semi-leptonic decay (~30% with e and μ), the orange-red section shows the dilepton channel (~4% without τ contribution). The contributions coming from τ final states are shown in gray. Figure taken from [101].

1.4 Success and limitations of standard model

The standard model of particle physics is an elegant theoretical framework based on non-abelian gauge field theory where the exact nature of the particle interactions is determined by the local gauges. The SM describes ordinary matter particles and their interactions. The predictions have been tested in different particle physics experiments over several decades. The measured values of all SM parameters and the cross-sections of different reactions are being measured with very high precision. Figure 1.7 shows a summary of measured cross-section values of several SM processes by the ATLAS experiment [102]. These measurements are done at a center-of-mass energy of $\sqrt{s} = 7, 8$, and 13 TeV using proton-proton collision data. Most of the measured cross-sections agree very well with the theoretical calculations. The third and fourth row from the top shows the Z boson, and $t\bar{t}$ pair production cross-sections, respectively. These two processes are highly relevant to the searches presented in this dissertation. Current precision test results and the discovery of the Higgs boson established the validity of the SM up to the electroweak scale.

One problem with the SM is that it has many free parameters. The values of these free parameters are not predicted by the SM and are chosen by various assumptions and measurements. If we consider neutrinos are Dirac fermions, then there are a total of 26 free parameters in the SM. The parameters are:

- Twelve Yukawa couplings for a total of twelve fermions of the SM. They can be expressed as the mass of the fermions.
- Three coupling constants of the gauge interactions: g_s, g, g'. These quantities are often expressed as strong coupling α_s, QED coupling α and Fermi coupling for weak interaction G_F.
- Two parameters describing the Higgs potential, the VEV (v) and the Higgs mass (m_{H}).
- A total of eight mixing angles for the CKM and PMNS matrices.
- The θ_{QCD} parameter, usually assumed to be zero.

If we put aside the θ_{QCD} parameter, of the 25 SM parameters, 14 are related to the Higgs sector, eight with the flavor sector, and three with the gauge interactions. With so many parameters, the SM sometimes appears to be a model where the parameters are chosen to match observed data. The mechanism of neutrino mass generation is still not known to us, and the neutrino masses are not experimentally measured yet. Furthermore, all the fermion masses are free parameters in the SM. So the fermion mass generation looks ad hoc, and it is not clear why the SM does not predict them. There are also some underlying structures in the fermion masses of different generations, which is unlikely to be incidental. The three gauge interactions are also of a similar order of magnitude, indicating they might be a lower energy manifestation of a parameter of a Grand Unified Theory (GUT).

Despite the success, there are several issues with the standard model and it can not explain many experimental observations. One of the famous examples is the overwhelming evidence of the existence of dark matter over a very large range of astrophysical scales. The SM does not describe this non-luminous dark matter (DM). Another problem is that we see more matter than anti-matter in the universe. If the early universe started from matter-antimatter equality then there has to be CP violation to explain the current asymmetry. Although CP is violated in the SM through the CKM phase, this effect is too small to explain the observed baryon asymmetry. So,

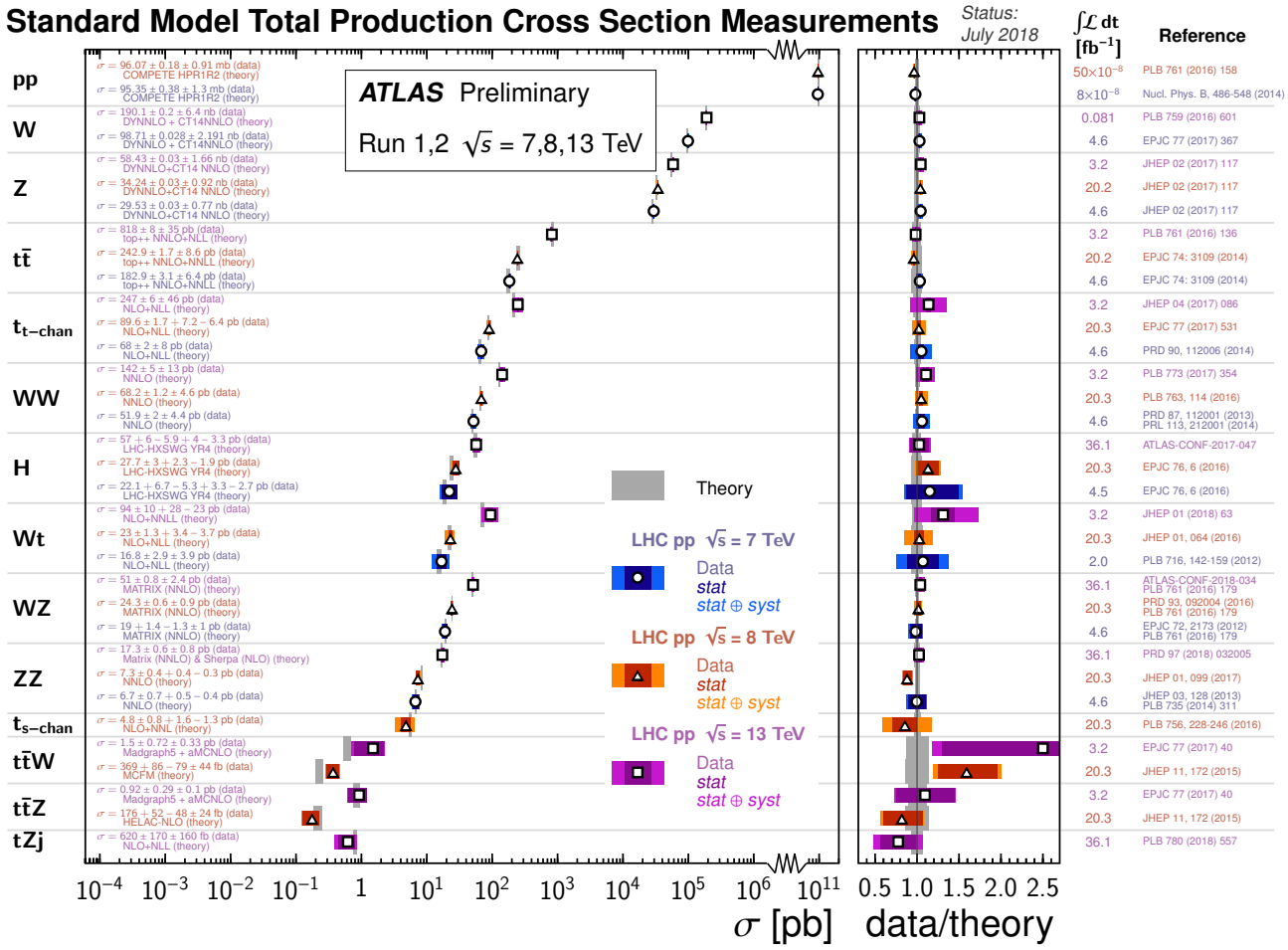


Figure 1.7: Summary of several Standard Model total production cross section measurements, corrected for leptonic branching fractions, compared to the corresponding theoretical expectations. All theoretical expectations were calculated at NLO or higher. The dark-color error bar represents the statistical uncertainty. The lighter-color error bar represents the full uncertainty, including systematics and luminosity uncertainties. The data/theory ratio, luminosity used and reference for each measurement are also shown. Uncertainties for the theoretical predictions are quoted from the original ATLAS papers. They were not always evaluated using the same prescriptions for PDFs and scales. Not all measurements are statistically significant yet. Figure taken from [102].

far we have not found such CP-violating processes. The origin of neutrino mass is another unsolved problem and SM does not include any description of the neutrino mass generation. After the discovery of the Higgs boson by the ATLAS [103] and CMS [104] collaborations the mechanism of electroweak symmetry breaking became more concrete. But the existence of the 125 GeV Higgs boson introduces a conceptual problem associated with the quantum correction of its mass term. The Higgs boson mass gets corrected by loops of other particles. In the SM, the dominant contribution comes from the top quark loop to the Higgs propagator. If the momentum of these virtual particles, corresponding to the loops, are cut off at a scale Λ, then the correction to the observed Higgs mass, relative to the bare boson mass with no loop correction, is proportional to Λ^2. Λ, usually, corresponds to the relevant scale of the theory. So, the Higgs mass correction is not

a problem up to the electroweak scale. But if we assume that the SM is a part of a more complete theory valid up to a very high energy scale of a GUT, $\Lambda_{GUT} \sim 10^{16}$ GeV or Planck scale, $\Lambda_P \sim 10^{19}$ GeV, then these mass corrections become very large. This problem is known as *electroweak hierarchy problem* or *fine-tuning problem*. Many extensions to the SM predicts new particles that cancel the mass corrections. Furthermore, there are also stability issues of the SM vacuum. We have not directly measured the Higgs quartic coupling, λ, yet. The Higgs quartic coupling tells us the exact shape of the Higgs potential. Based on current data we know that the SM vacuum is in a meta-stable state [105]. Besides, the SM does not provide us any description of gravity. There are several other problems discussed in details in [106, 107]. To address all these problems, plenty of theories Beyond the SM (BSM) were developed. A few of those theories will be summarized in the next section.

1.5 Beyond standard model

Many BSM theories have being developed over the past three decades. In general, these GUT theories extend the gauge of the SM and embed the SM into a larger symmetry group. Thus, these theories introduce a lot more particles and interactions. A new force carrier could be heavy with a mass in the TeV range, hence measurable in the LHC. Models with an alternative mechanism of electroweak symmetry breaking were popular before the Higgs discovery, but some of the modified versions of those theories are still relevant for LHC physics. There are also models with an extra dimension that includes gravity and solves the hierarchy problem. Some of the BSM models are very generic minimal extensions of the SM with an additional $U(1)$ gauge. Hence they predict a new gauge boson. In the dissertation, searches for such new heavy gauge bosons are described. The models are assumed to have non-zero coupling to the SM particles, thus can decay into SM particles. Two classes of searches are presented in this dissertations, where the new particle decays into two leptons (electrons and muons) or a top anti-top quark pair.

If the new vector boson has properties similar to other SM gauge bosons, then we expect it to decay into SM particles. Thus, one simple final state will be a decay to two leptons like electrons or muon. The existence of such new particles will, most of the time, enhance the production rate of SM particles. Thus, if we can estimate the di-lepton production from all the SM processes, we can see if there is any hint of new physics by comparing our estimate with data. In this dissertation, a direct search for new high mass particles is performed with LHC data collected in the year 2015 and 2016. A high mass particle with significant coupling to the leptons will manifest as a resonance. Several benchmark BSM signal models are used to interpret the results of this search. Two such models are a GUT model based on E_6 algebra and Sequential Standard Model (SSM).

However, the new physics could be leptophobic and could have very minimal coupling to leptons. Many theories assume that the leptophobic BSM sector could couple strongly with the SM quarks. In that case, an absence of any new physics in dilepton pair production will not rule out these theories. Such a boson can instead form a di-quark resonance. Particularly the final states with top-quarks are interesting. Due to its uniques properties, the top quark stands out among all other SM particles. Its mass is a few orders of magnitude higher than the first generation quarks. The top-quark mass is close to the VEV of the Higgs field, making its Yukawa coupling close to unity. The highest contribution to the Higgs mass correction comes from the top-quark. Thus, it is common to have new physics associated with the top-quark in the models addressing naturalness. In models where Higgs is a pseudo-scalar, like Little Higgs models, a top-quark partner t' with a different electroweak coupling is predicted. These t' and third-generation particles of supersym-

metric models are expected to have large coupling to the SM top-quarks. So, their existence will show up in the top-quark finals states. Some other models predict heavy boson, Z', with preferential coupling to top-quarks. The propagator and its couplings to top quarks themselves might be sensitive to any such new heavy particles. Thus, precision measurements of top-quark properties can be considered as indirect searches for new physics. Almost in all the top measurements, the data is found to be consistent with the theoretical predictions. In this dissertation, direct searches for resonant signals in the $t\bar{t}$ final state are presented. The results are interpreted with benchmark Topcolor assisted Technicolor model and simplified dark matter models.

1.5.1 GUT model based on E_6 algebra

This kind of theory is motivated by superstring models and can be reduced to the SM after symmetry breaking.

$$E_6 \rightarrow SO(10) \times U(1)_\psi \rightarrow SU(5) \times U(1)_\chi \times U(1)_\psi$$

where $SU(5)$ would break down into G_{SM}, and the additional local $U(1)$ symmetries will be associated with new gauge bosons. Most studies assume only one Z' boson, with coupling being the linear combination

$$Z'(\theta_{E_6}) = Z'_\chi \cos\theta_{E_6} + Z'_\psi \sin\theta_{E_6} \tag{1.31}$$

where $\theta_{E_6} \in [-\pi, \pi]$ is the mixing angle. Different values of the mixing angle correspond to different state of Z' as described in [108] and summarized in Table 1.5.

Table 1.5: List of commonly used Z' models.

Model Name	Z' **State**	θ_{E_6} **Value**
ψ model	Z'_ψ	$\pi/2$
η model	Z'_η	$\pi - \tan^{-1}(\sqrt{\frac{5}{3}})$
Neutral N model	Z'_N	$\tan^{-1}(\sqrt{15})$
Insert model	Z'_I	$\tan^{-1}(\sqrt{\frac{3}{5}})$
Secluded sector model	Z'_S	$\tan^{-1}(\sqrt{\frac{15}{9}})$
χ model	Z'_χ	0

1.5.2 Sequential Standard Model

The sequential standard model (SSM) [109] is the simplest extension of SM with an additional spin-1 heavy gauge bosons generally called Z', W'. Z'_{SSM} boson has the same coupling to fermions as the SM Z boson. It serves as a useful reference case when comparing constraints from various sources. Many of the LHC analyses, including the ones presented here, use this model as a benchmark for setting limits and generating simulated events. The Z'_{SSM} can decay into both two leptons and quarks. Unlike the SM, the self-coupling of these new bosons are assumed to be zero. Also, their coupling to the SM W and Z boson is assumed to be zero.

1.5.3 Topcolor assisted Technicolor model

Several theories were introduced to solve the electroweak hierarchy problem introduced by the electroweak symmetry breaking (EWSM). One idea proposes that the Higgs boson is a fermion bound state, where the fermions are bound by a new strong interaction. In such strongly-coupled systems, the electroweak symmetry breaking could be dynamic, analogous to the spontaneous breaking of approximate chiral symmetry breaking in QCD. The possibilities of using simple QCD for EWSB were largely ruled out after discovery of a SM-like Higgs boson. A scaled-up version of this idea is based on a new non-Abelian gague group, G_{TC} together with additional *techniquarks* that are charged under both G_{TC} and electroweak. Such theories are generally called Technicolor theories [110, 111]. The initial QCD-like Technicolor was introduced by Weinberg [112] and Susskind [113] in the late 1970's. In the Technicolor theories, EWSB is achieved by adding a new technicolor gauge force that becomes strong near 100 GeV. Technicolor models have several problems, but two important features are: all the fermions are massless and the existence of Higgs-like scaler is not clear. To fix some of the issues of the Technicolor model, SM gauge group (G_{SM}) and Technicolor gauge group are connected. These models are called Extended Technicolor model (ETC) [111]. In ETC, G_{SM} and G_{TC} are embedded in a larger gauge group $G_{ETC} \supset G_{SM} \times G_{TC}$, and the SM fermions are connected with the techniquarks within representations of G_{ETC}. It can generate nonzero fermion mass, but the third generation fermion masses, particularly the top-quark mass, were too small. To fix top-quark mass, *topcolor* models were connected to Technicolor models.

The topcolor model is a dynamic EWSB model where the large top-quark mass gets generated through the formation of a dynamic $t\bar{t}$ condensate. The fermion condensates are formed by a new strong gauge force, which preferentially couples with the third generation quarks. In the topcolor scheme the QCD gauge group, $SU(3)_C$ is embedded into a minimal $SU(3)_1 \times SU(3)_2$ group. The first gauge interaction, $SU(3)_1$, is stronger and only acts on the third generation quarks. While the second gauge interaction, $SU(3)_2$ is weaker and acts on the first and second generation quarks. A color octet of massive bosons, *topgluons*, is produced when the $SU(3)_1 \times SU(3)_2 \rightarrow SU(3)_C$ symmetry is broken. These bosons couple mainly to $t\bar{t}$ and $b\bar{b}$ condensates and produce degenerate top- and bottom-quarks around 600 GeV [114]. To get the top-quark mass scale correct, the topcolor model needs to be combined with some other theory. Topcolor assisted Technicolor (TC2) [115, 116] model is one such development.

To make the masses of the two quarks different, a tilting mechanism is implemented in this model by introducing an additional strong $U(1)'$ interactions attractive to $t\bar{t}$, but repulsive to $b\bar{b}$. This additional symmetry corresponds to a neutral gauge boson Z'_{TC2}. The simplest approach is to embed the weak hypercharge, $U(1) \rightarrow U(1)_1 \times U(1)_2$. The up- and down-type quarks are given different $U(1)$ charges to achieve the tilting. So the gauge group structure of the topcolor with the $U(1)'$ tilting mechanism takes the following form

$$SU(3)_1 \times SU(3)_2 \times U(1)_1 \times U(1)_2 \times SU(2)_L \rightarrow SU(3)_C \times U(1)_Q$$

There are several models based on the scheme of the tilting mechanism. The model IV referred to in the article [114] represents one special case, which couples strongly to the first and third-generation quarks and has almost zero couplings to leptons. The predicted cross-sections of such leptophobic topophyllic Z'_{TC2} are large enough to be experimentally accessible at the LHC [117]. This model is parametrized by the topcolor fitting parameter, $\cos\theta_H$, which controls the width and production cross-section. The tilting parameters are denoted as f_1 and f_2, which control the

coupling to up-type and down-type quarks, respectively. The TC2 model used in this dissertation (model IV) for $t\bar{t}$ resonance searches has $f_1 = 1$ and $f_2 = 0$. These values are chosen to maximize the $t\bar{t}$ branching ratio.

After the discovery of the Higgs boson, a wide range of Technicolor models are no longer considered as an alternative theory of EWSB, but it is demonstrated that a subset of these models with *walking technicolor* can accommodate a Higgs-like scalar. Then the Higgs has to be a composite particle. However, the motivation of using the TC2 model in the $t\bar{t}$ resonance search does not fully rely on the validity of such. Since the signal widths are narrower than the experimental resolution, the derived upper limits on production cross-section are valid for any narrow width Z' signal. Moreover, this particular benchmark model was used in previous searches conducted by ATLAS and other experiments. The choice of this benchmark model thus allows for direct comparisons with previous results.

1.5.4 Dark matter mediator models

The dark matter (DM) mediator model is one of simplified benchmark models with a s-channel mediator, Z'_{DM}, described in the LHC Dark Matter working group report [118]. This is a simple extension to the SM with an additional $U(1)$ gauge symmetry, in which a dark matter candidate particle has charge only under this group. In this model, some of the SM particles also have change under this group, hence this new gauge boson can mediate interactions between the SM and DM particles. Two models are considered where the spin-1 Z' has vector and and axial-vector couplings with the SM and DM fields. The dark matter particle χ in this model is considered to be a Dirac fermion with mass m_χ and gets produced via the exchange of the Z'_{DM} mediator. The simple Lagrangian for the vector and axial-vector mediators can be written as

$$\mathcal{L}_{\text{vector}} = g_q \sum_{q=u,d,s,c,t,b} Z'_\mu \bar{q}\gamma^\mu q + g_\chi Z'_\mu \bar{\chi}\gamma^\mu \chi \tag{1.32}$$

$$\mathcal{L}_{\text{axial-vector}} = g_q \sum_{q=u,d,s,c,t,b} Z'_\mu \bar{q}\gamma^\mu\gamma^5 q + g_\chi Z'_\mu \bar{\chi}\gamma^\mu\gamma^5 \chi\,, \tag{1.33}$$

where g_q and g_χ are the quark and DM coupling, respectively. These simplified models have five free parameters, g_q, g_χ, lepton coupling g_l, m_χ and mediator mass $M_{Z'}$. For the models considered in this analysis, we specifically assume the mediator does not couple to lepton, $g_l = 0$. The quark coupling is assumed to be universal for all cases, thus it takes the same value for all quarks. For the vector and axial-vector mediator model we assumed $g_\chi = 1$ and $g_q = 0.25$ following the benchmark values recommended by the LHC Dark Matter Working Group [119, 120]. Many different values of m_χ and $M_{Z'}$ are used for the new physics searches described in this dissertation.

CHAPTER

TWO

Experimental Setup

The research works described in this dissertation, is performed using the data from one of the experiments of the European Organization for Nuclear Research, known as CERN[5]. The laboratory is on the French-Swiss border outside Geneva. CERN is the largest particle physics laboratory in the world. CERN has an international working environment with representatives of 110 nationalities, and researches are spread around 1000 institutes or universities in 70 different countries. There are several experiments around different beam-lines in the CERN complex as shown in Figure 2.1. I worked on one of the experiments located around the Large Hadron Collider.

2.1 The Large Hadron Collider

The Large Hadron Collider (LHC) [122] is a hadron accelerator, which most of the time collides protons with protons. It is installed in a 26.7 km long underground tunnel, which lies between 45 m and 170 m below the surface. The LHC uses the same tunnel previously used by the Large Electron Proton (LEP) collider [123]. There are two rings with counter-rotating beams. The beams are allowed to cross at the four interaction points, where the four LHC experiments, ATLAS [124], CMS [125], LHCb [126] and ALICE [127] are located. The LHC is designed to collide proton beams at the center-of-mass energy of $\sqrt{s} = 14$ TeV with an instantaneous luminosity of 1×10^{34} cm^{-2}s^{-1}. This dissertation uses only the proton-proton collision data.

2.1.1 The LHC acceleration chain

The LHC relies on a series of low energy accelerators. For proton-proton collision runs, the process starts with Hydrogen atoms. Protons are created by stripping the electron out of Hydrogen atoms using an electric field. The resulting protons are accelerated to an energy of 50 MeV using the linear accelerator, LINAC2. The protons are then sent to the Proton Synchrotron Booster (PSB) and get accelerated to 1.4 GeV. The booster then injects the beam to the Proton Synchrotron (PS), which increases the energy to 25 GeV. The protons then enter the Super Proton Synchrotron (SPS), where they are accelerated to 450 GeV. Finally, protons enter the large ring of the LHC. The LHC has

[5]The acronym CERN was historically derived from *Conseil européen pour la recherche nucl´eaire*. Nowadays, 'CERN' has become a standalone name for the lab itself and is currently referred to as the *Organisation européenne pour la recherche nucléeaire*; or, in English: the 'European Organization for Nuclear Research.'

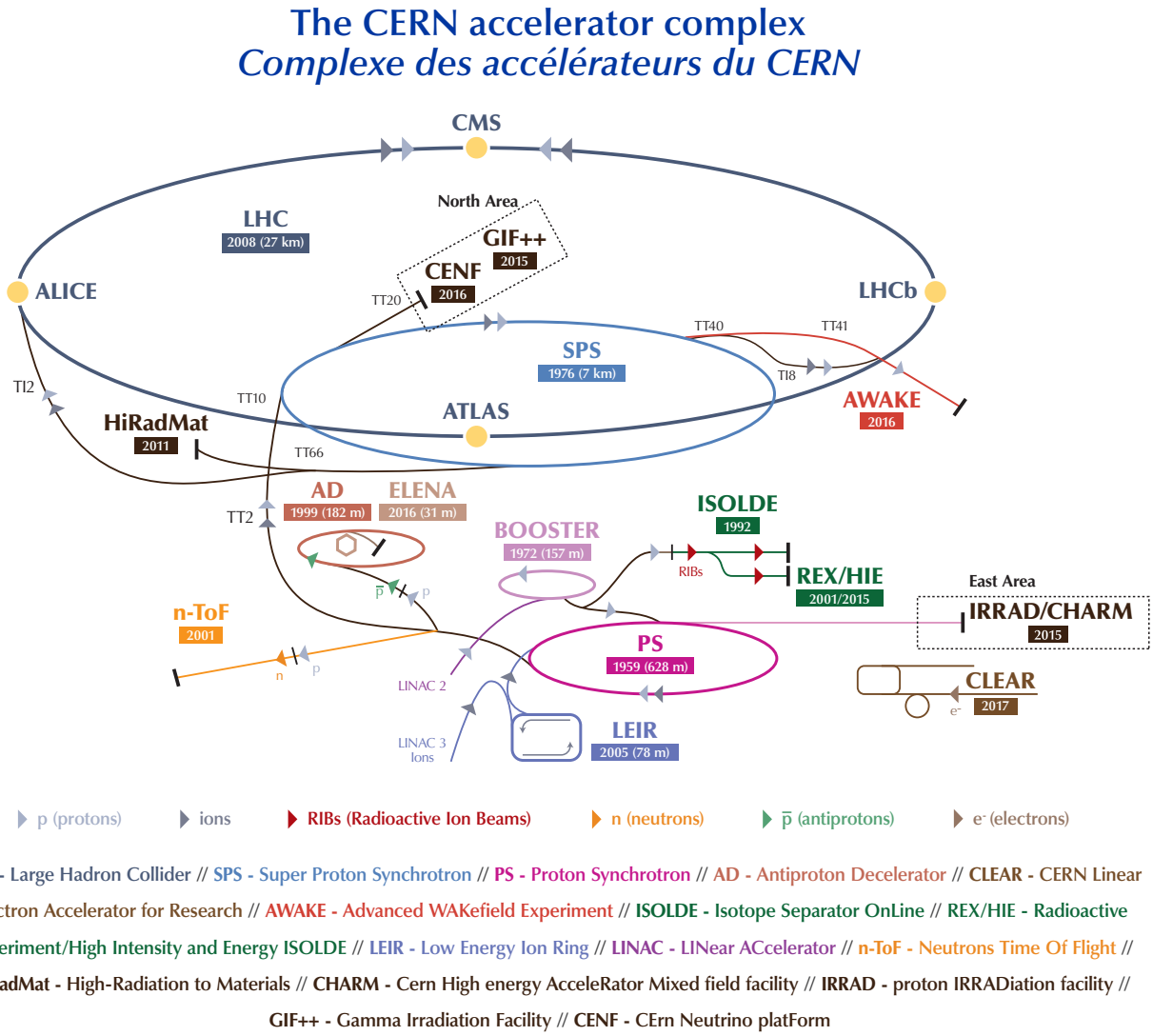


Figure 2.1: Illustration of the various beam lines and experimental facilities of the CERN accelerator complex. The protons that circulate through the LHC, and that are eventually made to collide inside the ATLAS detector, follow the path: LINAC2 → BOOSTER → Proton Synchrotron (PS) → Super Proton Synchrotron (SPS) → LHC. Figure taken from [121].

two vacuum pipes, where two beams circulate in opposite directions. It takes roughly 4 minutes 20 seconds to fill each LHC ring. Although the LHC is designed to collide proton beams with a center-of-mass energy of 14 TeV, due to technical limitations it was operated at $\sqrt{s} = 7$ TeV (2011) and $\sqrt{s} = 8$ TeV (2012) during the first run. During the second run (2015-2018), the center-of-mass energy was increased to 13 TeV. The beams of charged particles at LHC are accelerated using radio frequency (RF) cavities and guided around the ring using superconducting magnets. The LHC uses eight RF cavities per beam, each delivering 2 MV (an accelerating filed of 5 MV/m). The RF cavities increase the beam energy by more than 14 times. It takes roughly 20 minutes for the proton beams to reach their maximum energy of 6.5 TeV. The electric fields in the RF cavities are

made to oscillate at 400 MHz and are operated at a temperature of 4.5 K temperature.

2.1.2 Beam structure

The proton beams circulating around the LHC rings are not continuous proton streams; instead, they are made up of proton *branches*. It the LHC operates at its designed luminosity, each bunch contains $\sim 1.15 \times 10^{11}$ protons at the start of a nominal fill. Each proton beam contains 2808 bunches at its full intensity [128]. The bunches are spaced by 25 ns (or 7.5 m), and each bunch is roughly 7.7 cm long. All possible locations of bunches (called buckets) in a proton beam at LHC are not filled in order to provide the necessary gaps for several actions to take place like the beam dumps, the injection from the SPS to LHC, and the rise of the magnetic field of the kicker magnets. The LHC bunch structure is discussed in detail in reference [129].

2.1.3 Luminosity and pileup

The event rate of a process with production cross-section σ_p can be written as

$$\frac{dN_p}{dt} = \mathcal{L} \cdot \sigma_p \, . \tag{2.1}$$

The proportionality factor, $\mathcal{L}$ is known as *instantaneous luminosity*. If the shape of the LHC beams are approximated to be Gaussian, $\mathcal{L}$ can be given by

$$\mathcal{L} = \frac{N_1 N_2 f_{rev} N_b}{4\pi\sigma_x\sigma_y} S \tag{2.2}$$

where:

- N_b is the number of bunches in one beam,
- N_1 and N_2 are the number of protons in the colliding bunches (intensity of the colliding bunches),
- f_{rev} is the bunch revolution frequency,
- σ_x and σ_y are the transverse beam width assuming Gaussian beam shape,
- S is the luminosity reduction factor coming from the geometric effects and non-zero bunch crossing angle.

For LHC typical numbers are: $N_b = 2808, N_1 = N_2 = 1.15 \times 10^{11}, f_{rev} = 11.2455$ kHz. The beam sizes at the crossing point are 16.7μm. S can be estimated roughly 0.808. With these numbers the designed instantaneous luminosity of the LHC at the beginning of a fill becomes:

$$\mathcal{L} = 1.21 \times 10^{34}\ \mathrm{cm}^{-2}\mathrm{s}^{-1}\ \times\ 0.808 \approx 1 \times 10^{34}\ \mathrm{cm}^{-2}\mathrm{s}^{-1}.$$

The instantaneous luminosity can be integrated to get the total luminosity, $L = \int \mathcal{L} dt$. The integrated luminosity is expressed in *inverse femto-barn* (fb^{-1}) unit, where 1 b = $10^{-28}\mathrm{m}^2$. Thus, the total number of expected events of a process (p) with cross section σ_p can be estimated from the total integrated luminosity, $N_p^{\text{event}} = \sigma_p \cdot L$

At the LHC, many more than one head-on collision happens per bunch crossing. The majority of them are low energy inelastic collisions. These additional collisions other than the collision of interest are referred to as *pileup*. The in-time pileup comes from the additional collisions of the same bunch in which the hard scattering happens. Out-of-time pileup is from the remnant of collisions happening in the previous or later bunch crossings relative to the collision of interest. The detector readout time is larger than the spaces between the bunches; hence the out-of-time pileup events create additional responses in the detector making the measurements more difficult. Pileup is parameterized with the mean number of interactions per bunch crossing. The number of interactions in a bunch crossing, μ, is assumed to obey Poisson distribution [130, 131]. μ is proportional to the inelastic scattering cross-section, $\sigma_{\text{inelastic}}$ and can be written as

$$\mu = \frac{\mathcal{L}_{\text{bunch}}\sigma_{\text{inelastic}}}{f_{rev}} , \tag{2.3}$$

where $\mathcal{L}_{\text{bunch}}$ is the luminosity of a bunch.

2.1.4 LHC Run-2 performance

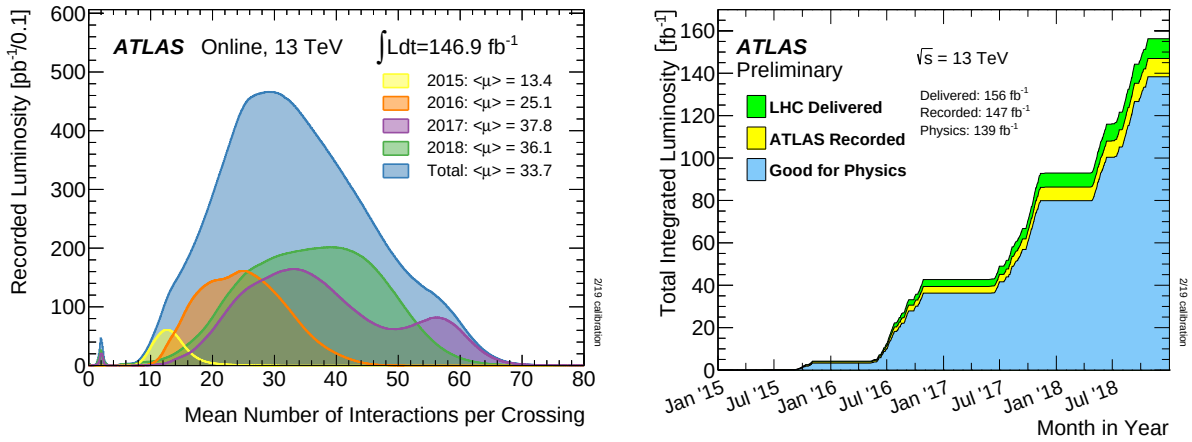
LHC data-taking period starting from the year of 2015 to the end of 2018 is referred to as Run-2. In this dissertation, LHC Run-2 data are used for the searches for the new particles. At the very beginning of 2015, bunches were spaced by 50 ns. But the bunch spacing changed to 25 ns soon after, and data with 25 ns bunch spacing are used for physics analyses presented in this dissertation. LHC increased the intensity of the proton bunches during Run-2 operation; hence, the instantaneous luminosity changed with time. LHC reached the designed instantaneous luminosity of $\sim 1 \times 10^{34}\text{cm}^{-2}\text{s}^{-1}$ for first time in 2016. Later in 2017, the luminosity was increased further and reached twice the design value. Higher instantaneous luminosity resulted in a higher number of interactions per bunch crossing. The average number of interactions per bunch crossing $< \mu >= 33.7$ during Run-2, compared to $< \mu > \approx 20$ during Run-1. The distributions of the mean number of interactions per bunch crossing, μ, are shown in Figure 2.2a for different data collection years. A total of 156 fb^{-1} of pp collision data was delivered by LHC, where ATLAS recording 147 fb^{-1} data of that 139 fb^{-1} are rated as 'good' based on the data quality requirements and used for physics analysis. The integrated luminosity collected in Run-2 is shown in Figure 2.2b as a function of time. Almost 95% of the collected data was used for physics analyses; which is a great achievement. The dilepton search results, discussed in chapter 6 of this dissertation, are produced using 36 fb^{-1} data collected in 2015 and 2016. In the top-antitop resonance searches, discussed in chapter 7 and chapter 8, the full Run-2 data of 139 fb^{-1} is used.

2.2 The ATLAS detector

ATLAS is a general-purpose particle detector [124] located at one of the LHC collision points. It investigates a wide range of physics, from precise measurements of SM quantities to searches for a variety of BSM physics. This section contains a brief overview of the ATLAS detector.

2.2.1 Coordinate system

ATLAS uses a right-handed coordinate system where the beam direction defines the z-axis. The $x-y$ plane is perpendicular to the beam axis. The positive x direction is defined by the axis pointing towards the centre of the LHC ring starting from the interaction point and the y points



(a) (b)

Figure 2.2: (a) The luminosity weighted distribution of mean number of interactions per bunch crossing (pp) for the 2015 to 2018 data taking periods. The average μ values for each year are shown. (b) The cumulative luminosity as a function of time is shown. The total delivered luminosity, ATLAS collected luminosity and luminosity used in the physics analysis are show in different colors. Figures taken from [132].

upwards. The A-side of the detector is located at the positive z direction, and the opposite side is called C-side. The azimuthal angle, ϕ, is defined around the beam axis in the $x-y$ plane, whereas the polar angle, θ, is measured from the beam axis. In ATLAS, generally the polar angle is not used directly. Instead a function of the polar angle called pseudorapidity, η, defined as $\eta = -\ln\tan(\theta/2)$is commonly used. Another similar variable is called rapidity and defined as

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

where E is the energy and p_z is the z-component of the momentum. The difference between the rapidities of two particles is invariant under the Lorentz boost along the z-axis. The same is true for pseudorapidity when the particle mass is zero. As a result, both variables are commonly used in collider physics. It is difficult to measure the total momentum of a highly relativistic particle, hence the measurement of rapidity becomes inaccurate. Thus η is a more commonly used variable in ATLAS. So the very high momentum limit or in the very low mass limit, rapidity is the same as pseudorapidity. For massive objects, rapidity is sometimes used in ATLAS. The transverse quantities like transverse momentum, $p_T = \sqrt{p_x^2 + p_y^2}$, transverse energy, $E_T = \sqrt{m^2 + p_T^2}$ and missing transverse energy, E_T^{miss} are defined in the transverse $(x-y)$ plane. The distance between two particles is generally defined as $\Delta R = \sqrt{\Delta\phi^2 + \Delta\eta^2}$. For massive objects often the distance is defined in the rapidity-azimuthal angle space like $\Delta R = \sqrt{\Delta\phi^2 + \Delta y^2}$.

2.2.2 Detector overview

The ATLAS detector has a cylindrical shape. It is 44 meters long, and the diameter is 25 meters. The total weight of the full detector is approximately 7000 metric tons. A cut-away view of the ATLAS detector is shown in Figure 2.3. It is symmetric with respect to the interaction point. It has

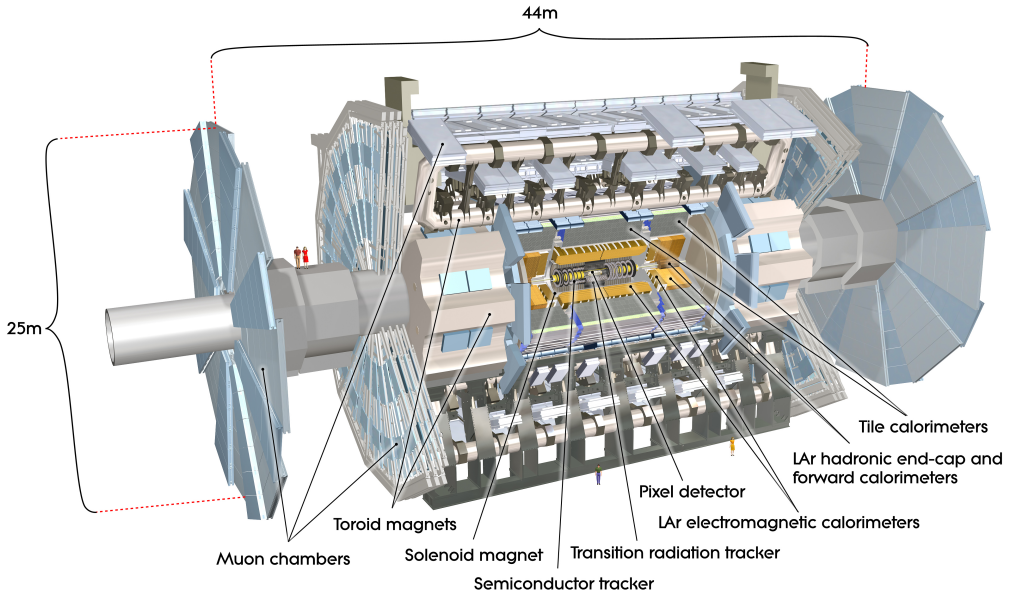


Figure 2.3: Cut-away view of the full ATLAS detector. Figure taken from [124].

an onion-like structure, where all the subdetectors are arranged in a concentric cylindrical shape in the central barrel region. In the forward and backward regions, the subdetectors are arranged in disk-like structures, known as the endcaps. There is a tracking detector in the core of the ATLAS detector. The tracking detectors are embedded in a strong solenoidal magnetic field. Outside the tracking detectors, there are calorimeters. Calorimeters are located outside the magnetic field and measure particle energy. Almost all the particles lose their energies in the calorimeters. But muons lose very little energy while passing through the calorimeters, so a set of dedicated muon detectors are installed in the outermost part of the ATLAS detector. The muon detectors are all tracking detectors and are embedded inside a toroidal magnetic field. The muon detectors are often collectively referred as muon spectrometer.

2.2.3 ATLAS magnets

Four superconducting magnet systems are used in ATLAS [133]. They are the central solenoid, the barrel toroid, and two endcap toroids. The central solenoid is aligned with the beam axis and provides a 2 T magnetic field. The solenoid magnet provides a magnetic field for the inner detector. Since the central solenoid is placed in front of the calorimeter system, the magnet system needed to be as thin as possible to reduce material interaction. Thus, the solenoid assembly contributes roughly 0.66 radiation length[6] at 90^0 incidence [134]. The single-layer coil is wound with a high-strength Aluminum-stabilized Niobium-Titanium conductor, specially developed to achieve a high field while keeping the thickness minimal.

The barrel and two endcap toroids provide roughly 0.5 T and 1 T magnetic field, respectively. The barrel toroid consists of eight coils and provides a magnetic field to the muon detector system. The conductor and coil-winding technology are the same in the barrel and endcap toroids; it is

[6]The radiation length is the average distance over which the energy of an electron is reduced by a factor of $1/e$ by bremsstrahlung.

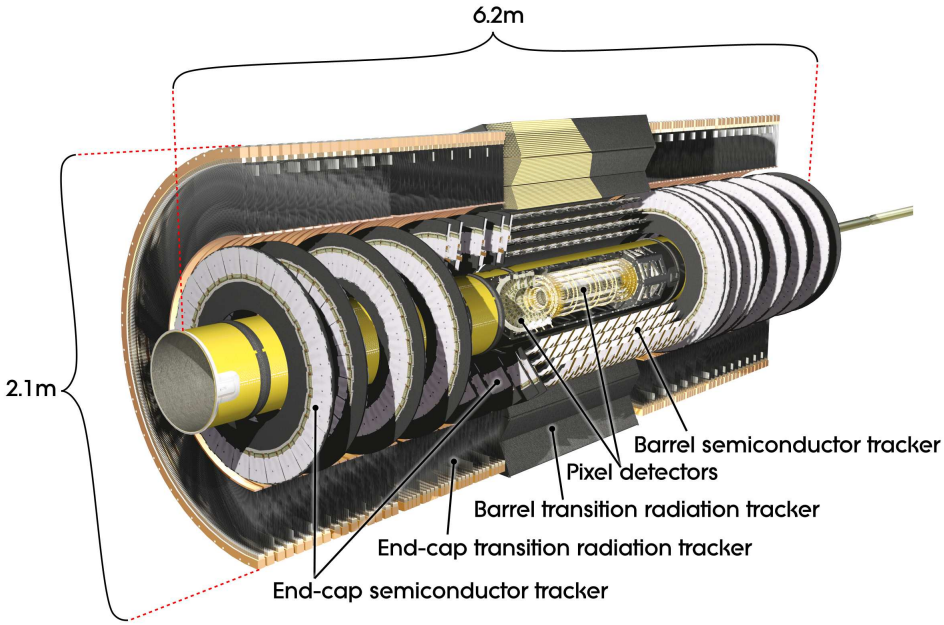


Figure 2.4: Cut-away view of the ATLAS inner detector. Figure taken from [124].

based on winding a pure Al-stabilized Nb/Ti/Cu conductor [124]. The two endcap toroids are also made of eight coils each.

2.2.4 Inner detector

The innermost part of the ATLAS detector is known as the inner detector (ID) [135]. The ID is a multi-technology tracking detector composed of three different sub-detectors: the Pixel Detector, the Semiconductor Tracker, and the Transition Radiation Tracker. It is embedded in a 2 T solenoid magnetic field, which bends the charged particles. The ID provides full ϕ coverage and extends up to $|\eta| = 2.5$. The detector is designed to provide a transverse momentum resolution of $\sigma_{p_T}/p_T = 0.05\% p_T$ GeV $\oplus 1\%$ [135] and a transverse impact parameter resolution of 10 μm in the central region for high momentum charged particles [124]. A cutaway view of the inner detector is shown in Figure 2.4. The different ID subdetectors are briefly discussed in the next section.

The Pixel detector and the Insertable B-Layer

The silicon pixel detector [136] is the innermost part of the ID. It consists of 1744 silicon pixel modules with 46,080 readout channels per module. There are three concentric cylindrical layers at radii of 50.5, 88.5, and 122.5 mm around the beam pipe, and three disks perpendicular to the beam pipe on each side, extending to 650 mm. The innermost pixel layer is called the *b-layer* and the remaining two layers are *layer-1* and *layer-2*. Typically pixels are 50 μm (400 μm) long in the transverse (longitudinal) direction. In between the readout chips the longitudinal pixel size is 600 μm due to design limitations. The sensors are 250 μm thick. There are almost 80 million pixels in the pixel detector.

During the long shutdown between LHC Run-1 and Run-2, the pixel detector was upgraded, and a new layer, the *Insertable B-layer* (IBL) [137, 138] was added. At present, IBL is the innermost layer and has a mean radius of 33 mm. A typical IBL pixel has a size of 50 μm (250 μm) in the

transverse (longitudinal) direction. Two different technologies are used in the IBL sensors. Planer silicon sensors are used in the central region, and 3D pixel sensors at high $|\eta|$. The thickness of the planner sensors is 200 μm, while the 3D sensors are 230 μm thick.

The Semiconductor Tracker

The pixel detector is surrounded by a silicon microstrip detector (SCT). SCT is arranged in four-barrel layers and two end caps with nine disks each. It consists of 4,088 two-sided modules with more than 6 million readout channels. SCT layers start at a radius of 299 mm and extend up to 514 mm [139]. All the SCT sensors in the barrel region have identical rectangular geometry, with 768 readout strips of 80 μm pitch (approximately 12 cm in length). Two such pairs are connected in a module and a second pair of identical sensors are glued back-to-back at a stereo angle of 40 mrad. The endcap sensors have a trapezoidal shape and the strip pitches vary from 56.9 μm to 94.2 μm. SCT has a hit resolution of ~ 17 μm in $r-\phi$ direction and 580 μm along z direction.

The Transition Radiation Tracker

The outermost part of the ID consists of the Transition Radiation Tracker (TRT) detector [140], which extends the track reconstruction further out to a radius of 1082 mm. It consists of polyimide drift (straw) tubes of 4 mm radius filled with a gas mixture of 70% Xe, 27% CO_2, and 3% O_2. Each tube has a 0.03 mm diameter gold-plated tungsten wire in the center, acting as an anode wire. TRT tubes are 144 (37) cm long in the barrel (endcap) regions. The tubes are arranged parallel (perpendicular) to z axis in the barrel (end-cap) region and covers up to $|\eta| = 2.0$. Charged particles passing through the tubes ionize the gas mixture. Ultra-relativistic charge particles produce transition radiation in the X-ray range while crossing the material boundary used in the TRT. The amount and pattern of transition radiation depend on the particle's mass, and lighter charged particles produce more radiation. So, electrons passing through the radiator will release a notably larger amount of transition radiation than heavier charged particles, such as pions. Therefore, the TRT plays an important role in electron identification and provides substantial discriminating power between electron and pions over the energy range between 1 and 200 GeV.

2.2.5 Calorimeter

The calorimeter system is outside the solenoid magnet and covers up to $|\eta| < 4.9$. Calorimeters are designed to measure the energy of particles either by completely stopping them or absorbing most of the energy of particles passing through. Typically a calorimeter consists of alternating high density "passive" absorber layers like lead and "active" medium such as solid lead-glass or liquid argon. If the absorber and the active material are the same, the calorimeter is a *homogeneous calorimeter*. Whereas in a *sampling calorimeter* the two materials are different. All the different calorimeter subdetectors used in ATLAS are sampling calorimeters. The energy resolution of a calorimeter can be written as

$$\frac{\sigma_E}{E} = \frac{a}{\sqrt{E}} \oplus \frac{b}{E} \oplus c \,, \tag{2.4}$$

Here a is the stochastic term and stands for the photoelectron statistics and accounts for the Poisson-like fluctuations. b is the noise term and is the energy equivalent of the electronic and other background noise like pileup noise. This term is dominant at the low energy. c is the constant term and it accounts for the non-uniformity of the detector like geometry, calibration. This term is dominant at higher energy. The three terms are added in quadrature indicated by the $\oplus$

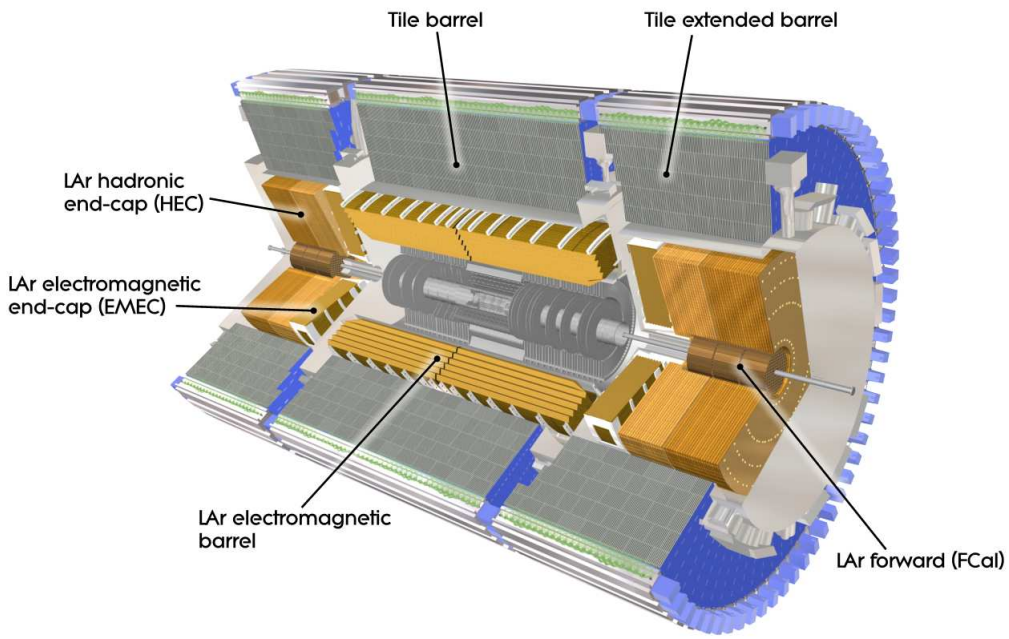


Figure 2.5: Cut-away view of the ATLAS calorimeter detector. Figure taken from [124].

symbol. Unlike the tracking detector, the relative uncertainty of calorimeter generally decreases with energy as the stochastic and noise term vanishes.

Two separate calorimeter systems are used in ATLAS, the electromagnetic calorimeter, and hadronic calorimeter. Figure 2.5 shows a cutaway view of the full ATLAs calorimeter systems.

Electromagnetic calorimeter

The electromagnetic calorimeter (EMCal) is designed to completely stop the electrons and photons. The ATLAS EMCal is a very high granularity liquid-argon (LAr) sampling calorimeter [141] in which lead is used as the absorber. LAr is a radiation hard material and offers a stable linear response. It is divided into one barrel (covering $|\eta| < 1.475$) and two endcaps (covering $1.375 < |\eta| < 3.2$). The barrel is further divided into 16 modules each covering $\Delta\phi = 22.5^0$. When an electron passes through lead, it emits bremsstrahlung photons on the other hand a high energy photon can pair produce e^-e^+. The cascade of bremsstrahlung and pair production creates an electromagnetic shower and ionizes the LAr. The electromagnetic interactions of high energy electrons and photons in matter are characterized by the radiation length, X_0. The radiation length is the average distance over which the energy of an electron is reduced by a factor of $1/e$ by bremsstrahlung. In the case of pair production, X_0 corresponds to 7/9 of the photon mean free path [142].

ATLAS EMCal depths are more than $22X_0$ in the barrel region and more than $24X_0$ in the endcap region. The depth increases with $|\eta|$. The are three longitudinal sampling layers for $|\eta| < 2.5$ region and two for $2.5|\eta| < 3.2$ region. The second layer has the highest granularity in $\Delta\eta \times \Delta\phi$ space, where the cell sizes are of 0.025×0.025. The fine segmentation helps to distinguish single photons from diphoton produced in neutral pion decay ($\pi^0 \to \gamma\gamma$).

Hadronic calorimeter

The hadronic calorimeter is designed to detect strongly interacting particles like hadrons. The hadron calorimeters function very similarly to electromagnetic calorimeter but the longitudinal shower development is determined by the nuclear interaction length, λ_I. The hadronic interaction length is the mean distance between hadronic interactions for relativistic hadrons. For most materials, the hadronic interaction length is significantly larger than the electromagnetic radiation length. The ATLAS hadronic (HCal) calorimeter has three components: the Tile calorimeter (TileCal) [143], LAr hadronic endcap calorimeter (HEC), and LAr forward calorimeter (FCal). The TileCal has one barrel (covering $|\eta| < 1.0$) and two extended barrels (covering $0.8 < |\eta| < 1.7$). There are three layers with different λ_I; at $\eta = 0$ the thickness is 9.7 λ_I. The first two layers have a granularity of 0.1×0.1 in $\eta \times \phi$ whereas the granularity is 0.2×0.1 in the last layer. The TileCal is a sampling calorimeter with scintillating plastic tiles as active material and steel as an absorber. The hadronic endcaps consist of two wheels on each side. The endcaps are LAr sampling calorimeters and use copper plates instead of lead as an absorber. The highest granularity is 0.1×0.1 in $\eta \times \phi$ in $|\eta| < 2.5$ and outside this region the cell size doubles. The target resolution of the electromagnetic calorimeter is [141]

$$\frac{\sigma_E}{E} = \frac{10\%}{\sqrt{E}} \oplus 0.7\% \,. \tag{2.5}$$

The pion energy resolution obtained from test beams using TileCal as well the LAr calorimeter [144] can be written as

$$\frac{\sigma_E}{E} = \frac{52\%}{\sqrt{E}} \oplus \frac{1.6\text{GeV}}{E} \oplus 3.0\% \,. \tag{2.6}$$

The FCal is designed to measure both electromagnetic and hadronic activities. It covers the very forward region, $3.1 < |\eta| < 4.9$, and consists of three layers. The first layer is a LAr calorimeter and uses copper as an absorber. It is optimized for electromagnetic measurement. The last two layers are LAr calorimeters using tungsten as an absorber and measure the energy of hadronic interaction. The depth of FCal is roughly $10\lambda_I$.

2.2.6 Muon spectrometer

The muon spectrometer (MS) is designed to detect muons with a pseudorapidity coverage up to $|\eta| \leq 2.7$. It measures the momentum of muons based on the deflection of charged particle tracks in the large toroid magnetic field. The toroid magnet provides an average field of 0.5 T, pointing in the ϕ direction. Thus the toroidal field is perpendicular to the muon tracks and deflects the muon tracks in the η direction. A cutaway view of the full muon system is shown in Figure 2.6. Four different types of detectors are used in the MS to do two types of tasks: precision tracking and triggering. The barrel muon chambers are installed between eight coils of the superconducting toroid magnets. In the two endcap regions, the chambers are in planes perpendicular to the beam axis. The MS chambers are arranged in three stations in the barrel and endcap, and there is one extra station in the transition region between barrel and endcap as shown in Figure 2.7. MS provide momentum measurements with a relative resolution below 3% over a wide muon p_T range increasing up to 10% at $p_T \approx 1$ TeV [145].

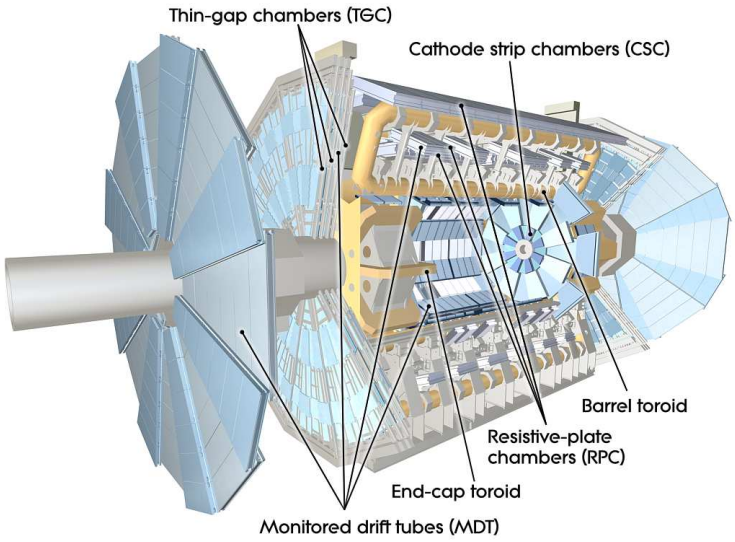


Figure 2.6: Cut-away view of the ATLAS muon spectrometer. Figure taken from [124].

Monitored Drift Tubes

The monitored drift tube (MDT) chambers are built from pressurized drift tubes with a diameter of 29.97 mm filled with a gas mixture of Argon and Carbon (93% and 7%). The electrons created from the ionization are collected in the central anode wire (tungsten-rhenium). MDT covers up to $|\eta| < 2.7$. MDT modules are robust and a single tube failure has only a negligible consequences on the module's overall performance. The average spatial resolution per tube is around 80 μm.

Cathode-strip chambers

The cathode-strip chambers (CSC) are installed in the forward region, and cover $1.0 < |\eta| < 2.7$. In the forward region, the MDTs are not capable of handling the high particle rate, so CSCs are used. The CSC is made up of multiwire proportional chambers where the anode wires are oriented in the radial direction. ATLAS CSCs are filled with a gas mixture of Argon and Carbon dioxide. The cathode plates are segmented into strips and arranged in the parallel and the perpendicular directions relative to the wires. CSC has higher radiation tolerance and can safely operate at 1000 Hz/cm^2.

Thin Gap Chambers

Thin Gap Cambers (TGC) are used in the endcaps covering $1.05 < |\eta| < 2.7$. TGCs are multi-wire proportional chambers like CSCs, and have very good timing resolution. Besides complementing the MDT measurements in the endcap region, the TGCs are used to trigger muons. TGCs are used for triggering events in the $|\eta| < 2.4$ region.

Resistive Plate Chambers

Resistive Plate Chambers (RPC) are gaseous parallel plate detectors. Two resistive plates are held at a high potential difference, and the gap between them is filled with a gas mixture. RPCs are

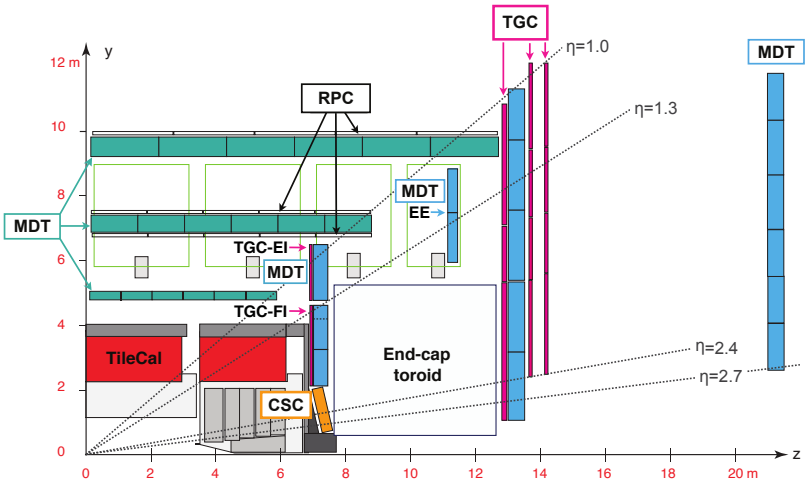


Figure 2.7: A schematic picture showing a quarter-section of the ATLAS muon system. Figure taken from [146].

using in the barrel region (covering $|\eta| < 1.05$) and their timing resolution is very good, about 2 ns. Thus RPCs are used for triggering along with providing η, ϕ position measurement.

2.2.7 Trigger and data acquisition

The ATLAS trigger and data acquisition (TDAQ) system is developed to select only a tiny fraction of the 40 million collision events happening every second. Due to the limited bandwidth and storage capability, it is impossible to record all the collision events. The trigger system is designed to identify interesting events out of all the collision events and store them. In Run-2, the ATLAS TDAQ system had two components, hardware-based low-level trigger or level-1 (L1) trigger and software-based high-level trigger (HLT). With the help of the two-step triggering method, the recorded event rate decreases to around 1 kHz [147]. A schematic diagram of the ATLAS TDAQ system used in Run-2 is shown in Figure 2.8.

Level-1 trigger

The level-1 trigger is a hardware-based trigger and uses the calorimeter and muon detectors. For each event, one or more regions of interest (ROI) are defined in the η space to identify interesting events. It selects events with high energy electrons, muons, photons, jets, and τ-leptons, large missing energy, and high transverse energy. L1 calorimeter trigger uses reduced granularity information from both electromagnetic and hadronic calorimeters. On the other hand, the L1 muon trigger system is used to identify high momentum muons in the barrel and two endcap regions using RPCs and TGCs respectively. A new topological trigger processor (L1Topo) is introduced at the beginning of the LHC Run-II. The topological trigger uses L1 trigger objects as inputs and applies topological selection combining the kinematic and angular information of > 1 object. Using L1Topo, we can compute quantities like angular separation, energy sum, invariant mass, and thus largely improve our background rejection ability. The L1 trigger reduces the event rate to roughly 100 kHz. The L1 decisions are made by the Central Trigger Processor (CTP), which gets information from L1Calo, L1Muon, and the L1Topo system. The system is optimized to provide a fast decision. The acceptance decision is made within 2.5 μs.

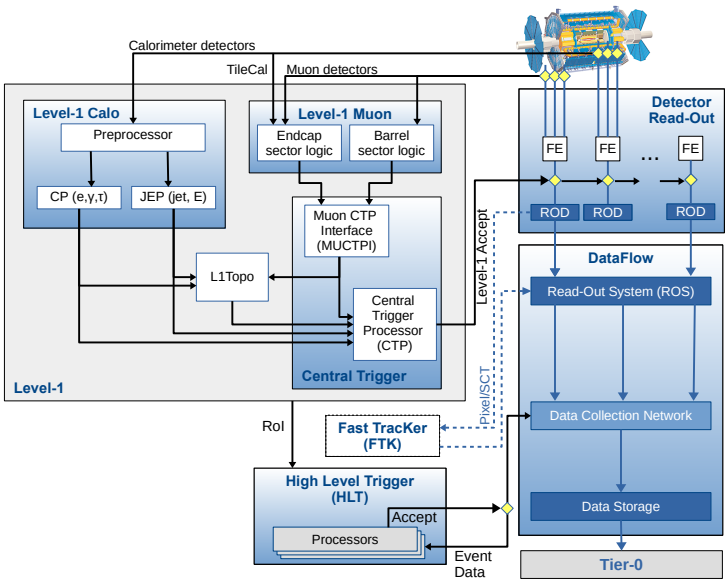


Figure 2.8: A schematic picture ATLAS trigger and data acquisition system using in Run-2. Figure taken from [147].

The L1 trigger has a read-out latency and needs some time to read out a triggered event. If a trigger gets fired while the previous event is still read out it needs to be stored in the memory (generally called read-out buffer). To cope up with the memory limit, a mechanism is set up to ignore the bunch crossings to limit the number of L1 accepts within the constraints on the detector read-out latency [148]. This is called *dead time*.

High Level Trigger

Events passing the L1 trigger go through the software-based high-level trigger (HLT). There are fast trigger algorithms to reject events in the earlier stage, followed by more involved and computationally heavy full reconstruction algorithms. The algorithms used in HLT are not always the same as the offline reconstruction used later on. One example is track reconstruction; track reconstruction involves several steps and tracking used in the HLT is a lighter and faster version of the offline algorithm. The HLT software is very similar to ATLAS offline reconstruction software Athena [149], which is based on the high energy physics data procession software, Gaudi [150]. HLT stores events with an average rate of 1 kHz.

2.2.8 Data quality

The data collection only starts after the LHC declares the collision of stable particles. Once all the subdetector high voltages are ramped up and ATLAS declares itself ready to collect data for physics, the data recording starts. During data collection, sometimes subdetector modules or components can stop working, for example due to a high voltage trip. The collected data during this transient detector stage is marked, such that they can be removed if needed during data analysis [151]. To ensure this, the data collected by each subdetector system are monitored by automatic procedures as well as manually during the data taking. This is the online data quality monitor.

There is also a more thorough off-line data quality monitor procedure, where collected data are properly checked, after full reconstruction, before storing on the disk. Finally, based on all the data quality checks, a good run list (GRL) of the collected data is defined to identify which datasets are good to use for physics analyses.

2.2.9 ATLAS detector simulation

A detailed simulation of the ATLAS detector is necessary in order to study the expected detector response in a wide range of physics processes. Simulation of the full ATLAS detector is done using the GEANT4 simulation toolkit [152, 153] as a part of the ATLAS simulation framework [154] which is integrated into the ATLAS Athena framework. Once the physics processes are simulated, all the stable particles are propagated through the full ATLAS detector simulated by GEANT4. The pile-up effects are simulated by overlaying the energy deposits (hits) on the detector. Trigger effects are also simulated and the events are reconstructed using the same reconstruction method as used for data. Simulating the full detector and pile-up effects take huge computing resources. In ATLAS, it is seen that almost 90% of simulation time is spent to simulate the calorimeter responses. this is mostly due to the complex calorimeter showers developed by the primary particles. To decrease the simulation time and computing overhead, a fast calorimeter simulation method is developed to parametrize the calorimeter responses. This simulation method is called ATLAS fast simulation or AFII [155, 156]. AFII works reasonably well, but cannot simulate the quantities relevant for jet substructure which will be needed for two of the searches presented in this thesis. Nonetheless, ATLAS fast simulation method is very useful when these kinetic variables are not used and saves lots of computing time.

CHAPTER

THREE

Pixel Cluster splitting with Mixture Density Network

3.1 Introduction

The ATLAS inner detector consists of a silicon pixel detector in the core surrounded by a silicon strip detector and a transition radiation tracker in the outer region, summarized in subsection 2.2.4. The pixel detector has four layers in the central region and two sets of three disks in the forward and backward regions [136, 138]. Charged particles ionize the silicon bulk of a pixel sensor, then the created charges drift towards the electrodes across the sensor due to the voltage difference. Typically charge from an ionizing particle gets deposited in multiple pixels as shown in Figure 3.1 (left) due to charge drift in presence of a B-field, electron-hole diffusion and δ-rays. The cluster formed by the activated pixels is called a hit. In a dense environment, where the average

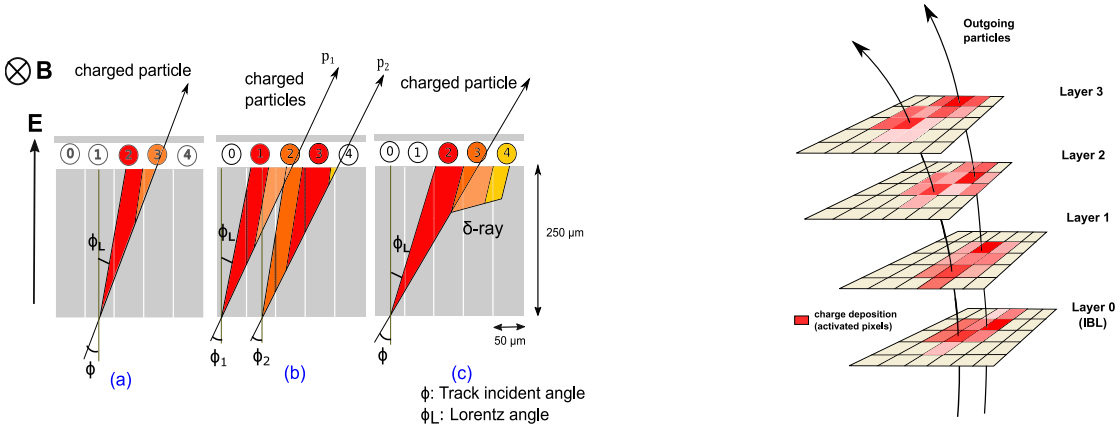


Figure 3.1: Left: (a) In presence of ATLAS solenoid magnet the charges get diffused and drifted to nearby pixels. (b) Charge clusters of two very close charge particles get merged. (c) δ rays extend the charge cluster size [157]. Right: a schematic diagram where the clusters from two nearby tracks get merged.

separation between particles becomes comparable to the detector granularity, their pixel clusters

can get merged as shown in Figure 3.1 (right). As a result, multiple tracks can get associated with one hit. It is crucial to understand the multiplicity of a hit i.e. how many charged particles went through that same area. In the past, a connected component analysis (CCA) [158] was used to group the neighboring pixels, and a charge interpolation technique was used to estimate the particle intersections. In this approach, no attempts were made to split the merged clusters. In 2011, a neural network (NN)-based approach was introduced in ATLAS [157] to identify clusters created from multiple charged particles and estimate the hit positions. With the use of these NNs, the number of shared measurements between tracks reduced almost by a factor of three, and the position estimate of the clusters improved significantly [157].

3.2 Pixel cluster-splitting neural networks: current approach

There are three sets of neural networks used in the current algorithm. There is a 3-class classifier, *number network*, to determine the particle multiplicity of a given pixel cluster. The outputs of the network can be $1, 2$ and ≥ 3. There are three regression networks, *position networks*, each corresponding to the outputs of the number network and whose goal is to determine the best estimate of the location where each charged particle crossed that pixel detector; this is called hit position. The hit position is estimated in the local x and local y coordinates[7]. There are two *error networks* for each position network to determine the associated uncertainties on the hit position. The error networks are classifiers and there are a total of six of them. In total there are thus ten different networks.

3.2.1 Network architectures

All the networks mentioned above are quite shallow with only two hidden layers. The layers are not very dense either, with ≤ 60 nodes in any layer. Hence, it is very fast to evaluate them. As these networks are called very frequently during track reconstruction, making the network evaluations faster is very important. There are a total of 60 input variables given to these networks which include the following:

- A 7×7 digitized charge matrix, obtained from the calibration of time-over-threshold values measured by the pixel sensors, centered on the charge centroid. The charge matrix has been flattened into a vector of 49 elements in row-major order;
- A length-7 vector of pixel pitches in the local y direction, in which the pixel pitch (size) is not constant;
- A binary variable encoding the inner detector region (endcap or barrel);
- An integer variable representing the cylinder (barrel) or disc (endcap) number;
- The incidence angles (θ, ϕ) of the track candidate(s).

More details about the network parameters and their performances can be found in the ATLAS note [159].

[7] The positions are measured in a frame of reference local to the pixel sensor considered, in which the local x and y directions correspond to the transverse and longitudinal directions with respect to the beamline, respectively.

3.2.2 Network training: least-squares formalism

Neural networks model a mapping between a set of training input variables, $\boldsymbol{x} = \{x_1, \ldots x_d\}$, to a set of output variables, $\mathbf{t} = \{t_1, \ldots t_c\}$. Usually the networks are trained on a finite set of training examples, $\{\boldsymbol{x}^q, \mathbf{t}^q\}$, where q labels a particular training example and runs from 1 to n. The most common method of neural network training involves the minimization of the sum-of-square error, defined over the training data set, of the form

$$E(\mathbf{w}) = \frac{1}{2} \sum_{q=1}^{n} \sum_{k=1}^{c} \left[f_k(\boldsymbol{x}^q; \mathbf{w}) - t_k^q \right]^2 , \tag{3.1}$$

where t_k denotes the components of the target vector and $f_k(\boldsymbol{x}; \mathbf{w})$ denotes the corresponding outputs of the network mapping function which is a function of network parameters, $\mathbf{w}$, called *weights* and *biases*. The error function is minimized w.r.t the weights to get the best prediction, $f_k(\boldsymbol{x}; \mathbf{w}^*)$; where $\mathbf{w}^*$ is the set of weight values which minimizes the error function. It turns out that $f_k(\boldsymbol{x}; \mathbf{w}^*)$ is the conditional average [8] of the target data, conditioned on the input vector. This method is used in the current algorithm.

3.3 Mixture Density Networks

The simple neural network-based algorithm performs well but there is scope for further improvement. Currently the hit positions and the associated uncertainties are estimated in two steps using two different types of network. One can try to reduce the network steps by estimating the position and uncertainties simultaneously. Furthermore, the performance of the error networks does not always match expectation [159]. This can be alleviated by using a more powerful algorithm. So an alternative algorithm, *mixture density networks (MDN)* [160], based on probabilistic learning is introduced which can estimate position and uncertainty together.

3.3.1 MDN algorithm

The main goal of network training is to learn the underlying data generator. The data generator can be quantified as the input-target joint probability distribution, $p(\boldsymbol{x}, \mathbf{t})$. The joint distribution can be written in terms of the conditional distribution as $p(\boldsymbol{x}, \mathbf{t}) = p(\mathbf{t}|\boldsymbol{x}) \cdot p(\boldsymbol{x})$. Since $p(\boldsymbol{x})$ is a property of the network inputs, the conditional probability is the quantity that gets modeled through the network training. As described in Equation 3.2.2 the conventional method is to use the least-squares formalism but if we assume that the conditional distribution of the target data is a Gaussian, then the least-squares formalism can be obtained using the maximum likelihood [160]. So one can start with the following conditional probability

$$p(\mathbf{t}|\boldsymbol{x}) = \prod_{k=1}^{c} p(t_k|\boldsymbol{x}) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{1}{2\sigma^2} \sum_{k=1}^{c} \{f_k(\boldsymbol{x}; \mathbf{w}) - t_k\}^2 \right]. \tag{3.2}$$

Here σ is the global variance and the target variable (t_k) mean is modeled with a very flexible network function, $f_k(\boldsymbol{x}; \mathbf{w})$. The values of the parameters $\mathbf{w}$ are determined from the finite set of training examples $\{\boldsymbol{x}^q, \mathbf{t}^q\}$ by maximizing the likelihood

$$\mathcal{L} = \prod_{q=1}^{n} p(\mathbf{t}^q, \boldsymbol{x}^q) = \prod_{q=1}^{n} p(\mathbf{t}^q|\boldsymbol{x}^q) p(\boldsymbol{x}^q). \tag{3.3}$$

[8] Conditional average of a quantity $Q(\mathbf{t})$ is defined as: $\langle Q|\boldsymbol{x}\rangle \equiv \int Q(\mathbf{t}) p(\mathbf{t}|\boldsymbol{x}) d\mathbf{t}$.

The only **w**-dependent term in $-\log(\mathcal{L})$ is the least-squares term. Thus minimizing it will give the conditional average as an estimate of the network prediction. It is seen that the conditional averages predicted by the networks are optimal for classification problems [160]. However in regression problems, the goal is to predict continuous variables and the conditional average represents a very limited statistic. Hence there can be significant benefit of using a more complete description of the target data distribution. A Mixture Density Network is one such method. In this algorithm, the Gaussian approximation is relaxed to a *mixture model*. The mixture model can be written as a linear combination of different kernel functions ($\phi_i(\mathbf{t}|\boldsymbol{x})$) as

$$p(\mathbf{t}^q|\boldsymbol{x}^q) = \sum_{i=1}^{m} \alpha_i(\boldsymbol{x})\phi_i(\mathbf{t}|\boldsymbol{x}) \,, \tag{3.4}$$

where m is the number of mixture components, and the parameters $\alpha_i(\boldsymbol{x})$ are called *mixing coefficients*. Only Gaussian kernels of the form

$$\phi_i(\mathbf{t}|\boldsymbol{x}) = \frac{1}{\sqrt{2\pi\sigma_i(\boldsymbol{x})^2}} \exp\left[-\frac{\|\mathbf{t}-\mu_i(\boldsymbol{x})\|^2}{2\sigma_i(\boldsymbol{x})^2}\right]. \tag{3.5}$$

are considered in this study since any density function can be approximated with this linear combination [161]. This is called a Gaussian mixture model (GMM). The parameters of the mixture model should be continuous functions of $\boldsymbol{x}$ and can be modeled using the outputs of a feed-forward network. The combination of the feed-forward network and the mixture model is called mixture density network [160]. Since the mixing coefficients α_i are probabilities they should add up to unity. To impose this condition α_i is defined as a *softmax* function [162, 163] of the network outputs (z_i) as

$$\alpha_i = \frac{\exp(z_i^\alpha)}{\sum_{j=1}^{M}\exp(z_j^\alpha)}. \tag{3.6}$$

The variances σ_i^2 of the kernel functions should be non-negative. So, they are defined as the absolute value of the network outputs as

$$\sigma_{ik}^2 = |z_{ik}^\sigma|. \tag{3.7}$$

The means, μ_i, represent location parameters and should be represented directly by the network outputs as

$$\mu_{ik} = z_{ik}^\mu \tag{3.8}$$

As before, the loss function is defined to be the $-\log(\mathcal{L})$, hence can be written as

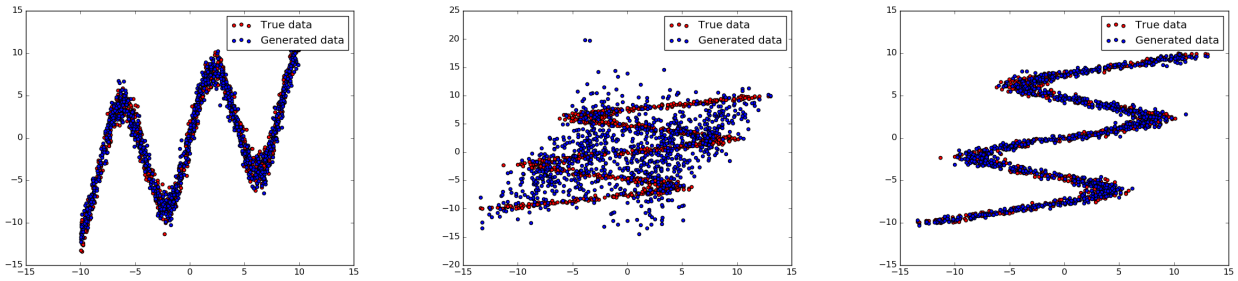
$$E = -\sum_q \ln\left[\sum_{i=1}^{m} \alpha_i(\boldsymbol{x}^q)\phi_i(\mathbf{t}^q|\boldsymbol{x}^q)\right]. \tag{3.9}$$

Finally, the loss gets minimized with respect to the network weights and biases during the training process. More detailed discussion on this method can be found in [160] and [164].

3.3.2 Toy Example: Inverse Problem

Before addressing the actual problem at hand, the code is tested using a toy example. One of the popular problems known to be solved by the MDN was chosen. Examples are generated from a sinusoidal function of the form

$$y = 7.0\sin(0.75x) + 0.5x + \epsilon \,, \tag{3.10}$$



(a) Fitting sinusoidal data

(b) Conventional method can not model inverse data

(c) MDN can model inverse data

Figure 3.2: Performance of the MDN in modeling an inverse problem.

where ϵ is a Gaussian distributed random number. First a conventional neural network is trained such that it learns the mapping from x to y. In this case, for each input variable there can only be a single output variable. The network with sum-of-square error minimization as well as with likelihood maximization under one Gaussian kernel approximation gives an excellent performance in modeling the underlying generator of the data. This is shown in Figure 3.2a. Now let's consider the corresponding inverse problem using the exact same data but this time try to find a mapping from y to x. So, essentially we want the network to learn the following relation between x and y

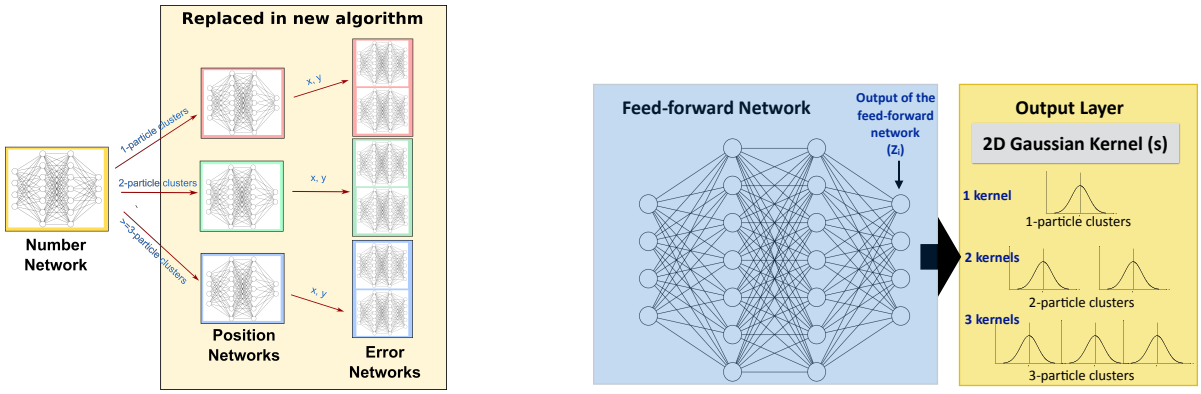
$$x = 7.0\sin(0.75y) + 0.5y + \epsilon. \tag{3.11}$$

In this case there are many possible output values for a single input value and neither the least-square method nor the one Gaussian kernel method can model the underlying generator, as shown in Figure 3.2b. But a MDN with multiple Gaussian kernels successfully models the underlying generator. The performance of a MDN with 10 one-dimensional Gaussian kernels is shown in Figure 3.2c; it is able to accurately reproduce the underlying relationship.

3.3.3 Building MDN

The network is implemented using the Keras [165] software package with Theano [166] as a backend. The feed-forward network is built using the Keras functional API with three dense hidden layers. The layers are activated using Rectifier-Liner (ReLU) activation function. The output layer contains one or more GMM unit. There are three different networks for different cases, one for each of solutions to the number network. The hit position estimates and the associated uncertainties can be estimated from the network kernel parameters. For simplicity, only one Gaussian kernel with dimension two is considered in this GMM. The kernels are parametrized with mean and precision[9]. Two dimensional kernels can predict the mean and uncertainty along both local x and y simultaneously. A schematic diagram of the workflow and the MDN structure is shown in Figure 3.3. A customized loss function is developed for this study to calculate the negative log-likelihood (Equation 3.9). The loss function uses the network outputs as described in Equation 3.6, Equation 3.7 and Equation 3.8 to estimate the model parameters while calculating the loss. The optimization is done using Adam optimizer [167]. Details of the network parameters are listed in Table 3.1.

[9]The precision is defined as inverse of variance.



(a) Pixel NN work flow

(b) MDN architecture

Figure 3.3: Schematic diagrams of (a) the pixel neural network work flow and (b) mixture density network.

Table 3.1: The network structure and the hyperparameters used to train the MDNs. The input and output layer size is denoted in parenthesis in the structure row of the table.

Hyperparamaters	MDN (1 particle)	MDN (2 particles)	MDN (3 particles)
Structure	(60)-100-50-50-(1-2-2)	(60)-100-80-50-(2-4-4)	(60)-100-80-50-(3-6-6)
Activation	ReLU	ReLU	ReLU
Output Layer	1 GMM	2 GMM	3 GMM
Output activation	(softmax-linear-absolute)	(softmax-linear-absolute)	(softmax-linear-absolute)
Learning rate	0.0001	0.0001	0.0001
L2 regularizer	0.0001	0.0001	0.0001
Batch Size	100	100	100
Gradient Clipping	clipnorm = 1	clipnorm = 1	clipnorm = 1
Loss Function	custom	custom	custom

3.3.4 MDN training

The training dataset is the same as described in subsection 3.2.1. The training, testing, and validation dataset were created using Monte Carlo simulation of dijet process where the truth-level jets have transverse momentum between 1.8 and 2.5 TeV. A total of 12 million training examples were generated, and they were split 9 : 1 into training and validation set. An independent set of 5 million examples is used to evaluate the performance of these networks. The training is done in small *mini-batches*, and the model parameters are updated after each mini-batch. The size of the mini-batch is a hyperparameter and is called batch size. A batch size of 100 is used for MDN training. MDN is very sensitive to several hyperparameters, and a number of technical issues were observed during the studies. Some of the issues are summarized in subsection 3.3.5 and subsection 3.3.6.

3.3.5 Overflow and Underflow problem

Two common issues of the MDN are the underflow and over flow problem. The underflow problem is referred to the situation when a number of interest is too small for the computer to represent it. Whereas the overflow problem occurs in presence of large numbers. These problems can lead

to inaccuracy and sometimes can lead to NaNs[10] in training or evaluation step. The underflow and overflow problem is very common when the log of sums of exponentials are involved. Let's consider the following equation

$$z = \ln\left(\sum_{n=1}^{N} \exp\{x_n\}\right), \tag{3.12}$$

where a number z is expressed as a logarithm of sum of exponentials of another numbers x_n. For large values of x_n the exponential, $\exp\{x_n\}$, becomes very large and gets interpreted as infinity by the computer. For example at $x_n = 1000$, $\exp(1000) = \infty$ due to this overflow problem. On the other hand for very small x_n the exponential becomes zero. When $x_n = -1000, \exp(-1000) = 0$ because of the underflow problem. The logarithm of these values thus becomes $+\infty$ or $-\infty$. This underflow and overflow problem can arise in our case since the log likelihood includes the summation of several exponential functions.

$$\begin{aligned} -\ln \mathcal{L}(\mathbf{t}|\boldsymbol{x}) &= -\ln\left\{\sum_{i=1}^{m} \alpha_i(\boldsymbol{x}^q)\phi_i(\mathbf{t}^q|\boldsymbol{x}^q)\right\} \\ &= -\ln\left\{\sum_{i=1}^{c} \frac{1}{(2\pi)^{d/2}\sigma_i(\boldsymbol{x})^d} \exp\left(-\frac{\|\mathbf{t} - \mu_i(\boldsymbol{x})\|^2}{2\sigma_i(\boldsymbol{x})^2}\right)\right\} \end{aligned} \tag{3.13}$$

A common trick to resolve the numerical instability is called *log-sum-exp trick*, which could be given by the following identity

$$\ln \sum_{n=1}^{N} \exp\{x_n\} = a + \ln \sum_{n=1}^{N} \exp\{x_n - a\}, \tag{3.14}$$

for any a. Typically a is taken as $\max\{x_n\}$ to make the exponent as small as possible. The *log-sum-exp trick* is applied in this study while implementing the loss function of the MDN.

3.3.6 The persistent NaN problem

Even after using the *log-sum-exp trick* the loss function sometimes became NaN due to several situations like:

1. $\sigma(\boldsymbol{x})$ is very small.
2. Big difference is seen between the predicted values($\mu(\boldsymbol{x})$ and the true values($\mathbf{t}$).

The second case is more common for this study. Sometimes the true values are outside the expected range of $[-3.5, 3.5]$ in the pixel index unit [11]. If the predicted position values are very close to zero and the target true position values are comparatively large, then the exponential becomes very small and the logarithm of that returns negative infinity. To resolve this issue three things could be done:

1. Pre-process the training set by throwing away the training examples where the true values are unexpectedly large.

[10] NaN means Not a Number

[11] One pixel is one unit and the centre of the cluster is (0,0).

2. Regulate the network prediction where it is too far off from the true values. For example the exponent could be constrained to be in $[-5, 5]$, $|\mathbf{t} - \mu_i(\boldsymbol{x})| <= 5.0$. And if the $|\mathbf{t} - \mu_i(\boldsymbol{x})| > 5.0$, it could be set to zero such that this particular example does not have any effect on the training process.

3. The calculated likelihood before taking the logarithm could also be regulated to keep the values in $[-10, 10]$. If the values are outside this range those could be set to zero.

The above mentioned tricks are not applied in the current network. Some other tricks which help stabilize the network training are also studied. Some of the tricks are used in the implemented MDN.

1. **Gradient Clipping:** Some times the gradient of the loss function becomes large and that eventually saturates the training loss and the network does not learn anything. This are resolved by the using Gradient clipping [168] technique to clip the gradient to a smaller range. The gradient clipping makes the training process stable across mini batches. Gradient clipping is applied in our network.

2. **Weight regularization:** Weight regularization [169, 170] is applied to make the training stable. Weight regularization is used in the MDN.

3. **Batch normalization:** On application of the batch normalization technique [171] on the hidden and output layers the network training becomes stable. The batch normalization acts like weight regularization, so in presence of batch normalization further weight regularization were not applied. In the final MDN used for the subsequent studies, batch normalization is not used.

3.4 Results

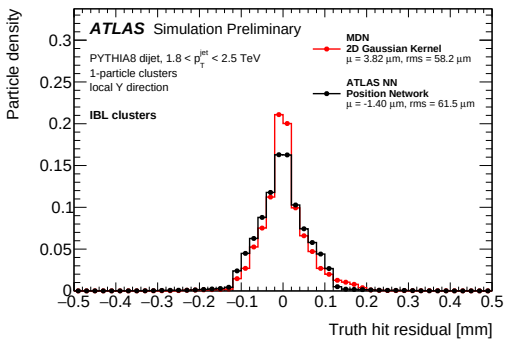
The networks were trained up to the point where the loss gets saturated. The trained models were evaluated on the testing set to get the predicted positions and uncertainties. Since MDNs do two different predictions, two metrics are defined to examine their performance. The metric for position estimation is the residual and defined as the difference between predicted and true position,

$$\text{residual} = x_{\text{pred}} - x_{\text{true}} \quad \text{or} \quad y_{\text{pred}} - y_{\text{true}} \,. \tag{3.15}$$

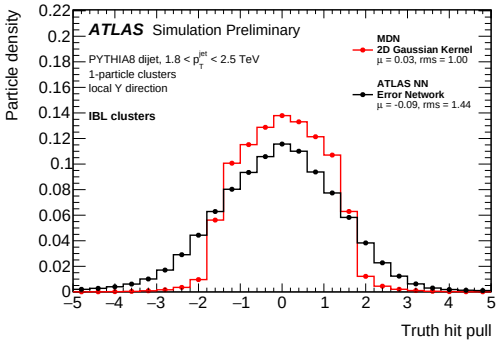
The other metric is the pull and is defined as the ratio of residual over predicted uncertainty:

$$\text{pull} = \frac{x_{\text{pred}} - x_{\text{true}}}{\sigma_{x,\text{pred}}} \quad \text{or} \quad \frac{y_{\text{pred}} - y_{\text{true}}}{\sigma_{y,\text{pred}}}$$

For all the MDNs these two quantities are evaluated and their performance is compared with the current networks. If the network predictions are very close to the true values the residuals are expected to follow narrow distributions centered around zero. In addition, if the networks predict uncertainties consistent with the predicted positions, then the pulls should follow a standard normal distribution. The performances are evaluated into different detector regions, barrel (central) and endcap (forward, backward). Since the first barrel layer (IBL) has different pixel pitches, IBL performance plots are made separately. The IBL-only performance plots in local y direction for 1-particle networks are shown in Figure 3.4. The residuals are narrower and the widths of the pull distributions are closer to unit width for the new MDN. Similar performance plots for 2-particle

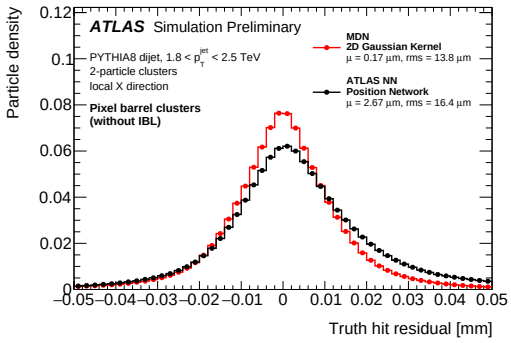


(a) Local y residuals

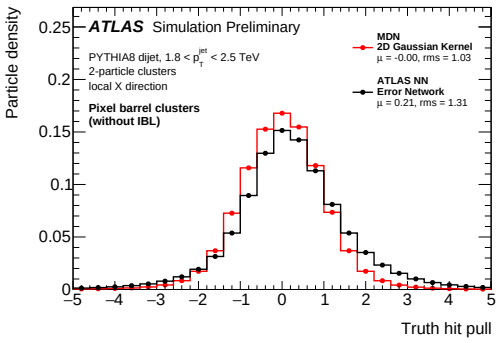


(b) Local y pulls

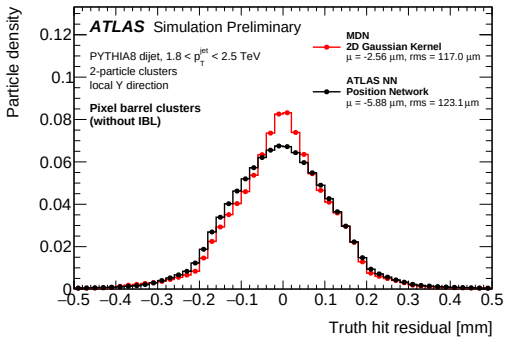
Figure 3.4: Comparison of residuals (a) and pulls (b) in the local y direction obtained from the ATLAS NN (black) and MDN (red) using 1-particle IBL clusters.



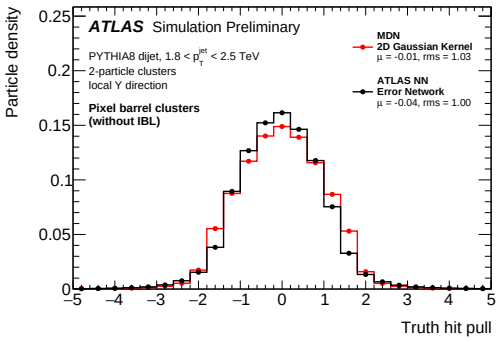
(a) Local x residuals



(b) Local x pulls



(c) Local y residuals

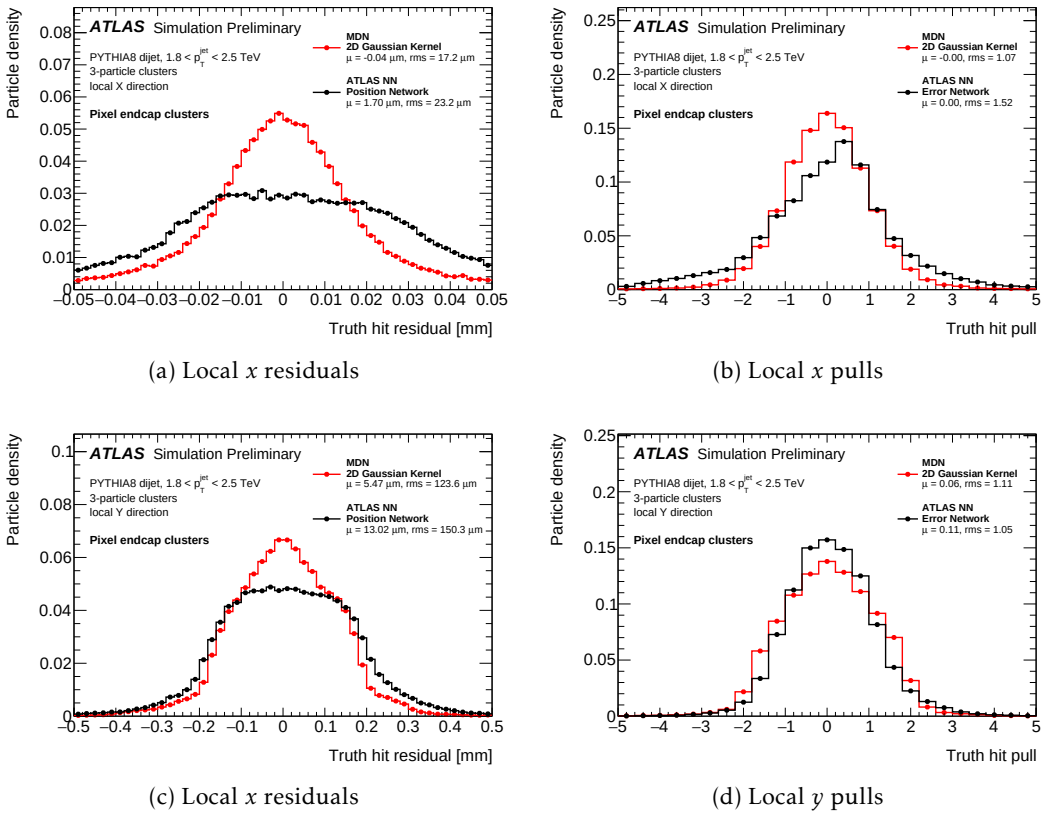


(d) Local y pulls

Figure 3.5: Comparison of residuals (a, c) and pulls (b, d) in the local x (a, b) and local y (c, d) direction obtained from the ATLAS NN (black) and MDN (red) using 2-particle barrel (without IBL) clusters.

networks are shown in Figure 3.5. The performance of 3-particle networks on endcap clusters are shown in **??**. MDN residuals are always more centered at zero, and have larger peaks. This indicates that the MDN predictions are less bias, and network predicted positions are closer to the

truth positions. In general MDN performs similarly or better than the current networks.



(a) Local x residuals

(b) Local x pulls

(c) Local x residuals

(d) Local y pulls

Figure 3.6: Comparison of residuals (a, c) and pulls (b, d) in the local x (a, b) and local y (c, d) direction obtained from the ATLAS NN (black) and MDN (red) using 3-particle endcap clusters.

3.5 Conclusions

An MDN-based algorithm can estimate the hit position and uncertainties at the same time. Only one MDN is required along with the number network. So this algorithm has the potential to be faster than the current one. MDNs also perform better than the current position and error networks. The residual distribution has a higher peak around zero indicating that more often the predicted values are close to the true values. MDN pulls are also much closer to standard normal distributions indicating that the uncertainties are better modeled. All these studies are done in standalone software set-ups, and we haven't been able to study yet the improvements to the final track reconstruction or establish the timing comparison. The MDN algorithm has now been added to the ATLAS track reconstruction software framework. So, we will be able to do more one-to-one comparisons of the results. In the future, the impact parameter resolution will be estimated using the MDN algorithm. It will further help to understand any potential effect on flavor tagging. Upon showing better performance on track reconstruction, MDN could be one of the powerful algorithms used for future tracking in high luminosity LHC. The MDN algorithm is now being studied actively by other members of the collaboration in the context LHC Run-3.

CHAPTER

FOUR

Object Reconstruction

This chapter provides an overview of the reconstruction of physics objects with the ATLAS detector relevant to this dissertation. An overview of the reconstruction is given in section 4.1. The basic components like tracks, vertices, and energy clusters, used for particle reconstruction and identification are described in section 4.2. Electrons and muons are the two charged leptons are used in this dissertation. Their reconstruction and identification methods are described in section 4.3 and section 4.4, respectively. Many different types of jets are used in the searches presented in this dissertation. The jet definitions and algorithms are summarized in section 4.5. The reconstruction and calibration of the small radius calorimeter jets (with radius parameter $R = 0.4$) are discussed in section 4.6. The truth jets are defined in section 4.7. Variable radius track jets where the radius parameter is jet p_T dependent are used in this dissertation to identify b-jets. These jets are described in section 4.8. The large radius calorimeter jets (with radius parameter $R = 1.0$) are described in section 4.9. A b-tagging algorithm is used to identify the jets coming from b-hadron decay. Different b-tagging methods used in this dissertation are summarized in section 4.10. A top-tagging technique is used to identify a large-R jet formed with the decay products of a high momentum top-quark decay. The top-tagging methods and their performances are described in section 4.11. Since the ATLAS detector cannot detect neutrinos, their momenta are usually reconstructed using the missing transverse energy (MET). The definition of MET is described towards the end of this chapter, in section 4.12.

4.1 Reconstruction overview

A schematic description of different particle interactions in the ATLAS detector is shown in Figure 4.1. All charged particles leave tracks in the ID. Besides leaving tracks in the ID, electrons produce electromagnetic shower in the electromagnetic calorimeter. Electrons deposit almost all their energy in the electromagnetic calorimeter. As a result, electrons are reconstructed combining the track and calorimeter responses. Muons, on the other hand, are very weakly interacting particles. They leave tracks in the ID but do not interact much with the calorimeter. They form another set of tracks in the muon spectrometer. Hence, information from the ID and MS are used to reconstruct muons. Being neutral, a photon does not leave any track in the ID but deposit their energy on the calorimeters. Hadrons develop a hadronic showers, mostly in the hadronic calorimeter. These

showers are less confined than the electromagnetic shower. In addition, charged hadrons like protons, pions, and kaons also create tracks in the ID. Neutral hadrons like neutral pions, kaons, and neutrons form similar showers in the calorimeter but do not have a charged track. Neutrinos go undetected through the ATLAS detector. However, their presence is inferred by calculating the missing transverse energy.

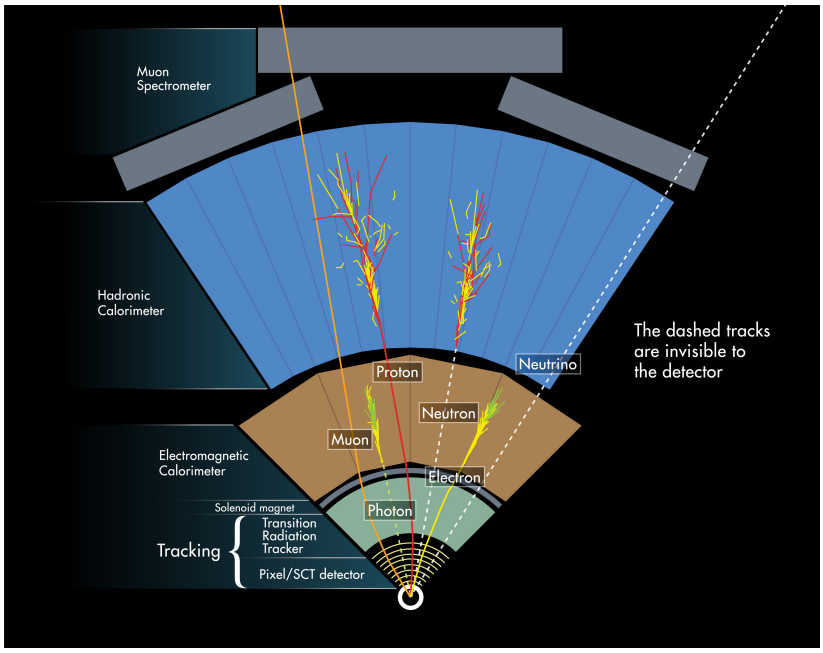


Figure 4.1: Schematic diagram showing the particle interactions with the ATLAS detector. Figure taken from [172].

4.2 Tracks, vertices and energy clusters

This section summarizes some of the basic components used in particle reconstruction. All charged particles create tracks in the ID or MS. The ID tracks are usually associated to a vertex. A vertex is defined as the reconstructed point of origin of two or more particles. Vertices are crucial for reconstructing almost all the particles and play an important role in identifying the b- and c-hadron decay. Many particles deposit their energy in the calorimeter and activate the calorimeter cells. The activated calorimeter energy cells are grouped to build energy clusters. Energy clusters play an important role in constructing particles like electrons, photons, and all the hadrons. This section will summarize the reconstruction method of tracks, vertices, and energy clusters.

4.2.1 Tracks

Charged particle trajectories are called tracks. All charged particles leave tracks in the ID, and muons leave tracks in the MS as well. This section briefly describes the ID tracking [173, 174]. When a charged particle passes through the ID, it ionizes the sensor material, thus producing electric charges in the sensors. A connected component analysis (CCA) algorithm groups the adjacent pixels and strips in a given sensor, where the deposited energy yields a charge above a threshold. Then three-dimensional measurements are built to represent the point where the charged particle

crosses the sensors. These three-dimensional charge clusters are called *space points*. These space points are connected to build tracks. First, the track seeds are formed by combining three adjacent space points. Then the track candidates are built from the chosen seeds by incorporating additional space points from different layers of pixel and SCT using a combinatorial Kalman filter. The baseline track finding algorithm uses an inside-out approach starting from seeds in the pixel or SCT and extrapolating to outer layers. In busy environments in which charged particles are so close to each other that their charge clusters merge forming one charge cluster, generally called a merged cluster. Thus multiple track candidates get associated with one merged cluster. Merged clusters make the track fitting more difficult since the track to cluster association becomes ambiguous. This is addressed by using a dedicated ambiguity solving stage. In the ambiguity solver, all the track candidates are scored based on their quality. Tracks candidates failing all the quality criteria are scored low and are only considered after the high-quality tracks are reconstructed. The track candidates with merged clusters are usually identified as low quality tracks. So the ambiguity solver splits the merged clusters using a neural network based algorithm. This method is described in more detail in chapter 3. Once the ambiguities are solved the track finding algorithm is run again to refit the tracks. Finally, the tracks are extended to the TRT. After finding the tracks, the tracks are refitted using a χ^2-fitter. There is also an outside-in algorithm that starts with TRT hits and extends the tracks inwards. This helps to recover any missing hits in one of the layers. In presence of the ID solenoid magnet the tracks follow helical trajectories in the transverse plane ($r-\phi$ plane) and can be fully parameterized by the following set of five parameters:

$$(d_0, z_0, \phi, \theta, q/|p|),$$

where d_0 and z_0 are the transverse and longitudinal impact parameters, respectively, ϕ and θ are the azimuthal and polar angle of the track, and $q/|p|$ is the ratio of particle charge and momentum. The track reconstruction method and its performance are described in more detail in reference [174].

4.2.2 Vertices

Vertices are referred to as particle interaction points. A proton-proton interaction point is defined as the primary vertex. The primary production vertex is important to measure the kinematic properties of a charged particle. The primary vertex reconstruction [175] process is divided into two stages: vertex finding where reconstructed tracks are associated with the vertex candidates and vertex fitting where the actual vertex position and uncertainty are estimated [176]. The process starts with a set of selected tracks and a seed vertex position. The vertex position is determined using an iterative primary vertex fitting algorithm with an annealing procedure [177]. Each input track is assigned a weight based on its compatibility with the vertex estimate. The track weights and the vertex position are updated iteratively. Tracks found incompatible after the last iteration are collected and used for another vertex reconstruction. The process continues until the algorithm can not find any new vertices. Finally, all vertices with at least two tracks are kept as valid vertex candidates. The vertex with the highest sum of squared track p_T is considered to correspond to the hardest proton-proton interaction point and is called the primary vertex. Vertices that are incompatible with the beam collision region are considered as secondary or displaced vertices. Secondary vertices usually originate from the decay of long-lived charged particles, such as b-hadrons. Secondary vertex reconstruction is very important for identifying b- or c-jets and are further discussed in section 4.10. Sometimes we also find displaced vertices coming from photon conversion and uncharged long-lived BSM particle decay. Hence, displaced vertex reconstruction is also very important for long-lived particle searches.

4.2.3 Energy clusters

Calorimeter cells with deposited energy are grouped to form clusters if they share common edges. There are two different clustering methods used in ATLAS. The *topo-cluster* reconstruction algorithm is always used to build the energy clusters used for jet reconstruction. The *sliding-window* algorithm was used to build the energy clusters used in electron and photon reconstruction until the end of 2016. Since 2017 the topo-cluster reconstruction algorithm is also used to build the EM clusters used for electron and photon reconstruction. The two methods are briefly described below.

Sliding-window algorithm

The *sliding-window* [178] algorithm combines the EM calorimeter cells within a fixed-size rectangular window and sums up their energies to identify electrons and photons. The sliding-window algorithm is based on three main steps: tower building, precluster (seed) finding and cluster filling.

1. The $\eta - \phi$ space of the EM calorimeter is divided into a grid of $N_\eta \times N_\phi = 200 \times 256$ elements of size $\Delta\eta \times \Delta\phi = 0.025 \times 0.025$. The element size corresponds to the second layer of the EM calorimeter. The energies of all the cells in the longitudinal direction inside these elements are summed to form a *tower* energy.

2. A sliding window of fixed size, 3×5 in $\eta\times\phi$ is used to search for seeds. A seed is then selected if the total transverse tower energy in window is more that 2.5 GeV. If two seed clusters are found within an area of $\Delta\eta\times\Delta\phi = 5\times9$, then the higher energy seed is kept. Hence duplicate seeds do not appear.

3. The clusters are formed around the seeds by taking all the cells within a rectangular window of $3\times7(5\times5)$ in the barrel (endcap) region.

Topo-cluster reconstruction algorithm

The topological cluster-based algorithm [179] is implemented for reconstructing electrons, photons, and jets in ATLAS. The fixed-sized window was used in the past for electron and photon reconstruction for clustering due to the technical difficulties of applying calibration to dynamically-sized clusters. By using a variable-sized cluster, it is possible to recover lower energy deposits coming from bremsstrahlung photons and add them back to the electron energy. The topological cluster-based algorithm is the default for jet reconstruction. This algorithm relies on the cell energy significance which can be defined as:

$$\zeta_{\text{cell}}^{\text{EM}} = \frac{E_{\text{cell}}^{\text{EM}}}{\sigma_{\text{noise, cell}}^{\text{EM}}} \tag{4.1}$$

Here, $E_{\text{cell}}^{\text{EM}}$ is the absolute cell energy at the EM-scale and $\sigma_{\text{noise, cell}}^{\text{EM}}$ is the expected cell noise coming from electronics and pileup. The clustering algorithm follows the following steps:

1. *Proto-clusters* are seeded from cells with $|\zeta_{\text{cell}}^{\text{EM}}| \geq 4$.

2. Neighboring cells with $|\zeta_{\text{cell}}^{\text{EM}}| \geq 2$ are added to the proto-clusters.

3. Neighboring proto-clusters sharing some cells are merged.

There topoclusters are known as "4-2-0" topoclusters. The "4-2-0" topoclusters having high EM fraction (ratio of ECAL and HCAL energy) are used to build the electron, photons and jet clusters. First the seed clusters are formed using the "4-2-0" topoclusters. Then soft energy clusters, mostly coming from bremsstrahlung photons, are identified as satellite clusters. Satellite clusters are added to the seed clusters to form the final superclusters if they pass the necessary selection criteria.

Next we review the reconstruction of the different physics objects like the leptons and the jets. In this dissertation, electrons and muons are used, so the discussion will be restricted to those two leptons.

4.3 Electrons

Most of the incoming particles deposit their energies in multiple calorimeter cells in both η and ϕ directions. Electrons are reconstructed in the central region of the ATLAS detector with $|\eta| < 2.47$. Electrons deposit almost all their energy in the EM calorimeter. Being charged particles, they also leave tracks in the ID. So, both ID and colorimeter information are used to reconstruct electrons.

4.3.1 Electron reconstruction

Electron reconstruction is a multi-step process and starts with the calorimeter cluster formation. A schematic diagram of electron reconstruction is shown in Figure 4.2. Clustering algorithms are designed to group the calorimeter cells. The efficiency to reconstruct EM-cluster candidates (localized energy deposits) in the electromagnetic calorimeter is given by the number of reconstructed EM calorimeter clusters N_{cluster} divided by the total number of electrons produced, N_{all}. The efficiency of this clustering algorithm depends on E_T and $|\eta|$ and ranges from 65% at $E_T = 4.5$ GeV, to 96% at $E_T = 7$ GeV, to more than 99% above $E_T = 15$ GeV as shown in Figure 4.3. Once the seed clusters are found, they are matched with the well-reconstructed ID tracks. If the matching fails the cluster is marked as an unconverted photon. If the matching is possible but the associated tracks are not coming from a primary vertex it is marked as a converted photon. The sliding-window algorithm is used in the dilepton resonance search described in chapter 6.

In the other method, the calorimeter clusters are formed using the topo-cluster reconstruction algorithm [181]. Once the electromagnetic calorimeter clusters are formed, regions of interest (ROI) are created. The standard track pattern reconstruction is performed within these ROIs. If this algorithm fails to form an ID track seed that is within ROI, a modified pattern reconstruction algorithm is performed based on a Kalman filter formalism. Next, the track candidates are fitted with a global χ^2 fit, which allows additional energy loss due to material interaction in case the Kalman filter-based algorithm fails. Finally, these tracks, loosely associated with a calorimeter, are refitted using a Gaussian Sum Filter (GSF) to improve track parameter estimation. The loosely matched tracks are then matched with EM clusters with a tighter requirement in η and ϕ to build the electron candidates.

4.3.2 Electron identification

A likelihood-based (LH) identification method is used to select the prompt electrons. The prompt electrons are defined as the signal. A combination of jets that mimic the signature of prompt electrons, electrons from photon conversions in the detector material, and non-prompt electrons

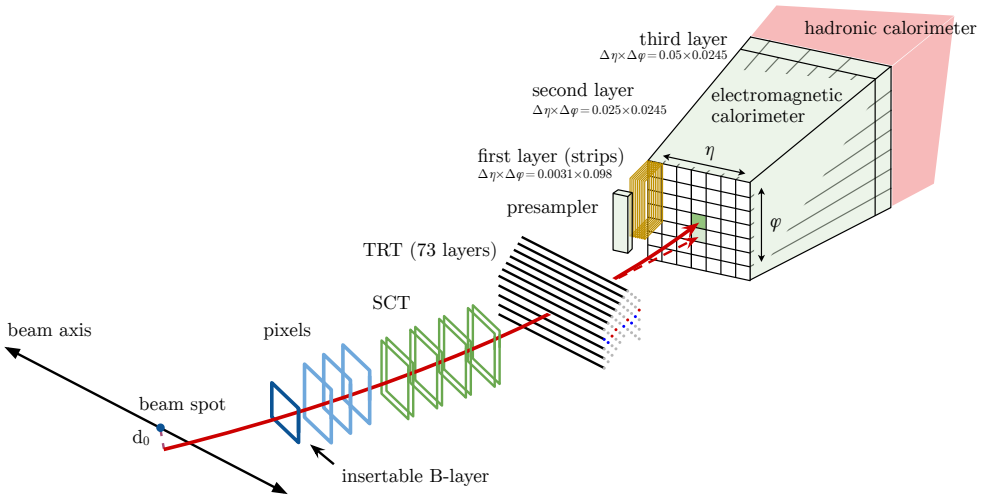
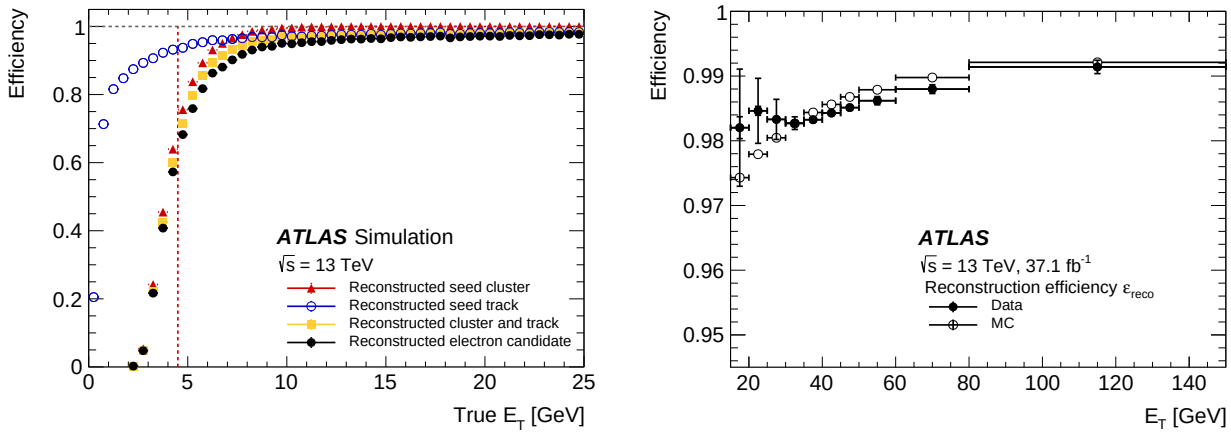


Figure 4.2: A schematic diagram of the path of an electron through the detector. The red trajectory shows the path of an electron, which originates close to the beam spot, first traverses the tracking system (pixel detectors, then silicon-strip detectors and lastly the TRT), and then enters the electromagnetic calorimeter. The dashed red trajectory indicates the path of a photon produced by the interaction of the electron with the material in the tracking system. Figure taken from [180]



(a) (b)

Figure 4.3: (a) Total electron reconstruction efficiency as a function of true electron E_T is calculated from single electron simulation sample. (b) Reconstructed efficiency relative to reconstructed clusters, given by the number of reconstructed electron candidates N_{reco}, divided by the number of EM-cluster candidates N_{cluster}, as a function of electron transverse energy, E_T calculated from $Z \to ee$ events. Figure taken from [180]

from the decay of hadrons containing heavy flavors are defined as the background. All the relevant quantities related to both the tracking system and calorimeter are included in the likelihood. The likelihood discriminant used in this method is a combination of the signal and background likelihoods built with the probability distribution functions of the quantities relevant for electron identification. The LH-based electron identification is more efficient compared to a simple cut-

based identification. In a cut-based method, an electron fails the identification criteria if it does not pass the cut on a specific quantity. On the other hand, in an LH-based identification, all the criteria are included in the LH and an electron does not fail the identification criteria just due to one quantity.

Based on the selection efficiency and background rejection, four values of the LH discriminant are used to define four operating points. Currently used operating points in ATLAS are *VeryLoose*, *Loose*, *Medium*, and *Tight*. The efficiency of identifying a prompt electron with E_T = 40 GeV is 93%, 88%, and 80% for Loose, Medium, and Tight operating points, respectively. LH-based identification operating points are optimized and studied using $Z \to ee$ and $J/\Psi \to ee$ events [180]. The identification efficiencies are shown in Figure 4.4. In this dissertation, Loose, Medium, and Tight operating points are used.

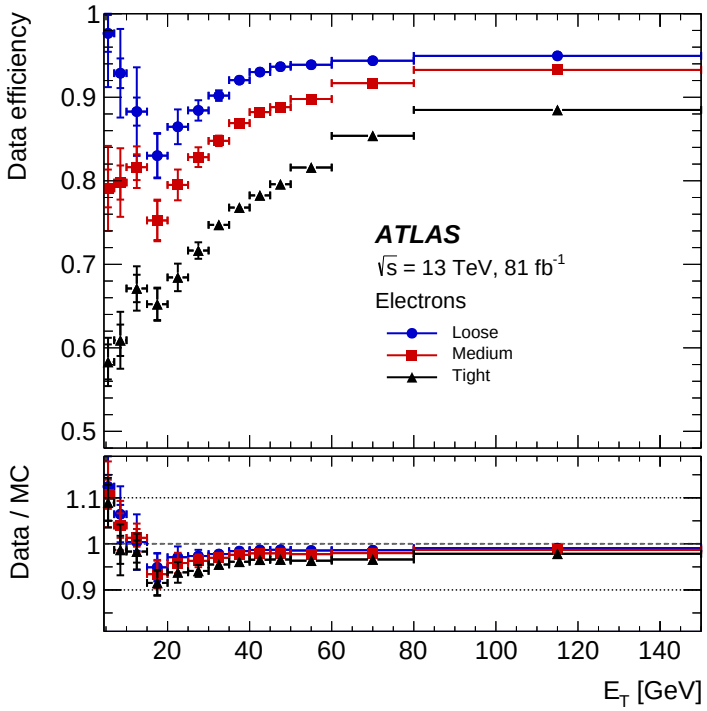


Figure 4.4: The electron identification efficiency in $Z \to ee$ events in data as a function of E_T for Loose, Medium, and Tight operating point. Figure taken from [181].

4.3.3 Electron isolation

The amount of detector activity near the candidate electron object is referred to as isolation. A characteristic signature of signal processes (electrons from hard scattering vertex, or from heavy particles such as H, Z, W decay) usually shows very little activity both in the calorimeter and the ID in an area $\Delta\eta \times \Delta\phi$ surrounding the candidate object. Isolation variables are defined to quantify the amount of activity in the vicinity of the candidate object, excluding the candidate itself. There are two types of isolation variables, track-based isolation variable, and calorimeter-based isolation variable.

- E_T^{cone20} : The raw calorimeter isolation variable, $E_{T,\text{raw}}^{\text{iso}}$ is the sum of the transverse energy of

positive-energy topological clusters whose barycenter fall within a cone of $\Delta R = 0.2$ around the electron cluster barycenter. E_T^{cone20} is then defined by subtracting the electron contribution from $E_{T,\text{raw}}^{\text{iso}}$ with some additional correction due to pile-up and leakage.

- $p_T^{\text{varcone20}}$: A variable radius cone is used to construct the track isolation variable. When an electron coming from a high momentum heavy particle other decay products can be close to the electron direction. $p_T^{\text{varcone20}}$ is computed by adding transverse momentum of all the tracks with $p_T > 1$ GeV in a cone of size $\Delta R = \min(10\text{ GeV}/p_T^e, 0.2)$ around the electron track.

Different isolation operating points (OP) are defined based on isolation selection and efficiency requirements. Isolation OPs are defined in three categories: track-only, calo-only and combined track and calo. In this dissertation, Loose OP is used in the dilepton resonance search which has around 99% efficiency in the calorimeter and track selection, but has a combined efficiency of 98%. In the $t\bar{t}$ resonance search, described in chapter 8, track-only isolation OP is used to maximize high momentum signal acceptance. In this track-only isolation method a new isolation variable, $p_T^{\text{varcone20}}\text{TTVA}_{p_T 1000}$ (slightly improved version of $p_T^{\text{varcone20}}$) is used. It is required to have $p_T^{\text{varcone20}}\text{TTVA}_{p_T 1000}/p_T < 0.06$.

4.4 Muons

Muons interact very weakly with the detector materials, so they deposit very little energy in the calorimeter. So, muons are reconstructed using the tracks in the MS and the ID.

4.4.1 Muon Reconstruction

Muon reconstruction first starts with the independent reconstruction of tracks in the MS and the ID. The tracks from the individual sub-detectors are then combined to form combined muon tracks extending to the full detector. The ID muon tracks are reconstructed like any other charged particle track. In the MS, muon reconstruction starts with a search for hit patterns inside each muon chamber to form track segments. The track candidates are then formed by using a combinatorial fit of the track segments. At least two track segments are required to build a muon track candidate, with the exception of the transition region between barrel and endcap where a single high-quality segment can be used. Once the muon track candidates are built, they are fitted with a global χ^2 fit. Track candidates are accepted based on χ^2 threshold. The track fitting is repeated several times by removing the inconsistent hits and using additional hits.

There are several algorithms used in ATLAS to combine ID and MS muon tracks. Four different types of muons [145] are defined depending on which sub-detector is used in the reconstruction:

- Combined (CB) muon: Combined muons are built by combining the MS and ID hits using a global refit. Outside-in pattern recognition is used for most of the muons, where the MS tracks are extrapolated towards the ID. MS hits can be added or removed during this process to improve the fit quality. The remaining muons are reconstructed by an inside-out manner, where the ID tracks are extrapolated out towards the MS.

- Segment-tagged (ST) muon: Segment-tagged muons are built from an ID track which is associated with at least one local track segment in the MDT or the CSC chamber. ST muons are used when muons traverse only one layer of the MS chambers.

- Calorimeter-tagged (CT) muon: Calorimeter-tagged muons are built from an ID track if it can be matched to an energy deposit in the calorimeter compatible with a minimum-ionizing particle. CT muons have the lowest purity but they recover acceptance in the $|\eta| < 0.1$ region where there is no MS coverage.

- Extrapolated (ME) muon: Extrapolated muons are built only with the MS tracks using a loose requirement on the track to interaction point association. ME muons are mainly used to extend the acceptance in to the region uncovered by ID, $2.5 < |\eta| < 2.7$.

The CB muons are used in the analyses described in this dissertation.

4.4.2 Muon Identification

Muon identification is performed by applying quality requirements in order to suppress the background mostly coming from pion and kaon decay and to select prompt muons with high efficiency. The following variables are used in muon identification for combined tracks:

- *q/p significance:* This is defined as the absolute value of the difference between the ratio of charge, q and the momentum, p measured from the track curvature in the ID and the MS divided by quadratic sum of the corresponding uncertainties.

- $\rho^{'}$: This is defined as the absolute value of the difference between the transverse momentum measurements in the ID and MS divided by the p_T of the combined track.

- The normalized χ^2 of the normalized fit. [145]

Muon identification operating points are defined based on the identification selections. There are four such operating points in ATLAS: *Loose*, *Medium*, *Tight*, *High-*p_T. Muon identification criteria are optimized using $Z \to \mu\mu$ and $J/\Psi \to \mu\mu$ decay [145] as shown in Figure 4.5. In this dissertation only *Loose* and *Medium* working points are used.

The medium identification criteria provide the default muon selection in ATLAS. This selection is optimized to minimize the systematic uncertainties associated with muon reconstruction and calibration. It uses only CB and ME tracks. For CB tracks at least three hits in at least two MDT layers are required. In the central region, $|\eta| < 0.1$, the criteria are relaxed and tracks are allowed with at least one MDT layer but no more than one MDT hole layer. Here a hole refers to an active sensor traversed by the track but containing no hits. The ME tracks need to have at least three MDT or CSC layers. In addition, the q/p significance is required to be < 7 to suppress non-prompt contamination. The muon reconstruction efficiency with this selection is around 96% for muons with $20 < p_T < 100$ GeV.

Loose identification criteria, on the other hand, are designed to maximize reconstruction efficiency while providing good quality muon tracks. All muon types are used in this selection. All the CB and ME tracks satisfying the medium selection also pass the loose selection. CT and ST muons are only used in the central region, $|\eta| < 0.1$, to improve the acceptance. The reconstruction efficiency of the loose muons with $20 < p_T < 100$ GeV.

4.4.3 Muon Isolation

Muon isolation refers to the measurement of detector activity around a muon candidate and is used to reject non-prompt muons. Prompt muons produced from heavy boson decay, such as W, Z, or H, are usually isolated from other particles. Non-prompt muons on the other hand are

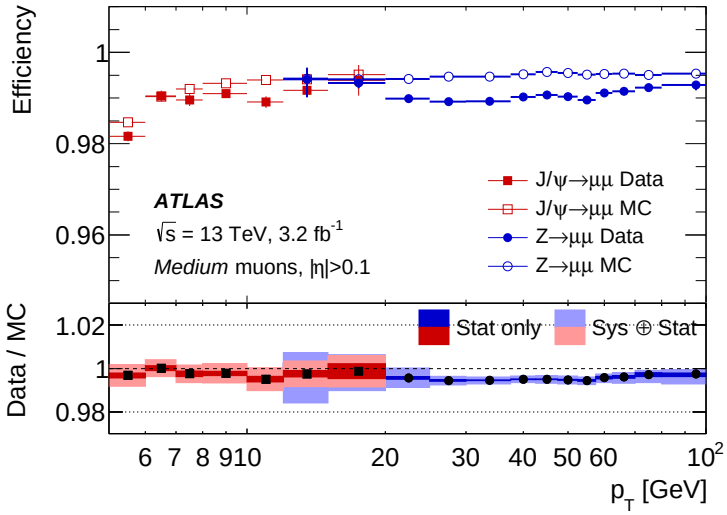


Figure 4.5: Reconstruction efficiency of the medium muon selection as a function of muon p_T, in the region $0.1 < |\eta| < 2.5$ obtained from $Z \to \mu\mu$ and $J/\Psi \to \mu\mu$ events. The error bars indicate statistical uncertainty. Figure taken from [145]

produced in a busy environment typically surrounded by other tracks and calorimeter deposits. Two variables are defined to assess muon isolation: a track-based and a calorimeter based variable.

- $p_T^{\text{varcone30}}$: This is a track-based isolation variable and is defined as the sum of the transverse momenta of the tracks with $p_T > 1$ GeV in a cone of size $\Delta R = \min(10\,\text{GeV}/p_T^{\mu}, 0.3)$ around the muon p_T, excluding the muon track. The p_T-dependent variable cone size helps to improve performance for the muons coming from high momentum particle decay.

- $E_T^{\text{topocone20}}$: This is a calorimeter-based variable and defined as the sum of the transverse energy of the topoclusters within a cone of size $\Delta R = 0.2$ around the muon after correcting the pileup effects and removing the muon energy deposits.

Based on the isolation selection, several isolation operating points are defined. The operating points are generally categorized into two groups: track-only isolation and combined track and calorimeter based isolation. In this dissertation track-only isolation operating point,*TightTrackOnly_VarRad*, is used for muon isolation. In this track-only selection a new isolation variable, $p_T^{\text{varcone30}}\text{TTVA}_{p_T1000}$ (slightly improved version of $p_T^{\text{varcone30}}$) is used. It is required to have $p_T^{\text{varcone30}}\text{TTVA}_{p_T1000}/p_T < 0.06$. This isolation criteria provided a very high acceptance for high p_T muons.

4.5 Jets

QCD quark confinement prevents partons from existing as isolated particles. Moreover due to the strength of the strong force, the final state partons produced in the proton-proton collision form showers of quarks and gluons. The quarks then hadronize and form color-neutral hadrons. The spray of these hadrons is collectively called a jet. Jets mostly contain pions (π^+, π^-, γ^0 coming from π^0 decay)) and roughly 60% of jet energy comes from the pions. Charged pions ($\pi^{\pm}$) are considered as stable since they they have $c\tau = 10$ mm. On the other hand, neutral pions (π^0) decay

instantaneously to a pair of photons. The other significant contributions come from kaons, light baryons (mostly protons and neutrons), and strange baryons. Thus the jets leave several tracks in the ID and deposit most of their energy in the calorimeters. In ATLAS jets are reconstructed from different input objects such as tracks and/or calorimeter clusters using different jet clustering algorithms. Thus a jet is a proxy of the initial parton and helps us to understand short-distance physics.

4.5.1 Jet reconstruction

At the LHC, jets are reconstructed with a bottom-up approach by clustering the constituents (tracks or energy clusters) using anti-k_T algorithm [182]. The anti-k_T algorithm is a sequential recombination algorithm, and can be formulated as described below:

1. Consider all the constituents in the initial object list.
2. Build two sets of distances from the list of objects:
   - *inter object distance* defined as
   $$d_{ij} = \min\left(p_{T,i}^{-2}, p_{T,j}^{-2}\right)\frac{\Delta_{ij}^2}{R^2} \tag{4.2}$$
   where $\Delta_{ij}^2 = (y_i - y_j)^2 + (\phi_i - \phi_j)^2$ is the angular distance between i^{th} and j^{th} object and R is the jet radius parameter.
   - *beam distance* defined as
   $$d_{iB} = p_{T,i}^{-2} \tag{4.3}$$
   where B refers to the beam.
3. Find the smallest distance among all d_{ij} and d_{iB}.
   - If d_{ij} is the smallest distance then object i and j are removed from the list and combined to a new object k. The new object k is then added to the object list.
   - If the smallest distance is d_{iB} then i^{th} object is called a jet and removed from the list.
4. Repeat the steps until there are no more objects left in the list.

A theoretically well-motivated jet algorithm should be both infrared and collinear safe. In an infrared-safe algorithm, the jet definition should be much less sensitive to the soft radiation of the initial parton. The anti-k_T algorithm starts clustering jets from the hardest objects and successively aggregates the soft objects around it. As a result, hard jets are less sensitive to surrounding energy deposits. So, the anti-k_T algorithm is infrared-safe by construction. Collinear-safe algorithms ensure that the formation of a jet is insensitive to an exact collinear split, i.e. the jet formation should not be affected if a single particle is replaced by two collinear particles of half the original energy. The anti-k_T algorithm is also collinear safe. The boundaries of a jet should be well-defined even in the case where two jets would overlap. In the anti-k_T algorithm, the constituents are uniquely assigned to one jet, so jet boundaries are well defined. Thus anti-k_T algorithm satisfies all the criteria to be a theoretically motivated jet algorithm. The anti-k_T (AKT) algorithm is used to reconstruct the different types of jets used in this dissertation. The AKT jets have a nearly circular shape in the $y - \phi$ plane, where y is the rapidity. A slight variation of the anti-k_T algorithm,

called the k_T algorithm where the exponent of p_T used in and is +2 instead of −2. In the k_T (KT) algorithm the clustering starts from the soft objects. This is useful for reconstructing low p_T jets and estimating contributions from pile-up or other soft radiations. The KT algorithm is used for pile-up mitigation and probing the substructures of the large radius jets ($R = 1.0$).

4.6 Truth jets

Particle-level truth jets are created in the simulated events by clustering the stable particles which have a lifetime $c\tau > 1cm$. In this process only particles with significant energy deposits in the calorimeter are used, so muons and neutrinos are excluded. Particles from pile-up interactions are not included when building the truth jets. Particle-level truth jets are built using the same jet finding algorithm with the same selection as the reconstructed jets. The particle-level truth jets are geometrically matched to the reconstructed jets using the angular distance ΔR. Particle-level jets or the truth jets are taken as the reference for simulation-based jet calibrations performed in ATLAS, and for studies of the jet energy and mass resolution.

4.7 Small-R jets

In this dissertation we call small-R jets the jets that are built from the calorimeter topoclusters and/or tracks, reconstructed with the anti-k_T algorithm with the jet radius parameter R set to 0.4. These jets are often referred to as AKT4 jets. Truth particle jets are geometrically matched to the reconstructed jets with the requirement $\Delta R < 0.3$. There are two different methods of building these jets in ATLAS. There are described in the next two sections.

4.7.1 EMTopo jets

EMTopo jets are built using only the topoclusters as jet constituents. The topoclusters are calibrated at the electromagnetic scale (EM-scale) which correctly accounts for the energy of electromagnetic showers. The topoclusters can further be calibrated using a local cell signal weighting calibration scheme (LCW) [179]. In this dissertation, small-R jets built from EM-scaled topoclusters are used [183]. EMTopo jets are used in the dilepton resonance search presented in chapter 6 and were studied initially in the context of the semi-leptonic $t\bar{t}$ resonance search presented in chapter 8 but finally were not used.

EMTopo jet calibration

Jets are calibrated to correct the energy, mass, and direction such that they correspond to the properties of the underlying truth particles. Thus jets need to be corrected to account for several effects like pile-up, non-compensating calorimeter response, differences between data and MC etc. The jet calibration work flow, shown in Figure 4.6, was developed in ATLAS at the beginning of Run-2 in 2015. This calibration work flow was used in the dilepton resonance search presented in chapter 6. The calibration steps are briefly discussed here, some of the steps are discussed in detail in Appendix A. A complete description can be found in [183]. To derive this calibration, reconstructed jets are geometrically matched to truth jets within $\Delta R = 0.3$. Only isolated jets with no other calorimeter jet with $p_T > 7$ GeV within $\Delta R = 0.6$, and with only one associated truth jet with $p_T^{\text{truth}} > 7$ GeV within $\Delta R = 1.0$ are used to avoid ambiguity in truth matching.

The pileup effects are further mitigated by using a jet vertex tagger (JVT) algorithm. The JVT algorithm uses a 2D likelihood to tag pileup jets [184]. Jets with a significant number of high momentum tracks not coming from the primary vertex are not used in the physics analysis.

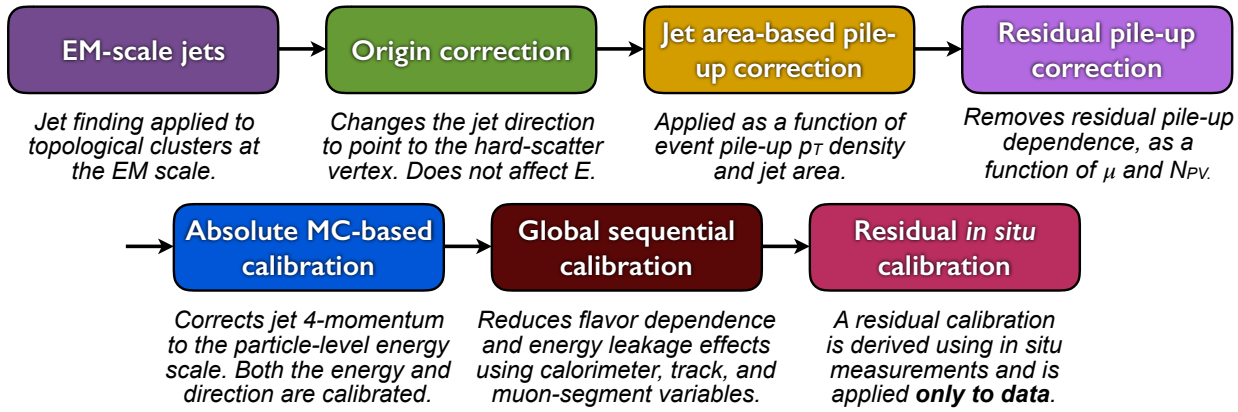


Figure 4.6: A schematic diagram shows the calibration stages for EM-scale jets. Other than the origin correction, each stage is applied to the jet four-vectors. Figure taken from [183].

The first four steps of the calibration are MC-based, while the in-situ calibration is done using data. In the first calibration step, the origin correction is applied to correct jet 4-momenta to the hard-scatter primary vertex while keeping the jet energy constant. Then the pileup correction is used to remove the contribution to the jet energy due to pileup effects. It is done in two steps: an area-based p_T density subtraction followed by the residual correction. Next, jet energy and η are corrected to the associated truth jets. The inverse of the energy response, $E^{\text{reco}}/E^{\text{truth}}$, is used as the correction factor. Then a global sequence calibration is applied to resolve the remaining differences between truth and reconstructed jets. The correction is applied sequentially to a set of jet observables without changing the average jet energy response. A residual in-situ calibration is applied to the jets in data to account for the differences between the jet responses in data and simulation. Some well-measured reference objects, including photons and Z bosons, are used to derive this correction.

4.7.2 PFlow small-R jets

PFlow jets use calorimeter as well as ID information. In this case, the jets are built using topoclusters and ID tracks. To avoid energy double-counting in PFlow reconstruction, the energy corresponding to the charged track is subtracted from the calorimeter energy. A cell-based energy subtraction algorithm is developed to remove overlaps between momentum and energy measurements. The topclusters used in the particle flow algorithm are calibrated at the EM scale as before. This algorithm provides a list of tracks, unmodified topoclusers, and a list of new topoclusters created after the energy subtraction procedure. The schematic flow chart showing different steps of the algorithm is shown in Figure 4.7.

PFlow jets are reconstructed using anti-k_T algorithm with $R = 0.4$ and is constructed from these energy-subtracted topoclusters and tracks that are matched to the primary vertex. These tracks are required to have $|z_0 \sin\theta| < 2$, where z_0 is the distance of the closest approach of a track from the primary vertex.

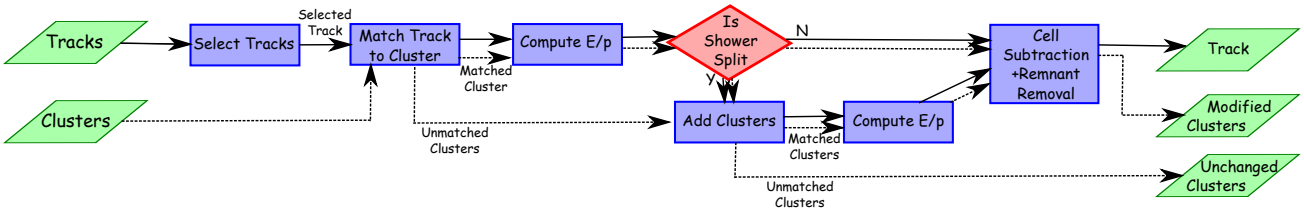


Figure 4.7: A flow chart of the particle flow algorithm. The algorithm takes ID tracks and topolocusters as input and returns a set of tracks, energy-subtracted topoclusters and unchanged topoclusters. Figure taken from [185].

PFlow jet calibration

The calibration of PFlow jets follows a similar scheme as the EMTopo jets, and the steps are the same. Some of the differences are mentioned here. The area-based pileup correction for PFlow jets is very similar to EMTopo jets. The pileup correction of EMTopo jets is done using transverse energy density computed from the calorimeter inputs. In the case of PFlow jets, the charged and neutral particle flow objects are treated differently and the correction is calculated accordingly. The MC-based jet energy correction is derived following the same method as done in the EM-Topo jets. In the global sequence correction a reduced number of variables is used. A detailed description of the PFlow jet calibration can be found in [185].

4.8 Variable-R track jets

Track jets are reconstructed from ID tracks using the anti-k_T algorithm. Only tracks in the region $|\eta| < 2.5$ with $p_T > 500$ MeV are selected as input to the anti-k_T algorithm. The input tracks need to also pass the track selection requirements mentioned in subsection 4.2.1. To improve the efficiency of selecting tracks coming from the primary vertex and reduce pileup effects, the tracks are required to have a longitudinal impact parameter $z_0 \sin\theta < 3$ mm. Until 2018, ATLAS used track jets of fixed radius, like $R = 0.2, 0.4$ etc. Since 2019 variable radius (VR) track jets are used in most physics analyses using track jets [186]. Unlike fixed radius track jets, the radius of a VR track jet is p_T dependent.

$$R_{\text{eff}}(p_T) = \begin{cases} R_{\text{max}}, & \text{for } \rho/p_T > R_{\text{max}} \\ \frac{\rho}{p_T}, & \text{for } R_{\text{min}} \leq \rho/p_T \leq R_{\text{max}} \\ R_{\text{min}}, & \text{for } \rho/p_T < R_{\text{min}} \end{cases} \tag{4.4}$$

where $\rho = 30$ GeV, $R_{\text{max}} = 0.4$ and $R_{\text{min}} = 0.02$. This algorithm allows very high p_T jets to be narrower, thus allowing close-by jets to be reconstructed with better accuracy. This strategy was particularly optimized for boosted $H \to bb$ decay where the fixed-radius track jets overlap with each other, and the two b's were being reconstructed as one track jet. Similarly, VR tracks are also useful in the boosted top-decay topology. In this dissertation, VR track jets are used to identify a b-hadron using a flavor tagging method that uses these VR track jets as inputs. The track jets, defined in the region $|\eta| < 2.5$, are required to have $p_T > 10$ GeV. The momentum and mass measurements of these jets are not used in this dissertation, so the calibration strategies are not discussed here.

4.9 Large-R jets

Final state jets coming from the hadronic decay of high momentum H, W, Z bosons, or top quarks become collimated along the decaying particle. As a result, the calorimeter signatures often overlap and can be reconstructed as a single jet with a large radius parameter. Large-R jets are built from the calorimeter topoclusters with the anti-k_T jet finding algorithm using the jet radius $R = 1.0$. These jets are sometimes referred to as AKT1 jets in this dissertation. The topoclusters used to build large-R jets are locally calibrated to the hadronic scale (sometimes refers as LC-scale) by using the Local Cell Weighting (LCW) scheme [179]. The large-R jets are further groomed with the trimming algorithm [187] to reduce effects of pileup, soft emissions, and the underlying event on jet substructure measurement. The trimming algorithm used in this dissertation reclusters the large-R jet constituents to form subjets using the k_T algorithm [188] with $R_{\text{sub}} = 0.2$. Soft subjets with $p_T^{\text{subjet}}/p_T^{\text{jet}} < 0.05$ are removed. Jets are redefined using the remaining constituents and the jet four-momenta are recalculated. In this dissertation, trimmed large-R jets with $p_T > 200$ GeV and $|\eta| < 2.5$ are used.

Large-R jet mass definitions

The large-R jet mass can be defined in three different ways.

1. *Calorimeter-based mass*, m^{calo}, is the jet mass calculated from the calorimeter clusters.

2. *Track-assisted mass*, m^{TA}, is defined as jet mass built from tracks only and then multiplied by the ratio of p_T of the calorimeter jet and the track jet:

$$m^{\text{TA}} = m^{\text{track}} \frac{p_T^{\text{calo}}}{p_T^{\text{track}}} .$$

   Track jets are reconstructed using tracks having $p_T > 500$ MeV and are also ghost-associated [189] to the large-R jet. Here m^{track} and p_T^{track} are the mass and p_T of the track jet obtained from the ID. p_T^{calo} is the p_T of the calorimeter jet (in this case large-R jet). The track-assisted mass measurement has better resolution for high p_T jets than the pure calorimeter one.

3. *Combined mass*, m^{comb}, is defined as the weighted sum of the calo mass and track-assisted mass. It can be written as:

$$m^{\text{comb}} = w_{\text{calo}} m^{\text{calo}} + w_{\text{TA}} m^{\text{TA}},$$

   where w_{calo} and w_{TA} are determined using the expected mass resolutions, σ_{calo} and σ_{TA}, of the calorimeter and track-assisted measurements respectively [190]. It is expressed as

$$w_{\text{calo}} = \frac{\sigma_{\text{calo}}^{-2}}{\sigma_{\text{calo}}^{-2} + \sigma_{\text{TA}}^{-2}}, \qquad w_{\text{TA}} = \frac{\sigma_{\text{TA}}^{-2}}{\sigma_{\text{calo}}^{-2} + \sigma_{\text{TA}}^{-2}} \tag{4.5}$$

   Thus the combined mass always has better resolution than either of the two inputs. In this dissertation the combined mass is used for large-R jets.

4.9.1 Large-R jet calibration

Large-R jets are calibrated in two steps like small radius jets. First, there is a MC simulation-based calibration, then an in-situ calibration is done using data. Figure 4.8 shows a flow chart of the

calibration steps used for full Run-2 data analyses. This calibration is used in the $t\bar{t}$ resonance searches presented in this dissertation. In the MC simulation-based calibration the truth jets are matched to the reconstructed large-R jets by using the requirement $\Delta R < 0.75$.

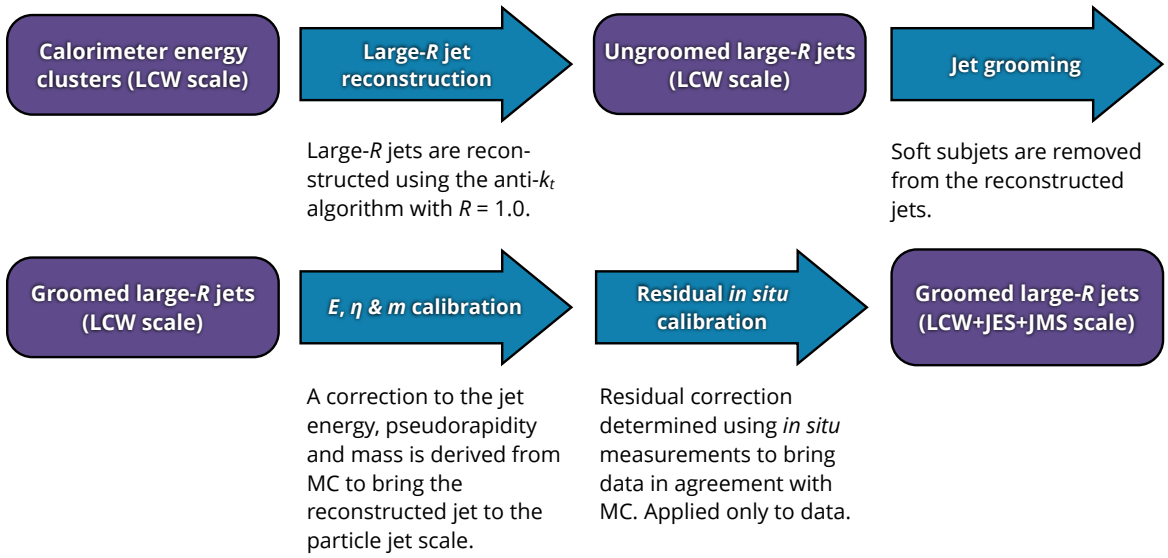


Figure 4.8: Overview of the large-R jet reconstruction and calibration. Figure taken from [190]

The MC-based calibration of the large-R jets follows similar steps as the small-R jets. Along with the energy calibration, a jet mass calibration is applied to the large-R jets. Since jet mass is more sensitive to soft, wide-angle contributions, cluster splitting and merging, and calorimeter geometry, it is important to properly calibrate it when jets are used in physics analysis. The in-situ calibration also includes the jet mass calibration. There are two different methods of jet mass calibration: the R_{trk} method and the forward folding method [190]. The R_{trk} method relies on the fact that the ATLAS detector gives two independent measurements of the large-R jet observables, like mass p_T. The jets constructed from the inner detector tracks only include the hits from their charged particle constituents. The calorimeter jets, on the other hand, are built using the energy clusters. The average calorimeter-to-track jet response

$$R_{\text{trk}} = \left\langle \frac{p_T^{\text{calo}}}{p_T^{\text{track}}} \right\rangle \tag{4.6}$$

should be proportional to the average calorimeter-to-truth jet response. The behavior should be similar in data and MC simulation. Thus the double ratio of R_{trk} should be unity for well-modeled observables and is used for in-situ calibration. This method is described in detail in [190, 191]. The R_{trk} method is used in the boosted all-hadronic $t\bar{t}$ resonance search presented in chapter 7. The forward folding method, on the other hand, uses high purity simulated samples with high-p_T large-R jets coming from hadronically-decaying W bosons and top-quarks. It is used to give the relative energy and mass scales and resolutions between data and simulation. The forward folding method is described in more detail in [190]. This method is used in the semi-leptonic $t\bar{t}$ resonance search presented in chapter 8 of this dissertation.

4.10 b-tagging

b-tagging is the method of identifying the jets which originate from b-partons and contains b-hadrons. Such jets are called b-jets. Their identification is important for the reconstruction of the top-quark decays presented in this dissertation. The top-quark decays very fast, before even forming hadrons, almost always to a W boson and a b-quark. The identification of b-jets is also crucial in Higgs physics since the Higgs boson decays to b-quarks almost 60% of the time.

The aim of b-tagging is to exploit the b-hadron decay signatures and lifetimes and discriminate the b-jets from other jets. In contrast to b-jets, c-jets originate from c-quarks and contains D-hadrons, whereas jets originating from gluons, first generation quarks and s-quarks are referred as light jets. b-hadrons are relatively long-lived and have lifetime, τ, ~ 1.5 ps ($c\tau \sim 450$ μm). So, they can travel a few millimeters from the production vertex before decaying, therefore creating a vertex displaced from the primary collision point, as shown in Figure 4.9. For example, a b-hadron of $p_T = 50$ GeV can travel around ~ 4.5 mm inside the detector before it decays into other charged particles.

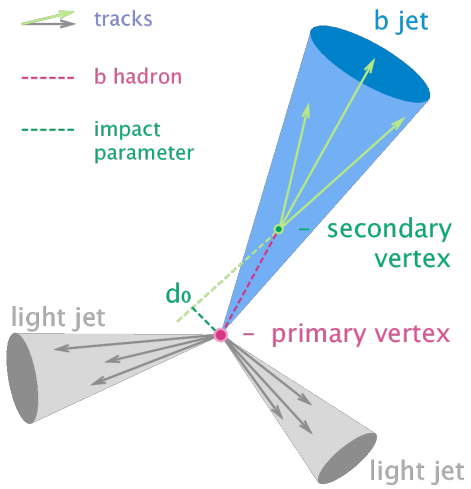


Figure 4.9: Schematic diagram comparing the b-jet topology with light jets. Important variables like primary vertex, secondary vertex, and impact parameter are shown in the diagram. Figure taken from [192].

4.10.1 Key ingredients for b-jet identification

The identification of b-jets requires the reconstruction of several objects which are briefly described here.

Tracks: ID tracks with $p_T > 500$ MeV are considered in b-tagging. Additional requirements are applied in some algorithms to reject fake and pile-up tracks.

Primary vertex (PV): PV reconstruction is crucial for b-tagging since the PV is the reference point from where the secondary vertex displacement is computed. At least one PV is required in each event. Displaced charged-particle tracks originating from b-hadron decays can then be selected by requiring large transverse and longitudinal impact parameters.

Track-jet matching: Track to jet matching is done using the angular separation ΔR (track, jet). Since the decay products of a high p_T b-hadron are expected to be collimated along with the jet momentum, ΔR (track, jet) is expected to be small. As a result, the ΔR (track, jet) requirement depends on the p_T of the jet. It is wider for low p_T jets (0.45 for jet p_T = 20 GeV) and narrower for high p_T jets (0.26 for jet p_T = 150 GeV) [193]. If multiple jets fulfill the matching criteria, the closest jet is chosen. The jet axis is used to assign the sign to the impact parameters. The impact parameter is assigned a positive sign if the track crosses the jet axis at a point upstream (or above) to the primary vertex along the jet axis. If the intersection point of the track and the jet axis is downstream (or behind) the primary vertex along the jet axis, the impact parameter is assigned a negative sign. Signed impact parameters are used in several b-tagging algorithms.

Jet flavor labeling: A flavor label is assigned to the jets in the simulation by matching them to at least one heavy hadron with $p_T > 5$ GeV within a cone of $\Delta R = 0.3$ around the jet axis. A jet is labeled as b-jet, if it is matches with at least one b-hadron. If no b-hadron is found, then the jet is tried to be matched with c-hadrons and then to a τ lepton. Jets matched to a c-hadron (τ-lepton) are labeled as c-jet (τ-jet). The remaining jets are labeled as light-jets.

4.10.2 b-tagging algorithms

Usually, b-tagging is done on the small radius jets with $R = 0.4$ or track jets with a variable radius. In this dissertation, b-tagging is done on both small-R jets and VR track jets in different searches. There are many algorithms for b-tagging, and all of them use the impact parameters and the vertices. These low-level algorithms evolved with time, and later in Run-2 outputs of some of these algorithms are combined to develop high-level algorithms. The high-level algorithms, in general, are better performant and are built based on multivariate discriminants. There are three different low-level b-tagging algorithms and two different types of high-level algorithms. A detailed description and performance of these algorithms are provided in reference [193].

Algorithms based on impact parameters: IP2D/IP3D

There are two complementary impact parameter-based algorithms, IP2D and IP3D [194]. IP2D uses only the signed transverse impact parameter significance, d_0/σ_{d_0} whereas the IP3D uses both signed transverse and longitudinal impact parameter significance, $z_0 \sin\theta/\sigma_{z_0 \sin\theta}$. Both of these algorithms are based on log-likelihood ratio (LLR) discriminant, separating b-jets from light-flavor jets. In addition to that, two more LLR functions are defined to separate b-jets from c-jets and c-jets from light jets. These three LLR discriminants for both IP2D and IP3D are used as inputs to the high-level taggers.

Secondary vertex finding algorithm: SV1

The secondary vertex finding algorithm (SV1) reconstructs a single displaced secondary vertex in a jet [195]. Then the algorithm identifies all the possible two-track vertices. Tracks compatible with long-lived particle (for example: K_s, Λ) decay, photon conversion, or hadronic interaction with the detector material are rejected. The SV1 algorithm runs iteratively until it finds a reasonable track-to-vertex association based on a χ^2 test and a vertex invariant mass less than 6 GeV. For a b-jet containing both b- and c-hadron decay vertices, this method merges the vertices into a single common vertex if they are close, otherwise reconstructs the vertex with largest jet multiplicity.

Finally, a LLR discriminant is constructed using the vertex mass, the energy fraction[12] and the number of two-track vertices. Eight discriminating variables related to the SV1 algorithm are used as inputs to the high-level taggers.

Topological multi-vertex finding algorithm: JetFitter

The topological multi-vertex finding algorithm, JetFitter [196], exploits the topological structure of the weak b- and c-hadron decays inside a jet. It uses a modified Kalman filter to find a common line containing the PV, and the b, or c vertices to approximate the b or c-hadron flight direction. This algorithm is responsible for vertex position and uncertainty reconstruction. Eight discriminating variables from the JetFitter algorithm are used as inputs to the high-level taggers.

MV2

MV2 algorithm is the first high-level b-tagging algorithm. The MV2 algorithm [194] consists of a boosted decision tree (BDT) algorithm that takes variables from the previously described low-level taggers. The BDT is trained on a hybrid sample of SM $t\bar{t}$ and $Z' \to b\bar{b}$ events. The BDT is trained using b-jets as the signal and a mixture of c-jets (7%) and light-jets (93%) as background. The output of this BDT is the MV2c10 score. The MV2 algorithm is used in this dissertation in the dilepton resonance search, described in chapter 6.

DL1

The other high level b-tagging algorithm is DL1 [194]. It is based on a deep neural network (DNN). The DNN is trained with the same hybrid sample $t\bar{t}$ and $Z' \to b\bar{b}$ events. DL1 algorithm uses a multi-class classifier that outputs the probabilities for a jet to be a b-jet, a c-jet, or a light-jet. Almost all the inputs used in the MV2 BDT are also used as DL1 DNN inputs along with some additional JetFitter c-tagging variables. The final DL1 b-tagging discriminant is defined by combining the b-jet, c-jet and light-jet probabilities as follows:

$$D_{\mathrm{DL1}} = \ln\left(\frac{p_b}{f_c \cdot p_c + (1 - f_c) \cdot p_{\mathrm{light}}}\right). \tag{4.7}$$

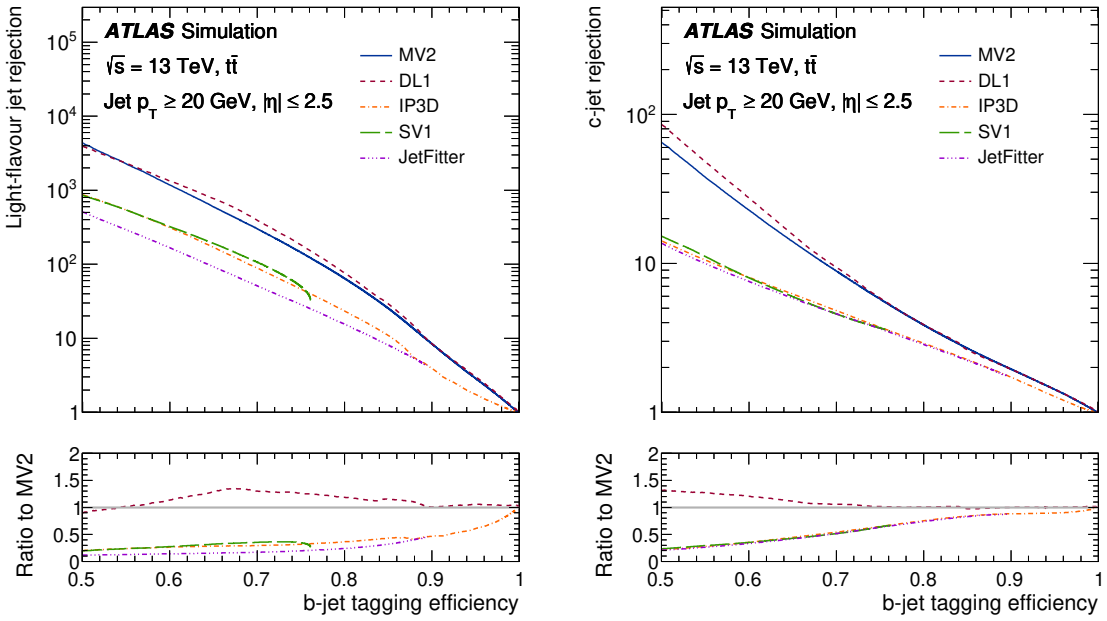
Here p_b, p_c, p_{light} and f_c are the b-jet, c-jet and light-jet probabilities, and the effective c-jet fraction in the background training sample. The DL1 algorithm is used in this dissertation, in the $t\bar{t}$ resonance searches described in chapter 7 and chapter 8.

Algorithm performance

The performance of the b-tagging algorithms is measured based on how correctly it identifies the jets coming from a real b-quark (b-jet efficiency, ϵ_b) compared to the mis-tag rate where a jet originating from a c-quark (c-jet mis-tag efficiency, ϵ_c) or a light-jet (light-jet mis-tag efficiency, ϵ_l) incorrectly identified as a b-jet. The fixed cut operating points (OPs) are defined based on fixed selection requirements on the discriminant of a b-tagging algorithm corresponding to a specific average b-jet tagging efficiency. The light-jet rejections (defined as $1/\epsilon_l$) and c-jet rejections (defined as $1/\epsilon_c$) are shown Figure 4.10 as a function of b-tagging efficiency. The plots show comparisons of the performances of the three low-level taggers and the two high-level taggers. The performance

[12] The energy fraction is the total energy of all the tracks associated with the secondary vertex divided by the energy of all the tracks associated with the jet.

of the MV2 and DL1 algorithms are very similar. DL1 has a slightly higher c-jet and light-jet rejection in the lower b-tag efficiency regions. But both algorithms perform much better relative to the low-level taggers.



(a) light-jet rejection

(b) c-jet rejection

Figure 4.10: The (a) light flavor jet rejections and (b) c-jet rejections are shown as a function of b-tagging efficiency for the IP3D, SV1, JetFitter, MV2 and DL1 b-tagging algorithms evaluated on simulated $t\bar{t}$ events. Figure taken from [193].

Four OPs are defined corresponding to b-tagging efficiencies of 60%, 70%, 77% and 85%. The corresponding discriminant cuts and c–jet, τ, and light-jet rejections are also summarized in Table 4.1, which is taken from [193]. In this dissertation, the 70% and 77% OPs are used in the $t\bar{t}$ resonance searches and the 85% OP is used in the dilepton resonance search.

Table 4.1: c-jet, τ and light-flavor jet rejections for the different b-jet tagging efficiency single-cut operating points and corresponding discriminant threshold values for the MV2 and the DL1 algorithms. These are evaluated using the baseline $t\bar{t}$ events. Table taken from [193].

ϵ_b	MV2				DL1			
	Selection	Rejection			Selection	Rejection		
		c-jet	τ	Light-flavour jet		c-jet	τ-jet	Light-flavour jet
60%	> 0.94	23	140	1200	> 2.74	27	220	1300
70%	> 0.83	8.9	36	300	> 2.02	9.4	43	390
77%	> 0.64	4.9	15	110	> 1.45	4.9	14	130
85%	> 0.11	2.7	6.1	25	> 0.46	2.6	3.9	29

4.10.3 b-tagging calibration

The b-tagging calibration is done to correct the simulation to match the tagger performance on data. The b-tagging efficiency is estimated using b-, c-, and light-jet enriched data samples. Then the b-tagging efficiency on a data sample is compared with that of the simulated events to derive a scale factor (SF), SF $= \epsilon_b^{\text{data}}/\epsilon_b^{\text{MC}}$, where $\epsilon_b^{\text{data (MC)}}$ is the measured efficiency in data (MC) for a jet to be b-tagged. The SFs depend on the truth label of the jet and are derived as a function of the p_T and η of the jet. Two types of jet weights are used depending on the truth labeling of the jet. If a b-labeled jet gets b-tagged, the b-tagging weight is simply the MC-to-data (efficiency) scale factor, $SF(p_T)$. On the other hand, if a b-labeled jet is not tagged as a b-jet the jet weight becomes:

$$w_{\text{jet}} = \frac{1-\epsilon_b^{\text{data}}(p_T)}{1-\epsilon_b^{\text{MC}}(p_T)} = \frac{1-SF(p_T)\epsilon_b^{\text{MC}}(p_T)}{1-\epsilon_b^{\text{MC}}(p_T)} \tag{4.8}$$

4.11 Top tagging

In a two-body decay, the decay products of the high momentum parent particles become very collimated along the direction of the parent particle because they get a high Lorentz boost. It can be shown that the angular distance between the daughter particles $\Delta R \sim 2m/p_T$, where m and p_T are the mass and transverse momentum of the decaying particle. Thus in a high-momentum top-quark decay, the outgoing W-boson and the b-quark become very close to each other. In the all-hadronic top decay topology, the W boson decays into two light quarks. So in a boosted (high momentum) hadronic top decay, the calorimeter clusters of the three final state objects merge and get contained within a single large-radius jet ($R = 1$). A schematic diagram of the resolved and the boosted top-decay topologies are shown in Figure 4.11.

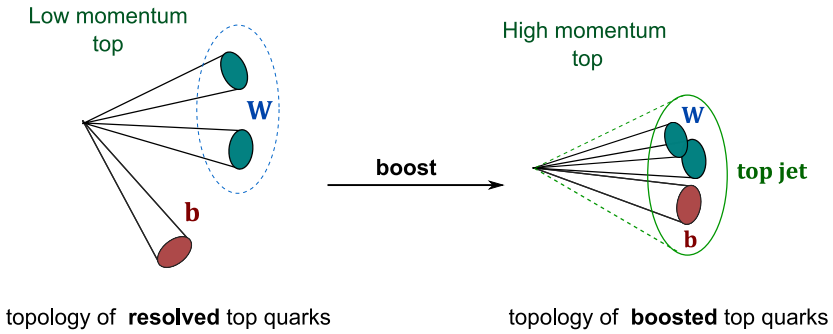


Figure 4.11: Schematic diagram showing resolved (left) and boosted (right) top quark decays. In the boosted case, all the decay products are close to each other being reconstructed as a single large-R jet in ATLAS.

Top-tagging is the method of identifying large-R jets that are coming from boosted hadronic top decay while rejecting other jets. Since a top jet is a combination of three different objects, various structures and patterns are visible inside it. Jet substructure variables are defined to probe the internal structure of the jets. These substructure variables of a top-jet are used to distinguish a top-jet from other QCD jets. Top-tagging methods use these jet substructure variables to identify a top-jet.

4.11.1 Jet substructure variables

Many jet substructure variables are used in top-tagging. Some important variables are highlighted in this section. The jet mass is an important substructure variable. Particularly, the mass of a trimmed large-R jet is less susceptible to pileup contamination. On average large-R jets containing top decay products have a higher mass than the background jets. The k_t splitting scale is another important substructure variable that comes from reversing the k_t clustering algorithm. Subjets merged in the last k_t clustering step provide the splitting scale $\sqrt{d_{12}}$. The other splitting scale, $\sqrt{d_{23}}$, corresponds to second-to-last merging. These splitting scales are proportional to jet masses relevant to those respective steps. For a hadronic top decay $\sqrt{d_{12}}$ is expected to be $m_t/2$, where m_t is the top quark mass. $\sqrt{d_{23}}$ is expected to be $m_W/2$. N-subjettiness, τ_N is a measure of how well the jet can be described with N or fewer subjets. The ratios of these variables like $\tau_{32} = \tau_3/\tau_2$ and $\tau_{21} = \tau_2/\tau_1$ acts as important substructure variables to distinguish hadronic top-jet with other jets. Another important variable is minimum pairwise invariant mass of the subjets, $Q_w = \min_{i,j}\{m_{ij}\}$, where i and j are the subjet indices. There are many other variables currently used and are all listed in reference [197].

4.11.2 Jet truth labeling

To evaluate the performance of the top taggers it is necessary to match the jets to truth top-partons. This is done in a three step process described below.

1. The reconstructed large-R jets are matched with a truth large-R jet computed using final state hadrons and electrons by requiring $\Delta R\,(j_{\text{true}}, j_{\text{reco}}) < 0.75$.

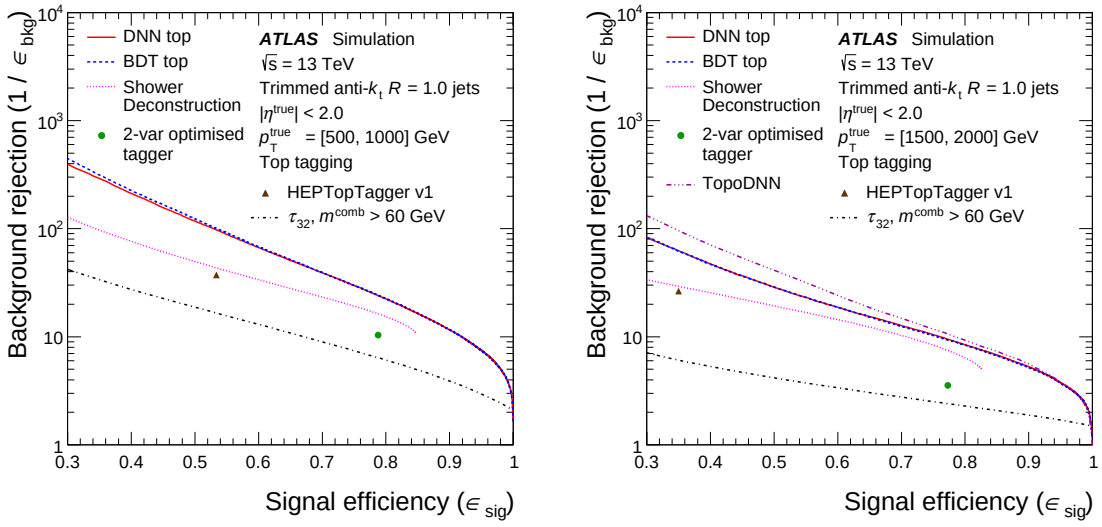
2. Those truth jets are matched with top-partons again with the same criteria $\Delta R\,(t_{\text{parton}}, j_{\text{true}}) < 0.75$.

3. Finally the decay products of top-parton needs to be within $\Delta R < 0.75 \cdot R_{\text{jet}}$ of the jet axis, where $R_{\text{jet}} = 1$ is the radius of the jet. Jets satisfying this criterion are called fully contained top-jets. This criterion is not absolutely required, an inclusive definition is also used without this requirement. These jets are also labeled as top-jets and generally called inclusive top-jets.

4.11.3 Top-tagging algorithms

Historically in ATLAS, different combinations of jet substructure variables along with the jet mass are used to identify the jets coming from top-quark decay. A DNN based top-tagging algorithm, developed in 2017, performs better than other methods. The DNN takes all the high-level substructure variables and some other jet-related variables as inputs and gives a DNN discriminant as the output. The tagger is trained on the jets obtained from simulated $Z' \to t\bar{t}$ events as signal and light-jets from other QCD processes as background. The algorithm tags the high momentum jets with p_T above 450 GeV and below 4 TeV. A jet is identified as a top-jet if the DNN discriminant value is above a certain threshold. Different working points (WP) are defined depending on the threshold values. The WPs are defined for a constant average efficiency of identifying a top jet in the simulation. Two commonly used working points are 50% and 80% WPs. In this dissertation, the DNN top-tagger with the 80% WP is used. The performance of different top-tagging algorithms is compared in Figure 4.12. The DNN top-tagger performs much better compared to the old two-variable taggers. Details about the algorithms and their performance are discussed

in reference [197]. Two different type of DNN top taggers are available which are trained with contained and inclusive top-jets, respectively. Both the taggers are explored in this dissertation.



(a) (b)

Figure 4.12: The performance comparison in simulation of different top-tagging algorithms at (a) low-p_T^{true} and (b) high-p_T^{true}. The performance of the DNN top-tagger is good in both the p_T regions. Figure taken from [197].

4.11.4 Top-tagging calibration

Top-tagging optimizations and performance studies are done on simulated events. To match up the performance on data, SFs are derived similarly to b-tagging. The SFs are applied as per-jet weights on the simulation. The calibration is done using $t\bar{t}$ events in the lepton+jets final state while γ+jet and multijet events are used to derive scale factors for quark and gluon-initiated background jets.

4.12 Missing Transverse Energy

From momentum conservation, it is expected that the vector sum of the momentum of all the final state particles produced in the proton-proton collision at the LHC should be equal to the sum of the momentum of the colliding partons. The initial momentum of colliding partons are not well measured, so it is not possible to use momentum conservation accurately in the longitudinal direction. Since the collisions are happening head-on, the initial transverse momentum can be assumed to be zero. And from the momentum conservation, we expect the final transverse momentum should add up to zero. Not all particles produced are measured detected by the ATLAS detector, for example, the neutrinos escape undetected. Thus the invisible (to ATLAS) particles create an imbalance in the energy flow. The transverse component of energy imbalance (missing energy) is called the *missing transverse energy* (MET). MET is usually parameterized by E_T (the magnitude) and the azimuthal angle, ϕ, in the transverse detector plane.

E_T^{miss} is estimated mostly using the reconstructed objects like leptons, photons, jets that have been discussed in this chapter. The second contribution to the E_T^{miss} reconstruction comes from soft particles, modeled via charged particle tracks that come from the primary vertex but are not matched with any hard objects [198]. Thus it can be written as

$$E_T^{\text{miss}} = -\sum_{\substack{\text{selected}\\ \text{electrons}}} p_T^e - \sum_{\substack{\text{accepted}\\ \text{photons}}} p_T^\gamma - \sum_{\substack{\text{accepted}\\ \tau\text{-leptons}}} p_T^{\tau_{had}} - \sum_{\substack{\text{selected}\\ \text{muons}}} p_T^\mu - \sum_{\substack{\text{accepted}\\ \text{jets}}} p_T^{\text{jet}} - \sum_{\substack{\text{unused}\\ \text{tracks}}} p_T^{\text{track}}. \tag{4.9}$$

The overlap between different objects is removed before calculating the E_T^{miss} in order to reduce double counting. MET plays an important role in several physics analyses in ATLAS and is in particular a very important handle for several BSM searches. In the dissertation, MET is used to calculate the neutrino four-momentum and to reconstruct leptonically decaying top quarks.

CHAPTER

FIVE

Statistical methods

This chapter summarizes the statistical techniques and methods used to understand and interpret the data for this dissertation. The frequentist approach, in which the probability is fundamentally associated with the frequencies of repeatable events, is used here. At the beginning of this chapter, the random variables are introduced, and then the relevant probability distributions are discussed. Then statistical modeling is briefly discussed. Finally, inference methods are discussed in detail. The likelihood-based inference is used in the searches for new particles described in this dissertation.

5.1 Random variables and probability distributions

A *random variable*, X, is a variable whose possible values are the numerical outcomes of a repeatable experiment. A *discrete random variable* takes a countable finite number of distinct values. For example, the number of defective detector modules in a box of ten modules is a discrete random variable. The probability distribution function, $p(x) = p\{X = x\}$, of a discrete random variable gives the probability of obtaining a particular value of the random variable $X = x$. The probability values should always be non-negative real numbers and they should add up to 1. The cumulative probability function, $F(x)$ can be written as

$$F(x) = p\{X \leq x\} = \sum_{x_i=0}^{x_i=x} p(x_i)\,, \tag{5.1}$$

where the random variable $X \in [0, x]$. On the other hand, a *continuous random variable* can take an infinite number of possible values. The probability of a particular outcome is always zero in this case, since infinite outcomes are possible. Instead, the probability is defined for a small range of possible outcomes. The probability of observing values in the range $[a, b]$ can be written as

$$p\{a \leq X \leq b\} = \int_a^b f(x)\, dx, \tag{5.2}$$

where $f(x)$ is the *probability density function (pdf)* which has dimension of $1/x$. The pdf should also be non-negative real numbers and is normalized to unity in the range $[-\infty, \infty]$. The cumulative probability functions (cdf) for a continuous random variable can be written as

$$F(x) = p\{X \leq x\} = \int_{-\infty}^{x} f(x)\, dx\,. \tag{5.3}$$

The expectation value of a function of a random variable, $g(X)$, is the weighted average of the function over the full x range. The expectation value of $g(X)$ can be written as

$$E\,[g(X)] = \sum_{\forall x} g(x)\, p(x)\,,$$

where X is a discrete random variable. For a continuous random variable, $E\,[g(X)]$ can be written as

$$E\,[g(X)] = \int_{-\infty}^{\infty} g(x)\, f(x)\, dx\,.$$

The mean, μ, of the random variable is given by the expectation value of X. The variance of the random variable is the expectation value of the distance to the mean, $\mathrm{Var}(X) = E\,[(X-\mu)^2]$. The standard deviation is defined as the square root of the variance. The probability distribution of two random variables, X and Y, can be described together by a joint distribution $f(x,y)$. Integrating the joint distribution over one of the variables gives the marginal distribution. Joint variability of the two random variables can be measured by the covariance. The covariance is defined as

$$V_{xy} = \mathrm{cov}(X,Y) = E\,[(X - E[X])\,(Y - E[Y])] = E\,[XY] - E\,[X]\,E\,[Y] = E\,[XY] - \mu_x\mu_y\,, \tag{5.4}$$

where μ_x and μ_y are the means of X and Y, respectively. For $X = Y$, $\mathrm{cov}(X,X) = \sigma_x^2$ where σ_x is the standard deviation of x. In multivariate analyses, understanding the relationship between two variables often gives useful insight. The correlation of two random variables can be measured by the correlation coefficient (ρ) defined as

$$\text{Correlation} = \rho = \frac{\mathrm{cov}(X,Y)}{\sigma_x\ \sigma_y}\,. \tag{5.5}$$

If there are more than two variables, $X_1, X_2, \ldots, X_N$, the covariance and correlation matrices can be defined as

$$V_{ij} = \mathrm{cov}(X_i, X_j) = E[X_i X_j] - E[X_i]E[X_j], \tag{5.6}$$

$$\rho_{ij} = \frac{V_{ij}}{\sigma_i\ \sigma_j}. \tag{5.7}$$

5.1.1 Poisson distribution

The Poisson distribution function is defined for a discrete random variable and the probability mass function is given by

$$f(x;\lambda) = \frac{e^{-\lambda}\lambda^x}{x!}, \quad \text{for } \ x = 0, 1, 2, \ldots \ \text{ and } \ \lambda > 0. \tag{5.8}$$

If X is a random variable which represents the number of independent events occurring in a fixed time, then X follows a Poisson distribution. The only parameter of the Poisson distribution, λ, represents the average number of events occurring within the time interval. λ is also the variance

of the Poisson distribution. Poisson distribution is a count distribution; it is used in this dissertation in counting events in the histogram bins of the kinematic variables. Each bin of a kinematic quantity can be thought of as an independent counting experiment, so the number of events in a histogram bin is assumed to be Poisson-distributed. The means of these independent Poisson-distributed random variables are estimated from the simulated or data-driven background distributions. Finally, a bin content observed in data is assumed to be a single draw from the same Poisson-distributed random variable.

5.1.2 Gaussian distribution

The Gaussian, or normal distribution, is a very widely used probability distribution. One of the major reasons for this is the Central Limit Theorem. The Gaussian distribution function is defined for a continuous random variable, and is given by

$$f(x;\mu,\sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-(x-\mu)^2}{2\sigma^2}}. \tag{5.9}$$

A Gaussian distribution has two parameters, the mean (μ) and the standard deviation (σ). The Gaussian distribution is an important distribution and used in many studies presented here. As an example, the parameters corresponding to the systematic uncertainties presented in this dissertation are assumed to be Gaussian distributed. Standard normal random variables follow Gaussian distributions with $\mu = 0$ and $\sigma = 1$. If X follows a Gaussian distribution with mean μ and σ, then $Z = (X - \mu)/\sigma$ follows the standard normal distribution.

5.1.3 Chi-squared (χ^2) distribution

If $Y_1, Y_2, \ldots Y_n$ are independent standard normal random variables then

$$x = \sum_{i=1}^{n} Y_i^2 \tag{5.10}$$

follows a χ^2 distribution function with n degrees of freedom, denoted as χ^2_n. The chi-squared distribution is defined as

$$f(x;n) = \frac{x^{\frac{n}{2}-1} e^{\frac{x}{2}}}{2^{\frac{n}{2}}\Gamma\left(\frac{n}{2}\right)}. \tag{5.11}$$

The gamma function $\Gamma(n) = (n-1)!$ for integer $n > 0$. For $n = 1$, the chi-squared distribution can be written as

$$\chi^2_1(x) = \frac{1}{\sqrt{2\pi}} \frac{1}{\sqrt{x}} e^{-\frac{x}{2}}. \tag{5.12}$$

The chi-squared value is used as a measure of goodness of fit in this dissertation, where Y = (data - fit).

5.1.4 Log-Normal distribution

If the logarithm of a random variable, X, follows a Gaussian distribution with mean μ and standard deviation σ, then X follows a log-normal distribution. The distribution function can be written as

$$f(x;\mu,\sigma) = \frac{1}{x} \frac{1}{\sigma\sqrt{2\pi}} e^{-\left[\frac{(\ln x-\mu)^2}{2\sigma^2}\right]} \tag{5.13}$$

It is a continuous probability distribution function. The mean of the log-normal distribution is $\exp\left(\mu + \frac{\sigma^2}{2}\right)$ and the variance can be written as $\left(\exp(\sigma^2) - 1\right)\exp\left(e\mu + \sigma^2\right)$.

Any variable that is the product of a large number of random factors where none of the factors dominates can be described by the log-normal distribution. A variable which is a product of a large number of independent, identically-distributed variables follows a log-normal distribution based on the central limit theorem in the log domain.

For example, the signal registered by an electron in the calorimeter may be described by a log-normal distribution, since an electron loses its energy in many ways: to dead material, neutron production, lost photons, and so on.

5.2 Likelihood function

The probability distributions discussed in section 5.1 are used to model the data. Let us suppose we have N different measurements, $x = \{x_1, x_2, \ldots, x_N\}$, where the measurements are statistically independent. Each x_i follows the probability distribution $f(x;\boldsymbol{\alpha})$, where $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \ldots, \alpha_m)$ are the m parameters of the statistical model. These parameters could be the parameters of the probability distributions as well as some other additional free parameters used to model the data. If we consider $f(x;\boldsymbol{\alpha})$ as a function of x_i for specified $\boldsymbol{\alpha}$ it gives the probability of all possible samples determined by the specified $\boldsymbol{\alpha}$. This is relevant before the experimental results are unblinded or before an experiment takes data, i.e. where there is no observation. But it does not allow us to learn anything new about the real world or the model. The other way is to consider $f(x;\boldsymbol{\alpha})$ as a function of $\boldsymbol{\alpha}$ for an observed x_i. Note, $f(x;\boldsymbol{\alpha})$ is not a probability density function of $\boldsymbol{\alpha}$ as it does not normalize to unity. This way of expressing $f(x_i;\boldsymbol{\alpha})$ is called likelihood. The likelihood function constructed with all the measurements is numerically equivalent to the joint distribution with specified values of x and can be written as

$$L(\boldsymbol{\theta}) = \prod_{i=1}^{N} f(x_i;\boldsymbol{\alpha})\,. \tag{5.14}$$

Likelihood-based inference enables us to determine the model parameters using the observed data in the experiment. In this dissertation, we always used likelihood-based statistical inference. The measurements are always binned into histograms. Suppose we are measuring a variable x. The event counts in each bin are considered as independent counting experiments. Let's denote them by $\nu = \{\nu_1, \nu_2, \ldots, \nu_N\}$, for a N bin measurement, and denote the expected number of SM background events in those N bins as $b = \{b_1, b_2, \ldots, b_N\}$. Similarly, the expected BSM signal events can be written as $s = \{s_1, s_2, \ldots, s_N\}$. Since data can have contributions from both background and signal, the expected value ν_i in the i^{th} bin can be written as

$$\nu_i = b_i + \mu s_i\,, \tag{5.15}$$

where, the signal strength, $\mu = \sigma_{\text{obs}}/\sigma_{\text{th}}$, is given by the ratio of a measured cross-section and the theory cross-section. Hence, μ is one of the likelihood parameters used in this dissertation. So, in this case $\boldsymbol{\alpha} = \{\mu\}$. Now, the probability of finding n_i data events in a bin where $\nu_i(\mu)$ events are expected can be written by the Poisson probability. Thus the likelihood of μ with N measurements, $n = \{n_1, n_2, \ldots, n_N\}$, can be written as

$$L(\mu) = p(n|\mu) = \prod_{i=1}^{N} \text{Poisson}(n_i|\nu_i(\mu))\,, \tag{5.16}$$

where, p denotes the probability distribution function. In our statistical analysis, μ is the main parameter of interest (POI).

5.2.1 Systematic uncertainties

Typically, the expected number of counts in a bin is affected by different sources of systematic uncertainty. To model the effect of the systematic uncertainty, the expected number of background and signal counts are parametrized with additional parameters, $\boldsymbol{\theta} = \{\theta_1, \theta_2, \ldots, \theta_m\}$. So, we rewrite Equation 5.15 as

$$\nu_i(\mu, \boldsymbol{\theta}) = b_i(\boldsymbol{\theta}) + \mu s_i(\boldsymbol{\theta}) \,. \tag{5.17}$$

With this new parametrization the likelihood becomes a function of these additional parameter, $\boldsymbol{\theta}$. We do not particularly want to measure these parameters, so there are called the *nuisance parameters*. Since these nuisance parameters are part of the likelihood, they have an impact on the estimated uncertainty of the POI. To reduce the parameter space and constrain the nuisance parameters, *auxiliary measurements*, y, are needed. One can build a joint model by using the auxiliary measurements and the combined likelihood can be written as

$$L(\mu, \boldsymbol{\theta}) = p(n|\mu; \boldsymbol{\theta}) \cdot p(y|\boldsymbol{\theta}) \,. \tag{5.18}$$

Here it is assumed that the auxiliary measurements follow the probability distribution $p(y|\boldsymbol{\theta})$. The auxiliary measurements could be measurements in a background region where they are expected to be Poisson-distributed over the bins. Often the auxiliary measurements for the systematic uncertainties are approximated by Gaussian constraints. Sometimes Gamma and log-normal distributions are also used to constrain the nuisance parameters. In our studies, the nuisance parameters are assumed to follow Gaussian distributions with mean θ_k^0 and standard deviation σ_{θ_k}, where σ_{θ_k} corresponds to a 1σ variation of the k^{th} systematics. The final likelihood is written as:

$$L(\mu, \boldsymbol{\theta}) = p(n|\mu; \boldsymbol{\theta}) = \prod_{i=1}^{N} \text{Poisson}(n_i \mid \nu_i(\mu, \boldsymbol{\theta})) \prod_{k \in \text{syst}} \text{Gaus}(\theta_k \mid \theta_k^0; \sigma_{\theta_k}) \tag{5.19}$$

The effect of limited statistics on simulation or background estimation methods is almost always included in the likelihood with a set of additional parameters for each bin. Those parameter variations are modelled with Gamma functions and are generally denoted as γ-stat.

5.3 Statistical inference

The inference methods used in the statistical analysis are summarized in this section. The two main inferences methods used in this thesis are parameter estimation and hypothesis testing.

5.3.1 Parameter estimation

Two popular methods of parameter estimation are *maximum likelihood* and *least square*. Maximum likelihood estimation is used in this dissertation. The maximum-likelihood estimator (MLE) of the model parameter $\boldsymbol{\theta} = \{\theta_1, \theta_2, \ldots, \theta_m\}$ are referred to as $\hat{\boldsymbol{\theta}}$. They are obtained by finding the global maximum of the likelihood function $L(\boldsymbol{\theta})$. For convenience and numerical stability, we use

the logarithm of the likelihood function, called the *log-likelihood*, and find the minimum of the negative log-likelihood defined as

$$-\ln L(\boldsymbol{\theta};x) = -\sum_{i=i}^{N} \ln f(x_i;\boldsymbol{\theta}). \tag{5.20}$$

The following equation is solved to find the MLEs of the model parameters

$$-\frac{\partial \ln L(\boldsymbol{\theta};x)}{\partial \theta_j} = 0\,, \quad \text{for} \quad j = 1,2,\ldots,m. \tag{5.21}$$

The estimated parameters are denoted as $\hat{\boldsymbol{\theta}}$ and they are asymptotically normally distributed. The covariance matrix of the estimated parameters, $V(\hat{\boldsymbol{\theta}})$, is asymptotically equal to the inverse of the observed information matrix and can be written as

$$\hat{V}_{ij}(\hat{\boldsymbol{\theta}}) = \left[\mathbf{I}(\hat{\theta})\right]^{-1} = -\left[\left.\frac{\partial^2 \ln L(\boldsymbol{\theta};x)}{\partial\theta_i \partial\theta_j}\right|_{\theta=\hat{\theta}}\right]^{-1}. \tag{5.22}$$

The matrix of the second derivative of the negative log-likelihood function is also called the Hessian matrix, $\mathbf{H}$ of the negative log-likelihood function. The standard deviation of a parameter θ_j can be estimated from the diagonal elements of this matrix and can be written as

$$\hat{\sigma}_{\hat{\theta}_j} = \sqrt{\hat{V}_{jj}(\hat{\boldsymbol{\theta}})}\,. \tag{5.23}$$

The solutions to Equation 5.21 and Equation 5.22 are found numerically using ROOT and MINUIT [199] software package in this dissertation.

Another way to estimate the uncertainty on a model parameter is to look at the change of the log-likelihood. We can find two points $\theta_{\text{low}} = \hat{\theta} - \Delta\hat{\theta}_-$ and $\theta_{\text{up}} = \hat{\theta} + \Delta\hat{\theta}_+$ such that $\Delta \ln L = -0.5$. $[\theta_{\text{low}}, \theta_{\text{up}}]$ will be the 68% confidence interval for θ. This method is not used in any of the studies presented in this dissertation.

5.3.2 Hypothesis testing

Hypothesis testing is used to compare different hypotheses. In this dissertation, only models with one parameter of interest (POI) μ and a set of nuisance parameters θ. They can be denoted together by $\alpha = (\mu, \theta)$, are studied. The POI is used to distinguish the hypotheses used in this dissertation. The null hypothesis, H_0, corresponds to $\mu = 0$ (SM prediction) while the alternative hypothesis is a simple hypothesis where $\mu > 0$ and can take a positive value (BSM prediction for our studies). In order to support or reject H_0, we need to define a test statistic. The test statistic is a lower dimensional summary of the data with discriminatory power. The Neyman-Pearson lemma [200] says the likelihood ratio

$$\lambda(x) = \frac{L(H_0|x)}{L(H_1|x)} \tag{5.24}$$

maximizes the statistical power to reject H_0 in favour of H_1, where x denotes observed data. Thus the likelihood ratio can be considered to be the most powerful test statistic to search for new physics. Test statistics in general are denoted as t or q in this chapter. The test statistic can take

values in the range $[0,1]$. In order to evaluate the level of agreement between data and the null hypothesis, the p-value is defined as

$$p = \int_{t_{\text{obs}}}^{\infty} f(t|H_0)\, dt, \tag{5.25}$$

where t_{obs} is the test statistic calculated from data and $f(t|H_0)$ is the pdf of the test statistic. In general it is not easy to find a functional form of the test statistic distribution. Figure 5.1a shows an example test statistic distribution and the shaded area shows the p-value of the observation. In particle physics, the p-values are generally expressed in terms of significance, Z, defined as

$$Z = \Phi^{-1}(1-p)\,, \tag{5.26}$$

where Φ^{-1} is the inverse cumulative function of a unit Gaussian pdf with mean zero and unit standard deviation. The Z values are defined from one-sided tail probabilities of a unit Gaussian distribution. Figure 5.1b illustrates the relationship between Z and the p-value. p-values can also be expressed in terms of Z values using the following expression

$$p = \frac{1 - \text{erf}(Z/\sqrt{2})}{2}\,, \tag{5.27}$$

where erf denotes the error function. It is defined as

$$\text{erf}\, z = \frac{2}{\pi}\int_0^z e^{-t^2} dt\,. \tag{5.28}$$

In the particle physics community it is commonly accepted that in order to claim discovery of a new particle a significance of at least $Z = 5$ (generally referred as 5σ) is required to reject the background hypothesis (H_0) . The corresponding p-value is $p = 2.87 \times 10^{-7}$. On the other hand, a threshold of $p = 0.05$ (corresponds to 95% confidence level) is used for the purpose of excluding a given signal hypothesis (H_1); the corresponding Z value is 1.64. In the studies presented in this dissertation, the hypothesis is always defined based on the value of the POI, which is the signal strength (μ).

5.3.3 Profile likelihood ratio

As described in the previous section, the likelihood ratio is the most powerful test statistic for rejecting the null hypothesis. The likelihood ratio can be expressed as a function of likelihood parameters $\boldsymbol{\alpha}$ as

$$\lambda(\boldsymbol{\alpha}) = \frac{L(\boldsymbol{\alpha}; x)}{L(\hat{\boldsymbol{\alpha}}; x)}\,, \tag{5.29}$$

where x represents the data and $\hat{\boldsymbol{\alpha}}$ are the MLEs. In the large sample limit, when the likelihood approaches a Gaussian, according to Wilk's theorem [201] $-2\ln\lambda(\boldsymbol{\theta})$ follows a chi-squared distribution (χ^2) with m degrees of freedom, where m corresponds to the number of parameters of the model. In this limit the quantiles $\chi^2_{1-\alpha}$ of the χ^2 distribution can be used to evaluate $100\cdot(1-\alpha)\%$ confidence intervals.

In the case where the likelihood depends on many parameters, but one is interested only in one parameter μ, and its uncertainty, one can use a *profile likelihood ratio* defined as:

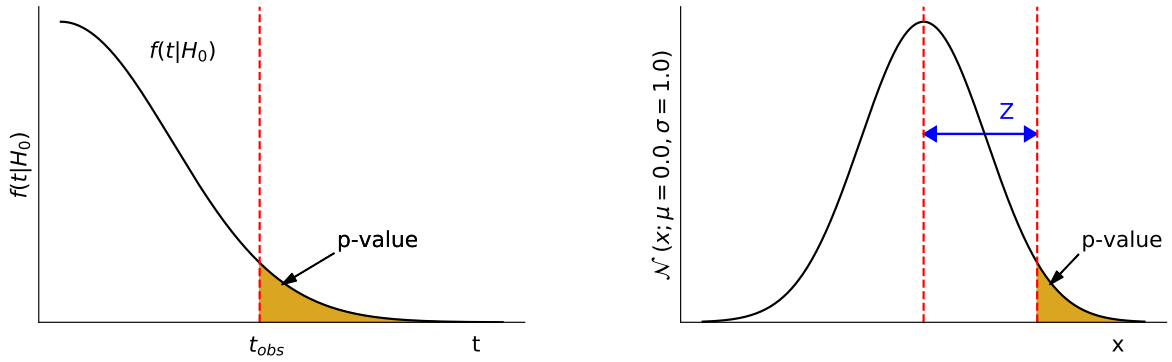


Figure 5.1: (a) Demonstration of how to calculate the p-value after knowing the observed test statistic value, t_{obs}. The p-value is the integral of the right-hand tail of the test statistic distribution. (b) The relationship between the p-value and the significance Z is shown using a standard normal distribution.

$$\lambda(\mu) = \frac{L(\mu, \hat{\hat{\theta}}_\mu)}{L(\hat{\mu}, \hat{\theta})} \tag{5.30}$$

where $\hat{\hat{\theta}}_\mu$ in the numerator is the conditional MLE of θ that maximizes L for a given value of μ. The denominator is the unconditional maximized likelihood function, i.e., $\hat{\mu}$ and $\hat{\theta}$ maximize the likelihood, L. The profile likelihood ratio also follows a χ^2 distribution in the limit of high statistics. The likelihood ratio is used for some studies presented in this dissertation, and the profile likelihood ratio is always used in the hypothesis test done for significance and limit calculation.

5.3.4 Test statistic

The profile likelihood ratio is used as a test statistic in our statistical analysis. For convenience, we use the log of the profile likelihood ratio, and it is multiplied by negative 2 by convention. So the exact test statistic is defined as,

$$t_\mu = -2 \ln \lambda(\mu) \,. \tag{5.31}$$

The profile likelihood ratio takes values between 0 and 1. Higher values of t_μ indicate larger disagreement between data and μ. To quantify the level of disagreement, the p-value can be computed. Most of the discussions related to the test statistics are adopted from [202].

Discovery test statistic

To test for the discovery of a new signal, the background-only hypothesis with $\mu = 0$ is used. The one-sided test statistic used is

$$q_0 = \begin{cases} t_0 = -2 \ln \lambda(0) & \hat{\mu} \geq 0 \\ 0 & \hat{\mu} < 0 \,. \end{cases} \tag{5.32}$$

Here $\lambda(0)$ is the profile likelihood ratio as defined in Equation 5.30 computed for $\mu = 0$. It is used to reject the $\mu = 0$ hypothesis against a positive signal hypothesis, $\mu \geq 0$. The discovery significance can be given by the p-value computed in a similar manner as Equation 5.25

$$p_0 = \int_{q_{0,\,\mathrm{obs}}}^{\infty} f(q_0|0)\, dq_0 \,. \tag{5.33}$$

$f(q_0|0)$ is the pdf of q_0 under the background-only ($\mu = 0$) hypothesis.

Test statistic for upper limit

For setting an upper limit on the signal strength parameter, μ, the hypothesis test is inverted. The signal plus background hypothesis is tested against the background-only hypothesis (H_{s+b} or $H_{\mu=1}$). In the case where the data is not consistent with the H_{s+b} hypothesis, we do not reject it, rather we set an upper limit on the signal strength. The test statistic is defined as:

$$q_\mu = \begin{cases} t_\mu = -2\ln\lambda(\mu) & \hat{\mu} \leq \mu \\ 0 & \hat{\mu} > \mu \,. \end{cases} \tag{5.34}$$

In this case the p-value (p_μ) can be calculated in a similar fashion :

$$p_\mu = \int_{q_{\mu,\mathrm{obs}}}^{\infty} f(q_\mu \mid \mu) dt_\mu \,, \tag{5.35}$$

where $q_{\mu,\mathrm{obs}}$ is the value of the test statistic q_μ obtained from data and $f(q_\mu \mid \mu)$ denotes the pdf of q_μ under the hypothesis of signal strength μ. The upper limit on μ is the largest μ that still satisfies $p_\mu \leq \alpha$, where α is the specified threshold ($\alpha = 0.05$ for 95% CL).

5.3.5 Asymptotic approximation

As shown in the last section, to calculate the p-values, a test statistic pdf is needed and it is not trivial to model the distributions. The pdf $f(t_\mu|H_\mu)$ is estimated by repeatedly generating expected datasets according to a particular hypothesis, H_μ (referred to as toys or pseudoexperiments). Generating toys is computationally very expensive. To overcome this problem, an approximation for the sampling distribution provided by [202] is used in this dissertation. The method is briefly summarized in this section.

Let us consider a test of the signal strength parameter μ, and data distributed according to a strength parameter μ'. If the estimated signal strength $\hat{\mu}$ follows a Gaussian distribution centered around μ' with a standard deviation σ, then based on Wald's approximation [203], it can be shown that for a single parameter of interest

$$-2\ln\lambda(\mu) = \frac{(\mu - \hat{\mu})^2}{\sigma^2} + \mathcal{O}(1/\sqrt{N}) \,, \tag{5.36}$$

where N is the sample size and μ is the hypothesized value of the signal strength parameter. For large N the higher order terms, $\mathcal{O}(1/\sqrt{N})$, can be ignored and it can be shown that the test statistic $t_\mu = -2\ln\lambda(\mu)$ follows a *non-central chi-square* distribution with one degree of freedom,

$$f(t_\mu;\Lambda) = \frac{1}{2\sqrt{t_\mu}} \frac{1}{\sqrt{2\pi}} \left[\exp\left(-\frac{1}{2}\left(\sqrt{t_\mu} + \sqrt{\Lambda} \right)^2 \right) + \exp\left(-\frac{1}{2}\left(\sqrt{t_\mu} - \sqrt{\Lambda} \right)^2 \right) \right], \tag{5.37}$$

where the non-centrality parameter Λ is

$$\Lambda = \frac{(\mu - \mu')^2}{\sigma^2} \,. \tag{5.38}$$

It can be denoted as $f(t_\mu;\Lambda) \approx \frac{1}{2}\chi^2_1(\Lambda)$. For an important special case where $\mu' = \mu$, $\Lambda = 0$ and the test statistic, $-2\ln\lambda(\mu)$, follows a chi-squared distribution with one degree of freedom, $f(t_\mu;0) \approx \frac{1}{2}\chi^2(0)$. This result was first shown by Wilks [201]. Figure 5.2 shows the test statistic distribution of a toy example. In the toy example a two bin distribution of data, background and signal is assumed. Here $\mu' = \mu = 1.0$ was chosen, and the test statistic is created using the $H_{\mu=1}$ hypothesis, also denoted as H_{s+b}. Figure 5.2 shows that the test statistic distribution follows a χ^2 distribution with one degree of freedom.

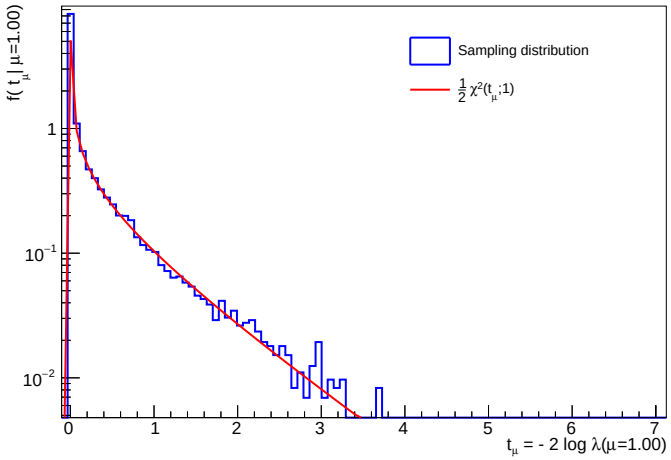


Figure 5.2: Test statistic distribution created from toy data containing two bins for signal, background and data. The sampling distribution shown in the blue histogram, is created with hypothesis H_{s+b}. Since $\mu' = \mu = 1.0$, the sampling distribution follows a χ^2 distribution with one degree of freedom shown in the red curve.

Discovery test statistic

Assuming the validity of the asymptotic approximation (Equation 5.36), the discovery test statistic can be written as

$$q_0 = \begin{cases} \frac{\hat{\mu}^2}{\sigma^2} & \hat{\mu} \geq 0 \,, \\ 0 & \hat{\mu} < 0 \,, \end{cases} \tag{5.39}$$

where $\hat{\mu}$ follows a Gaussian distribution with mean μ' and standard deviation σ. The pdf of q_0 takes the following form

$$f(q_0|\mu') = \left(1 - \Phi\left(\frac{\mu'}{\sigma}\right)\right)\delta(q_0) + \frac{1}{2}\frac{1}{\sqrt{2\pi}}\frac{1}{\sqrt{q_0}}\exp\left[-\frac{1}{2}\left(\sqrt{q_0} - \frac{\mu'}{\sigma}\right)^2\right] . \tag{5.40}$$

For the special case of $\mu' = 0$, it reduces to the half chi-squared distribution $\frac{1}{2}\chi^2_1$ with one degree of freedom

$$f(q_0|0) = \frac{1}{2}\delta(q_0) + \frac{1}{2}\frac{1}{\sqrt{2\pi}}\frac{1}{\sqrt{q_0}}e^{-\frac{q_0}{2}} \,. \tag{5.41}$$

Under the asymptotic approximation the significance can be written with a very simple expression

$$Z_0 = \Phi^{-1}(1 - p_0) = \sqrt{q_0}\,. \tag{5.42}$$

Test statistic for upper limit

Assuming the validity of the asymptotic approximation (Equation 5.36), the test statistic used for setting an upper limit can be written as

$$q_\mu = \begin{cases} \Lambda = \frac{(\mu-\hat{\mu})^2}{\sigma^2} & \hat{\mu} \le \mu\,, \\ 0 & \hat{\mu} > \mu\,, \end{cases} \tag{5.43}$$

where $\hat{\mu}$ follows a Gaussian distribution with mean μ' and standard deviation σ. The pdf of $f(q_\mu|\mu')$ takes the following form

$$f(q_\mu|\mu') = \Phi\left(\frac{\mu'-\mu}{\sigma}\right)\delta(q_\mu) + \frac{1}{2}\frac{1}{\sqrt{2\pi}}\frac{1}{\sqrt{q_\mu}}\exp\left[-\frac{1}{2}\left(\sqrt{q_\mu} - \frac{\mu-\mu'}{\sigma}\right)^2\right]. \tag{5.44}$$

For the special case of $\mu' = \mu$, it reduces to the half chi-squared distribution,

$$f(q_\mu|\mu) = \frac{1}{2}\delta(q_\mu) + \frac{1}{2}\frac{1}{\sqrt{2\pi}}\frac{1}{\sqrt{q_\mu}}e^{-\frac{q_\mu}{2}}\,. \tag{5.45}$$

The corresponding significance for the hypothesized signal strength, μ, is

$$Z_\mu = \Phi^{-1}(1 - p_\mu) = \sqrt{q_\mu}\,. \tag{5.46}$$

If the p-value is below a specified threshold, α ($\alpha = 0.05$ for the 95% confidence level), the upper limit on μ can be written as

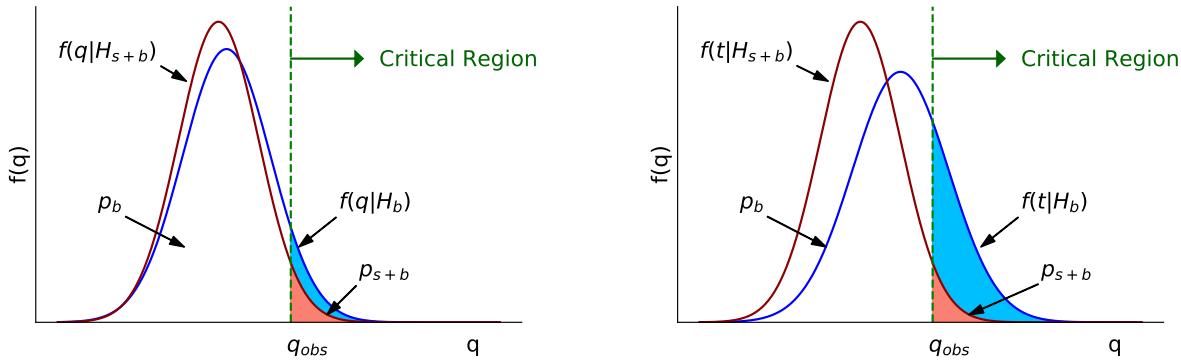
$$\mu_{\text{up}} = \hat{\mu} + \sigma\Phi^{-1}(1 - \alpha)\,. \tag{5.47}$$

5.3.6 CL_s method

In the standard hypothesis testing for estimating the upper limit, we test the $s+b$ hypothesis H_{s+b} against the background-only hypothesis H_b . The p-value, p_{s+b}, is defined as the probability of finding a value q equally or less compatible with H_{s+b} relative to what is obtained in data, q_{obs}. In the standard "CL_{s+b}" method, we carry out the test based on p_{s+b}. H_{s+b} is excluded at a confidence level $(1-\alpha)$ if we find $p_{s+b} < \alpha$.

The problem with only considering the $s+b$ p-value, p_{s+b}, is that one can exclude a signal hypothesis with probability close to α even if the sensitivity to that signal model is very low. Such a scenario is illustrated in the schematic diagram Figure 5.3b where the q distribution under b and $s+b$ hypothesis almost overlap each other. But based on the p_{s+b}, one might exclude the $s+b$ hypothesis. In general, if the expected number of signal events (s) are very low compared to the background events b and the observed data has downward fluctuation, H_{s+b} will be excluded. Intuitively, in the limit where $s << b$, $H_{s+b} \sim H_b$, hence the exclusion probability should go to zero. But in CL_{s+b} method it approaches $\alpha \cdot 100\%$.

Here, one can also build the test statistic distribution under the b hypothesis, H_b, and calculate the corresponding p-value p_b. Figure 5.3a shows a schematic diagram showing the distributions



(a) (b)

Figure 5.3: (a) Distributions of the test statistic q under the $s+b$ and b hypotheses in an example where sensitivity to the signal model is very small. (b) Schematic diagram showing the test statistic distribution under b and $s+b$ hypotheses. The observed test statistic value is indicated by the dotted line. The s+b p-value is shown by the red filled area, whereas the area filled with skyblue shows (1-p_b).

and the p-values. In order to protect against this situation, most modern particle physics experiments use the "CL_s" method [204]. Instead of using p_{s+b} directly, we use "CL_s" which is defined as

$$\mathrm{CLs} = \frac{p_{s+b}}{1-p_b}\,. \tag{5.48}$$

Essentially the p_{s+b}-value is penalized by dividing it by the power of the test, $1-p_b$. If the test statistic distributions under H_{s+b} and H_b are well separated, then $(1-p_b) \gg p_{s+b}$ and the exclusion based on CL_s will be very similar to the exclusion based on p_{s+b}. However, if the two distributions almost overlap on each other, then $(1-p_b) \approx p_{s+b}$. As a result, CL_s becomes approximately equal to one, and the ability to exclude the $s+b$ hypothesis gets reduced. The exclusion limit on the signal strength (μ) is set with 95% confidence level by finding the μ for which $CL_s < \alpha$, where $\alpha = 0.05$.

5.3.7 Asimov data

The asymptotic approximation discussed in subsection 5.3.5 is valid in the large sample limit. To determine λ appearing in Equation 5.37, we need to estimate σ, the standard deviation of the $\hat{\mu}$ distribution. It can be estimated using a special artificial dataset called the Asimov dataset. The Asimov dataset is defined in such a way that estimators of all the parameters $\alpha = \{\mu, \theta\}$ evaluated using this dataset are equal to the true or expected parameter values [202]. In practice, the Asimov dataset is defined as the one in which the bin content for each bin of a distribution is set to simulated values and the uncertainty is set to the Poisson or Gaussian error depending on the expected number of events. In this dissertation, the Asimov dataset is created from the MC simulations of signal and background processes, and data-driven background estimates. If a signal strength μ' is used for the generation of the Asimov dataset, then the corresponding profile likelihood ratio,

$\lambda_A(\mu)$, is

$$\lambda_A(\mu) = \frac{L_A(\mu, \hat{\hat{\boldsymbol{\theta}}}_\mu)}{L_A(\hat{\mu}, \hat{\boldsymbol{\theta}})} = \frac{L_A(\mu, \hat{\hat{\boldsymbol{\theta}}}_\mu)}{L_A(\mu', \boldsymbol{\theta})} \,. \tag{5.49}$$

L_A is the Asimov likelihood and $\boldsymbol{\theta}$ are the true or expected values of the parameters. Since the Asimov dataset corresponds to a signal strength $\hat{\mu} = \mu'$, Equation 5.36 becomes

$$q_{\mu,A} = -2 \ln \lambda_A(\mu) \approx \frac{(\mu - \mu')^2}{\sigma_A^2} = \Lambda \,. \tag{5.50}$$

Then the variance can be approximated as

$$\sigma_A^2 = \frac{(\mu - \mu')^2}{q_{\mu,A}} \approx \frac{(\mu - \mu')^2}{\Lambda} \,. \tag{5.51}$$

It is important to note that using the Asimov dataset we get an estimate of the noncentrality parameter Λ that characterizes the $f(q_\mu|\mu')$ distribution. The test statistics q_0 and q_μ are monotonically related to $\hat{\mu}$. Therefore the median of the test statistic distributions can be found using the median of $\hat{\mu}$, which is μ'. If σ_A is used then the median of $f(q_\mu|\mu')$ is $q_{\mu,A}$. Furthermore, assuming a signal strength μ', the discovery significance can be given by $\text{med}[Z_0|\mu'] = \sqrt{q_{0,A}}$. Similarly the median exclusion significance assuming $\mu' = 0$ can be written as $\text{med}\left[Z_\mu|0\right] = \sqrt{q_{\mu,A}}$.

5.3.8 Toy example

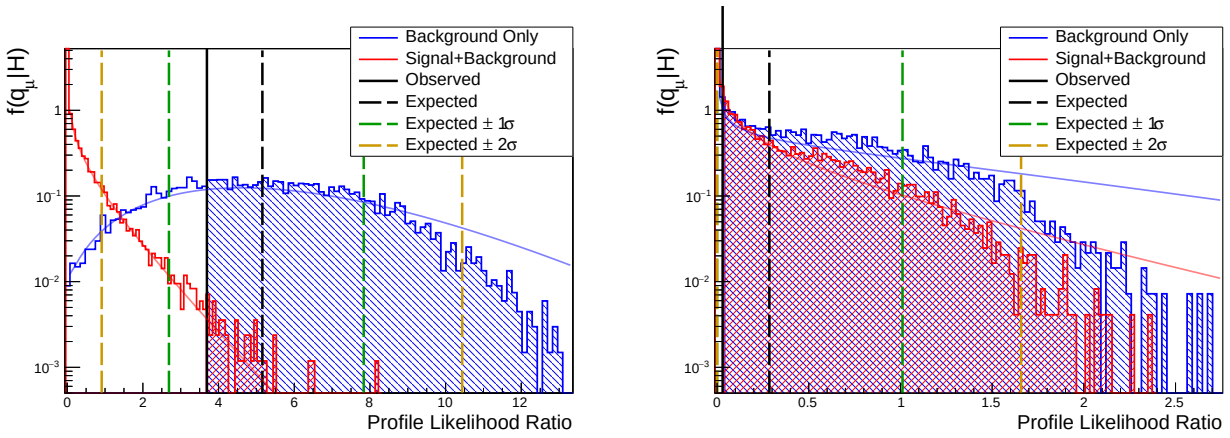
To give a visualization of the hypothesis testing performed in this dissertation, a simple Poisson model is created as a toy model. The background event count (b) is set to 10, with a 20% uncertainty assumed to be Gaussian distributed. The data count (d) is set to 12. The parameter of interest is the signal count, denoted as $\mu \cdot s$. An inverted hypothesis test is performed here to set an upper limit on the signal count. The test statistic distributions for b-only and $s+b$ hypotheses are shown in Figure 5.4a. The blue histogram is the sampling distribution of the test statistic under the b-only hypothesis, $f(q_\mu|0)$. In this case, the hypothesized signal count ($\mu \cdot s$) is 15. The x axis is the negative log profile likelihood ratio (q_μ). The blue curve is drawn using the χ^2 distribution given by Equation 5.44 with $\mu' = 0$ and $\mu = 15$ in the asymptotic limit. The red histogram is the sampling histogram of the test $s+b$ statistic distribution under the $s+b$ hypothesis, $f(q_\mu|\mu)$. If the asymptotic approximation is valid, the distribution should follow the χ^2_1 distribution with one degree of freedom, shown by the red curve. While the majority of the distribution is well modeled by the asymptotic approximation, it is observed that the asymptotic formula overshoots the higher tail of the q_μ, where the Wald approximation does not hold. The use of the asymptotic approximation results in higher values of p_{s+b} than the true distribution and usually results in a conservative limit on the signal count. The observed value of the test statistic obtained from the data, q_{obs}, is shown by the solid black vertical line. The observed test statistic is compared with the expected test statistic obtained from the median of q_μ under the background-only hypothesis shown by the dashed black vertical line. To know how much the data distribution should vary from the nominal background-only prediction due to statistical fluctuations and systematic variations, the $\pm 1\sigma$ (dashed green vertical line) and $\pm 2\sigma$ (dashed orange vertical line) quantiles are drawn around the expected median. Figure 5.4b shows a similar distribution where the hypothesised signal count $\mu \cdot s = 3$. The test-statistic distributions under the b-only and $s+b$ hypotheses look very similar, and the chi-squared approximation does not hold throughout. In this case, the negative 1σ and 2σ

quantiles are impossible to calculate from the inverse cumulative distribution since we are doing a one-sided test here. And q_μ cannot take values below zero. The $s+b$ p-value, p_{s+b}, is shown by the red shaded area in the distributions of Figure 5.4. The blue shaded area on the same plot shows (1-p_b), where p_b is the p-value under the background-only hypothesis. So, the CL_s value is the ratio of the red shaded area and the blue shaded area shown in the distributions of Figure 5.4.

Limit calculation from toy model

The toy model described in subsection 5.3.8 is also used to calculate exclusion limits. The limit setting is done using the CL_s method described in subsection 5.3.6 with 95% confidence level ($\alpha = 0.05$). To calculate the exclusion limits, the POI needs to vary until the CL_s becomes 0.05. In this toy example, a scan over the signal count ($\mu \cdot s$) is done in the range $[0, 21]$ in 21 points. For each scan, the test statistic distribution, similar to Figure 5.4, is estimated. The observed and expected p-values are calculated, including their $\pm 1\sigma$ and $\pm 2\sigma$ values. p-values are calculated both with the pseudoexperiements as well as using the asymptotic formula for all 21 points.

Figure 5.5a shows the CL_s scan using the asymptotic approximation, whereas Figure 5.5b shows a similar scan with pseudoexperiements. CL_b on these plots represents the background only p-value, p_b. The upper limit on the signal strength is obtained from the crossing point of the horizontal red line (corresponding to p-value = 0.05) and the relevant CL_s curve. The median expected CL_s values are shown by the black dotted curve. The green band shows the expected $\pm 1\sigma$ limit, and the yellow band shows expected $\pm 2\sigma$ limit. The positive 1σ and 2σ test statistic values highlighted in Figure 5.4 give stronger limits on the signal counts, hence they represent the negative bands in the two plots of Figure 5.5. Similarly, the negative 1σ and 2σ test statistic

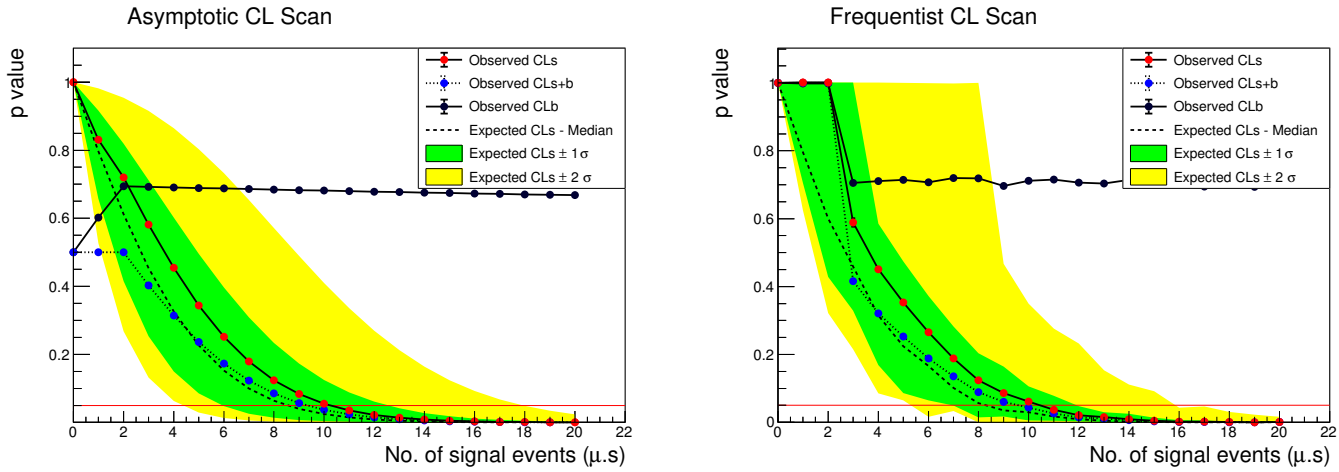


(a) $\mu \cdot s = 15$ (b) $\mu \cdot s = 3$

Figure 5.4: Test statistic distribution for upper limit based on a toy Poisson counting model. Background and data counts are given by $b = 10 \pm 2$, $d = 12$. The signal counts are (a) $\mu \cdot s = 15$ (b) $\mu \cdot s = 3$. The blue and red histograms are sampling distributions, generated through pseudoexperiments, of the test statistic (q_μ) under the b-only and $s+b$ hypothesis (denoted by H), respectively. The blue and red curves show the distribution obtained from the asymptotic approximation. The observed test statistic is shown by the solid black vertical line. The background-only distribution is used to get the median and the $\pm 1\sigma$ and $\pm 2\sigma$ values shown by the dashed black, green, and orange lines respectively.

values shown in Figure 5.4 correspond to the positive bands in the CL_s scan plots. The positive bands as well as the median values are not properly calculable for some cases because of the one-sided test statistic used here. This feature can be seen in the example case shown in Figure 5.4b, where $\mu \cdot s = 3$. The consequence of this feature can be seen in Figure 5.5b. The positive 1σ band is not properly calculated for $\mu \cdot s \leq 4$, whereas the positive 2σ band is not properly defined for $\mu \cdot s \leq 9$. Nonetheless, the computed observed and expected upper limits are very similar for the two methods. The observed upper limit on the signal count with the asymptotic approximation $\mu \cdot s^{\text{up}}_{\text{obs, asym}} = 10.23$, whereas the observed upper limit with toys is $\mu \cdot s^{\text{up}}_{\text{obs, toys}} = 10.46$. So in this case, asymptotic approximation works nicely.

In ATLAS in general the asymptotic approximation is used wherever possible as it saves considerable computing resources. Nonetheless, if the event counts are low then the results are often validated by calculating the limit with toys. In the dilepton resonance search (chapter 6) the asymptotic formula was used. In the $t\bar{t}$ resonance search (chapter 7) the final results are made with the asymptotic formula, but the limits on the signal strength are also computed with toys and compared.



(a) (b)

Figure 5.5: The CL_s scans over a range of signal counts ($\mu \cdot s$). The toy Poisson counting model with $d = 12$ and $b = 10$ with 20% background uncertainty is used. The p-values are calculated from the test statistic distributions (a) sampled with pseudoexperiments (toys) (b) approximated with asymptotic formula.

5.3.9 Wilks' test

Wilks' test is another a likelihood ratio-based test often used to determine how many parameters are required to fit data. The test statistic is given by

$$t = -2 \cdot \ln\left(\frac{L_m}{L_n}\right), \quad \text{with} \quad m < n, \tag{5.52}$$

where L_m is the likelihood of the model with m parameters, and L_n is the likelihood of the model with n parameters. It is important to note that the two models are nested and the m-dimensional

parameter space is a subset of the n-dimensional parameter space. furthermore, according to Wilks' theorem, in the asymptotic limit of a large data sample, the pdf of t is a chi-square distribution with $n-m$ degrees of freedom. The p-value can be calculated from the cumulative density function of χ^2_{n-m}. If the p-value is more than the threshold α, the fit with the fewer parameters, in this case with m parameters, is favoured since adding more parameters to the model does not make it describe the data better. Wilks' test is used in the boosted all-hadronic $t\bar{t}$ search to choose the number of parameters describing the background distributions, discussed in subsection 7.9.4.

5.4 BumpHunter search

The BumpHunter algorithm [205] is a model-independent statistical search method. The algorithm searches for a statistically significant deviation of data from the expected background. It can search for either excesses or deficits. The algorithm scans the full spectrum of the variable of interest in all possible search intervals. The search starts with the left-most minimum interval of the distribution, and the search interval position is shifted by one bin. The search interval size is then increased by one bin and the scan starts from the left again. The search method is demonstrated in the schematic diagram of Figure 5.6, where the minimum search interval is one bin. The sizes of the minimum and maximum search intervals are user-defined.

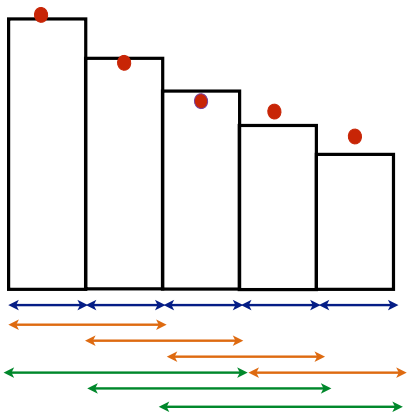


Figure 5.6: Diagram showing the intervals of the BumpHunter search process. The background expectation is depicted as black rectangles, and the observed data is shown as red dots. The arrows along the bottom of the diagram indicate the method where the BumpHunter searches over the bins, going from a minimum interval width of one to a maximum interval width of three.

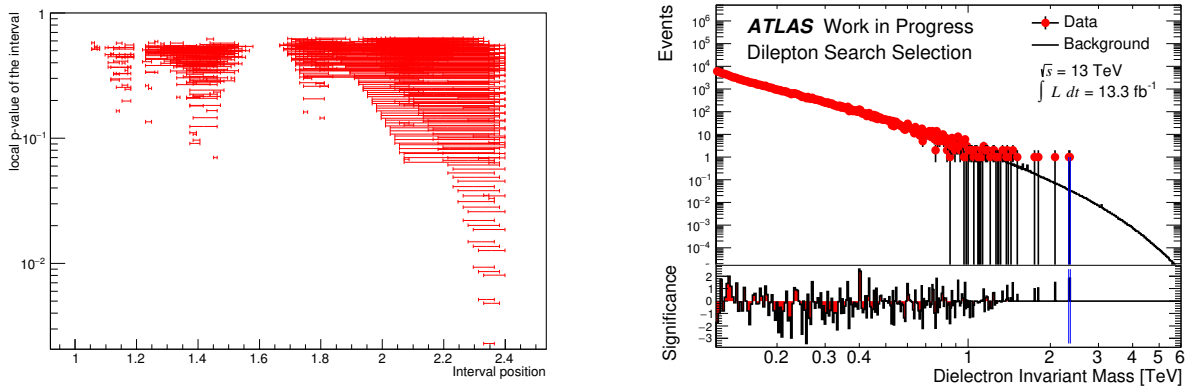
The local p-value is evaluated by counting the number of events in the observed data d and summing the number of expected background events b in each interval. The local p-value gives the probability of observing data (n) that is more deviant from the background (b) than the measured data (d), in a particular window. It is expressed in terms of an upper-incomplete gamma function $\Gamma(d,b)$, the cdf of a chi-squared distribution:

$$\text{local p-value} = p(n \geq d|b) = \Gamma(d,b) = \sum_{n\geq d}^{\infty} \frac{b^n}{n!} e^{-b} \,. \tag{5.53}$$

The interval with smallest local p-value is marked as the interval with most significant deviation. The negative log of this smallest local p-value is taken as the BumpHunter test statistic.

$$\text{BH-statistic} = -\ln(\text{smallest local } p\text{-value}) \,. \tag{5.54}$$

Example plots are shown in Figure 5.7 to demonstrate the algorithm. All the subsequent plots in this section are shown in the context of the dilepton resonance search with two electrons and two muons final state. In these examples, the dilepton invariant mass spectrum ($m^2_{\ell\ell} = p^\ell_1 + p^\ell_2$) is used to search for bumps. Figure 5.7a shows the scanned windows marked by red horizontal lines of variable lengths and the local p-value of each window. The interval in which this p-value is the smallest is flagged as the most deviant interval and the negative log of the smallest p-value is the BumpHunter test statistic. In Figure 5.7b the most deviant interval is marked with blue vertical lines. In this analysis searches are done only for excesses (bumps); deficits (dips) are not considered.



(a) Tomography plot

(b) BH scan plot

Figure 5.7: (a) Tomography plot: shows the test statistic of each window scanned by BumpHunter. The horizontal axis is showing the search intervals. (b) The interval with most significant deviation found by BumpHunter marked by two blue vertical lines.

The local p-values and the BH-statistic are not very meaningful. Any form of background fluctuation appearing in data will also contribute to the p-value. To estimate the BH-statistic coming from the background fluctuation, we need to obtain a distribution of the BH-statistic. The distribution is estimated by calculating the BH-statistic from pseudo-experiments obtained by Poisson fluctuations of the expected background spectrum. Then the BH-statistic value observed in data is compared with the background-only BH-statistic distribution. Hence a global p-value of the BumpHunter algorithm is defined using the following steps:

- Pseudo-experiments are generated by taking Poisson fluctuation of the nominal background spectrum. Each bin is fluctuated independently with Poisson distributed random numbers.
- The BH-statistic is calculated for each pseudo-experiment to get the background-only BH-statistic distribution.
- The BH-statistic is calculated in the data.

- Finally the BH global p-value is defined as

$$\text{BH global } p\text{-value} = \frac{\text{No. of PEs where (background-only statistic > observed BH-statistic)}}{\text{No. of all pseudo-experiments}} \tag{5.55}$$

Figure 5.8 shows the distribution of the BH-statistic from the pseudo-experiments. The blue line shows the observed BH-statistic value in data. The BH global p-value is the ratio of the integral of the distribution to the right of the blue arrow and the full yellow distribution.

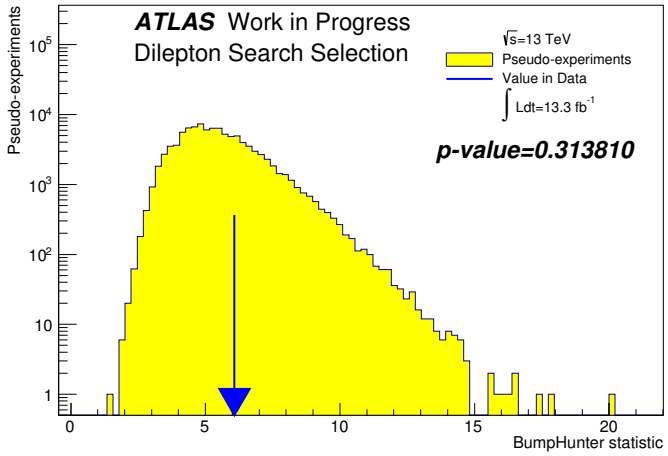


Figure 5.8: BumpHunter statistic distribution obtained with the background-only pseudo-experiments. The blue line corresponds to the BH-statistic obtained in data. The global p-value is the ratio of the right integral of the distribution staring from the blue line over the total area under the distribution.

5.4.1 Adding systematic uncertainties to the BumpHunter framework

The BumpHunter algorithm used in the ATLAS Collaboration did not have a consistent method for including the effect of systematic uncertainties. Several possibilities are studied for including systematic uncertainties in the BumpHunter framework. The most consistent method is a likelihood-based method with the systematic uncertainties added as likelihood parameters. The following steps are done:

- The background template is defined as the nominal background template plus the linear combination of the systematic variations. Thus the background count in the k^{th} interval can be written as

$$b_k(\theta) = b_k^0 \left(1 + \sum_{i=1}^{N_{\text{sys}}} \theta_i \sigma_{ik}\right). \tag{5.56}$$

  Here σ_{ik} is the relative uncertainty coming from i^{th} systematic variation in the k^{th} interval, and θ_i is the i^{th} nuisance parameter.

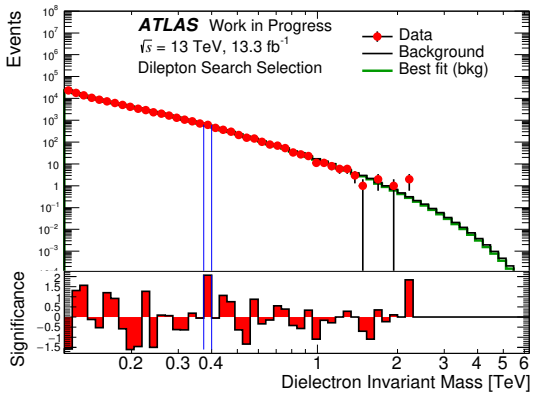
- Next the likelihood is built using the bins outside the search interval, often called the sideband region. The likelihood can be written as

$$\mathcal{L}(\theta) = \prod_{k=1}^{N_{\text{bins}}} \frac{b_k^{n_k}}{n_k!} e^{b_k} \prod_{i=1}^{N_{\text{sys}}} G(\theta_i; 0, 1). \tag{5.57}$$

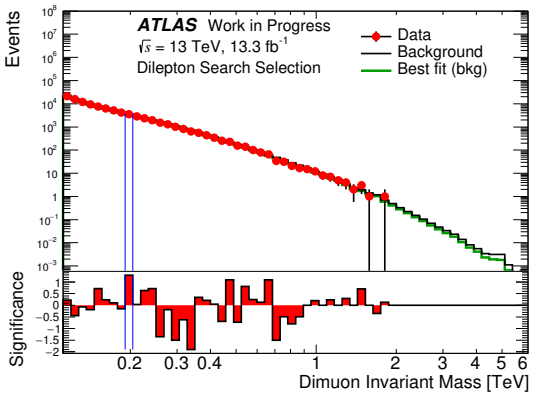
  Here n_k is the number of data event in the k^{th} interval and $G(\theta_i; 0, 1)$ is the standard normal Gaussian for the i^{th} nuisance parameter. The nuisance parameters are assumed to be Gaussian-distributed and the Gaussian constraints are added to the likelihood to ensure the likelihood has an upper bound.

- The negative log-likelihood is minimized in the sideband region to get the best-fit nuisance parameter values. Those estimated values of the nuisance parameters are used to define the background count in the search interval. Then the BH-statistic is calculated using Equation 5.54.

The effect of the systematic uncertainties is also considered while generating the background-only pseudo-experiments. Figure 5.9 shows the dilepton mass distributions in the electron and muon channels. The nominal background histogram is shown as a black histogram, and the green histograms are obtained by using the best-fit nuisance parameter values obtained from likelihood fit in the sideband region outside the most significant interval, marked by the vertical blue lines. The local p-values are calculated including the systematic uncertainties.



(a) Dielectron channel



(b) Dimuon channel

Figure 5.9: BumpHunter scan on the dilepton mass distribution in (a) dielectron and (b) dimuon channel in presence of systematic uncertainties. The observed data are shown in red points and the total nominal background is shown in the back histograms, where all the nuisance parameters are set to zero. The green histogram is obtained by setting the nuisance parameters to their best-fit values obtained from the profile likelihood fit of the sideband region. The lower panel shows bin-by-bin local significance. The most significant interval is indicated by the vertical blue lines.

CHAPTER

SIX

Search for high mass resonances in di-lepton final state

This chapter summarizes a search for high mass resonance (Z') which decays into two leptons (e^+e^- and $\mu^+\mu^-$) with the ATLAS detector at the LHC operating at $\sqrt{s}$ = 13 TeV. The search was performed with 36.1 fb^{-1} of Run-2 proton-proton collision data collected in 2015 and 2016. The analysis considers narrow width resonances, corresponding new high mass boson particles (Z') decaying to two leptons. Two benchmark models, Sequential Standard Model (SSM) and E_6 extended symmetry model, are used to interpret the results. The models are described in section 1.5. The results presented in this chapter were published in 2017 [206]. I was one of the main analyzers in the electron channel, so the final states with two electrons will be discussed in more detail. I optimized the event selections in the electron channel and performed the model-independent statistical search in both electron and muon channels. Previously, the search was carried out in ATLAS with the 3.2 fb^{-1} LHC Run-2 proton-proton collision data collected in 2015. Data was consistent with SM prediction and no evidence for new physics was found. A narrow width (3%) SSM Z'_{SSM} boson was excluded up to 3.4 TeV with 95% confidence level [207].

The chapter is organized as follows. An overview of the full analysis is given in section 6.1. The reconstructed objects involved in this search are defined in section 6.2. The two-lepton final state involves a small set of objects: electrons, muons, and jets. After the objects are defined, a set of selection criteria are developed to select events with two high momentum leptons. These event selection criteria are summarized in section 6.3. The background processes are described in section 6.4. Details of how the signal processes are simulated are discussed in section 6.5. Then the reconstruction of the main discriminant used in this analysis is described in section 6.6. The systematic uncertainties used in this analysis are summarized in section 6.7. The statistical search results are summarized in section 6.8 and the final limit results are shown in section 6.9.

6.1 Analysis overview

The analysis targets final states with two same flavor leptons, either two electrons, or two muons. Figure 6.1 shows a Feynman diagram of the signal process ($Z' \to e^+e^-$ or $\mu^+\mu^-$). The analysis

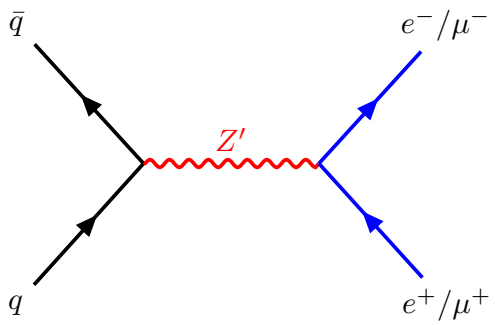


Figure 6.1: Feynman diagram of the Z' signal process decaying into two electrons or muons.

searches in the invariant mass spectrum of two leptons. The new Z' boson is expected to appear as a narrow bump over the dilepton invariant mass distribution of the SM. The invariant mass is built by combining the momenta of the two highest p_T electrons in an event, and defined as,

$$m_{ee}^2 = (p_1^e + p_2^e)^2. \tag{6.1}$$

Here p_1^e and p_2^e are the momentum 4-vectors of the two highest p_T electrons. The major background contribution comes from the irreducible Drell-Yan processes with two lepton final states. The other reducible background contribution comes from the SM $t\bar{t}$, single top, diboson, W+jets, and multijet processes. Multijet and W+jets processes are estimated using a data-driven technique, whereas all other processes are modeled with Monte Carlo simulation. The statistical searches are performed using a model-independent bump-hunting algorithm as well as with a model-dependent profile likelihood ratio-based hypothesis test. The searches are performed on the two channels independently first, and then a combined search is also performed. No trace of new physics is observed in the data. So in the absence of new physics, the upper limit on the signal cross-section times branching ratio is set with 95% CL.

6.2 Object definition

Electrons

In the two-electron final state, the main objects needed are the electrons. The electron reconstruction is described in section 4.3. In this analysis, the electrons are required to be within $|\eta| < 2.47$. The electrons falling in the transition region of the barrel calorimeter and the endcap calorimeter, which is located at $1.37 < |\eta| < 1.52$, are not used. The electrons coming from bad calorimeter clusters are also rejected. The transverse energy (E_T) of the electrons is required to be greater than 30 GeV since we are looking for high momentum leptons. To ensure that the electrons are associated with the correct primary vertex, requirements are applied to the impact parameters. The d_0^{BL}-significance is required to be within $[-5, 5]$, $|d_0^{\mathrm{BL}}/\sigma_{d_0}| < 5.0$. The longitudinal impact parameter with respect to the beamline is required to be within -0.5 mm and 0.5 mm of the primary vertex, $|z_0 \sin\theta| < 0.5$ mm. Finally, the electrons are required to pass the medium likelihood-based identification criteria and loose isolation criteria. The medium identification criteria has roughly 95% efficiency for electrons with E_T between 30 GeV and 500 GeV and it decreases at high E_T. The loose isolation uses the track isolation variable $p_T^{\mathrm{varcone0.2}}$ and calorimeter isolation variable

Table 6.1: Summary of the electron object definition.

Selection	Criteria
η coverage	$\lvert\eta\rvert < 2.47$ without $1.37 < \lvert\eta\rvert < 1.52$
Cleaning	Remove bad calorimeter clusters
Energy	$E_T > 30$ GeV
d_0-significance	$\lvert d_0^{\mathrm{BL}}/\sigma_{d_0}\rvert < 5.0$
$z_0 \sin\theta$	$\lvert z_0 \sin\theta\rvert < 0.5$ mm
Identification	LH medium
Isolation	Loose

$E_T^{\mathrm{cone0.2}}$. Isolation selection efficiency of the loose working point is around 98%. The electron definition is summarized in Table 6.1. Calibrations are applied to the electron candidates to correctly estimate the electron energies from the calorimeter signals and to match energy scale and resolution in the simulation to the data. Electron energy scale and resolution is calibrated up to $E_T = 1$ TeV using Run-1 data collected at $\sqrt{s} = 8$ TeV and Run-2 data collected at $\sqrt{s} = 13$ TeV. The energy resolution is extrapolated beyond $E_T = 1$ TeV and is estimated to be around 1%.

Small radius calorimeter jets

Jets are also used to study some of the background processes. Calorimeter small–R jets of $R = 0.4$, AntiKt4EMTopoJets, are using for these studies. The jets are required to have $p_T > 20$ GeV and $\lvert\eta\rvert < 2.5$. Jet cleaning is applied to remove the calorimeter noises and badly reconstructed jets. To suppress pile-up jets, a JVT cut is applied and the medium operating point is used. The medium operating points requires a jet JVT > 0.59 if the jet $p_T < 60$ GeV and $\lvert\eta\rvert < 2.4$. It has 92% average selection efficiency. The jets are also required not to overlap with electrons and it is required that the ΔR[13](el, jet) > 0.4. The jet definitions are summarized in Table 7.6.

Table 6.2: Summary of the jet object definition.

Selection	Criteria
η coverage	$\lvert\eta\rvert < 2.5$
Jet cleaning	Loose
Momentum	$p_T > 20$ GeV
JVT	Medium, JVT> 0.59
Overlap removal	ΔR(el, jet) > 0.4

6.3 Event selection

To select events with two electrons, a series of selection requirements are applied. Each event should belong to the Good Run List as described in subsection 2.2.8. Next the event needs to pass the di-electron trigger, `2e17_lhloose` trigger. Then an event cleaning is applied to remove

[13] $\Delta R = \sqrt{(\Delta\phi)^2 + (\Delta\eta)^2}$

events flagged as being incomplete. The events are further required to pass LAr and Tile error requirements to discard the noise bursts and corrupted data in the ECAL and HCAL. The SCT corrupted events are also removed. Each event needs to have at least one reconstructed proton-proton interaction vertex. Then the events are required to have at least two good quality well-reconstructed isolated electrons as defined in section 6.2. If there are more than two electrons in an event, the two highest p_T electrons are selected. Opposite charge requirement is not applied while selecting the events to avoid the consequences of charge misidentification for high momentum electrons. Finally, the invariant mass of the di-electron system is required to be greater than 80 GeV to reduce the photon-induced contribution. The di-electron event selection is summarized in Table 6.3. The acceptance times efficiency after all the selection criteria is around 70% for Z'_χ mass above 1 TeV as shown in Figure 6.2.

Table 6.3: Summary of the event selection

Selection	**Criteria**
GRL	Data events in the Good Run List
Di-electron trigger	2e17_lhloose
Event cleaning	Remove calorimeter noise, LAr, SCT errors, noise burst, corrupted error
At least 2 electrons	≥ 2 high p_T electrons
Dielectron mass cut	$m_{ee} > 80$ GeV

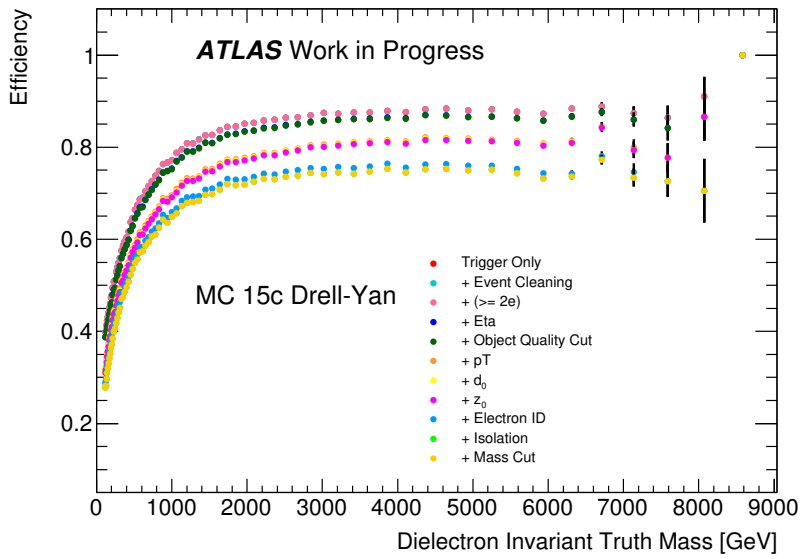


Figure 6.2: Signal acceptance times efficiency as a fucntion of dielectron truth mass. Different curves (with different colors) are showing the acceptance time efficiency after applying each selection cut.

6.4 Background modeling

The modeling of the background processes is summarized in this section. Most of the background processes are simulated by using Monte Carlo (MC) simulation. The QCD multijet and W+jets backgrounds are estimated using a data-driven technique.

6.4.1 Background processes

The SM background to a dilepton search consists of an irreducible component due to Drell-Yan ($q\bar{q} \to Z/\gamma^* \to \ell^+\ell^-$) as well as a small contribution from photon-induced events ($\gamma\gamma \to \ell^+\ell^-$, through $t-$ and $u-$ channel processes), and reducible components. The reducible background components are $t\bar{t}$, multijet & W+jets, and diboson backgrounds. The multijet & W+jets background contains events with a maximum of one real lepton and one or more non-prompt leptons or jets faking leptons. These contributions are non-negligible in the electron channel. Most of the processes are modeled using simulation, but it does not work well for the multijet & W+jets background. Therefore, a data-driven (DD) approach is used to determine the multijet & W+jets background in the electron channel. This background component is very negligible in the di-muon final state, therefore not considered.

Drell-Yan background

The expected contribution from the SM Drell-Yan process is modeled with the next-to-leading-order (NLO) Powheg [208] event generator implementing the CT10 [209] Parton Distribution Function (PDF) in conjunction with Pythia 8.186 [210] for showering, and the ATLAS AZNLO tune [211]. To ensure that the statistical uncertainty remains small in the search region 19 mass-sliced samples were created with parton-level dilepton invariant masses ranging between 120 GeV and > 5000 GeV. To cover the region below 120 GeV, an inclusive sample which only has a lower mass threshold of 60 GeV is used. The inclusive sample and the mass-sliced samples are added together to get the full invariant mass spectrum. To correct the DY cross-section from NLO (CT10) to next-to-next-to-leading-order (NNLO) using the CT14NNLO [212] PDF, a dilepton mass-dependent k-factor is calculated with VRAP 0.9 [213] for QCD corrections, and Mcsanc [214] 1.20 for EW corrections. The NNLO QCD correction factor is around ~ 0.98 at a dilepton (m_{ll}) mass of 3 TeV whereas the NLO EW correction is ~ 0.86 at $m_{ll} = 3$ TeV. The contribution from Photon-Induced (PI) events to the Drell-Yan process is not taken into account in the generation. So, it is described with a k-factor using MRST2004QED [215]. On top of these corrections, the DY background is smoothed in the range 120GeV < $m_{\ell\ell}$ < 1TeV to minimize the statistical fluctuations due to limited Monte Carlo sample size.

Diboson (WW, WZ, ZZ) processes

The diboson processes (WW, WZ, ZZ) are generated using Sherpa 2.1.1 [216] at NLO with the CT10 PDF. These simulated samples are generated with a series of mass thresholds, ensuring a small statistical uncertainty across the entire mass spectrum of interest to this search. As in Drell-Yan, all these mass-binned samples are added together to get the full distribution.

$t\bar{t}$ and single-top processes

The $t\bar{t}$ background is generated at NLO using Powheg with CT10 PDF and Perugia 2012 tune [217], in conjunction with Pythia 6 for the event showering. The single-top background (associated $W-t$

production) is generated at NLO using POWHEG with the CT10 PDF and Perugia 2012 tune, in conjunction with PYTHIA 6 for the event showering. These samples are generated for a top quark mass of 172.5 GeV, and with high enough statistics to ensure a small statistical uncertainty.

These samples are normalized to the higher-order cross-sections calculated with Top++ 2.0 [218]. It uses a global k-factor of 1.1949 (1.054), which brings the $t\bar{t}$ (single-top) cross-section values from NLO to NNLO in QCD, including resummation of next-to-next-to-leading logarithmic (NNLL) soft gluon terms. Details of these major SM background simulations are summarized in Table 6.4.

Table 6.4: Summary of the simulation information used to model the background processes in the dilepton resonance search. The columns from left to right give the process of interest, generator, matrix element order, correction to order of calculation, parton shower program, and PDF utilized.

Process	Generator	Order	Corrected	Parton Shower	PDF
$Z/\gamma^* \to \ell^+\ell^-$	POWHEG v2 [208]	NLO	NNLO	PYTHIA 8.186 [210]	CT10 [209]
$t\bar{t} \to \ell X$, $Wt \to X$	POWHEG v2 [208]	NLO	NNLO	PYTHIA 6.428 [219]	CT10 [209]
$WW, WZ, ZZ \to \ell X/\ell\nu/\ell\ell$	SHERPA 2.1.1 [216]	NLO	NLO	SHERPA 2.1.1 [216]	CT10 [209]

Multi-jet and W+jets background estimation

At most one real lepton gets produced in the multijet or W+jets process. These processes contribute mainly due to one or more jets being reconstructed as leptons. The background contribution coming from the multijet process is not possible to model accurately using simulation. Whereas the simulated W+jets process does not have enough statistics to model the background contribution in the signal region. Instead, these processes are estimated using data with a data-driven technique called *matrix method*. The matrix method relies on two important quantities, the *real rate* and the *fake rate*. Two different lepton selections are implemented for calculating the real and fake rates. The tighter selection is called *tight (T)*, and the looser one is called *loose (L)*. The tight selection is identical to the nominal signal selection described earlier in this chapter. The loose selection needs to pass the *Loose* electron identification criteria, and no isolation criteria are required. These two definitions are inclusive i.e. the set of objects passing tight selection is a subset of those passing the loose selection.

Real rate The real rate, r, is defined as the probability of a real electron passing the loose selection to satisfy also the tight (nominal) selection. To estimate r, a sample enriched with a real electron is needed. Hence, the real rate is estimated from simulated Drell-Yan samples in different E_T and $|\eta|$ regions. The real rate is around 95% for the low p_T electrons and increases with p_T.

Fake rate The fake rate, f, is defined as the probability of a fake electron satisfying the loose selection to also pass the tight (nominal) selection. The fake rate is estimated using fake electrons in data. The fake enriched data sample is obtained for different regions of E_T and $|\eta|$ using the events satisfying a single electron trigger with *Loose* or *Very Loose* identification criteria. The phase space is divided into separate p_T regions, and in each region, a specific trigger is used. The p_T regions and the triggers are summarized in Table 6.5 and Table 6.6. The real and fake rates can be written as:

$$r = \frac{N_{\mathrm{R}}^{\mathrm{tight}}}{N_{\mathrm{R}}^{\mathrm{loose}}}, \qquad f = \frac{N_{\mathrm{F}}^{\mathrm{tight}}}{N_{\mathrm{F}}^{\mathrm{loose}}}. \tag{6.2}$$

$N_{\mathrm{R}}^{\mathrm{tight}}\left(N_{\mathrm{R}}^{\mathrm{loose}}\right)$ is the number of real electrons passing the *tight* (*loose*) selection. $N_{\mathrm{F}}^{\mathrm{tight}}\left(N_{\mathrm{F}}^{\mathrm{loose}}\right)$ is the number of fake electrons passing the *tight* (*loose*) selection.

Table 6.5: Trigger scheme used for the selection of a fake-enriched sample for 2015 data.

Trigger	Pre-scale corrected luminosity [pb^{-1}]	$p_{\mathrm{T,low}}$ [GeV]	$p_{\mathrm{T,high}}$ [GeV]
HLT_e24_lhvloose_L1EM20VH	91.7987	29	65
HLT_e60_lhvloose	267.746	65	125
HLT_e120_lhloose	3212.96	125	∞

Table 6.6: Trigger scheme used for the selection of a fake-enriched sample for 2016 data.

Trigger	Pre-scale corrected luminosity [pb^{-1}]	$p_{\mathrm{T,low}}$ [GeV]	$p_{\mathrm{T,high}}$ [GeV]
HLT_e24_lhvloose_L1EM20VH	138.772	29	31
HLT_e26_lhvloose_L1EM20VH	208.158	31	65
HLT_e60_lhvloose_nod0	436.698	65	125
HLT_e120_lhloose	2202.35	125	145
HLT_e140_lhloose	10064.3	145	∞

The real electron contamination coming from W and Z decay is restricted by applying a few specific criteria. The W decay contribution is suppressed by vetoing events with large missing transverse energy, $E_T^{\mathrm{miss}} > 25$ GeV. Real electrons coming from the Z decay are suppressed by vetoing events with two-electron candidates passing *Medium* identification criteria or by vetoing events with two *Loose* electrons where the reconstructed mass, m_{ee}^{reco}, is compatible with the Z mass by requiring $|m_Z^{\mathrm{PDG}} - m_{ee}^{reco}| > 20$ GeV. Even though these criteria reduce the real contamination, the real contribution in the fake enriched data sample can not be completely ignored. The expected real contamination is estimated using simulated Drell-Yan, W+jets, Di-Boson, and $t\bar{t}$ events. This contribution is then subtracted from the number of fake electrons passing either loose or tight selection. The fake rate is also determined as a function of p_T and $|\eta|$. At the very low p_T, the fake rate is estimated to be around 25%, then it decreases with p_T and goes down to $\sim$ 17% around 45 GeV. The fake rate eventually starts increasing for electrons above $p_T = 1$ TeV and reteaches up to 33% at higher p_T.

The analysis considers two-electron candidates in the final state, and the selected events can be categorized based on which selection each lepton passes. The number of events is denoted by N_{xy} with $x, y \in T, L$. This leads to four different quantities: $N_{TT}, N_{TL}, N_{LT}, N_{LL}$. The first index represents the leading object and the second one subleading ($E_T^x > E_T^y$). Equation 6.3 shows the relation between the reconstructed and the real quantities N_{ab} with $a, b \in R, F$. The subscript R and F refer to real and fake electron candidates, respectively.

$$\begin{pmatrix} N_{TT} \\ N_{TL} \\ N_{LT} \\ N_{LL} \end{pmatrix} = \begin{pmatrix} r^2 & rf & fr & f^2 \\ r(1-r) & r(1-f) & f(1-r) & f(1-f) \\ (1-r)r & (1-r)f & (1-f)r & (1-f)f \\ (1-r)^2 & (1-r)(1-f) & (1-f)(1-r) & (1-f)^2 \end{pmatrix} \begin{pmatrix} N_{RR} \\ N_{RF} \\ N_{FR} \\ N_{FF} \end{pmatrix} \tag{6.3}$$

The fake background can be estimated from the part of N_{TT} that originates from a pair of

objects containing at least one fake electron:

$$N_{TT}^{\text{Multi-jet \& W+jets}} = rf(N_{RF} + N_{FR}) + f^2 N_{FF} \tag{6.4}$$

The equation depends on non-reconstructible truth quantities like N_{RF}, N_{FR} and N_{FF}. These can be expressed in terms of measurable quantities $(N_{TT}, N_{TL}, N_{LT}, N_{LL})$ by inverting the matrix in Equation 6.3.

$$\begin{pmatrix} N_{RR} \\ N_{RF} \\ N_{FR} \\ N_{FF} \end{pmatrix} = \frac{1}{(r-f)^2} \begin{pmatrix} (f-1)^2 & (f-1)f & f(f-1) & f^2 \\ (f-1)(1-r) & (1-f)r & f(1-r) & -fr \\ (r-1)(1-f) & (1-r)f & r(1-f) & -fr \\ (1-r)^2 & (r-1)r & r(r-1) & r^2 \end{pmatrix} \begin{pmatrix} N_{TT} \\ N_{TL} \\ N_{LT} \\ N_{LL} \end{pmatrix} \tag{6.5}$$

The combined fake contributions to the signal selection is given by,

$$\begin{aligned} N_{TT}^{\text{Multi-jets \& W+jets}} &= \alpha\left[2rf(f-1)(1-r) + f^2(1-r)^2\right] N_{TT} \\ &+ \alpha f r^2 (1-f)(N_{TL} + N_{LT}) \\ &- \alpha f^2 r^2 N_{LL}\,, \end{aligned} \tag{6.6}$$

where

$$\alpha = \frac{1}{(r-f)^2}. \tag{6.7}$$

High mass extrapolation The Z peak region is excluded in the QCD estimation process to mitigate the instabilities due to limited fakes in that region. At the Z peak, the estimated values become negative due to the overcorrection of the real electrons failing tight selection. So the matrix method is applied above 125 GeV. The fakes estimation suffers at a higher mass tail due to limited data events. A functional fit is performed to overcome this issue and extrapolate the fakes estimation to the highest mass region. The following polynomial function is used

$$f(x) = p_1(1-x)^{p_2} \cdot x^{p_3 + p_4 \ln x + p_5 (\ln x)^2}\,, \tag{6.8}$$

where $x = m_{ee}/\sqrt{s}$. This function is commonly known as the *di-jet function* and is widely used to estimate di-jet invariant mass. The fit is performed up to m_{ee} = 13 TeV. In order to estimate uncertainties on the result, the fit is repeated eight times, excluding each time one more bin from the low edge of the range. The weighted mean of all these fits is used as the central value of fakes estimation. The envelope estimated from the variation of the fits is used as the fit uncertainty.

6.5 Signal modeling

The benchmark signal models are simulated using Monte Carlo (MC) simulation. The Z' signal samples were generated at leading-order (LO) using Pythia 8186 with the NNPDF23LO PDF set [220] and the ATLAS A14 set of tuned parameters for parton shower and hadronization. Signal interference with SM Drell-Yan (DY) is not considered for the Z' signal of SSM and E_6 models due to their large model dependence. Higher-order QCD corrections in the signal process are calculated in the same way as for the DY background. The higher-order electroweak (EW) corrections are not applied to the Z' signal. Samples were produced for the E_6 Z'_χ model at a few representative pole masses (1,2,3,4,5 and 6 TeV). Z'_{SSM} was also generated for a few pole masses. The fine

grid signal models are created reweighting the LO Drell-Yan sample. Each DY event is reweighted to get the desired signal at a particular pole mass. The reweighting factor is calculated by taking the ratio of the BSM and SM cross-section for each event. The summary of the signal processes used in this search is presented in Table 6.7.

Table 6.7: SSummary of the simulation process used to generate the signal events used in the dilepton resonance search. The columns from left to right give the process of interest, generator, matrix element order, correction to order of calculation, parton shower program, and PDF utilized.

Process	Generator	Order	Corrected	Parton Shower	PDF
$Z/\gamma^* \to \ell^+\ell^-$	Pythia 8.186 [210]	LO	NNLO	Pythia 8.186 [210]	NNPDF23LO [220]
$Z' \to \ell^+\ell^-$	Pythia 8.186 [210]	LO	NNLO	Pythia 8.186 [210]	NNPDF23LO [220]

6.6 Dilepton invariant mass

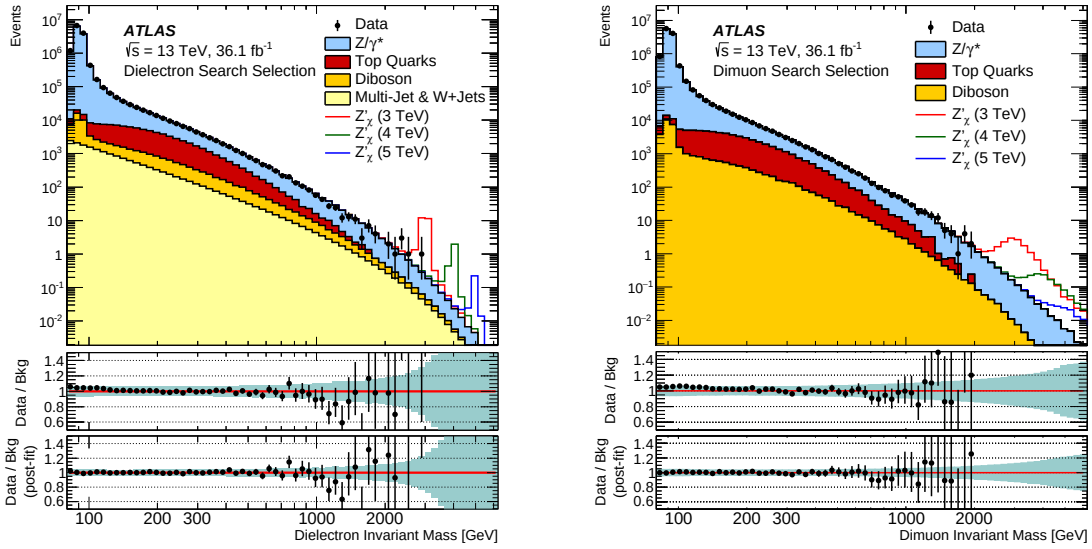
The m_{ee} distribution is estimated for all the major background processes and binned. A total of 66 bins, starting from 120 GeV till 5 TeV, are used. The binning choice is optimized for resonances with 3% width. A different binning based on the resolution in the di-electron channel is used in the search phase. The bin width varies from 10 GeV at m_{ee} = 1 TeV to 15 GeV at m_{ee} = 2 TeV and 20 GeV at m_{ee} = 3 TeV in the search phase. The invariant mass spectrum is shown in Figure 6.3 for both the Figure 6.3a di-electron and Figure 6.3b di-muon channels. The colored stacked histograms show the various SM backgrounds properly scaled by the production cross-section and data luminosity. The expected signal distributions predicted by the E_6 model are overlaid on the m_{ee} distribution for Z'_χ pole mass of 3, 4 and 5 TeV. There are two panels at the bottom showing the ratio of data and background. The sky-blue band shows the total contribution coming from the systematic uncertainties (discussed later), where all the uncertainties are added in quadrature. The upper bottom panel shows the ratio of data and nominal background, whereas the lower panel shows the ratio of data and the post-fit background. The post-fit background is obtained by adjusting the nominal background using the best-fit nuisance parameter values. Figure 6.3 shows that the data is consistent with the SM prediction, and there is no visible excess of events. The ratio of data and background does not have any particular shape either.

6.6.1 Differences between electron and muon channels

The mass resolution is much better in the electron channel. This is clearly visible in Figure 6.3 and Figure 6.4. The Z' signals are narrower in the two-electron final state. Due to worse reconstruction resolution, the signal $m_{\mu\mu}$ distribution is much wider than m_{ee}. The reconstruction efficiency is also lower in the muon channel. High p_T electrons are reconstructed more efficiently compared to high p_T muons. The signal acceptance times efficiency is around 70% in the electron channel, whereas it is only around 40% in the muon channel.

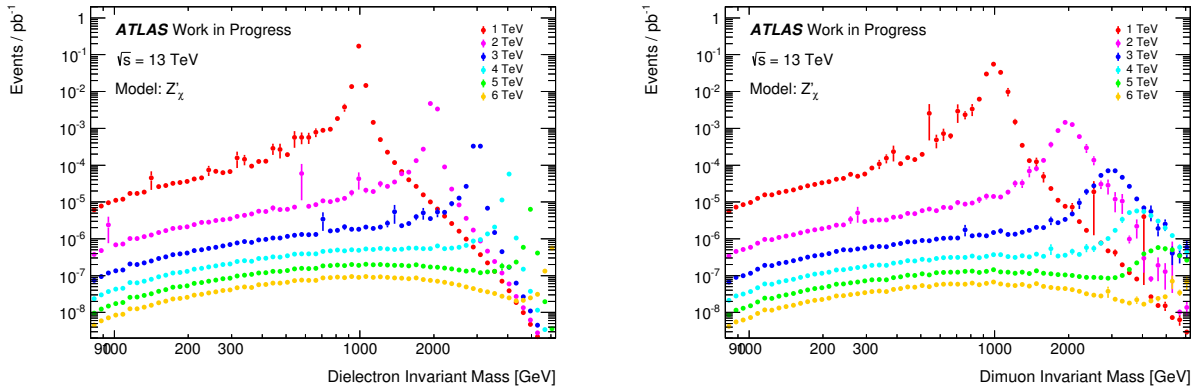
6.6.2 Event yields

The expected and observed event yields in different mass regions are presented in Table 6.8 for the di-electron channel. The expected background yields are split into different background processes,



(a) Dielectron channel (b) Dimuon channel

Figure 6.3: Invariant mass distribution in the (a) di-electron and (b) di-muon channels. The background processes are shown as filled histograms, and the Z'_χ signal distributions for pole mass 3, 4, 5 TeV are overlaid. The upper ratio plot shows the ratio of data and total predicted background. The lower ratio plot shows the ratio of data and post-fit background. The cyan bands in the two ratio plots show the total systematic uncertainty [206].



(a) Dielectron channel (b) Dimuon channel

Figure 6.4: Simulated signal templates for the Z'_χ signal in the (a) di-electron and (b) di-muon channels are presented for six representative pole masses. The widths of the di-electron mass distributions are much narrower, since the mass resolution is better in the electron channel.

and the yields for two signal scenarios are also added. In general, there is a very good agreement between data and the SM prediction.

Table 6.8: Expected and observed event yields in the dielectron channel in different dilepton mass intervals. The quoted errors correspond to the combined statistical, theoretical, and experimental systematic uncertainties. Expected event yields are reported for the Z'_χ model, for two values of the pole mass. All numbers shown are obtained before the marginalisation procedure.

m_{ee} [GeV]	80–120	120–250	250–400	400–500	500–700
Drell–Yan	11 800 000 ± 700 000	216 000 ± 11 000	17 230 ± 1000	2640 ± 180	1620 ± 120
Top quarks	28 600 ± 1800	44 600 ± 2900	8300 ± 600	1130 ± 80	560 ± 40
Dibosons	31 400 ± 3300	7000 ± 700	1300 ± 140	228 ± 25	146 ± 16
Multi-jet & W+jets	11 000 ± 9000	5600 ± 2000	780 ± 80	151 ± 21	113 ± 17
Total SM	11 900 000 ± 700 000	273 000 ± 12 000	27 600 ± 1100	4150 ± 200	2440 ± 130
Data	12 415 434	275 711	27 538	4140	2390
Z'_χ (4 TeV)	0.00635 ± 0.00021	0.0390 ± 0.0015	0.0564 ± 0.0025	0.0334 ± 0.0027	0.064 ± 0.004
Z'_χ (5 TeV)	0.00305 ± 0.00012	0.0165 ± 0.0006	0.0225 ± 0.0010	0.0139 ± 0.0007	0.0275 ± 0.0015
m_{ee} [GeV]	700–900	900–1200	1200–1800	1800–3000	3000–6000
Drell–Yan	421 ± 34	176 ± 17	62 ± 7	8.7 ± 1.3	0.34 ± 0.07
Top quarks	94 ± 8	27.9 ± 2.8	5.1 ± 0.7	< 0.001	< 0.001
Dibosons	39 ± 4	16.9 ± 2.1	5.8 ± 0.8	0.74 ± 0.11	0.028 ± 0.004
Multi-jet & W+jets	39 ± 6	16.1 ± 2.0	7.9 ± 2.3	1.6 ± 1.2	0.08 ± 0.27
Total SM	590 ± 40	237 ± 17	81 ± 7	11.0 ± 1.8	0.45 ± 0.28
Data	589	209	61	10	0
Z'_χ (4 TeV)	0.0585 ± 0.0035	0.074 ± 0.005	0.121 ± 0.011	0.172 ± 0.017	2.57 ± 0.27
Z'_χ (5 TeV)	0.0218 ± 0.0013	0.0295 ± 0.0021	0.040 ± 0.004	0.040 ± 0.004	0.280 ± 0.030

6.7 Systematic uncertainties

The dilepton resonance analysis is affected by several sources systematic uncertainties. All the systematic uncertainties with a non-negligible effect on the final results are considered and modeled in this analysis. The systematic uncertainties come from both theoretical and experimental sources. Both theoretical and experimental sources of uncertainties are considered for the background processes, whereas only the experimental effects are considered for the signal process. Both up and down systematic variations ($\pm 1\sigma$) are modeled where possible. The uncertainties are then symmetrized, taking the largest of the up and down variations. The effect of each systematic uncertainty on the di-electron invariant mass is estimated. The estimated uncertainties are used in the statistical search and limit setting along with the nominal m_{ee}.

6.7.1 Theory uncertainties

Theory uncertainties in background modeling mostly come from the DY background. The main sources are the variations of the nominal parton distribution function (PDF) set, PDF scales, the strong coupling (α_S), and electro-weak corrections as well as photon-induced (PI) corrections. The DY production cross-section also changes if a different PDF is chosen. This effect is also considered in this analysis. The theoretical uncertainties are the same in both electron and muon channels at generator lever, but the effect on the reconstructed dilepton mass spectrum becomes different in the two channels due to different reconstruction resolutions and efficiencies.

PDF variation systematic The variation of parton distribution functions has an effect on the DY cross section. Each PDF has a set of independent parameters associated to it, known as the eigenvectors of the PDF set in the functional space. These eigenvectors form an orthogonal basis and can be varied independently to quantify the systematic uncertainties associated to the PDF variations. Total 28 ± eigenvector error set variations of the CT14NNLO PDF with respect to its central value are used. For each eigenvector variation at 90% CL in the CT14NNLO parametrization [221], the DY cross section is calculated at NNLO as a function of $m_{\ell\ell}$ using using VRAP program [213]. The relative differences compared to the central CT14NNLO PDF values are used as the uncertainties on the cross-section due to PDF variations. The asymmetric uncertainty at each mass point is calculated using following equations

$$\Delta\,\sigma^{+} = \sqrt{\sum_{i=1}^{n}\left[\max\left\{\sigma_i^{+}-\sigma_0,\sigma_i^{-}-\sigma_0,0\right\}\right]^2}, \quad \text{and} \quad \Delta\,\sigma^{-} = \sqrt{\sum_{i=1}^{n}\left[\max\left\{\sigma_0-\sigma_i^{+},\sigma_0-\sigma_i^{-},0\right\}\right]^2}, \tag{6.9}$$

where n is the number of PDF eigenvectors, and σ_i^{+} (σ_i^{-}) is the cross-section for the higher (lower) value of the i^{th} PDF eigenvector. Using one nuisance parameter corresponding to these 28 variations might underestimate the effect since the variations have different impacts at different $m_{\ell\ell}$ region. So the variations with similar mass dependence are bundled together and re-diagonalized to a set of seven PDF eigenvectors [222]. Then seven different nuisance parameter are used corresponding to these seven variations. The larger of the positive and negative variation is taken as the systematic uncertainty on the signal cross-section. The seven PDF eigenvector variations are shown in Figure 6.5.

PDF scale and α_S uncertainties The PDF scale and α_S systematic uncertainties at NNLO for the Drell-Yan background are estimated using VRAP program [213]. The PDF scale uncertainty is estimated by varying the nominal CT14NNLO PDF renormalization and factorization scales by a factor of two following the PDF4LHC prescription [223]. The resulting maximum variations with respect to the nominal setting are used as PDF scale uncertainty.

The α_S variation is computed using VRAP program by varying the nominal input value of $\alpha_S = 0.118$ by an uncertainty of ± 0.003. The difference between the resulting spectrum and nominal spectrum is used as the systematics uncertainty.

PDF choice The nominal PDF used for the Drell-Yan background is CT14NNLO. The uncertainty due to this choice is investigated by comparing the central value of CT14NNLO with those from similar other PDFs. In his search, two other alternative PDFs, MMHT14 [224] and NNPDF3.0 [225] are used. The maximum absolute difference from the envelope of these comparisons is used as the PDF choice uncertainty. This background is only applied on the Drell-Yan background. This variation is shown in Figure 6.6.

Photon induced effects The photon induced dilepton cross-section is calculated at LO QCD using MRST2004qed. This background can be estimated with the analysis cuts at $|\eta| < 2.5$ and $p_T > 25$ GeV for both the leptons. To reduce this contribution an invariant mass cut is applied on in the analysis selection. The uncertainty on the PI is large since the photon radiation changes when choosing different quark masses for the photon PDF. The PI uncertainty accounts for the effect of the quark mass uncertainty on photon PDF. The uncertainty coming from photon induced effect is shown in Figure 6.6.

Higher order EW and QCD correction In the analysis, the both EW and QCD corrections are applied by adding the corrections on the DY background. The combined correction on the nominal spectrum can be written by $(1 + \delta_{\mathrm{EW}} + \delta_{\mathrm{QCD}})$. The EW correction systematics is estimated by comparing this additive approach with the multiplicative approximation $(1 + \delta_{\mathrm{EW}})(1 + \delta_{\mathrm{QCD}})$ treatment of the EW correction while combining the effect of the higher order QCD effects. The uncertainty is the difference from the nominal kinematic distribution and the spectrum obtained in the multiplicative treatment. It is shown in Figure 6.6.

Top quark and Diboson background uncertainties The theoretical uncertainties on the estimation of top quark and diboson backgrounds are derived by varying the factorization scale (μ_{F}), renormalization scale (μ_{R}), PDF and α_S are using the PDF4LHC prescription [223]. The normalization uncertainties are less than 1% of the total background as shown in Figure 6.6.

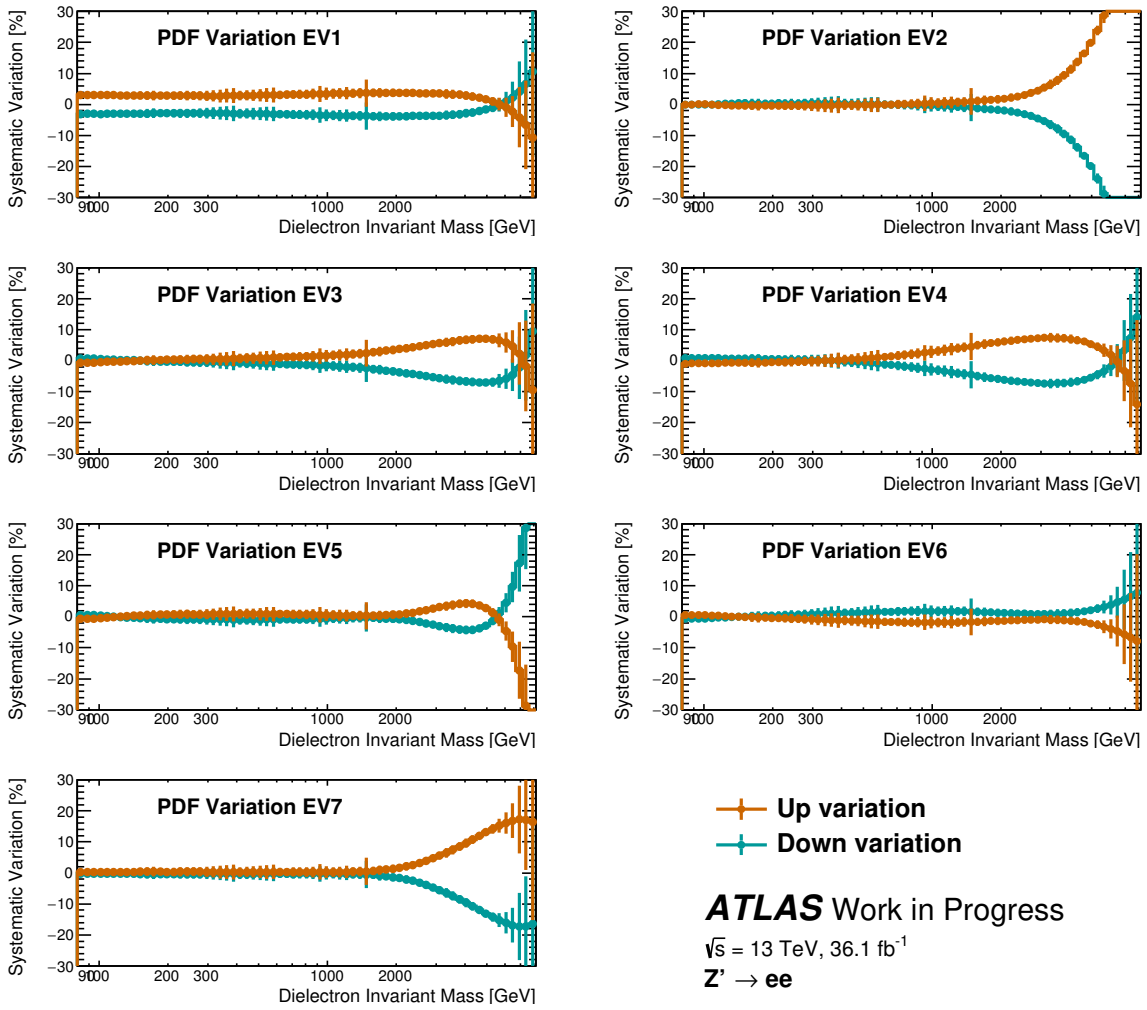


Figure 6.5: The seven PDF eigenvector up (down) variations are shown as a function of the dielectron invariant mass shown in orange (cyan).

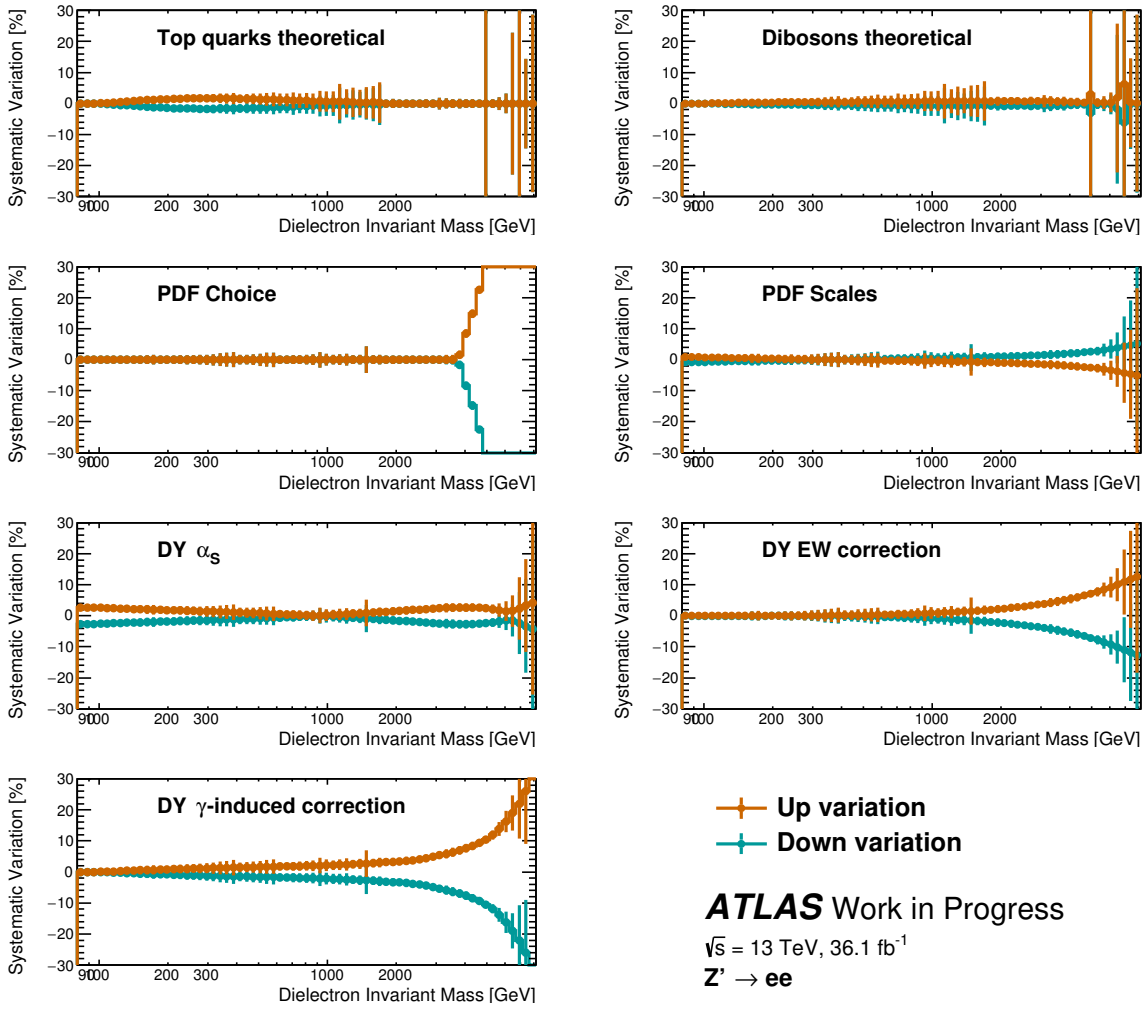


Figure 6.6: Theoretical uncertainties on the top-quark and diboson backgrounds are shown in the top row. PDF choice and scale uncertainties are shown in the second row from the top. Systematic uncertainties corresponding to α_S variation, electroweak correction, and photon-induced correction on the Drell-Yan background are shown in the bottom two rows. The up (down) variations are shown in orange (cyan) points.

6.7.2 Experimental uncertainties

The experimental uncertainties mostly come from lepton trigger, identification, isolation, and reconstruction efficiencies, lepton energy scale and resolution, pileup effects. Luminosity and beam energy uncertainties are also considered. Same sources of uncertainties are considered on all the background processes as well as on the signal. All systematic variations coming from the experimental sources are shown in Figure 6.7.

Electron isolation and identification uncertainties The electron identification efficiency is evaluated from the $Z \to ee$ events and extrapolated to higher masses. The uncertainty in the electron identification is estimated from the shower shape differences in the EM calorimeters between data and MC simulation in the $Z \to ee$ peak and propagated to higher E_T region. The effect of then

electron identification efficiency is independent of E_T above E_T = 150 GeV and is found to be around 2%. Isolation efficiencies on the other hand depend on the electron E_T and the uncertainties are estimated for electrons with $150 < p_T < 500$ GeV and above 500 GeV separately. Large electron isolation uncertainties occur in the combined calorimeter and track based isolation. The track-only isolations have smaller uncertainties.

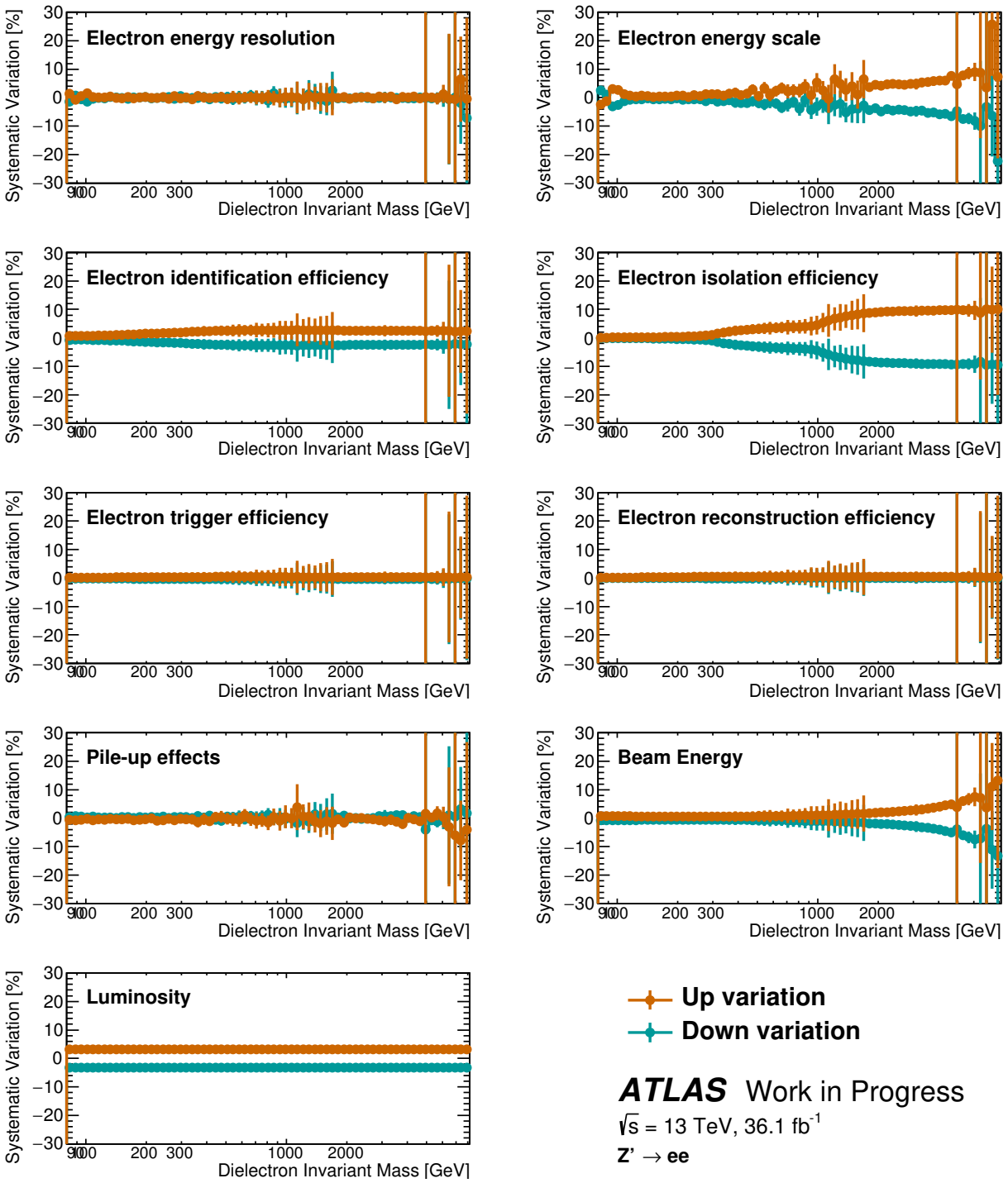


Figure 6.7: Experimental systematic uncertainties as a function of dielectron invariant mass. Top three rows shows the electron energy resolution, scale, identification efficiency correction, isolation efficiency correction, trigger efficiency correction and reconstruction efficiency correction uncertainties. Bottom two rows show the uncertainties coming from pile-up, beam energy and luminosity.

Pileup, Luminosity and Beam energy uncertainties The pileup uncertainties are obtained by varying the pileup reweighting in the MC events. The pileup uncertainty is estimated to be very small. An uncertainty of 0.65% due to beam energy is included. A flat luminosity uncertainty of 3.2% is used for the combined 2015 and 2016 dataset.

6.7.3 Systematics summary

All the systematic uncertainties and their contributions are summarized in Table 6.9. The electron isolation uncertainty contributes the most out of all the experimental uncertainties. It contributes around 9.1% (9.7%) at m_{ee} = 2 TeV (m_{ee} = 4 TeV). The dominant theory uncertainty is the PDF eigenvector variation uncertainty coming from the DY background. It is around 8.7% (19%) at m_{ee} = 2 TeV (m_{ee} = 4 TeV).

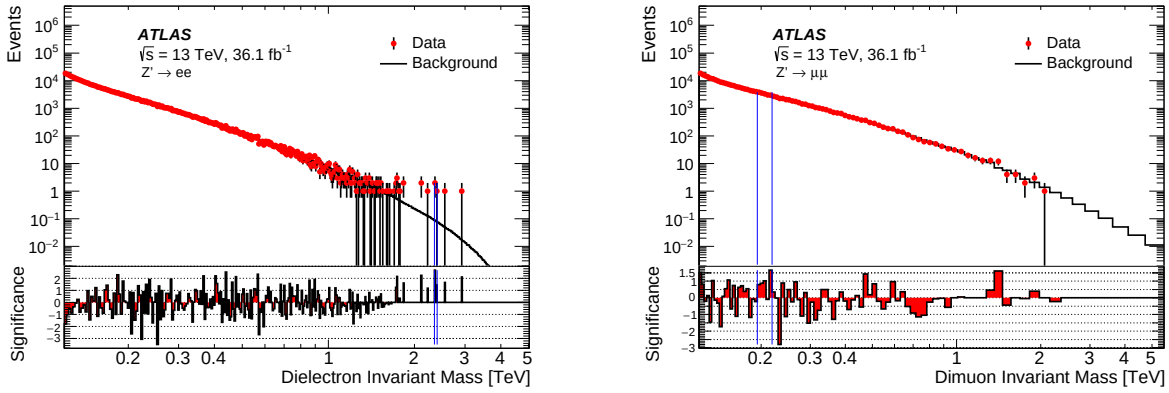
Table 6.9: Summary of the pre-marginalised relative systematic uncertainties in the expected number of events at dilepton masses of 2 TeV and 4 TeV. The values reported in parenthesis correspond to the 4 TeV mass. The reported values for the background represent the relative change in yields in the corresponding $m_{\ell\ell}$ histogram bin containing $m_{\ell\ell}$ = 2 TeV (4 TeV). For the signal uncertainties the values were computed using a Z'_χ signal model with a pole mass of 2 TeV (4 TeV) by comparing yields in the core of the mass peak (within the full width at half maximum) between the distribution varied due to a given uncertainty and the nominal distribution. "-" represents cases where the uncertainty is not applicable.

Source	**Dielectron channel** [%]	
	Signal	Background
Luminosity	3.2 (3.2)	3.2 (3.2)
MC statistical	<1.0 (<1.0)	<1.0 (<1.0)
Beam energy	2.0 (4.1)	2.0 (4.1)
Pile-up effects	<1.0 (<1.0)	<1.0 (<1.0)
DY PDF choice	-	<1.0 (8.4)
DY PDF variation	-	8.7 (19)
DY PDF scales	-	1.0 (2.0)
DY α_S	-	1.6 (2.7)
DY EW corrections	-	2.4 (5.5)
DY γ-induced corrections	-	3.4 (7.6)
Top quarks theoretical	-	<1.0 (<1.0)
Dibosons theoretical	-	<1.0 (<1.0)
Reconstruction efficiency	<1.0 (<1.0)	<1.0 (<1.0)
Isolation efficiency	9.1 (9.7)	9.1 (9.7)
Trigger efficiency	<1.0 (<1.0)	<1.0 (<1.0)
Identification efficiency	2.6 (2.4)	2.6 (2.4)
Lepton energy scale	<1.0 (<1.0)	4.1 (6.1)
Lepton energy resolution	<1.0 (<1.0)	<1.0 (<1.0)
Multi-jet & W+jets	-	10 (129)
Total	10 (11)	18 (132)

6.8 BumpHunter search method

The model independent search in the di-lepton analysis is performed using the BumpHunter described in section 5.4. The search is performed on the invariant mass spectrum of di-electron and di-muon, separately. In this resonance search, signal is expected to appear as a bump on the background spectrum, so the BumpHunter algorithm is used to search for a potential excess. The minimum search interval is two bins and the maximum search interval is half of the invariant mass spectrum. Some additional studies related to the BumpHunter algorithm is done in the context of dilepton resonance search and one such study is summarized in subsection 5.4.1.

The final search results in both di-electron and di-muon channels are shown in Figure 6.8. Different binning are used in the two channels. Since the dilepton mass resolution is higher in the electron channel, finer bins are used in the electron channel. The local p-values are always less than 3σ. The global p-values are calculated using 10,000 pseudo-experiments. The global p-values are 0.71 σ and 0.94 σ in the electron and muon channel, respectively. So, we can conclude that there is no significant deviation in the observed data over the SM background hypothesis.



(a) Dielectron channel

(b) Dielectron channel

Figure 6.8: Dilepton mass distributions in the (a) di-electron and (b) di-muon channel are shown. The data are shown in red points and the total nominal background is shown in back histogram. The lower panel shows the bin-by-bin local significance of the BumpHunter test statistic. The most significant interval is indicated by the vertical blue lines. The local significance values are almost always less than 3σ, indicating no significant deviation [206].

6.9 Limit setting

In absence of a new signal, limits are set at 95% CL on the BSM Z' signal cross-section × branching ratio (BR). The upper limits on the cross-section × BR (σ B) are interpreted as the lower limits on the Z' pole mass using the Z' theory cross-sections. A binned likelihood is built from the dilepton invariant mass distributions. The $m_{\ell\ell}$ region from 80 GeV to 6 TeV is included in the likelihood. The 80 GeV $< m_{t\bar{t}} <$ 120 GeV region is used as a control region and included in the likelihood as a single bin. This region is used to fix the normalization of the DY background and help constrain the systematic uncertainties. All the systematic uncertainties, described in subsection 5.2.1, are included in the likelihood as nuisance parameters. The signal strength, μ,

defined as the ratio of the observed signal cross-section times branching ratio (σ B) and its values predicted by theoretical models, is used as the main parameter of interest. Upper limits on σ B are set using Bayesian approach and the calculations are performed with the Bayesian Analysis Toolkit (BAT) [226]. BAT uses a Markov Chain Monte Carlo (MCMC) technique to compute the marginal posterior probability density of the parameter of interest. An uniform prior probability for the parameter of interest was used. The upper limits on the Z' σ B as a function of Z' mass ($M_{Z'}$) for the di-electron channel is shown in Figure 6.9. Theory cross-sections of two E_6 models, Z'_χ, Z'_ψ and the Z'_{SSM} are shown on the limit plot. The σ B limit becomes weaker above 3.5 TeV due to rapidly falling signal cross-section and increasing proportion of signal being produced off-shell. All the lower limits on Z' mass are summarized in Table 6.10. The limits are slightly weaker in the di-muon channel due to worse mass resolution than the electron channel. Combined dilepton mass limit on Z'_χ and Z'_{SSM} model are 4 TeV and 4.5 TeV respectively.

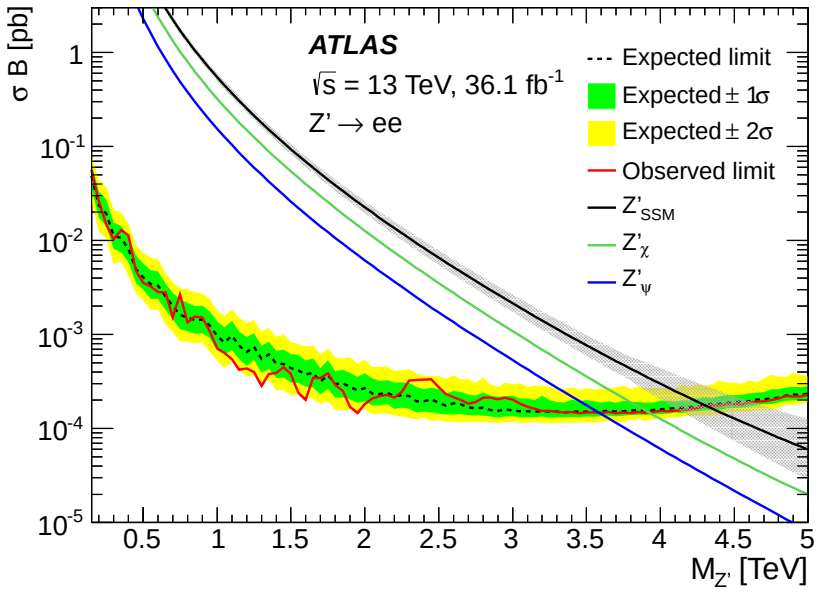


Figure 6.9: Upper 95% CL limit on Z' signal cross-section × branching ratio as a function of Z' pole mass ($M_{Z'}$). The results are shown for the di-electron channel. The red line represents the observed limit whereas the black dotted line shows the expected limit. The green and yellow bands represent the expected limit $\pm 1\sigma$ and $\pm 2\sigma$ bands respectively. The theory cross-sections lines for three different signal models are shown in solid black, blue and green lines [206].

6.10 Conclusion

A search for a new high mass resonance decaying into a pair of electrons is performed with partial Run-2 proton-proton collision data corresponding to 36 fb^{-1}. Events with two high p_T good quality isolated electrons are selected. The invariant mass of the two-electron system is used as the discriminant in this analysis. The narrow width BSM signals are expected to appear as a bump on the SM dilepton mass spectrum. We do not find any hint of a new resonant signal in the data; the observed di-electron mass spectrum is consistent with the SM prediction within the statistical and

Table 6.10: Observed and expected 95% CL lower mass limits for different Z' narrow width models. The widths are quoted as a percentage of the resonance mass. Limits are shown for di-electron, di-muon and combined dilepton channel [206].

Model	Width [%]	θ_{E_6} [rad]	Lower limits on $M_{Z'}$ [TeV]					
			ee		$\mu\mu$		$\ell\ell$	
			Obs	Exp	Obs	Exp	Obs	Exp
Z'_{SSM}	3.0	-	4.3	4.3	4.0	3.9	4.5	4.5
Z'_{χ}	1.2	0.50 π	3.9	3.9	3.6	3.6	4.1	4.0
Z'_{S}	1.2	0.63 π	3.9	3.8	3.6	3.5	4.0	4.0
Z'_{I}	1.1	0.71 π	3.8	3.8	3.5	3.4	4.0	3.9
Z'_{η}	0.6	0.21 π	3.7	3.7	3.4	3.3	3.9	3.8
Z'_{N}	0.6	$-0.08\ \pi$	3.6	3.6	3.4	3.3	3.8	3.8
Z'_{ψ}	0.5	0 π	3.6	3.6	3.3	3.2	3.8	3.7

systematic uncertainties. So, upper limits on the signal production cross-section × branching ratio are set with 95% CL. These cross-section limits are interpreted as the lower limit on the pole mass of the BSM signal. A lower limit on the Z'_{SSM} (3% width) mass is set at 4.3 TeV in the di-electron channel. The limit on Z'_{SSM} mass was improved by 1.1 TeV in this analysis. The mass limits on the E_6 model are weaker than the Z'_{SSM}. The lower limit on the Z'_{ψ} (0.5% width) mass is 3.6 TeV. The events with two electrons in the final states give a clean detector signature, and the reconstruction efficiency is high. So, there is less room for improvement. Hence, the analysis sensitivity increase with full Run-2 data was expected to be mostly driven by the increase in luminosity. So, I did not continue working on this search. The dilepton resonance search results with the full Run-2 data are available now [227]. The mass limit on the Z'_{SSM} improved roughly by 600 GeV. The new analysis used several new techniques and estimated the total SM background contribution using a data-driven approach. Overall the sensitivity improved by a factor of 1.5 above the m_{ee} of 3.5 TeV. But below 3.5 TeV the sensitivity is comparable to the analysis presented in this chapter.

CHAPTER

SEVEN

Search for high mass resonances in boosted all-hadronic top-antitop final sate

This chapter summarizes a $t\bar{t}$ resonance search in the all-hadronic final state, where both the top-quarks decay hadronically. This search focuses on the final state with high momentum (boosted) top quarks. The search is performed with full Run-2 proton-proton collision data (139 fb^{-1}) collected during the years 2015 – 2018 by the ATLAS detector at the LHC operating at $\sqrt{s}$ = 13 TeV. The $t\bar{t}$ final state is promising as a potential site of new physics as described in section 1.5. Among the different $t\bar{t}$ decay modes, the all-hadronic final state has the highest branching ratio, hence it is one of the favourite channels to search for new physics. The analysis is looking for narrow width resonances decaying into $t\bar{t}$ pairs and the results are interpreted in terms of the benchmark BSM models described in section 1.5. The results were published in 2020 [228]. No evidence for new physics was observed in the data, and limits were set on the Z'_{TC2} production cross-section × branching ratio (BR) as well as on the mass of Z'_{TC2} with a 95% confidence level (CL).

This chapter is organized as follows. A brief overview of the analysis is provided in section 7.1. The all-hadronic top-decay topology is discussed in section 7.2. The modeling of the SM background processes and BSM signal processes are described in section 7.3 and section 7.4, respectively. The analysis involves different physics objects like jets, electrons, muons, and missing transverse energy (MET). All these physics objects are defined in section 7.5. The event selections are summarized in (section 7.6). The categorization of the selected events in different signal and control regions is described in section 7.7. Modelling of the signal shape with analytic functions is described in section 7.8. The final background estimation is done using data, and the data-driven method is described in section 7.9. The observed distribution of the final discriminant is presented in section 7.10. The systematic uncertainties considered in this search are described in section 7.11. The final statistical analysis results are presented in section 7.12 and section 7.13. The search results are also interpreted using a dark matter mediator model and the results are summarized in section 7.14.

7.1 Analysis overview

The analysis presented in this chapter targets the boosted all-hadronic $t\bar{t}$ final state with two high momentum top-quarks. The high momentum top-quarks form two large-radius jets as described in section 7.2. An example Feynman diagram of the signal process ($Z' \to t\bar{t} \to qqbqqb$) is shown in Figure 7.1. In our case, the Z' would manifest as a narrow resonance through its decay into two top quarks which is expected to appear as a narrow bump on the $t\bar{t}$ invariant mass spectrum. That's why the analysis uses the invariant mass spectrum of two top quarks as the main discriminant. The invariant mass is defined as follows

$$m^2_{t\bar{t}} = (p_{J_1} + p_{J_2})^2. \tag{7.1}$$

Here p_{J_1} and p_{J_2} are the momentum 4-vectors of the two highest p_T top-tagged large-R jets passing all the selection criteria. In this search, the major background contributions come from other high cross-section SM multijet processes and the SM $t\bar{t}$ process. In this chapter, 'Multijet background' will refer to the background originating from the multijet production but not resulting in the production of top quarks. The hadronically decaying boosted top-quarks are identified using a DNN-based top-tagger. To further suppress the non-$t\bar{t}$ backgrounds and increase the sensitivity, b-quarks coming from the top-quark decay are identified using a b-tagging technique. The signal regions are defined based on the number of b-tagged jets within a top-tagged jet. A fully data-driven approach is taken in this analysis to model the backgrounds. The invariant mass spectrum of the SM background processes is obtained by fitting an analytic function to the data. Then a model-independent bump-hunting algorithm is used to search for signals on the invariant mass spectrum. A profile-likelihood based signal model-dependent statistical search is also performed. Since no trace of new physics is observed, upper limits on the signal cross-section (σ) × branching ratio are set at 95% CL using the profile-likelihood based approach. A previous version of this search was performed with 36.1 fb^{-1} LHC Run-2 proton-proton collision data and a narrow width Z'_{TC2} ($\Gamma/m = 1.2\%$) boson of mass less than 3.1 TeV was excluded at 95% CL [229].

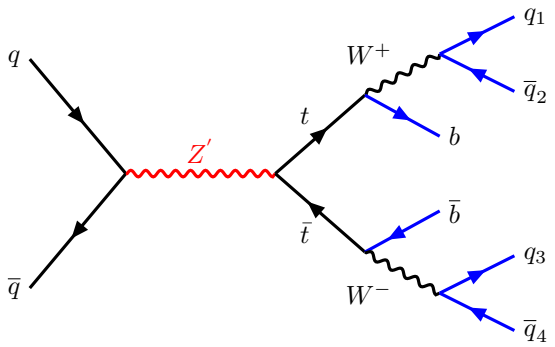


Figure 7.1: Feynman diagram of the Z' signal process where both the top quarks decays hadronically.

7.2 Top quark decay topology

The fully hadronic channel, where both W bosons decay hadronically, has a branching fraction of 46%. The six quarks in the final state appear as six jets in the ATLAS detector, but the event

topology depends on the momenta of the decaying top-quarks. For low transverse momentum top-quarks, the decay products are spread out, and the jets are far from each other. In this case six distinct small-radius ($R = 0.4$) jets are formed in the ATLAS calorimeter. This topology is called the *resolved topology*. If the top quark has high momentum, then the daughter particles become collimated along the top momentum due to the Lorentz boost. Hence, the hadrons coming from the top-decay become very close to each other, and their responses overlap in the calorimeter forming a large-radius jet with radius parameter (R) of 1. This is called the *boosted topology*. These two top decay topologies are shown schematically in Figure 7.2. This analysis focuses on the boosted top-decay topology where two large-R jets are expected in the final state.

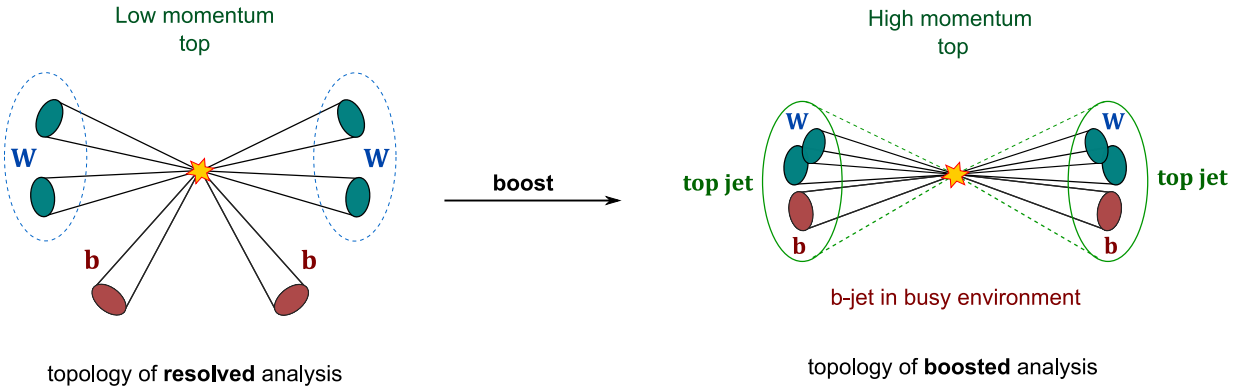


Figure 7.2: Schematic diagram showing the all-hadronic $t\bar{t}$ final state in resolved and boosted regime.

7.3 Background processes

In this search, the total background is estimated directly from data using a functional fit to the data. The functional form and other relevant studies are done on the simulated background as a proxy to the data. This section summarizes the background processes and how they are simulated. The major background contribution comes from standard model $t\bar{t}$ and multijet processes. The contributions from other processes like single top ($t+W$), boson (V) +jets and $t\bar{t}+W$ are negligible and not considered in this search. The modeling of the two major backgrounds is described below.

7.3.1 Standard model $t\bar{t}$ process

Standard model $t\bar{t}$ events are produced using the next-to-leading-order (NLO) generator Powheg-Box v2 [208, 230, 231] with the NLO NNPDF3.0 [225] PDF set in the matrix element calculations. The next-to-next-to-leading-order (NNLO) $t\bar{t}$ production cross-section in QCD is calculated using Top++2.0 [98, 218, 232–236]. It includes the resummation of next-to-next-to-leading logarithmic (NNLL) soft gluon terms. Parton showering, hadronization, and the underlying event are simulated using Pythia v8.230 with the LO NNPDF2.3 PDF set and the A14 set of tuned parameters. The h_{damp} parameter, which controls the transverse momentum of the first additional parton emission beyond the Born level, was set equal to the top-quark mass [237]. The main effect of this parameter is to regulate the high-p_T emission against which the $t\bar{t}$ system recoils. The top-quark kinematics in $t\bar{t}$ events are further corrected to account for electroweak (EW) higher-order effects [238]. This correction is used as an event weight, so the generated events are weighted by

this correction factor. The weights are derived as a function of the flavor and center-of-mass energy of the initial partons, as well as the decay angle of the top quarks in the center-of-mass frame of the initial partons. The value of the EW correction is 0.98 at $m_{t\bar{t}} = 0.4$ TeV and increases to 0.87 at $m_{t\bar{t}} = 3.5$ TeV.

7.3.2 Standard model multijet processes

The SM multijet process is simulated using Pythia 8.186 [210] with two jets in the hard process at leading order (LO) with the NNPDF23LO [220] PDF set and the A14 set of tuned parameters for the underlying event. The samples are produced in 13 slices of the leading truth jet p_T to ensure sufficient MC statistics over the wider range of p_T. The modeling of the multijet process in the MC simulation is not accurate and it is also estimated using a data-driven approach as described in section B.1.

Simulated multijet and $t\bar{t}$ processes are used to study different data-driven methods to estimate all or part of the background. The simulations are not used to estimate the final background used in this analysis. Details of these major SM background simulations are summarized in Table 7.1.

Table 7.1: Summary of MC sample information for background processes used in the boosted all-hadronic $t\bar{t}$ resonance search.

Process	Generator	Order	Corrected	Parton Shower	PDF
$t\bar{t} \to q_1 q_2 b\ q_3 q_4 \bar{b}$	Powheg v2 [208]	NLO	NNLO	Pythia 8.230 [210]	NNPDF3.0 NLO [225]
$pp \to jj$	Pythia 8.186 [210]	LO	LO	Pythia 8.186 [210]	NNPDF23LO [220]

7.4 Signal processes

The benchmark signal models are simulated using Monte-Carlo simulation. Signal interference with SM $t\bar{t}$ is not considered for the Z' signal model due to its large model dependence and minimal effect. The primary benchmark signal model is the Top-color assisted-Technicolor model [115, 116]. The other benchmark model is the Heavy Vector Triplet model [239], and it is used for combining the results with other searches in ATLAS. A summary of both the signal models is given in this section.

7.4.1 Topcolor-assisted-Technicolor model

Topcolor-assisted-technicolor (TC2) $Z'_{\mathrm{TC2}} \to t\bar{t}$ signal [115, 116] is simulated using Pythia v8.186 [210] with the default settings as a generic Sequential Standard Model (SSM) Z'_{SSM} process. The cross-sections are rescaled to the TC2 values while using them in the analysis. The samples are generated for a set of Z' masses ranging from 400 GeV to 5 TeV, using the LO NNPDF2.3 [220] PDF set with the A14 set of tuned parameters for parton shower and hadronization. The generated sample width is 1.2% of its mass, which ensures that the generated width is narrower than the reconstructed resolution. The leading-order (LO) cross-sections at 13 TeV used for the Z'_{TC2} are presented in Table 7.2. A multiplicative k-factor of 1.3 is used following [117] to bring the cross-sections to the next-to-leading order (NLO). The same method was used in the $t\bar{t}$ resonance search

with 36 fb^{-1} data [229]. The interference with SM $t\bar{t}$ is negligible and not considered in this analysis. In this analysis, only Z'_{TC2} samples above mass 1.4 TeV are used.

Table 7.2: Leading-order theoretical cross-sections for the $Z'_{\mathrm{TC2}} \to t\bar{t}$ signal, before applying the k-factor of 1.3.

Z'_{TC2} mass	LO cross-section [pb]
400 GeV	70.3
500 GeV	40.1
750 GeV	10.7
1 TeV	3.70
1.25 TeV	1.51
1.5 TeV	0.684
1.75 TeV	0.334
2. TeV	0.172
2.25 TeV	0.0924
2.5 TeV	0.0511
2.75 TeV	0.0289
3 TeV	0.0167
4 TeV	0.00213
5 TeV	0.000331

7.4.2 Heavy Vector Triplet model

Heavy Vector Triplet (HVT) $Z'_{\mathrm{HVT}} \to t\bar{t}$ signal [239] is another signal model which will be used for the combination of the results from different $t\bar{t}$ search channels. The HVT signal process is simulated using MadGraph5_aMC@NLO v2_6_0 [240] with the LO NNPDF2.3 PDF set, interfaced with Pythia v8.212 with the A14 set of tuned parameters for parton shower and hadronization. The Z'_{HVT} signal cross sections are given in Table 7.3. The total decay widths are around 2.5% for the mass range considered in this study. The HVT samples are generated only for three pole masses. To get the intermediate mass points, matrix-element (ME) reweighting using Z'_{TC2} samples as inputs is studied. Although the two processes share the same initial and final states the total width of the Z'_{HVT} is consistently smaller. The same PDFs are used for the two processes so we only need to consider ME reweighting. Different event generators are used for these processes, Z'_{TC2} signal are simulated using Pythia8, whereas Z'_{HVT} MadGraph5_aMC@NLO. Different approximations are used in the two different generators, making the reweighting difficult.

A summary of the signal processes used in this search is presented in Table 7.4.

7.5 Object definition

Jets are the most relevant physics objects in this search. They are used to reconstruct hadronically decaying top-quarks and to identify the b-quark coming from top-quark decay. Leptons, on the other hand, are not directly used in this search. They are used to veto events to guarantee that the sample of selected events is orthogonal to that used in the $t\bar{t}$ resonance search in the lepton+jets channel described in chapter 8. All these physics objects are described in this section.

Table 7.3: Leading-order theoretical cross-sections for only the fully hadronic $t\bar{t}$ final state, $Z'_{\mathrm{HVT}} \to t\bar{t} \to q\bar{q}'b + q\bar{q}'b$ signal including all branching fractions.

Z'_{HVT} mass	LO cross-section [pb]
1 TeV	0.441
4 TeV	2.08×10^{-4}
8 TeV	4.91×10^{-8}

Table 7.4: Summary of Monte Carlo sample information for signal processes used in the boosted all-hadronic $t\bar{t}$ resonance search. The columns from left to right give the process of interest, generator, matrix element order, parton shower program, and PDF utilized.

Process	Generator	Order	Corrected	Parton Shower	PDF
$Z'_{\mathrm{SSM}} \to t\bar{t}$	Pythia 8.186 [210]	LO	NNLO	Pythia 8.186 [210]	NNPDF23LO [220]
$Z'_{\mathrm{HVT}} \to t\bar{t}$	MG5_aMC@NLO v2_6_0 [240]	LO	LO	Pythia 8.186 [210]	NNPDF23LO [220]

Large radius calorimeter jets

Large radius jets with $R = 1$ are used in this analysis to reconstruct the boosted top quarks. Large-R jets are built using three-dimensional topological energy clusters calibrated with the local cluster weighting (LCW) procedure, described in section 4.2.3. The reconstructed jets are trimmed by building the subjets coming from the soft radiations using the radius parameter $R_{\mathrm{trim}} = 0.2$, and the subjets with p_T less than 5% of the p_T of the parent large-R jet (f_{cut}) are discarded. Large-R jets are required to have p_T greater than 200 GeV, $|\eta| < 2.0$ and mass greater than 40 GeV. The large-R jet definition is summarized in Table 7.5. Large-R jets were further used in this analysis for identifying top jets using the DNN-based top-tagging algorithm. Details of the top-tagging method were described in subsection 4.11.3.

Table 7.5: Summary of the large-R jet object definition.

Selection	**Criteria**
η coverage	$\|\eta\| < 2.5$
Input constituent	LCTopo
Grooming algorithm	Trimming
Trimming radius	$R_{\mathrm{trim}} = 0.2$
Trimming fraction	$f_{\mathrm{cut}} = 0.05$
Momentum	$p_T > 200$ GeV
Overlap removal	Not applied

Small radius calorimeter jets

Small-R jets of $R = 0.4$ are built using the three-dimensional topological clusters, calibrated to the electromagnetic energy scale. Small-R jets are required to have $p_T > 25$ GeV and $|\eta| < 2.5$. Jet cleaning is applied to remove the beam backgrounds and detector problems. A high-efficiency working point, called the *loose* working point [241], is used for jet cleaning. To suppress pile-up

jets, a JVT cut is applied and the medium operating point is used. The medium operating point requires a jet JVT > 0.59 if the jet $p_T < 60$ GeV and $|\eta| < 2.4$. The medium operating point has 92% average selection efficiency.

Table 7.6: Summary of the small-R jet object definition.

Selection	Criteria
η coverage	$\|\eta\| < 2.5$
Input constituent	EMTopo
Jet cleaning	Loose
Momentum	$p_T > 25$ GeV
JVT	Medium, JVT> 0.59
Overlap removal	$\Delta R(\text{e}, \text{jet}) > 0.4$
Overlap removal	$\Delta R(\mu, \text{jet}) > \min\left(0.4, 0.04 + 10\text{ GeV}/p_T^{\mu}\right)$

Variable radius (VR) track jets:

Track jets are built using the anti-k$_t$ algorithm using the charged particle tracks in the inner detector. The tracks used in the jet reconstruction are required to be associated with a primary vertex, and to have $p_T > 500$ MeV and $|\eta| < 2.5$ as described in section 4.8. Instead of using a fixed radius parameter, a p_T dependent radius parameter is used as defined in Equation 4.4 with $\rho = 30$ GeV, $R_{\text{max}} = 0.4$ and $R_{\text{min}} = 0.02$. By construction VR track jets can overlap on each other. The overlapping VR track jets in an event are removed. The deep learning based b-tagging technique is used to identify the VR tracks jets to identify the b jets.

Table 7.7: Summary of the variable-R track jet object definition.

Selection	Criteria
η coverage	$\|\eta\| < 2.5$
Input constituent	tracks
Momentum	$p_T > 10$ GeV
Minimum radius	$R_{\text{min}} = 0.2$
Maximum radius	$R_{\text{max}} = 0.4$

The small-R jets are further required not to overlap with electrons and muons defined below. It is required that the ΔR (e, jet) > 0.4 [14] and $\Delta R(\mu, \text{jet}) > \min\left(0.4, 0.04 + 10\text{ GeV}/p_T^{\mu}\right)$. The jet definitions are summarized in Table 7.6.

Electrons

Electrons are only used to veto events. The reconstruction of an electron object is described in section 4.3. The electrons are reconstructed from EM calorimeter clusters matched to an ID track within $|\eta| < 2.47$ excluding the transition region between the barrel and the endcap calorimeter located at $1.37 < |\eta| < 1.52$. Electrons coming from the bad calorimeter clusters and noise bursts are removed. The transverse energy (E_T) of the electrons is required to be greater than 25 GeV.

[14] $\Delta R = \sqrt{(\Delta\phi)^2 + (\Delta y)^2}$

To ensure that the electron candidates are associated with the correct primary vertex, requirements are applied to the impact parameters. The d_0^{BL}-significance is required to be within $[-5, 5]$, $|d_0^{\mathrm{BL}}/\sigma_{d_0}| < 5.0$. The longitudinal impact parameter with respect to the beamline is required to be within -0.5 mm and 0.5 mm, $|z_0 \sin\theta| < 0.5$ mm. Then the electrons are required to pass the tight likelihood-based identification criteria. The electron candidates are not required to be isolated from the hadronic activity in the event. This choice is made to be orthogonal to the $t\bar{t}$ resonance search in the lepton + jets final state. The efficiency of the tight identification criteria varies from 67% at $E_T = 25$ GeV to 99% at $E_T = 100$ GeV. The electron definition is summarized in Table 6.1.

Table 7.8: Summary of the electron object definition.

Selection	**Criteria**
η coverage	$\lvert\eta\rvert < 2.47$ without $1.37 < \lvert\eta\rvert < 1.52$
Cleaning	Remove bad calorimeter clusters
Energy	$E_T > 25$ GeV
d_0-significance	$\lvert d_0^{\mathrm{BL}}/\sigma_{d_0}\rvert < 5.0$
$z_0 \sin\theta$	$\lvert z_0 \sin\theta\rvert < 0.5$ mm
Identification	LH Tight
Isolation	Not applied

Muons

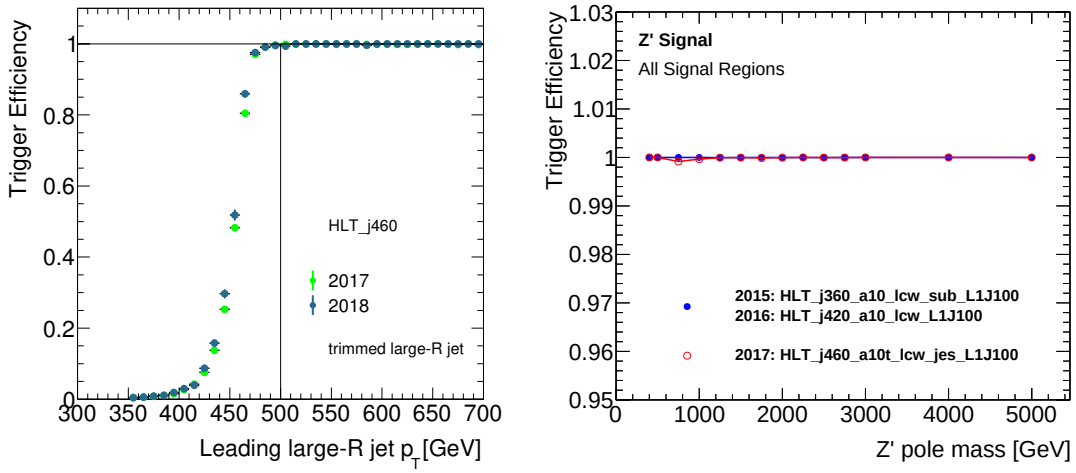
Muons are also used only to veto events. The reconstruction of a muon object in ATLAS is described in section 4.4. Combined muons are used in this analysis and are reconstructed using both ID and MS tracks within $|\eta| < 2.5$. Similar to the electrons, the muons are required to have p_T greater than 25 GeV. Conditions are imposed on the impact parameters to ensure that muons are associated with the correct primary vertex. The d_0^{BL}-significance is required to be within $|d_0^{\mathrm{BL}}/\sigma_{d_0}| < 3.0$. The longitudinal impact parameter with respect to the beamline is required to be within $|z_0 \sin\theta| < 0.5$ mm. The muons are required to satisfy the medium identification criteria. The muons are further required to be isolated from hadronic activities. In this case, track-only tight isolation criteria are used for muons. The muon identification efficiency is around 99% for muons with $p_T = 100$ GeV. The muon definition is summarized in Table 7.9.

Table 7.9: Summary of the muon object definition.

Selection	**Criteria**
η coverage	$\lvert\eta\rvert < 2.5$
Momentum	$p_T > 25$ GeV
d_0-significance	$\lvert d_0^{\mathrm{BL}}/\sigma_{d_0}\rvert < 3.0$
$z_0 \sin\theta$	$\lvert z_0 \sin\theta\rvert < 0.5$ mm
Identification	Medium
Isolation	Tight track-only

7.6 Event selection

The data events are required to pass the good run list. The events are then cleaned by removing detector noise. As we are searching for hadronically decaying top quarks pairs, two high p_T large-R jets are required. The selected events are required to pass the single large-R jet trigger. The single large-R jet trigger requires at least one large-R jet with $p_T > 360 - 460$ GeV depending on data taking periods. In 2015 and 2016, the untrimmed jets were used for triggering. During the later years, trimmed jet triggers were available. The analysis uses trimmed large-R jets to reconstruct the top-quark, so triggering on trimmed jets is preferred. To make sure the selected events are in the trigger plateau, the leading p_T large-R jet is required to have $p_T > 500$ GeV. The trigger efficiency of individual jets as a function of jet p_T and the trigger efficiency as a function of Z' pole mass is shown in Figure 7.3. These plots are showing that the selected events are in the trigger plateau.



(a) Leading Large-R jet p_T

(b) Z'_{TC2} pole mass

Figure 7.3: (a) Per-jet trigger efficiency as a function of trimmed large-R jet p_T for 2017 and 2018 data. Credit: Josu Cantero. (b) Trigger efficiency as a function of Z'_{TC2} pole mass. Almost 99.9% efficiency observed after applying leading large-R jet $p_T > 500$ GeV.

The two highest p_T large-R jets, J_1 and J_2 are required to be top-tagged by the DNN top tagging algorithm with 80% signal identification efficiency. The top-tagger is trained for a fully contained top decay topology, where all three partons coming from the top-quark decay are contained within the radius of the large-R jet. The top-tagger requires the large-R jet to have $p_T > 350$ GeV and mass > 40 GeV. As a result, the second-highest p_T jet is required to have $p_T > 350$ GeV. The tagger is calibrated for jet momentum up to 4 TeV. The scale factors are extrapolated outside the range of the calibration for the jets with higher p_T.

Trigger plateau

The physics objects used for triggering events are built on the fly during data taking. The reconstruction of these objects is not the same as the off-line reconstruction. The jet triggers used in this analysis operate above a certain p_T threshold, the *on-line threshold*. The triggers are fired if the identified jet has p_T greater than the on-line threshold. In this search, we use large-R jet triggers. In the schematic diagram of Figure 7.4 the trigger efficiency is shown as a function of large-R jet p_T. Due to jet p_T resolution effects and differences between on-line and off-line jets, the efficiency does not have a vertical turn-on at the on-line threshold value. Instead, it takes an S shape, and the trigger efficiency increases up to a point and then flattens out. The point where the trigger efficiency starts to flatten out is called the *off-line threshold*. The region above the off-line threshold, where the trigger efficiency is stable, is called the trigger *efficiency plateau*.

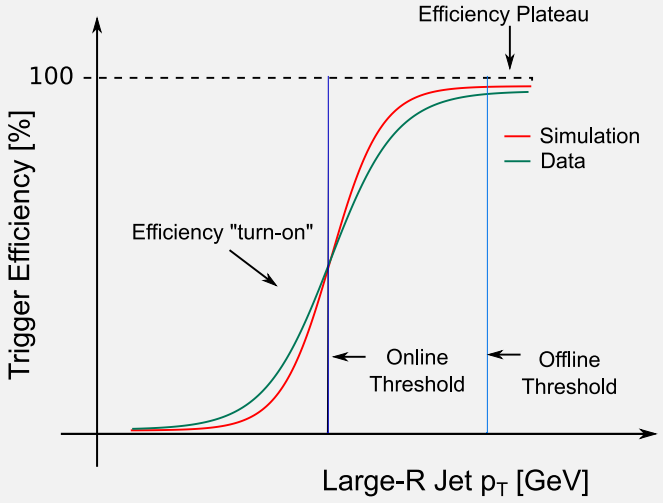


Figure 7.4: Schematic diagram illustrating the trigger efficiency turn-on. The trigger efficiency is shown as a function of large-R jet p_T. The diagram shows the efficiency curve for data (red) and simulation (green). Its usually easy to calibrate the trigger effects in the plateau, so physics analyses often need to use an off-line p_T threshold higher than the on-line trigger threshold.

The region where the trigger efficiency starts increasing from zero to its maximum is called the trigger *turn-on* region. It is difficult to model the trigger turn-on in simulation, so the calibrations are not typically derived for this region. In general, analyses require the triggered objects to have p_T greater than the off-line threshold. Thus it becomes easier to calculate scale-factors to account for differences in the measured trigger efficiencies in data and MC simulation.

In a signal event, the large-R jets are expected to be back-to-back in the transverse detector plane, as they are coming from a heavy particle decay. Hence, the jets are required to have $\Delta\phi(J_1, J_2)$ greater than 1.6. In this analysis, we are searching for s-channel signals, and to suppress the t-channel $t\bar{t}$ events and the other multijet events a rapidity cut, $\Delta\eta(J_1, J_2) < 1.8$, is applied. The events are also required to have at least one b-tagged track jet inside the large-R jet. Finally, the search region is defined to have the invariant mass of the two leading top-tagged large-R jets

be greater than 1400 GeV. This cut is applied to make sure the shape of the mass spectrum is strictly falling, and this is further discussed in section 7.9. The event selections are summarized in Table 8.3.

Table 7.10: Summary of the event selection for the $t\bar{t}$ resonance search in the boosted all-hadronic final state.

Selection	Criteria
GRL	Data events in the Good Run List
Single large-R jet trigger	p_T > 360-460 GeV (depending on run period)
Event cleaning	Remove calorimeter noise, LAr, SCT errors, noise burst
At least 2 large-R jets	$p_{T,J_1} > 500$ GeV and $p_{T,J_2} > 350$ GeV
Back-to-back	$\Delta\phi(J_1, J_2) > 1.6$
Suppress t-channel production	$\Delta\eta(J_1, J_2) < 1.8$
top-tagging	Both J_1 and J_2 with 80% DNN top-tagger (for signal region)
b-tagging	≥ 1 b-tagged track jets (for signal region)
Di-jet mass cut	$m_{J_1 J_2} > 1400$ GeV

7.7 Event categorization

Selected events are further categorized into three orthogonal regions based on b-tagging. Exactly two b quarks are expected in the $t\bar{t}$ signal, and hence our signal regions are created requiring at least 1 b-tagged track jet within the top-jets. Figure 7.5 shows a schematic diagram of the event topologies in the signal and control regions.

7.7.1 Control region (CR0b)

The control region (CR) is defined to validate the multijet process. Events in this region are required to have zero b-tagged track jets within $dR < 1.0$ of the top-tagged large-R jets. Two types of event can contribute in this CR. In the first case, there can be exactly zero b-tagged track jets in the event. Whereas, in the other situation there could be b-tagged jets in the event but none of them are close to the top-jets i.e. within $dR < 1.0$ of the jet axis. These two situations are shown in the bottom row of Figure 7.5 and marked as category '0' and '-1'. This control region is used as a template region for the data-driven estimate of the multijet process described in section B.1 of Appendix B.

7.7.2 Signal regions

There are two signal regions based on the number of b-jets close to the top-tagged large-R jets.

1 b-tag signal region (SR1b)

This signal region is defined to collect events with just one b-tagged track jet within $dR < 1.0$ of the top-tagged large-R jets. Although two b jets are expected in the $t\bar{t}$ process, due to b-tagging inefficiency in the dense environment within a top-jet we often get only one b-tagged jet. This region is created to capture all those cases. Due to looser b-tagging requirement, many multijet events enter in this region. SR1b is shown as category '1' and '2' in Figure 7.5.

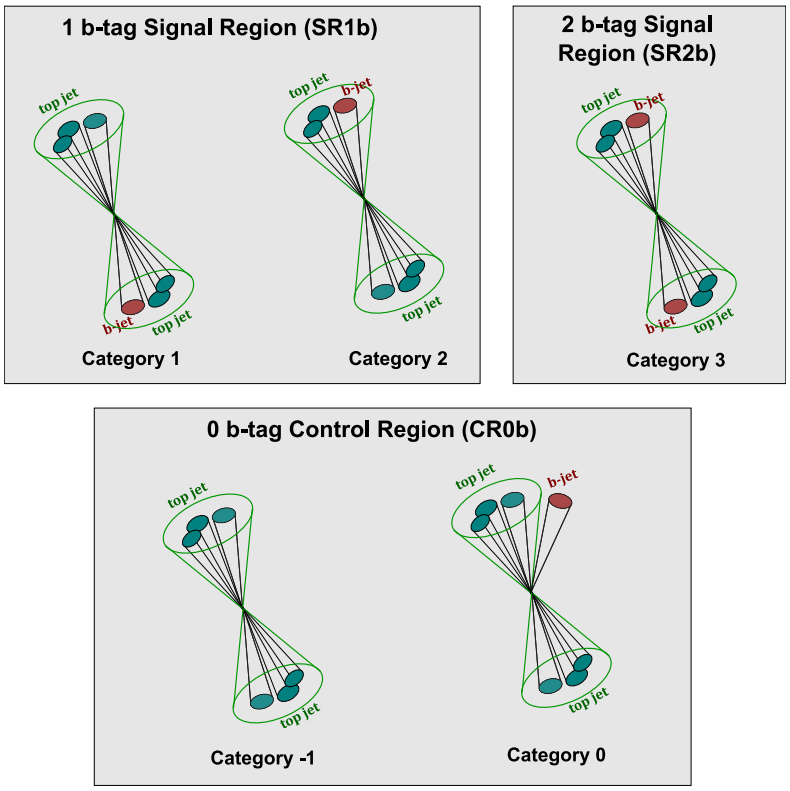


Figure 7.5: Schematic diagram showing the signal and control region definitions with different event categories.

2 b-tag signal region (SR2b)

The 2 b-tag signal region is created to collect the events where both the top-tagged large-R jets are associated with at least one b-tagged track jet. This region is enriched in $t\bar{t}$ events and has the highest sensitivity. It is shown as category '3' in Figure 7.5.

The relative fraction of multijet background is shown in Figure 7.6. Contribution from non-allhadronic top decay (with one or two leptons) is also considered. The contribution is very small, roughly 1-3% of total background. CR0b is mostly dominated by multijet background ($\sim$ 97.4%). SR1b has almost 77% multijet contribution whereas SR2b has the highest $t\bar{t}$ purity, and is enriched to $\sim$ 76% $t\bar{t}$ process. Signal contributions to the signal and control regions are shown in Figure 7.7. Here Z'_{TC2} signal of pole mass 3 TeV (Figure 7.7a) and 4 TeV (Figure 7.7b) is used. SR2b has higher sensitivity as expected.

7.7.3 Validation regions

Validation regions are defined in order to validate the background estimation. The 1 b-tag and 2 b-tag regions with high rapidity difference $|\Delta y(J_1, J_2)| > 1.8$ are very close to the signal regions. The relative fraction of the $t\bar{t}$ and multijet background in these regions are expected to be similar to the signal regions. The signal contamination coming from s-channel processes such as the targeted signal are relatively low. Thus these regions can be used to validate the total background estimation. These two regions are called 1 b-tag validation region (VR1b) and 2 b-tag validation

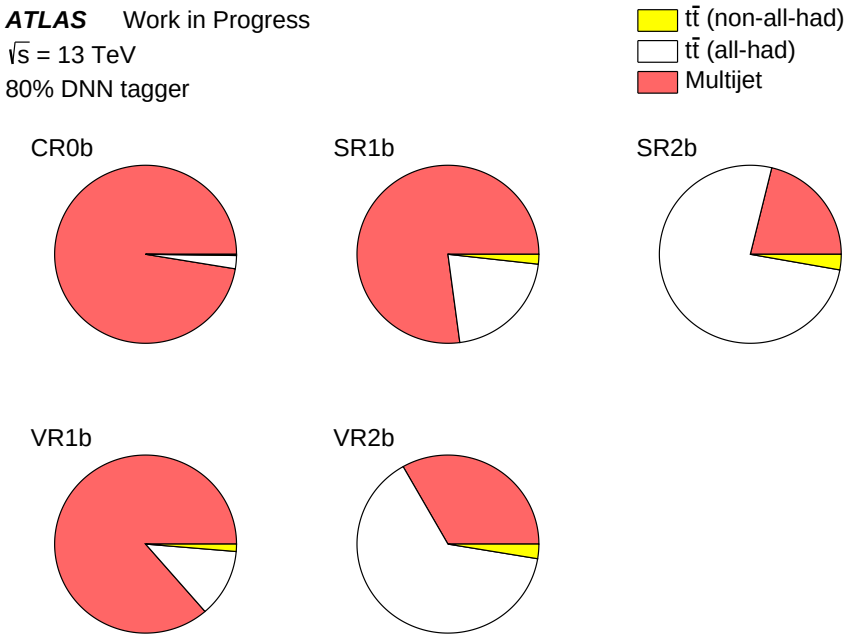


Figure 7.6: Background composition in control, validation and signal regions. The pie chart shows relative contributions per process in the different regions.

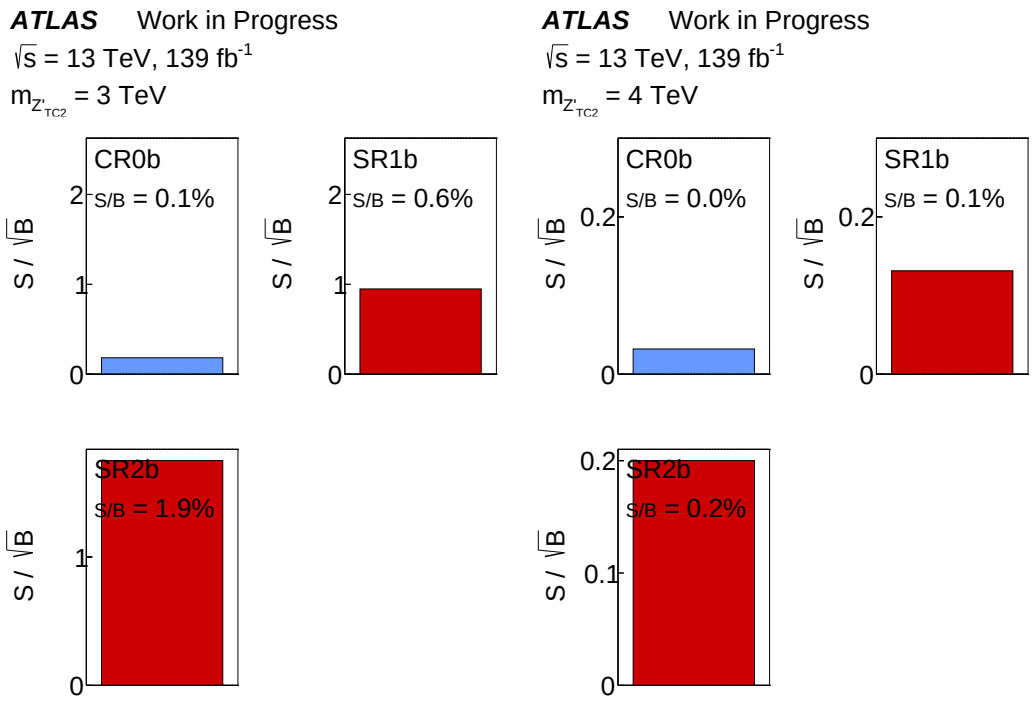
region (VR2b). The background composition of the validation regions are also shown in Figure 7.6.

Table 7.11: Summary of kinematic selections and categorization of signal, control and validation regions. The top-tagging requirement is applied to the two leading large-R jets for all the regions.

Kinematics				Number of associated b-jets		
$p_T(J_1)$	$p_T(J_2)$	$\lvert\Delta\phi(J_1,J_2)\rvert$	$\lvert\Delta y(J_1,J_2)\rvert$	0b	1b	2b
>500 GeV	>350 GeV	>1.6	>1.8	Multijet CR	$t\bar{t}$+ Multijet VR	$t\bar{t}$+ Multijet VR
			<1.8	Multijet CR	SR	SR

7.7.4 Addition regions for multijet estimation

In addition to these regions, 16 categories are defined to estimate the multijet background in the signal regions. These 16 categories are defined according to whether the leading and subleading large-R jets are top-tagged or b-tagged respectively. All these categories are shown in Table 7.12. The attributes in the row correspond to the leading large-R jet (highest p_T) and those in the column correspond to the sub-leading (second highest p_T) large-R jet. If the large-R jet is top-tagged, it is denoted by t, and otherwise by $\not{t}$, as indicated in the left column or the bottom row. Similarly, if the large-R jet is b-tagged, it is denoted by b, and otherwise by $\not{b}$. A total of 13 regions out of the 16 are used in this analysis. The data-driven method to estimate the multijet background is described in section B.1 of Appendix B.



(a) Z'_{TC2} signal: mass = 3 TeV

(b) Z'_{TC2} signal: mass = 4 TeV

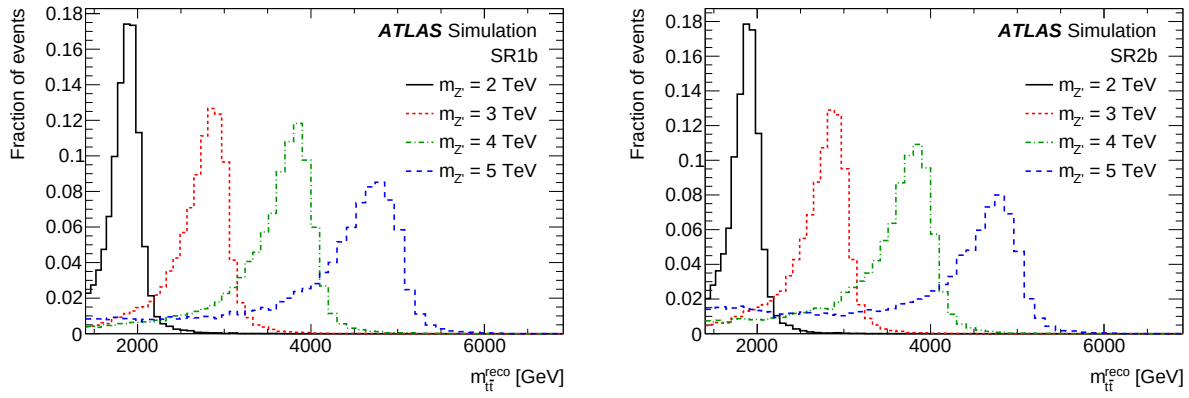
Figure 7.7: Signal contribution in each signal regions, calculated with the expected number of (a) 3 TeV and (b) 4 TeV Z'_{TC2} signal events (S) and total background events (B) in each region. The histograms show $S/\sqrt{B}$ with blue bar in the control region and red bars in the signal regions.

Table 7.12: Event categorization used to model the multijet background from data according to whether the leading and subleading large-R jets are top-tagged or b-tagged. Top(b)-tagged large-R jets are denoted by t (b), otherwise they are denoted by $\not{t}$ ($\not{b}$). The percentages in parentheses show the expected fractions of SM $t\bar{t}$ events obtained using the $t\bar{t}$ and multijet simulation samples. Non-$t\bar{t}$ or non-multijet background events are negligible. The analysis signal regions, SR1b and SR2b, are split into three regions and are shown in light orange color. The template region (TR) is shown in light green and the rest of the control regions A-I are shown in light blue [228].

Subleading large-R jet \ Leading large-R jet	$t\not{b}$	$\not{t}\not{b}$	$\not{t}b$	tb
tb	A (6.1%)		SR1b (23%)	SR2b (90%)
$t\not{b}$	B (0.5%)	E (1.8%)	TR (2.6%)	SR1b (28%)
$\not{t}b$	C (0.4%)		G (2.3%)	
$\not{t}\not{b}$	D (< 0.1%)	F (0.3%)	H (0.4%)	I (6.7%)

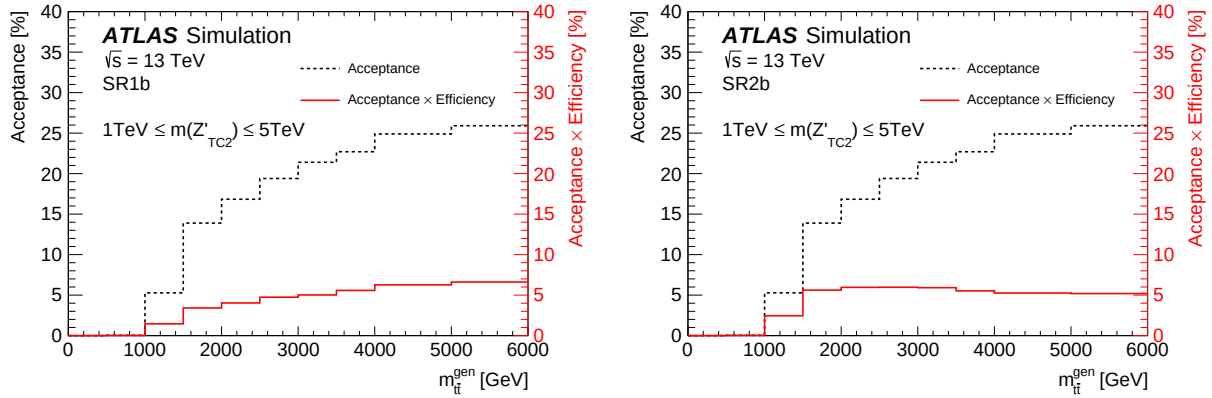
7.7.5 Signal efficiencies

The signal shapes for a few representative mass points in the two signal regions are shown as a function of reconstructed $t\bar{t}$ mass, $m_{t\bar{t}}^{\text{reco}}$, in Figure 7.8. The area under each curve is normalized to unity. So the distributions represent the signal shapes at the different mass points. We can notice that the higher mass signals have a large tail in the low mass region. This tail is coming from the low energy partons produced in the collision event. The signal acceptance × efficiency is estimated as a function of the invariant mass of the top quark pair at the generator level, $m_{t\bar{t}}^{\text{gen}}$, and shown in Figure 7.9. The branching fractions of the $t\bar{t}$ into all possible final states are included in the acceptance calculation. The acceptance is measured as the fraction of events with two lead-



(a) SR1b (b) SR2b

Figure 7.8: The $m_{t\bar{t}}^{\text{reco}}$ distributions of the simulated $Z' \to t\bar{t}$ signal events for Z' masses of 2, 3, 4 and 5 TeV in (a) SR1b and (b) SR2b [228].



(a) (b)

Figure 7.9: Acceptance (dashed histograms) and acceptance times selection efficiency (solid histograms) as a function of $m_{t\bar{t}}^{\text{gen}}$ in (a) SR1b and (b) SR2b. The $m_{t\bar{t}}^{\text{gen}}$ is calculated from the momenta of top and anti-top quarks at the generator (parton) level before final-state radiation. The branching fractions of the $t\bar{t}$ into all possible final states are included in the acceptance calculation [228].

ing truth-contained large-R jets, both satisfying the kinematic requirements used in this analysis,

and does not contain generator-level electrons or muons. The acceptance × efficiency is calculated using the events passing full analysis selections. The values of acceptance × efficiency are similar in the two signal regions. The minor differences are mainly due to the different b-tagging requirements.

7.7.6 Event yields

Expected and observed event yields in different signal regions are presented in Table 7.13 for the signal and control regions. The estimated background yield is very close to the observed data, as shown in the table.

Table 7.13: Event yields in CR0b, SR1b and SR2b regions in 2015-2018 data.

Type	CR0b	SR1b	SR2b
$t\bar{t}$ **(all-had)**	1117.5 ± 1191.3 (2.4%)	5967.1 ± 43.2 (21.1%)	6195.7 ± 38.4 (76.0%)
$t\bar{t}$ **(non-all-had)**	102.2 ± 99.1 (0.2%)	500.7 ± 9.9 (1.8%)	229.2 ± 5.2 (2.8%)
Multijet	45 466.5 ± 1329.2 (97.4%)	21 764.2 ± 98.6 (77.1%)	1729.1 ± 11.1 (21.2%)
Total	46 686.3 ± 488.3	28 232 ± 108	8154.1 ± 40.4
Data	46 454 ± 215.5	26 964 ± 164.2	8160 ± 90.3

7.8 Signal shape modeling using function fit

Z'_{TC2} signal samples are generated for some mass points between 1 TeV and 5 TeV. Above 3 TeV, there is a difference of 1 TeV between the mass points. So the signal samples in the intermediate region are created by morphing the shapes from the existing mass points. The shape of the signal $m_{t\bar{t}}$ distribution is modeled using a functional fit. The following three functions are studied to model the signal shapes:

- Reverse Landau + a Gaussian function
- Crystal Ball + a Gaussian function
- Crystal Ball + two Gaussian functions

A combination of a reverse Landau and a Gaussian function fit the $m_{t\bar{t}}$ shape quite well, but it did not work well while morphing the intermediate-mass points. The other two options worked relatively well. In the end, signal shapes are modeled using a combination of a Crystal Ball function and one Gaussian function because of its simplicity. The CB function is expected to model the truth decay width of the Z'_{TC2} process, and the Gaussian function is added to smear the distribution to model the detector effect. A Crystal Ball (CB) function of the form

$$h(x;\alpha_{\mathrm{CB}},n_{\mathrm{CB}},\mu_{\mathrm{CB}},\sigma_{\mathrm{CB}}) = N\cdot\begin{cases}\exp(-\frac{(x-\mu_{\mathrm{CB}})^2}{2\sigma_{\mathrm{CB}}^2}), & \text{for } \frac{x-\mu_{\mathrm{CB}}}{\sigma_{\mathrm{CB}}} > -\alpha_{\mathrm{CB}}\\ A\cdot(B-\frac{x-\mu_{\mathrm{CB}}}{\sigma_{\mathrm{CB}}})^{-n_{\mathrm{CB}}}, & \text{for } \frac{x-\mu_{\mathrm{CB}}}{\sigma_{\mathrm{CB}}} \leqslant -\alpha_{\mathrm{CB}}\end{cases} \tag{7.2}$$

where

$$A = \left(\frac{n_{\mathrm{CB}}}{|\alpha_{\mathrm{CB}}|}\right)^{n_{\mathrm{CB}}} \cdot \exp\left(-\frac{|\alpha_{\mathrm{CB}}|^2}{2}\right) \tag{7.3}$$

$$B = \frac{n_{\mathrm{CB}}}{|\alpha_{\mathrm{CB}}|} - |\alpha_{\mathrm{CB}}| \tag{7.4}$$

$$N = \frac{1}{\sigma_{\mathrm{CB}}(C+D)} \tag{7.5}$$

$$C = \frac{n_{\mathrm{CB}}}{|\alpha_{\mathrm{CB}}|} \cdot \frac{1}{n_{\mathrm{CB}}-1} \cdot \exp\left(-\frac{|\alpha_{\mathrm{CB}}|^2}{2}\right) \tag{7.6}$$

$$D = \sqrt{\frac{\pi}{2}}\left(1 + \mathrm{erf}\left(\frac{|\alpha_{\mathrm{CB}}|}{\sqrt{2}}\right)\right). \tag{7.7}$$

is used, where erf is the error function. The Crystal Ball function has four free parameters, α_{CB}, n_{CB}, μ_{CB} and σ_{CB}. The CB function is used with a Gaussian function of the form

$$g(x;\mu,\sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \tag{7.8}$$

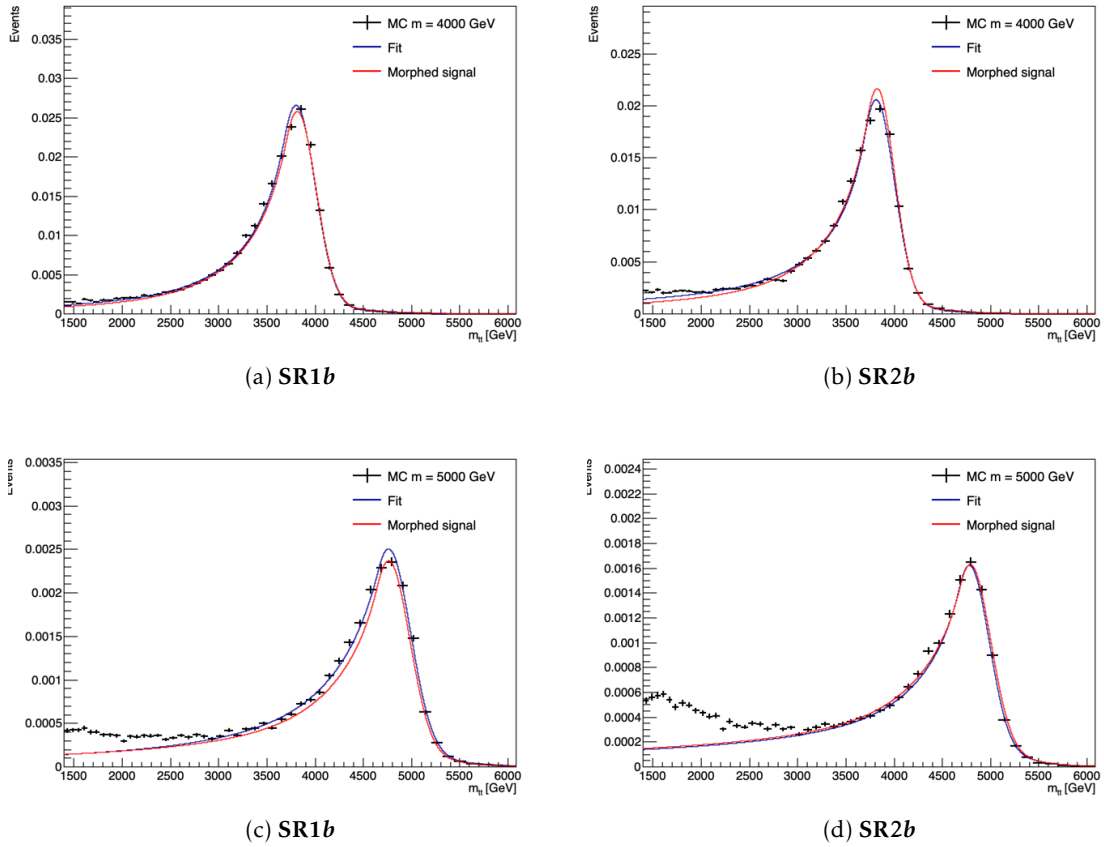
where μ is the mean and σ is the standard deviation. So, the signal fit function looks like

$$J(x;f,\mu,\sigma,\alpha_{\mathrm{CB}},n_{\mathrm{CB}},\mu_{\mathrm{CB}},\sigma_{\mathrm{CB}}) = f \cdot g(x;\mu,\sigma) + (1-f)\,h(x;\alpha_{\mathrm{CB}},n_{\mathrm{CB}},\mu_{\mathrm{CB}},\sigma_{\mathrm{CB}}) \tag{7.9}$$

where $0 < f < 1$ is a fractional coefficient. The fit function has seven shape parameters and one normalization parameter. The Gaussian μ is restricted to be bigger than μ_{CB}. This is done to ensure that the Gaussian does not move to low $m_{t\bar{t}}$ values to fit the low-mass tail for the higher mass signals (mostly above 4 TeV). The signal fits in the various regions are shown in Figure 7.10 for 4 and 5 TeV Z'_{TC2} signals. It is seen that the low mass tails are not well fitted by the CB + Gaussian signal fit function. The effect of this on the statistical analysis is estimated to be negligible. So, the Crystal Ball + Gaussian functional form is used to model the signal shapes.

7.9 Background modeling using function fit

The main background contribution in this analysis comes from SM $t\bar{t}$ and multijet processes. Both of them have smoothly falling (monotonically decreasing) $m_{t\bar{t}}$ distributions and give rise to a total background with similar properties. Section 7.3 describes modeling of individual background processes using Monte Carlo simulation. But in the analysis, the total background is estimated using data with functional fits while performing the statistical searches and calculating limits. The functional form and number of parameters of the fit function are chosen based on several studies with background templates. It is assumed that the shape of the background templates in the signal regions is similar to the data distributions. The fit quality is ensured to be good by requiring the χ^2-p-value and BumpHunter p-value to be greater than certain threshold values. It is also checked that the best function gives low "spurious signal". Spurious signal is the estimated false signal that could be found based on a functional fit in absence of actual signal. Once the background functional form is chosen, it is checked if the signal + background (s+b) fit can extract the signal pseudo-data. This study is done by injecting a known signal on top of the background template. All these studies are described in details in this section.



(a) **SR1*b***

(b) **SR2*b***

(c) **SR1*b***

(d) **SR2*b***

Figure 7.10: Simulated signal $m_{t\bar{t}}$ distribution fitted with the Crystal Ball and a Gaussian function. The red curve shows the functional fit the to the simulated signal template whereas the blue curve shows the morphed signal obtained from interpolating the fit parameters from nearby mass points. The $m_{t\bar{t}}$ distributions of a 4 TeV Z'_{TC2} are shown in (a) SR1b and (b) SR2b, and similar distributions for 5 TeV Z'_{TC2} are shown in (c) SR1b and (d) SR2b.

7.9.1 Why data-driven background estimation?

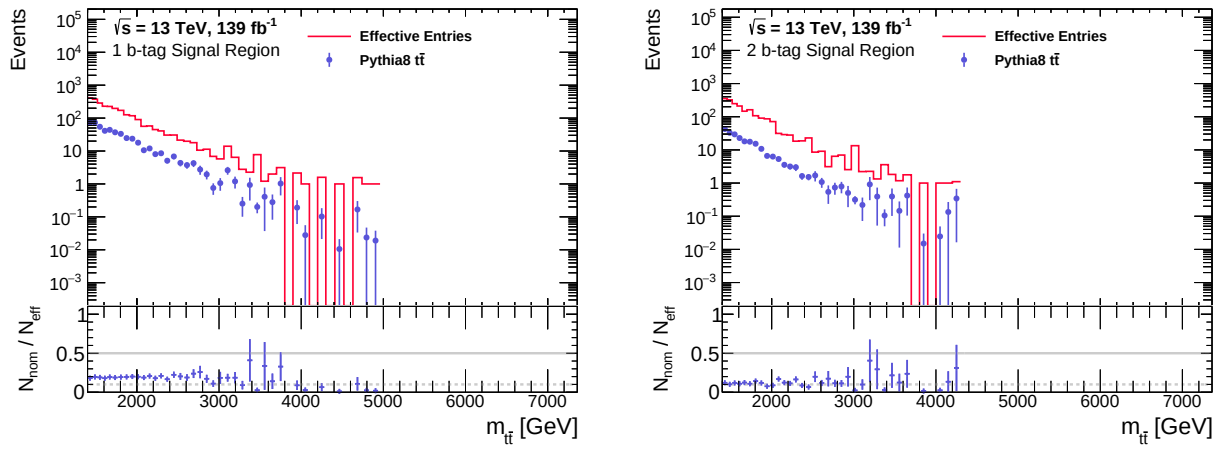
We were not satisfied with the level of agreement when using the MC modeling of the multijet process. Hence, a data-driven estimation process was developed. One such method is described in section B.1 of Appendix B. Moreover, since the dijet mass spectrum follows a smoothly falling distribution, it is possible to describe its shape with analytic functions. Nonetheless, modeling issues are also present in the $t\bar{t}$ simulations. As a result, in the searches and measurements involving $t\bar{t}$ processes, $t\bar{t}$ modeling uncertainties have a large impact on the final results. The invariant mass spectrum of the $t\bar{t}$ process follows a trend similar to the multijet process and can also be described by analytic functions. These polynomial functions can be obtained by fitting the data. The ability to estimate the total background from data allows us to circumvent the background modeling uncertainties. So, this way of estimating background is extensively studied in the analysis and is the method used to obtain the final results.

7.9.2 Monte Carlo effective entries

To choose and validate the background fit function, we need an $m_{t\bar{t}}$ distribution similar to data. Such spectra are built by combining the simulated SM $t\bar{t}$ and the multijet estimates. The statistical precision of the background distributions must be better than the data. Large statistical fluctuations will make the fit function unstable. So, first, the statistical uncertainty of the background spectra are compared with a data-like Poisson uncertainty. To check whether the simulated processes have sufficient statistics we look at the effective entries, which is defined as

$$N_{\text{effective}} = \left(\frac{\text{bin content}}{\text{bin error}}\right)^2 \tag{7.10}$$

Data is assumed to be Poisson distributed in each bin. So for bins with more than a few entries, bin error $\simeq \sqrt{\text{bin content}}$ and the $N_{\text{effective}}$ = bin content. If the $N_{\text{effective}}$ is larger than the background bin (scaled to expected data luminosity), then the simulated background has at least as good statistics as that of data. Hence those distributions can be used to determine the background fit function. The effective entries are checked for the individual background components. Figure 7.11 shows the effective entries for the MC $t\bar{t}$ spectrum. The spectrum is scaled to the expected data luminosity. It is seen that there is sufficient MC statistics over the whole range



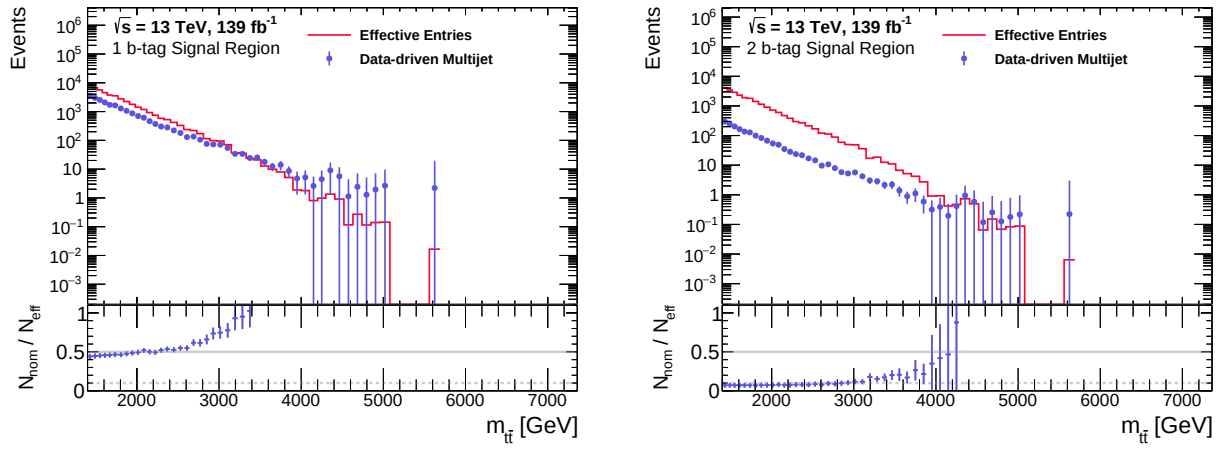
(a) **SR1*b***

(b) **SR2*b***

Figure 7.11: The effective entries in the $m_{t\bar{t}}$ distribution of the Monte Carlo simulated $t\bar{t}$ events are compared with the nominal prediction (black) in (a) SR1b and (b) SR2b. The distributions are scaled to the expected luminosity in data.

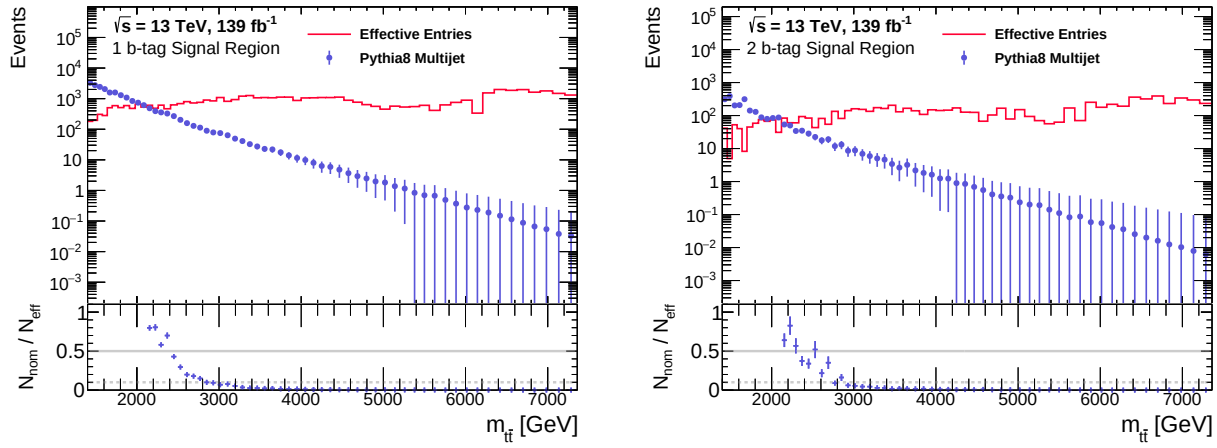
The data-driven multijet background describes the data well in the bulk of the distribution. But due to the limited number of data points at a very high $m_{t\bar{t}}$, it is not possible to estimate the multijet background there. So, the statistics in the data-driven multijet background become poor at high $m_{t\bar{t}}$. The statistics are found to be insufficient above 3.3 TeV in the 1b signal region and above 4.3 TeV in the 2b signal region as shown in Figure 7.12. The MC multijet spectrum, on the other hand, has better statistics in the high $m_{t\bar{t}}$ region, above 2.4 TeV, as is shown in Figure 7.13.

Based on the effective entries, neither the data-driven, nor the MC multijet spectrum is good enough for the background estimation studies. To get a sample with sufficient statistics at all



(a) **SR1*b*** (b) **SR2*b***

Figure 7.12: The effective entries in the data-driven multi-jet $m_{t\bar{t}}$ distributions are compared with the nominal estimates (black) in (a) SR1b and (b) SR2b. The distributions are scaled to the expected luminosity in data.

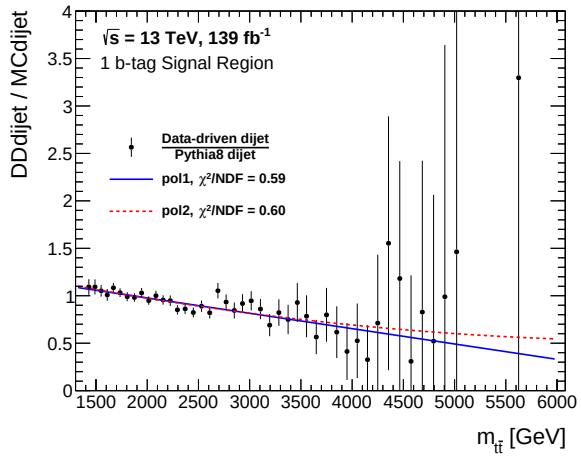


(a) **SR1*b*** (b) **SR2*b***

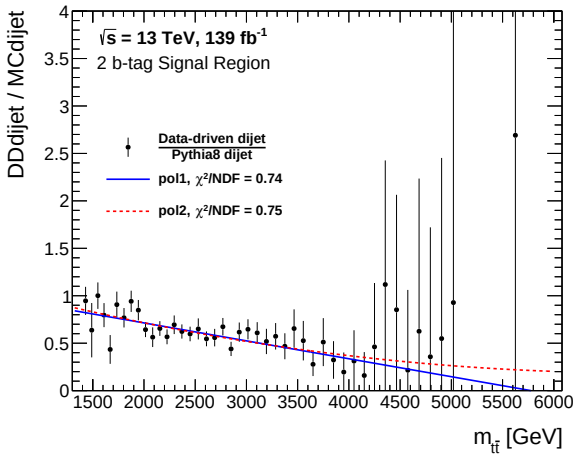
Figure 7.13: The effective entries in the $m_{t\bar{t}}$ distribution of the Monte Carlo simulated multi-jet events are compared with the nominal prediction (black) in (a) SR1b and (b) SR2b. The distributions are scaled to the expected luminosity in data.

masses, the MC and data-driven spectra are combined in the following way. The data-driven estimate can be used at low mass, whereas at higher mass, the MC sample has much better statistics. However, the data-driven sample has a shape that resembles the expected data better than the MC simulation. Therefore, the MC multijet spectrum needs to be corrected with a factor derived from the ratio of the data-driven sample and the MC sample. This correction factor is found by fitting the ratio of the two spectra. Both a 1^{st} and 2^{nd} order polynomial are tried (Figure 7.14). There is no difference between 1^{st} and 2^{nd} order polynomial correction, so the 1^{st} order polynomial is

used in this analysis. Figure 7.15 shows the effective entries of the combined multijet sample and MC $t\bar{t}$ combined. It is evident that the statistics is now sufficient and this spectrum is therefore well-suited to test the background estimation procedure.

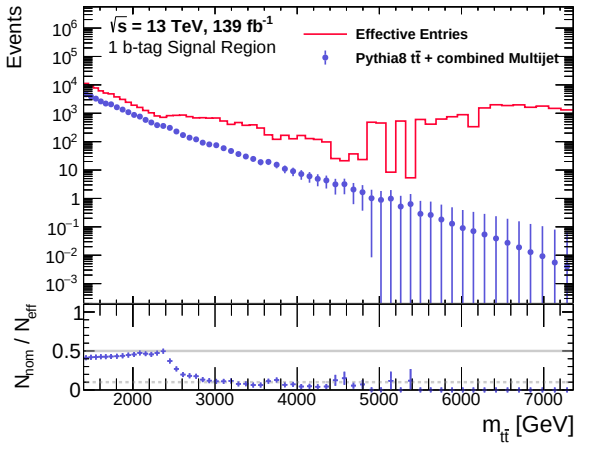


(a) **SR1*b***

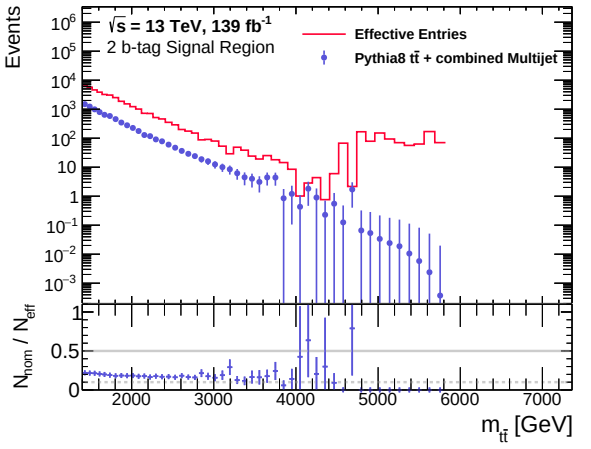


(b) **SR2*b***

Figure 7.14: The ratio of the MC multijet and data-driven multijet distributions in the (a) 1b signal region and (b) 2b signal region. The ratios are fit with 1st and 2nd order polynomial functions. The two fits look very similar, but for simplicity a 1st order polynominal is used.



(a) **SR1*b***



(b) **SR2*b***

Figure 7.15: The effective entries (red) in the $m_{t\bar{t}}$ distribution of the combined multijet and MC $t\bar{t}$ distribution are compared with the combined multijet and MC $t\bar{t}$ background (black) in (a) SR1b and (b) SR2b. The distributions are scaled to the expected luminosity in data. The effective entries are always more than total background count.

7.9.3 Fitting range and binning

We want to maximize the number of bins in order to be able to detect as narrow resonances as possible. However, we do not want the binning to be finer than the detector resolution. Therefore, the detector resolution is estimated using the truth and reconstructed mass of the di-jet system m_{JJ}, and the bins are chosen accordingly. To find the detector resolution, the ratio $m_{jj}^{\text{reco}}/m_{jj}^{\text{truth}}$ is plotted for several m_{jj}^{truth} ranges. Here m_{jj}^{reco} is the reconstructed invariant mass of the two highest p_T top-tagged large-R jets and m_{jj}^{truth} is the invariant mass of the truth large-R jets matched to the reconstructed jets. The ratios are fitted with a Gaussian distribution. The width divided by the mean of the Gaussian, σ/μ, is taken as the detector resolution. The measured detector resolution is plotted as a function of m_{jj}^{truth}. The x-value for each m_{jj}^{truth} range is set to the mean value of m_{jj}^{truth} in that range and the uncertainty on the y-axis is propagated from the Gaussian fits. The points are fitted with a function of the form $\frac{a}{m_{jj}^{\text{truth}}} + \frac{b}{\sqrt{m_{jj}^{\text{truth}}}} + c$. The bins are then derived from the function making sure the bin resolution (= bin center / bin width) is always greater than the detector resolution. Moreover, it was decided to round the bin edges to multiples of 10 GeV. The binning starts from 1400 GeV to ensure that the triggers are fully efficient and the m_{tt} spectrum is not being impacted by the leading large-R jet p_T cut of 500 GeV. There are 57 bins considered in the end. The binning used for the $m_{t\bar{t}}$ distribution is as follows:

1400, 1460, 1520, 1580, 1640, 1700, 1770, 1840, 1910, 1980, 2050, 2120, 2190, 2260, 2330, 2410, 2490, 2570, 2650, 2730, 2810, 2890, 2970, 3060, 3150, 3240, 3330, 3420, 3510, 3600, 3700, 3800, 3900, 4000, 4100, 4200, 4300, 4410, 4520, 4630, 4740, 4850, 4960, 5080, 5200, 5320, 5440, 5560, 5690, 5820, 5950, 6080, 6210, 6350, 6490, 6630, 6770, and 6910 GeV.

The background fitting range is chosen to be 1400 to 6910 GeV, beyond which the number of events in both SRs is expected to be insufficient to constrain the fit function. The expected number of events is below 0.01 at 5000 GeV in SR1b and 4100 GeV in SR2b. So, it is extremely unlikely to get any events above 6910 GeV in data. However, the effect of changing the fit range is studied. It is found to be very small and is covered by the background statistical uncertainty.

7.9.4 Background modeling function

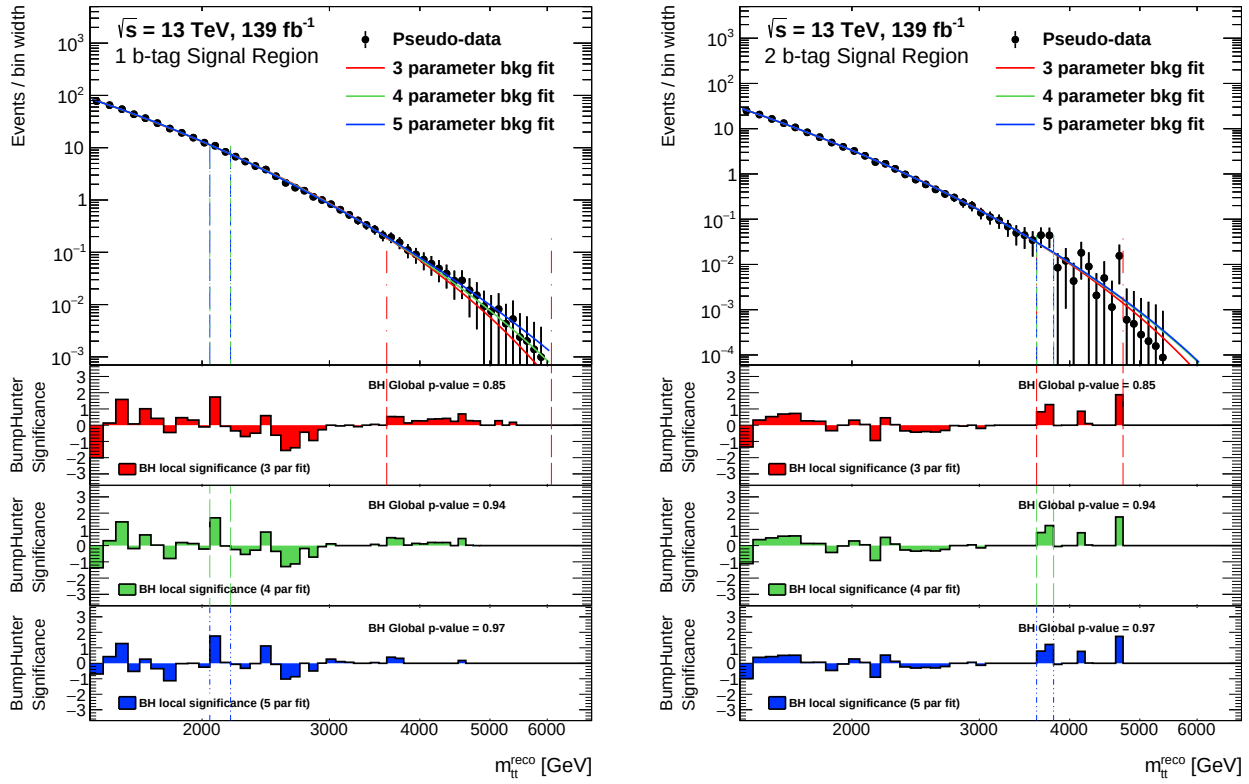
A general functional form is studied, which was previously used in the dijet search in ATLAS [242] and other experiments:

$$f(x) = p_0(1-x)^{p_1} x^{p_2+p_3 \ln x + p_4(\ln x)^2 + \cdots} \tag{7.11}$$

where $x = m_{t\bar{t}}/\sqrt{s}$, and p_i are free parameters. The $(1-x)$ factor usually tries to model the lower x region, and the $x^{p_1+\cdots}$ term controls the tail of the distribution.

The combined multijet (MC + data-driven) and MC $t\bar{t}$ distribution is used to check how many parameters are needed to describe the background distribution. To make the background distribution more "data-like" the errors are set to the Poisson error. Figure 7.16 shows fits with 3-, 4- and 5-parameters in the 1b and 2b signal region. Fit qualities are determined by the χ^2 p-value and BumpHunter (BH) p-value. The threshold for χ^2 p-value is set to be 0.05; χ^2 p-value greater than 0.05 is considered acceptable. The BH p-value threshold is set to 0.01. These thresholds are set based on previous studies done in the dijet resonance search in ATLAS [242]. The χ^2 p-values are very close to 1 for all three fits in the 1b signal region, but there is a visible improvement in BH p-value going from 3-parameters to 4-parameters. In the 2b signal region, the χ^2 p-values for 4-

and 5-parameter fits are very similar and close to 1, whereas, for the 3-parameter fit, it is around 0.9. The BH p-values also change visibly from 3-parameters to 4-parameters. Based on these observations, the 4- and 5-parameter fits are preferable over the 3-parameter fit. In the end, Wilks' test is used to choose the preferred one.



(a) **SR1*b***

(b) **SR2*b***

Figure 7.16: The $m_{t\bar{t}}$ distributions fitted with 3 (red), 4 (green) and 5 (blue) parameter function in the two signal regions: (a) SR1b, and (b) SR2b. The $m_{t\bar{t}}$ distribution is obtained from combined multijet and MC $t\bar{t}$ samples with the error set to the Poisson error. The BH p-values are also shown. The vertical lines indicate the interval with most significant deviation.

7.9.5 Wilks' test

Wilks' test checks whether adding additional parameter to the fit makes it significantly better at describing the distribution. The test statistic is described in subsection 5.3.9. The threshold α is chosen to be 0.1, for this study. The same threshold was used in the boosted bb plus ISR search [243]. The ATLAS dijet search [242] uses $\alpha = 0.05$, which would however not change our conclusion.

In this case three models are compared, namely, 3-, 4- and 5-parameter fit functions . It is observed that adding the fourth parameter causes a visible improvement in χ^2 and BH p-values, while adding the fifth one does not. Wilks' p-values are calculated comparing these models pairwise. Wilks' p-values are shown in Table 7.14. According to Wilks' test the 4-parameter fit is preferred in SR1b with a Wilks' p-value of 0.037, whereas the 3-parameter fit is good enough in

the $2b$ signal region with a Wilks' p-value of 0.234. However, this result comes from the background spectrum. To be more conclusive, Wilks' test is performed on the pseudo-experiments created from this nominal template (see 7.9.7). The background function is required to have low spurious signal contribution, as will be discussed in subsection 7.9.9.

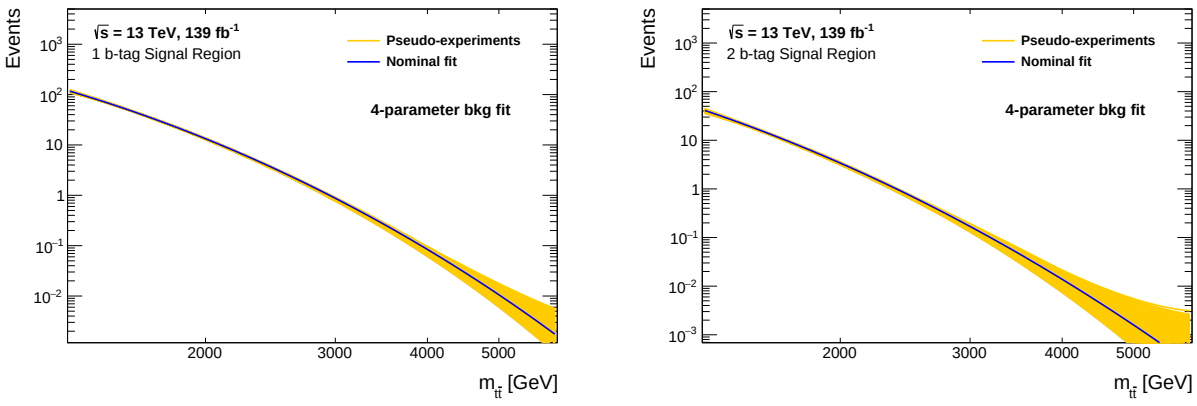
Table 7.14: Table showing the Wilks' p-values calculated by comparing Model 1 and Model 2.

Region	Model 1	Model 2	Wilks' p-value
SR1b	3-parameter	4-parameter	0.037
	4-parameter	5-parameter	0.637
SR2b	3-parameter	4-parameter	0.234
	4-parameter	5-parameter	1.00

7.9.6 Background fitting with pseudo-experiments

So far, we had been using the nominal background distribution, which is a combination of the multijet sample and the MC $t\bar{t}$ sample with Poisson errors to validate the fits. However, data usually fluctuates around the nominal background distribution. So to get a better idea of the robustness of our fitting method, it is tested on pseudo-experiments drawn from this nominal background distribution. A set of 1000 pseudo-experiments is created. In a pseudo-experiment, each bin value is fluctuated by a Poisson distributed random number with mean equal to the original bin content. Then, the error in the bin is set as the square root of the new bin content.

Figure 7.17 shows the 4-parameter fits to the 1000 pseudo-experiments in each signal region. The fits on the pseudoexperiments, shown in red lines, look very similar to the fit done on the nominal spectrum, shown in blue, in both regions except for a few 4-parameter fits in the $2b$ signal region.



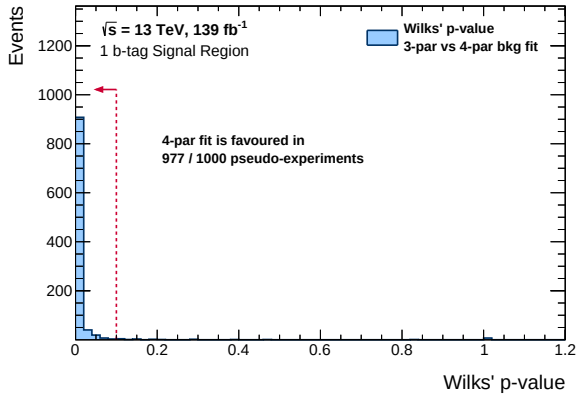
(a) **SR1*b***

(b) **SR2*b***

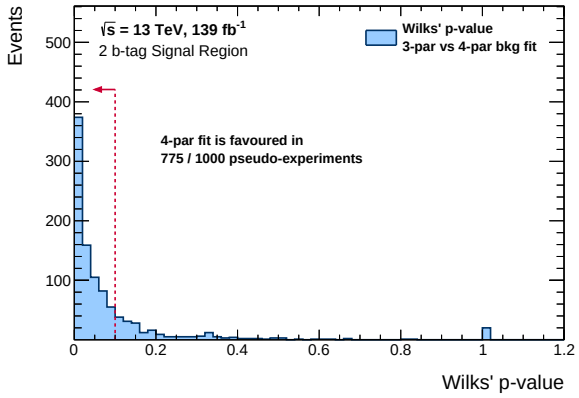
Figure 7.17: 4-parameter fits performed on the 1000 pseudo-experiments (orange) in (a) SR1b and (b) SR2b. The blue line shows the fit to the nominal spectrum.

7.9.7 Wilks' test on pseudo-experiment

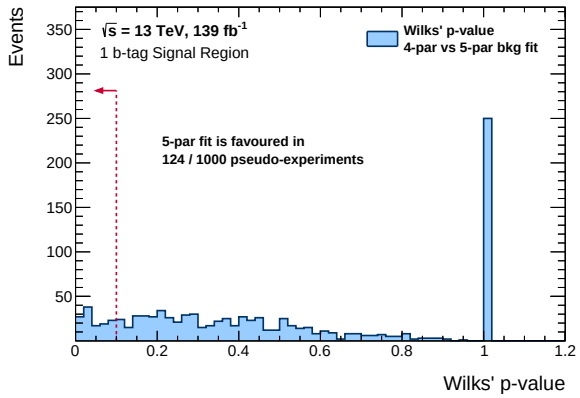
Previously (subsection 7.9.5), Wilks' test was done for the fits to the nominal spectrum. In this section, Wilks' p-value is calculated for all 1000 pseudo-experiments to get a more conclusive picture. The p-value distributions are shown in Figure 7.18. Like before, if the p-value is less than 0.1, the fit with more parameters is preferred. It is seen that for the 1b signal region, 3-parameters are not enough. Whereas the 4-parameter function is preferred in many cases. In the 2b signal region, the result is not as clear. However, in 77% of the cases, the 4-parameter fit is favoured over the 3-parameter fit. But in almost 34% of the pseudo-experiments, the 5-parameter fit is favoured over the 4-parameter fit. It is, therefore, difficult to conclude how many parameters are needed in the data in the 2b signal region.



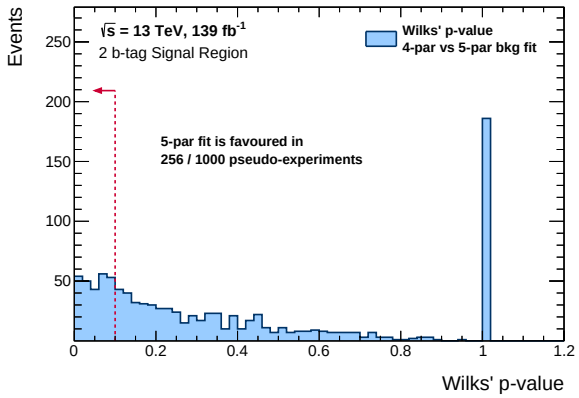
(a) **SR1*b***



(b) **SR2*b***



(c) **SR1*b***



(d) **SR2*b***

Figure 7.18: Wilks' p-value for the pseudo-experiments comparing fits with (a, c) 3 and 4 parameters and (b, d) 4 and 5 parameters in SR1b (upper) and SR2b (lower) are shown. The red arrow indicates the pseudo-experiments where the higher number of parameters are favoured.

7.9.8 Final choice of fit functions

In SR1b, both 4-parameter and 5-parameter fit-functions give a similar spurious signal pull. But the 4-parameter fit-function is sufficient to describe the background shape according to Wilks'

test. In SR2b, the 3-parameter background fit-function is good enough to describe the background shape according to Wilks' test. But it tends to introduce a larger spurious signal. So, finally, the 4-parameter background function is chosen for the SR2b as well.

7.9.9 Spurious signal test

Wilks' test, χ^2 p-value, and BH p-value help us find the functions that can fit the background shape. Once an acceptable background fit function is found, its flexibility is tested using a spurious signal test. In the spurious signal test, the background distribution is fitted with a signal + background (s+b) model. Signal shapes are modeled using a combination of Gaussian and Crystal Ball function as described in section 7.8. Any signal fitted from this s+b fit would not be a real signal and is called a "spurious signal". A spurious signal often arises if the background function is not flexible enough to describe the background shape. The fit then adjusts the other fit-parameters to fit the background shape better with the s+b function. Figure 7.19 shows a schematic diagram to demonstrate the spurious signal test.

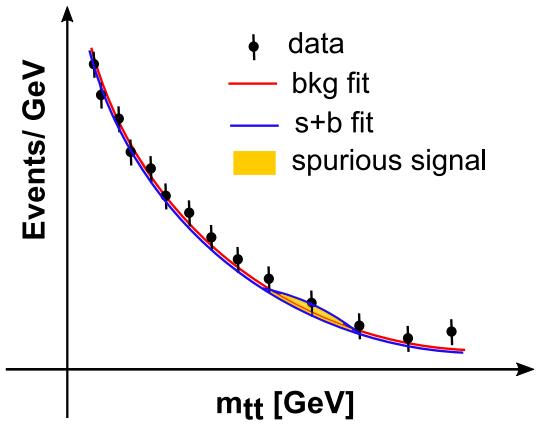


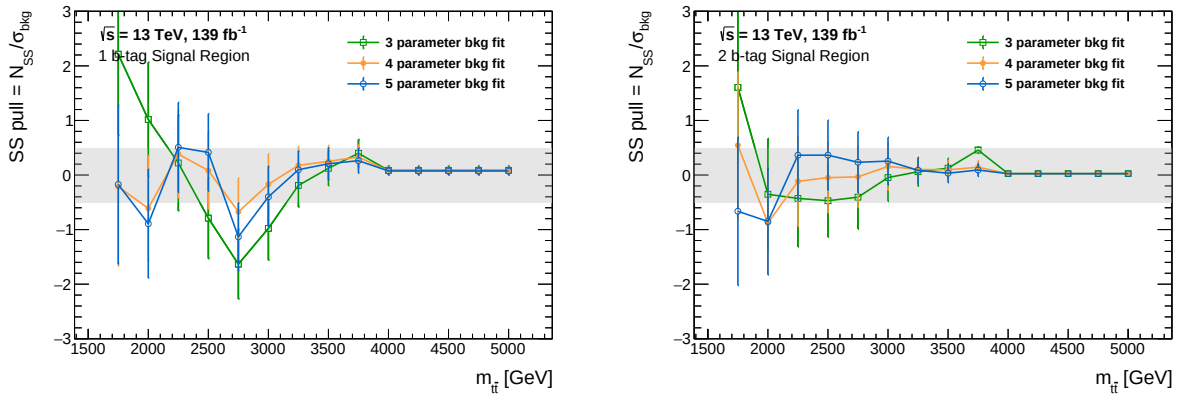
Figure 7.19: Schematic diagram showing the spurious signal test. The black dots represent simulated data created from background-only distributions. The red curve is the background fit and blue curve shows the s+b fit. It is shown that even in absence of real signal the background fit parameters could get adjusted and predict some signal (shown in yellow).

The spurious signal can also arise due to a rapid change in detector response or reconstruction efficiency as a function of the variable of interest leading to a jump in the event rate. This may come from different calibrations, where different efficiencies are used for different mass or p_T ranges. It can not be predicted without looking at the data. A fit function with too much flexibility where all the parameters are nicely constrained can also give a non-zero spurious signal. So, it is very important to make sure the amount of spurious signal is within a reasonable limit. To avoid false discovery coming from spurious signals, an optimal functional form is required not to introduce absolute spurious signal yields larger than 50% of the statistical uncertainty of the background prediction. From different functions with similar fit quality, the one with the least spurious signal is chosen.

The spurious signal yield is estimated from a signal plus background (s+b) fit to the background-only distribution. A $s+b$ fit function of the form

$$f(x) = f_{\text{bkg}}(x) \; + \; \mu_{\text{SS}} \cdot f_{\text{sig}}(x) \tag{7.12}$$

is used, where $x = m_{t\bar{t}}/\sqrt{s}$ and μ_{SS} is the spurious signal strength. Three different dijet functions of the form Equation 7.11 with 3, 4 and 5 parameters are considered as f_{bkg}. The signal function is defined in Equation 7.9. The spurious signal strength is the ratio of the number of expected spurious signal events and expected signal events from theory. Ideally, μ_{SS} should be zero for a background-only distribution, but in reality we observe non-zero values of μ_{SS}. From the fitted valued of μ_{SS}, the spurious signal signal yield, N_{SS} is calculated in a 20% window around the signal peak. In this study, the spurious signal yields are always compared with the background uncertainties, σ_{bkg}. The background uncertainty is calculated by taking the square root of the background event count within the aforementioned 20% window around the signal peak. The ratio of the N_{SS} and σ_{bkg} is defined as the *spurious signal pull*. The spurious signal pulls (N_{SS}/σ_{bkg}) are shown in Figure 7.20 as a function of $m_{t\bar{t}}$ for three different background fit functions. Generally, the spurious signal pull is quite large for the 3-parameter background function in both signal regions and lower in the 4- and 5-parameter fits.



(a) **SR1*b***

(b) **SR2*b***

Figure 7.20: Spurious signal pull for background functions with 3-, 4- and 5-parameters in (a) SR1b and (b) SR2b. The gray shaded area shows the ±50% around zero which is used for comparison.

7.9.10 Signal injection test

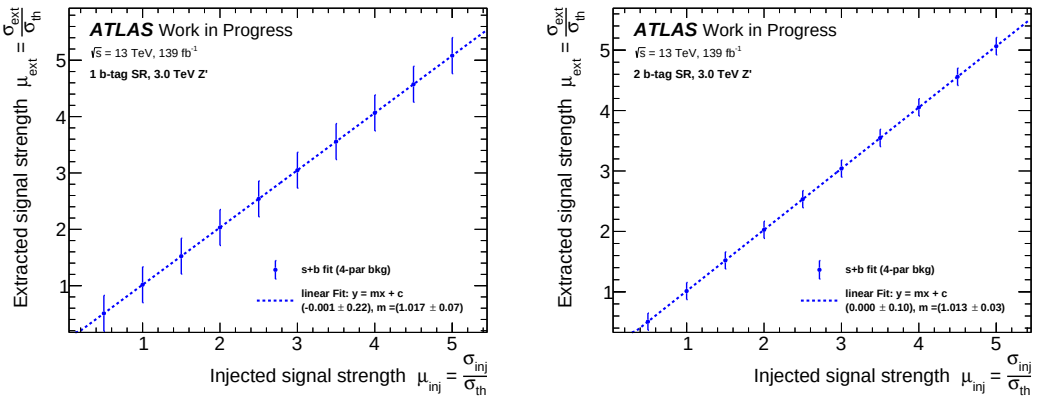
A signal injection test is performed to demonstrate the ability to extract the signal from the data. The 4-parameter background modeling function is used to create a nominal background $m_{t\bar{t}}$ histogram. The simulated signal events are used to get the signal histograms. The following steps are performed in this test:

1. Signal events are injected on top of the nominal background, estimated by the 4-parameter b-only fit function. The injected signal strengths (μ_{inj}) are varied from 0.5 to 5.0.

2. Each signal injected spectrum is fitted with a $s+b$ function of the form $f(x) = f_{bkg}(x) + \mu_{ext} \cdot f_{sig}(x)$. Here, $x = m_{t\bar{t}}/\sqrt{s}$ and μ_{ext} is the extracted signal strength.

The signal injection test is done for nine signal mass points, $m_{Z'}$ = 1.5, 1.75, 2.0, 2.25, 2.50, 2.75, 3.00, 4.00, and 5.00 TeV. Figure 7.21 shows the comparison of μ_{ext} and μ_{inj} for 3 TeV Z'_{TC2}

signal in both the signal regions. The data points are fitted with straight lines and the slopes are very close to 1.0. In both signal regions, the extracted signal strengths are proportional to the injected signal strengths. On average, the extracted signal is slightly overestimated, but the bias is very small.



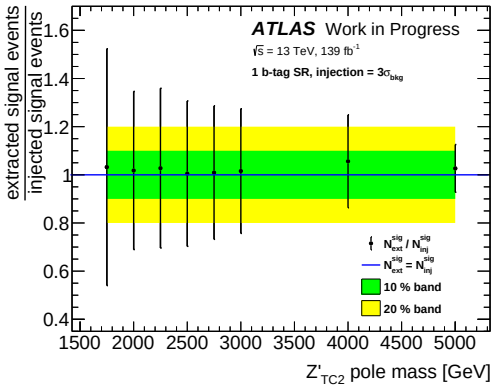
(a) **SR1*b*** (b) **SR2*b***

Figure 7.21: Extracted signal strength (μ_{ext}) as a function of injected signal strength (μ_{inj}) in (a) SR1b and (b) SR2b-tag signal. Z' signal of 3 TeV pole mass is injected with different signal strengths starting from 0.5 to 5.0. Only the 4-parameter background function is used in this study. The data points show a linear trend in both signal regions and the slopes are very close to 1.0.

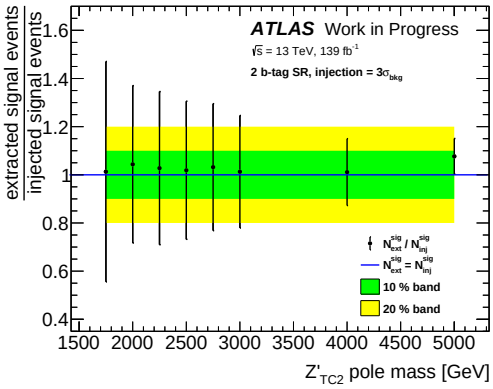
Further checks are done to see how much injected signal can be recovered (or extracted) with the $s + b$ fit. Here the number of injected signal events and extracted signal events are compared. The following steps are done in this study:

1. A window is defined around the signal peak. The signal width is defined as the standard deviation of the signal histogram. Window size = peak ±2× signal width.
2. The background uncertainty (σ_{bkg}) is defined as the square root of the background events under the signal peak (inside the window).
3. Signal is injected with $3 \cdot \sigma_{bkg}$ and $5 \cdot \sigma_{bkg}$ on the background spectrum.
4. A $s + b$ fit is performed to extract signal strength and the number of extracted signal events is calculated from there.

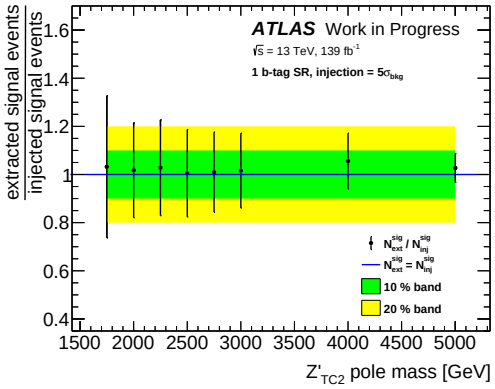
The ratio of injected and extracted signal events are shown in Figure 7.22 for $3\sigma_{bkg}$ and $5\sigma_{bkg}$ injection. The results show that the agreement is good between the number of injected and extracted signal events. Overall the extracted signal is slightly over estimated. At 5 TeV the difference is a bit larger but stays within 10%.



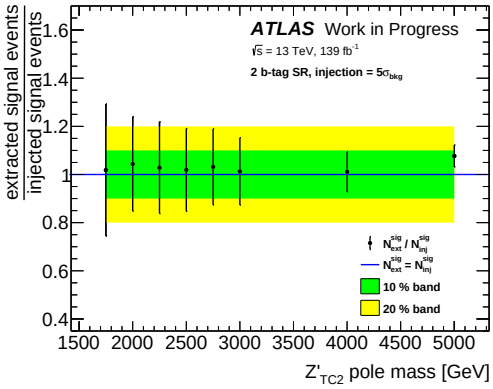
(a) 1b-tag Signal Region



(b) 2b-tag Signal Region



(c) 1b-tag Signal Region



(d) 2b-tag Signal Region

Figure 7.22: Ratios of extracted and injected signal events in (a,c) SR1b and (b,d) SR2b. The number of injected signal events at each mass point is equal to $3\sigma_{bkg}$ in the upper row and $3\sigma_{bkg}$ in the lower row. Extracted signals are overestimated but the difference is within 10% of the number of injected signal events.

Background-only fit on signal injected spectrum

The ability to recover signal if it is present is also tested with b-only fits. Pseudoexperiemnts generated from the signal injected spectrum are treated as data and the backgrounds are estimated using b-only fits. The expected 95% CL upper limit on signal cross section is then calculated. Figure 7.23 shows one such plot where signal is injected at 3 TeV with two times the theory cross section. The limit plot correctly shows that the signal injected mass region (3 TeV) cannot be excluded.

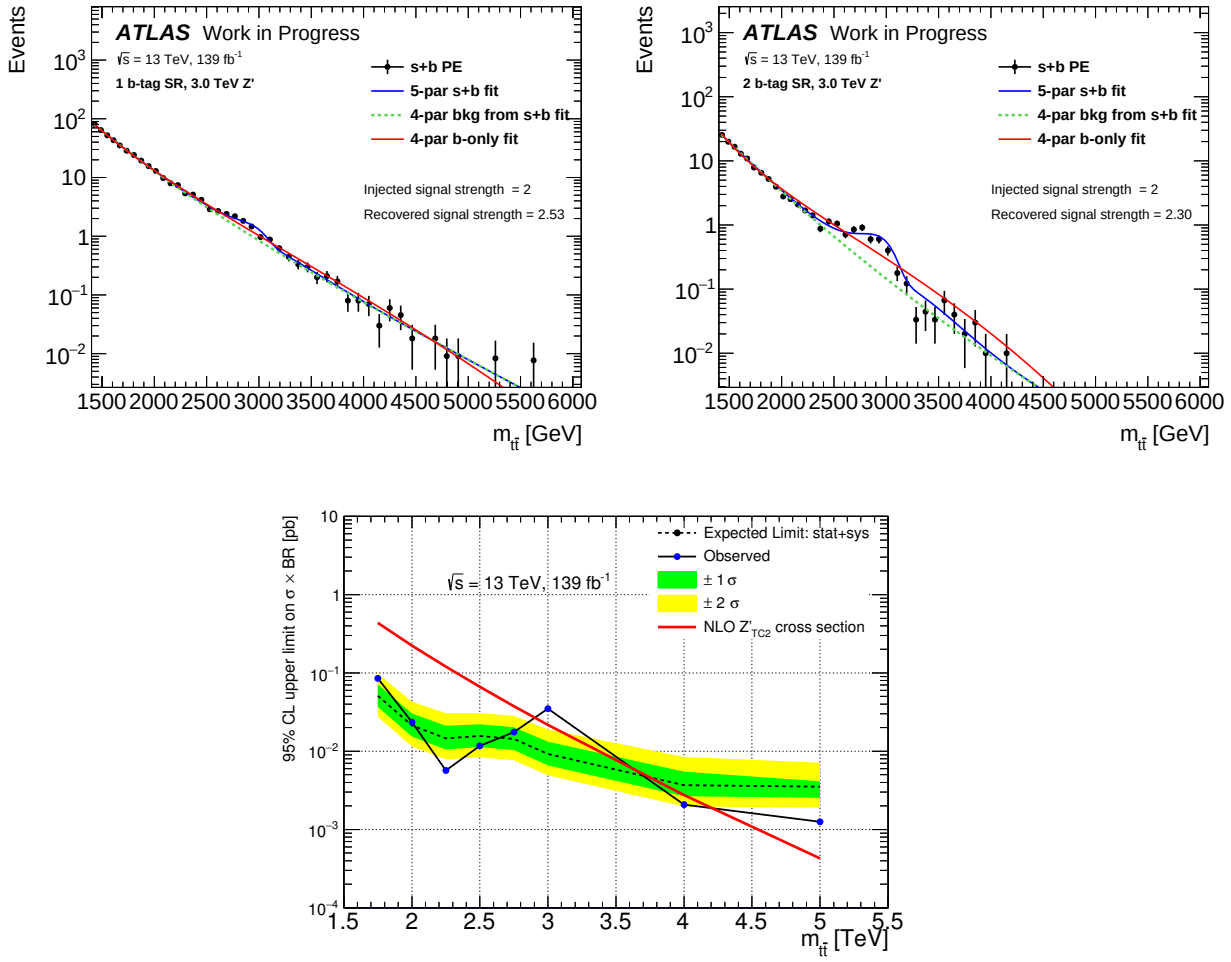


Figure 7.23: (top) PEs generated from the signal injected spectrum in both the signal regions. Signal is injected at 3 TeV with two times the theory cross section. The background is estimated from the 4-parameter b-only fit on the signal injected PE. (bottom) 95% CL upper limit on signal cross section is calculated from the signal injected PE and the estimated background. There is a significant deviation between the observed and expected limits around 3 TeV, as a result a 3 TeV Z' signal cannot be excluded.

7.10 $t\bar{t}$ invariant mass

The invariant mass of the $t\bar{t}$ system ($m_{t\bar{t}}$) is used as the main discriminant for this search. Any Z' signal should appear as a bump on the falling background $m_{t\bar{t}}$ spectrum. The observed $t\bar{t}$ invariant mass distributions in the two signal regions are shown in Figure 7.24. The background-only fits are shown in the red curves. The fit parameter uncertainties are shown in the cyan bands. Signal distributions predicted by the Z'_{TC2} model are overlaid on the observed $m_{t\bar{t}}$ distribution for pole mass of 2 and 4 TeV, shown scaled to 5 times the expected number of events. The BumpHunter local significances are shown in the lower panel. It shows that the data is consistent with the background and there is no visible deviation on data.

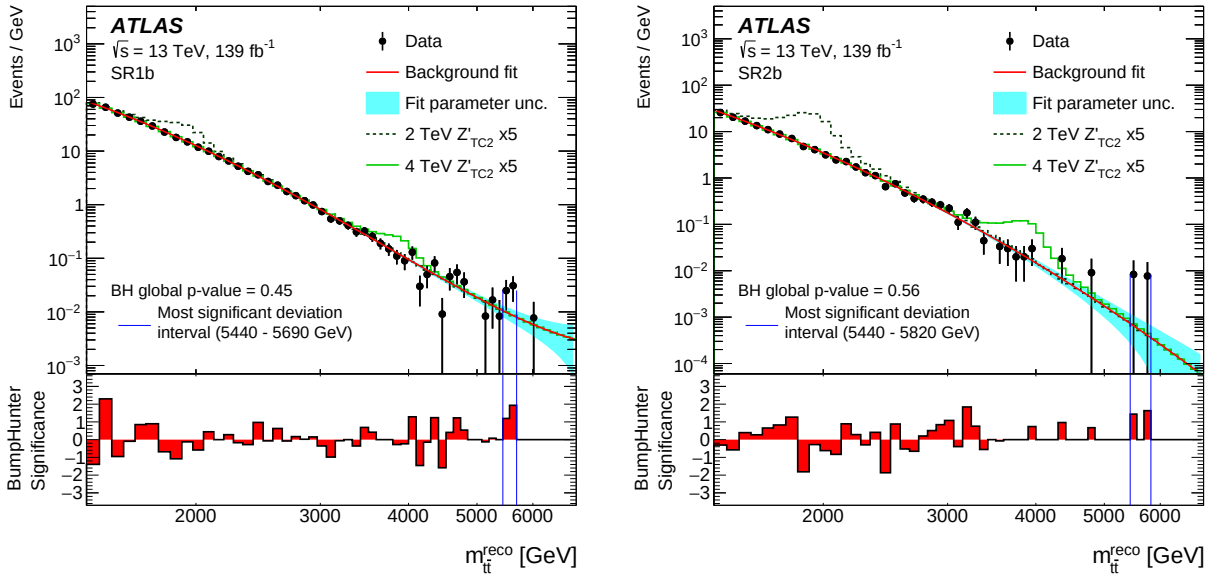


Figure 7.24: Observed $m_{t\bar{t}}^{\mathrm{reco}}$ distributions for (a) SR1b and (b) SR2b, shown together with the result of the fit with the 4-parameter function. The shaded bands around the fits indicate the effect of the fit parameter uncertainty on the background prediction. The bin width of the distributions is chosen to be the same as that used in the background parameterization. The predicted Z' signal distributions with masses of 2 and 4 TeV (multiplied by a factor of 5) are superimposed on the background prediction. The lower panel shows the significance of data with respect to the background prediction from the fit, calculated using BumpHunter. The two vertical lines extending between the upper and lower panel represent the most significant deviation interval. The global p-value of the interval is 0.45 (0.56) for SR1b (SR2b) [228].

7.11 Systematic uncertainties

As the background is modeled from the data using a functional fit, the relevant background modeling uncertainties are the fit-parameter uncertainties. The effect of the individual fit parameter uncertainties is estimated as an envelope uncertainty on the total background spectrum using the pseudo-experiments. This uncertainty is called fit-parameter uncertainty. The effect of the spurious signals is also quantified and included as an uncertainty. For the signal prediction, all the experimental and theoretical systematic uncertainties are considered. For each source of system-

atic uncertainty, the effect is estimated on the $t\bar{t}$ invariant mass. Both up and down variations ($\pm 1\sigma$) are modeled for most of the uncertainties. In principle, the up and down variations can be different, but we have taken a conservative approach of symmetrizing the uncertainties using the largest of the up or down variations. Thus symmetric up and down variations are derived as a function of $t\bar{t}$ invariant mass and used in the statistical interpretation. The following sections describe these uncertainties in more detail.

7.11.1 Experimental signal uncertainties

The largest experimental uncertainties come from jet calibration, top-tagging, and b-tagging calibrations. All these uncertainties are applied to the signal models. The effect of pileup, luminosity, and JVT uncertainties are also considered for the signal processes. All the major sources of uncertainty are listed in this section.

Luminosity, pileup and JVT

The absolute luminosity calibration is performed using beam-separation scans (van der Meer scans) during dedicated running periods in each year, which allow us to infer the luminosity from the direct measurements of beam parameters. The calibration is then extrapolated to the physics data-taking regime using complementary measurements from several ATLAS subdetectors. The total uncertainty on the integrated luminosities are 2.0-2.4% for each year of 2015-2018, and 1.7% on the full data sample [131].

All simulated signal events are re-weighted to get the same luminosity profile as observed in the data. The pileup uncertainty is estimated by varying the pileup reweighting. It is found to be around 0.1% on the 4 TeVsignal.

The JVT uncertainty is calculated from the JVT corrections (scale factor) derived for correcting the modeling in the simulation. The JVT scale factor uncertainty is also small and is always less than 0.2%.

Trigger efficiency

The p_T of the leading large-R is required to be > 500 GeV to ensure we are not affected by the trigger turn-on. The trigger was found to be fully efficient for large-R jets above this value, as shown in Figure 7.3. Almost 99.9% Z'_{TC2} signal events gets accepted. Therefore, we do not consider any trigger uncertainty in this analysis.

Jet energy and mass scale

The jet energy scale (JES) and mass scale (JMS) uncertainties are derived using the R_{trk} method [191]. The relative p_T uncertainty is determined using the ratio of the transverse momentum of the calorimeter jet, p_T^{jet} , to that of the track jets, $p_T^{\text{track-jet}}$ as described in [191]. The same procedure is followed for mass as well. Kinematic observables as measured in the calorimeter and the tracker are studied and the residual difference between data and MC is considered as a systematic contingency. This consists of a baseline component based on the difference between data and PYTHIA8, a modeling component using the difference between Pythia8 and Herwig7 or Pythia8 and Sherpa as well as an uncertainty component for tracks being used as a reference. A statistical uncertainty on the measurement is also assigned to the total uncertainty. In the following plots and tables, different components of the JES uncertainties are denoted as AKT10 JES, whereas the

JMS uncertainties are written as AKT10 JMS. The four components of of these uncertainties are denoted in the brackets.

Jet energy and mass resolution

The large-R jet p_T resolution uncertainty is determined by smearing the p_T distributions of the two leading top-tagged large-R jets. The large-R jet p_T is smeared using a Gaussian of width = 0.02 (an absolute 2% uncertainty), and the effects are propagated to the $m_{t\bar{t}}$ distribution. The difference between smeared and nominal $m_{t\bar{t}}$ distribution is defined as the jet energy resolution (JER) uncertainty. This uncertainty is named as the "up" variation and the corresponding "down" variation is defined as the mirror image of the up variation. In the following plots and tables, this uncertainty is labeled as AKT10 JER. The jet mass resolution (JMR) uncertainties are derived for top-jets by smearing the simulated jet mass using a Gaussian with a width corresponding to 20% relative uncertainty on the mass resolution of a jet. The JMR uncertainty is labeled as AKT10 JMR (m_{top} reso.) in the following plots and tables.

Top-tagging

The top-tagging efficiency for hadronically decaying top quarks is measured using both data and simulation samples, which are enriched in $t\bar{t}$ events with the lepton+jets final state. The ratio of the two efficiencies is applied to the Monte Carlo simulated samples as a scale factor. All the jet calibration uncertainties and some additional uncertainties affect the top-tagging scale factor. The uncertainty on the scale factor is estimated to be 10-15% per jet depending on the p_T of the jet. Since this search involves high p_T jets (up to 3-4 TeV) and the extraction of signal scale factors is done in a $t\bar{t}$ single-lepton topology, where there isn't sufficient data above a large-R jet p_T of 1 TeV, the signal scale factor uncertainty at higher p_T is extrapolated from the measurement at lower p_T. Additional extrapolation uncertainties are considered at high p_T. The components of jet p_T scale and top-tagging uncertainties associated with the same sources of systematic uncertainties are varied together in the statistical analysis procedure. Top-tagging uncertainty components are labeled as t-tag eff. in the following plots and tables, and the components of the top-tagging uncertainty are written in the brackets.

***b*-tagging**

b-tagging uncertainties are treated similarly to top-tagging. The *b*-tagging efficiencies in data and simulated two-lepton final state are compared and a correction factor (scale factor) is derived as the difference between efficiencies in data and simulation. The b-tagging uncertainties are propagated through this b-tagging scale factor. The uncertainties are derived for various kinematic regions and separately for *b*-, *c*-, and light-flavored jets. An additional *b*-tagging uncertainty due to high-p_T extrapolation is also applied.

7.11.2 Background modeling

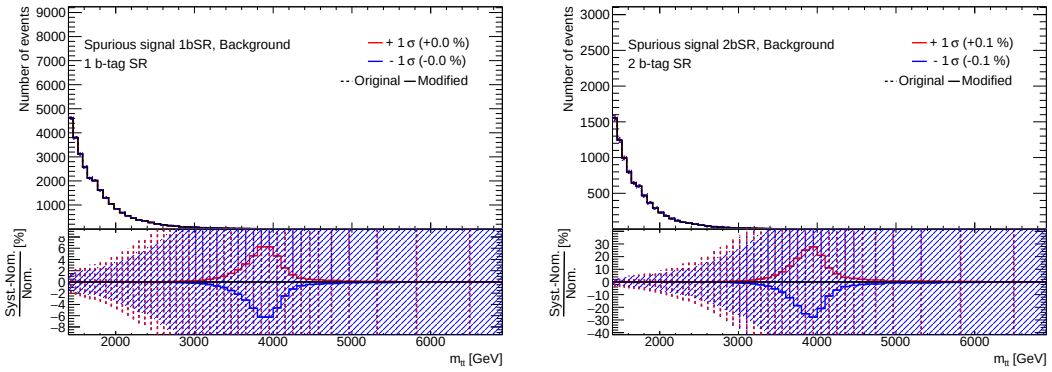
Spurious signal uncertainty and the fit parameter uncertainty are the two background uncertainties considered in this analysis. These two uncertainties are described in this section.

Spurious signal uncertainty

The spurious signal uncertainty is estimated to quantify the bias on the extracted signal yield caused by the choice of the background modeling function by fitting the background-only $m_{t\bar{t}}$ distribution with signal plus background (s+b) model of the form Equation 7.12. All the parameters of the signal function (Equation 7.9) are kept fixed in the fit, allowing only the parameters of the background function (4-parameter dijet function of the form Equation 7.11) and spurious signal strength, μ_{SS}, to vary. μ_{SS} is allowed to take both positive and negative values in the fit. The results of the spurious signal test described in subsection 7.9.9 are used to estimate the spurious signal uncertainties. The $s+b$ fit is done at different signal mass points and μ_S is calculated as a function of $m_{t\bar{t}}$. The spurious signal up (down) variation is calculated by adding (subtracting) the scaled signal template ($\mu_{SS} \cdot m_{t\bar{t}}^{\text{sig}}$) to the nominal background ($m_{t\bar{t}}^{\text{bkg}}$) like,

$$\text{SS variation} = m_{t\bar{t}}^{\text{bkg}} \pm \mu_{SS} \cdot m_{t\bar{t}}^{\text{sig}}. \tag{7.13}$$

So, the relative uncertainty has the same shape as the nominal signal model. Figure 7.25 shows the up and down variation of the spurious signal uncertainty in the two signal regions for the 4 TeV signal mass point. The SS uncertainty is added in the likelihood as a background uncertainty using a nuisance parameter, θ_{SS}.

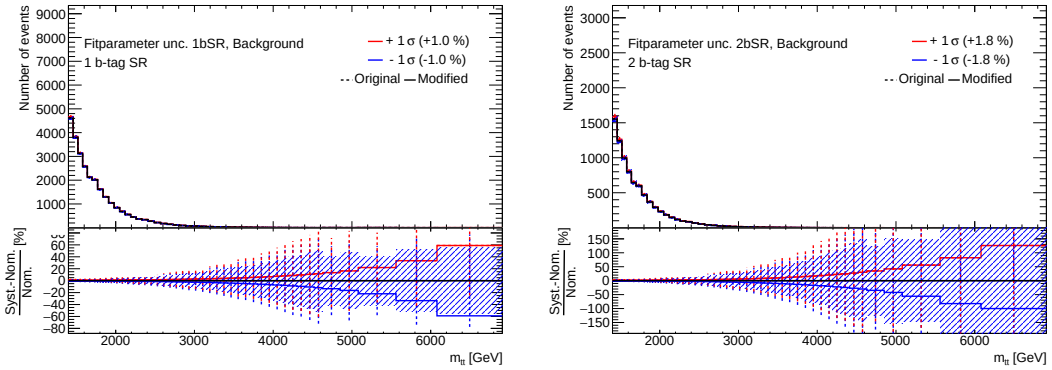


(a) SR1b (b) SR2b

Figure 7.25: Spurious signal uncertainty for the 4 TeV signal mass point in (a) 1 b-tag SR (SR1b) and (b) 2 b-tag SR (SR2b).

Fit parameter uncertainty

The fit parameter uncertainty is associated with the statistical fluctuation of data. To estimate the uncertainty, we generate pseudo-experiments from the nominal background using Poisson statistics and fit each pseudo-experiment. The uncertainty of each bin is defined to be the RMS of the function value in that bin of all pseudo-experiments. The overall uncertainty is taken as an envelope uncertainty over the full background spectrum as shown in Figure 7.26. In the profile likelihood it is defined as fully correlated across the bins and one nuisance parameter is used.



(a) SR1b (b) SR2b

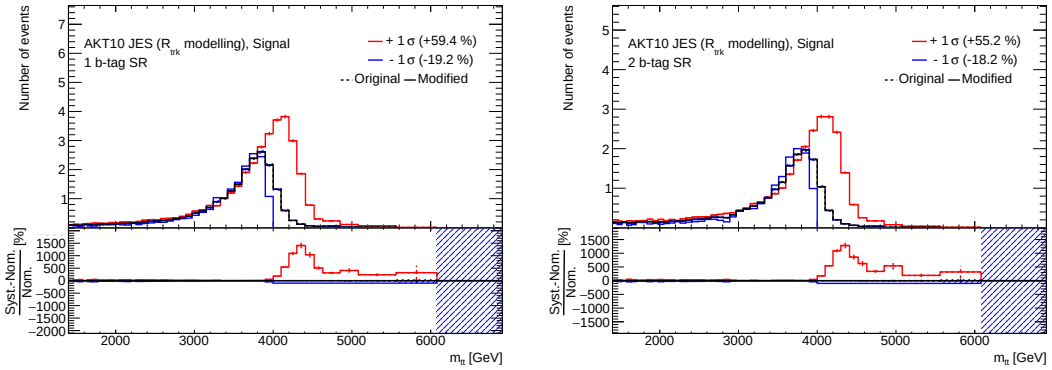
Figure 7.26: Fit parameter uncertainty in (a) 1 b-tag SR (SR1b) and (b) 2 b-tag SR (SR2b) as a function of $m_{t\bar{t}}$. The uncertainty increases with $m_{t\bar{t}}$ and becomes almost 35% around $m_{t\bar{t}} = 6$ TeV.

7.11.3 Theory uncertainties

Theory uncertainties for the signal process are not considered in this analysis. It is expected that the theory uncertainties will be much smaller than some of the experimental and modeling uncertainties like large-R jet calibration and spurious signal uncertainties.

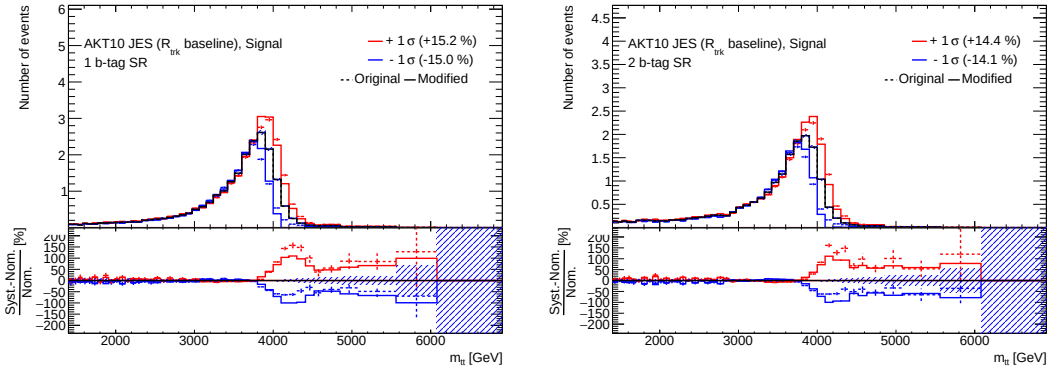
7.11.4 Systematics summary

The effect of different systematic uncertainties on the signal acceptance is summarized in Table 7.15 and Table 7.16 for SR1b and SR2b, respectively. The modeling component of the JES uncertainty has the largest contribution, it is around 60% (55%) for a 4 (3) TeV Z'_{TC2}. Figure 7.27 shows these variations as a function of $m_{t\bar{t}}$. The baseline and tracking components of the JES uncertainty (shown in Figure 7.28) are also large and contribute significantly. JES baseline is around 15% (10%) and JES tracking is around 9% (7%) in SR1b (SR2b). Other than the JES uncertainties, two eigenvector variations (E_0, E_1) of b-tagging efficiency scale factor uncertainties, the b-tagging extrapolation uncertainty, and an eigenvector variation of the light-jet mistag rate have large contributions.

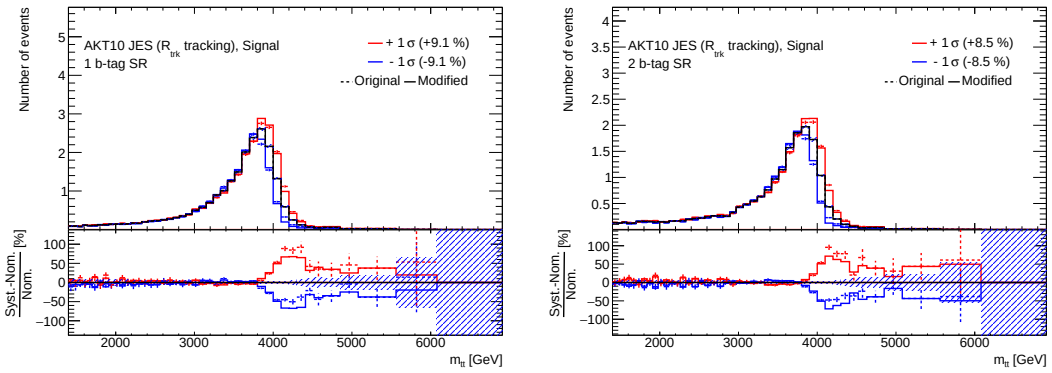


(a) SR1b (b) SR2b

Figure 7.27: Up and down variations of JES modeling uncertainty as a function of $m_{t\bar{t}}$ in (a) SR1b and (b) SR2b for the 4 TeV Z'_{TC2} signal. The variations are around 55-60%.



(a) SR1b (b) SR2b



(c) SR1b (d) SR2b

Figure 7.28: Up and down variations of JES baseline (a, b) and tracking (c, d) uncertainties as a function of $m_{t\bar{t}}$ in (a, c) SR1b and (b, d) SR2b for the 4 TeV Z'_{TC2} signal.

Table 7.15: Summary of the pre-fit relative systematic uncertainties in the expected number of events at Z'_{TC2} masses of 2 TeV and 4 TeV in 1 b-tag signal region (SR1b). The spurious signal uncertainty is different for different test signal masses. The values reported in parenthesis correspond to the 2 TeV mass, the others refer to 4 TeV. The values quoted for the background represent the relative change in the total expected number of events. For the signal uncertainties the values were computed using a Z'_{TC2} signal model with a pole mass of 2 TeV and 4 TeV by comparing the total yields.

Source	**Background (%)**	**4 TeV Z' (%)**	**2 TeV Z' (%)**
Statistical unc.	2.79	2.45	3.11
Fitparameter unc. SR1b	1.00	N/A	N/A
Spurious signal SR1b	0.027 (0.355)	N/A	N/A
AKT10 JES (R_{trk} baseline)	N/A	15.21	10.69
AKT10 JES (R_{trk} modelling)	N/A	50.97	39.04
AKT10 JES (R_{trk} totalstat)	N/A	3.20	0.37
AKT10 JES (R_{trk} tracking)	N/A	9.08	6.99
AKT10 JMS (R_{trk} baseline)	N/A	1.68	0.77
AKT10 JMS (R_{trk} modelling)	N/A	8.98	4.79
AKT10 JMS (R_{trk} totalstat)	N/A	1.60	0.10
AKT10 JMS (R_{trk} tracking)	N/A	1.99	1.34
AKT10 JER	N/A	0.03	0.02
AKT10 JMR (m_{top} reso.)	N/A	4.02	1.98
JVT	N/A	0.15	0.19
Luminosity	N/A	1.70	1.70
pileup modelling	N/A	0.14	0.07
b-tag (E0)	N/A	13.47	11.64
b-tag (E1)	N/A	9.00	5.86
b-tag (E2)	N/A	2.22	1.70
b-tag (E3)	N/A	1.20	0.25
c-mistag (E0)	N/A	0.23	0.32
c-mistag (E1)	N/A	0.35	0.44
c-mistag (E2)	N/A	0.14	0.09
b-tag extrap.	N/A	15.86	3.92
c-mistag extrap.	N/A	0.01	0.01
light-mistag (E0)	N/A	1.87	0.52
light-mistag (E1)	N/A	0.73	0.18
light-mistag (E2)	N/A	0.15	0.02
light-mistag (E3)	N/A	0.10	0.02
t-tag eff. (dijet modelling)	N/A	0.42	0.20
t-tag eff. (dijet stat.)	N/A	0.35	0.13
t-tag eff. (γ+jets modelling)	N/A	0.02	0.18
t-tag eff. (γ+jets Stat.)	N/A	0.01	0.09
t-tag eff. (b-tag unc.)	N/A	4.25	4.08
t-tag eff. ($t\bar{t}$ modelling, had.)	N/A	2.00	2.12
t-tag eff. ($t\bar{t}$ modelling, ME)	N/A	3.82	0.89
t-tag eff. ($t\bar{t}$ modelling, rad.)	N/A	1.10	1.28
t-tag eff. (high-p_T ext.)	N/A	6.63	6.27
t-tag eff. ($t\bar{t}$ stat.)	N/A	4.51	4.42
Total systematic unc.	1.00 (1.06)	57.39	44.59

Table 7.16: Summary of the pre-fit relative systematic uncertainties in the expected number of events at Z'_{TC2} masses of 2 TeV and 4 TeV in 2 b-tag signal region (SR2b). The spurious signal uncertainty is different for different test signal masses. The values reported in parenthesis correspond to the 2 TeV mass, others refer to 4 TeV. The values quoted for the background represent the relative change in the total expected number of events. For the signal uncertainties the values were computed using a Z'_{TC2} signal model with a pole mass of 2 TeV and 4 TeV by comparing the total yields.

Source	**Background (%)**	**4 TeV Z' (%)**	**2 TeV Z' (%)**
Statistical unc.	4.78	3.17	1.96
Fitparameter unc. SR2b	1.80	N/A	N/A
Spurious signal SR2b	0.044 (0.433)	N/A	N/A
AKT10 JES (R_{trk} baseline)	N/A	13.28	9.47
AKT10 JES (R_{trk} modelling)	N/A	50.01	41.74
AKT10 JES (R_{trk} totalstat)	N/A	1.67	0.21
AKT10 JES (R_{trk} tracking)	N/A	7.29	6.29
AKT10 JMS (R_{trk} baseline)	N/A	1.13	0.70
AKT10 JMS (R_{trk} modelling)	N/A	7.63	4.14
AKT10 JMS (R_{trk} totalstat)	N/A	1.16	0.13
AKT10 JMS (R_{trk} tracking)	N/A	1.52	1.14
AKT10 JER	N/A	0.01	0.03
AKT10 JMR (m_{top} reso.)	N/A	4.04	1.99
JVT	N/A	0.16	0.20
Luminosity	N/A	1.70	1.7 0
pileup modelling	N/A	0.11	0.11
b-tag (E0)	N/A	2.02	4.04
b-tag (E1)	N/A	1.29	2.02
b-tag (E2)	N/A	0.33	0.61
b-tag (E3)	N/A	0.16	0.11
c-mistag (E0)	N/A	0.45	0.23
c-mistag (E1)	N/A	0.73	0.39
c-mistag (E2)	N/A	0.33	0.16
b-tag extrap.	N/A	0.98	1.24
c-mistag extrap.	N/A	0.05	0.01
light-mistag (E0)	N/A	4.30	1.59
light-mistag (E1)	N/A	1.50	0.34
light-mistag (E2)	N/A	0.31	0.04
light-mistag (E3)	N/A	0.19	0.02
t-tag eff. (dijet modelling)	N/A	0.32	0.12
t-tag eff. (dijet stat.)	N/A	0.28	0.09
t-tag eff. (γ+jets modelling)	N/A	0.02	0.13
t-tag eff. (γ+jets Stat.)	N/A	0.01	0.07
t-tag eff. (b-tag unc.)	N/A	4.31	4.21
t-tag eff. ($t\bar{t}$ modelling, had.)	N/A	2.06	2.21
t-tag eff. ($t\bar{t}$ modelling, ME)	N/A	3.66	0.89
t-tag eff. ($t\bar{t}$ modelling, rad.)	N/A	1.14	1.34
t-tag eff. (high-p_T ext.)	N/A	6.72	6.46
t-tag eff. ($t\bar{t}$ stat.)	N/A	4.58	4.57
Total systematic unc.	1.80 (1.85)	53.68	44.28

7.12 Statistical analysis

The statistical analysis is carried out in three steps:

- Model-independent resonance search using BumpHunter. Any resonant signal produced by any sources will be identified in this step.
- Hypothesis test based on the benchmark signal. A profile-likelihood-ratio method is used to quantify the significance.
- If no significant deviation is found in steps 1 and 2, a Frequentist CL_s method [204] is used to compute observed and expected upper limits for benchmark signal masses.

The profile likelihood ratio introduced in subsection 5.3.3 is used to build the test statistic. Systematic uncertainties are added to the likelihood following the same method discussed in subsection 5.2.1. All the nuisance parameters, θ, are redefined for convenience such that $\theta_0 = 0$ and $\sigma_\theta = \Delta\theta = 1$ for the nominal hypothesis. One dedicated nuisance parameter, θ_{SS}, corresponding to the spurious signal uncertainty is added to the likelihood function shown in Equation 5.19. The the likelihood function takes the following form

$$L(\mu,\theta,\theta_{SS}) = p(n|\mu;\theta;\theta_{SS}) = \prod_{i=1}^{N} \text{Poisson}(n_i \mid \nu_i(\mu,\theta,\theta_{SS})) \prod_{k\in\text{syst}} \text{Gaus}(\theta_k \mid \theta_k^0;\sigma_{\theta_k}) \,. \tag{7.14}$$

The HistFactory [244] software is used to build the likelihood functions. The HistFactory software is based on RooFit [245] and RooStats [246] software. These tools and packages are used through a steering tool called TRExFitter. This package was developed for use in ATLAS. The POI is not constrained in the likelihood and is determined using the MINOS algorithm. The test statistic is built using the negative log of profile likelihood ratio as described in subsection 5.3.4. Two different types of profile likelihood fits are done:

- s + b fit: the likelihood includes both signal and background model.
- b-only fit: the likelihood includes only the background model.

Data is always required in ordered to define the likelihoods. The nuisance parameters including, θ_{SS} are constrained in the likelihood with Gaussian terms. The signal strength is the only unconstrained parameter in the likelihood. All signal systematics uncertainties are assumed to be fully correlated across the two signal regions. The background uncertainties are treated as fully uncorrelated in the two signal regions and they are defined separately for each signal region. In order to understand the nature the fit and behavior of the systematic uncertainties the profile likelihood fits are also performed on the Asimov dataset, created from background-only spectrum and signal+background spectrum. The fit results with the Asimov dataset can be found in section B.3 of the Appendix B. This section contains only fit results where actual data is used in the likelihood. In all these fits, the background is estimated with 4-parameter polynomial functions described in subsection 7.9.4. In the s+b fits, the 4 TeV Z'_{TC2} signal with 1.2% relative width is used. The binning of the $m_{t\bar{t}}$ distributions are made slight larger than what used for estimating background. It is done to increase the number of events in the high $m_{t\bar{t}}$ bins and make the profile likelihood fit stable.

7.12.1 b-only profile likelihood fit

The b-only fit only includes the background uncertainties in the likelihood. The best-fit values of the parameters, often referred to as *post-fit* parameter values, and their uncertainties give useful insight about the background uncertainties. The pull of a nuisance parameter is defined by comparing the best-fit value, $\hat{\theta}$, with its pre-fit value, θ_0, and dividing the different by its pre-fit uncertainty, $\Delta\theta$. So, the pull of a nuisance parameter can be written as

$$\text{pull} = \frac{\hat{\theta} - \theta_0}{\Delta\theta} .$$

If $\hat{\theta} = \theta_0$, then the pull is zero. The post-fit nuisance parameter pulls are shown in Figure 7.29. The error bars represent the post-fit uncertainty, $\Delta\hat{\theta}$. A post-fit error bar smaller than ±1, indicates that the nuisance parameter was constrained by the fit. In the b-only fit, the fit parameter uncertainties get constrained to ∼ 60% of the input value in both signal regions and get pulled by a small amount. The fit parameter uncertainty in $1b$ signal region is pulled by roughly 0.4σ. The constraints are very similar to those seen in the Asimov data results shown in Figure B.7 of Appendix B.

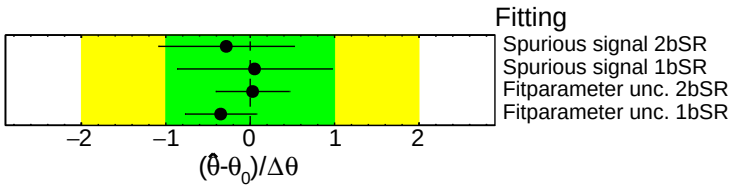


Figure 7.29: Post-fit (b-only) nuisance parameter (NP) pulls observed in data. All four background NPs can be constrained by the fit.

7.12.2 s+b profile likelihood fit

In this case, the signal model is added to the likelihood. So all the signal systematic uncertainties also get added to the likelihood. Along with all the nuisance parameters corresponding to the signal systematics, the signal strength, μ, is included in the likelihood. μ is an unconstrained parameter in the likelihood and is allowed to take both positive and negative values. The best-fit signal strength value is $\hat{\mu} = -0.24^{+0.47}_{-0.33}$, where the uncertainties include both statistical and systematics uncertainties. The correlations between the likelihood parameters in the fit are calculated using the Hessian matrix. The parameters with correlation above 0.1 are only added in the matrix shown in Figure 7.30. The top-tagging and b-tagging systematics have small correlations with each other, and also with μ. The fit parameter uncertainty and spurious signal uncertainty have a higher correlation with μ.

The signal strength is the parameter of interest in this study, and the impact of the nuisance parameters on the fit is evaluated by studying the change in the best-fit μ value when the nuisance parameters are fixed to certain values. The impact $\Delta\mu$ of a given uncertainty is the difference between the best-fit values in the fit with the the NP set to it's best-fit value, $\hat{\theta}$, and two other fits in which it is fixed to a value $\hat{\theta} \pm \sigma$, where σ is either pre-fit or post-fit uncertainty. The pre-fit impact of a NP is obtained by fixing it to $\hat{\theta} \pm \Delta\theta$, where $\Delta\theta = 1$. Whereas the post-fit impacts

are obtained by fixing the nuisance parameter value to $\hat{\theta} \pm \Delta\hat{\theta}$, where $\hat{\theta}$ is the post-fit uncertainty. The post-fit impacts can be smaller than the pre-fit impacts. The nuisance parameters are shown ranked on the size of their impact on the signal strength, $\Delta\mu$. One such ranking plot with the top twenty systematic uncertainties is shown in Figure 7.31. The pre-fit impacts are shown in open blue and cyan horizontal bars, whereas the filled blue and cyan bars show the post-fit impact. The spurious signal uncertainty in the $2b$-tag signal region has the highest post-fit impact and changes μ by 0.21. The JES modeling uncertainty has the second-highest post-fit impact. Other JES and JMR uncertainties are also large uncertainties along with the fit-parameter and spurious signal uncertainties.

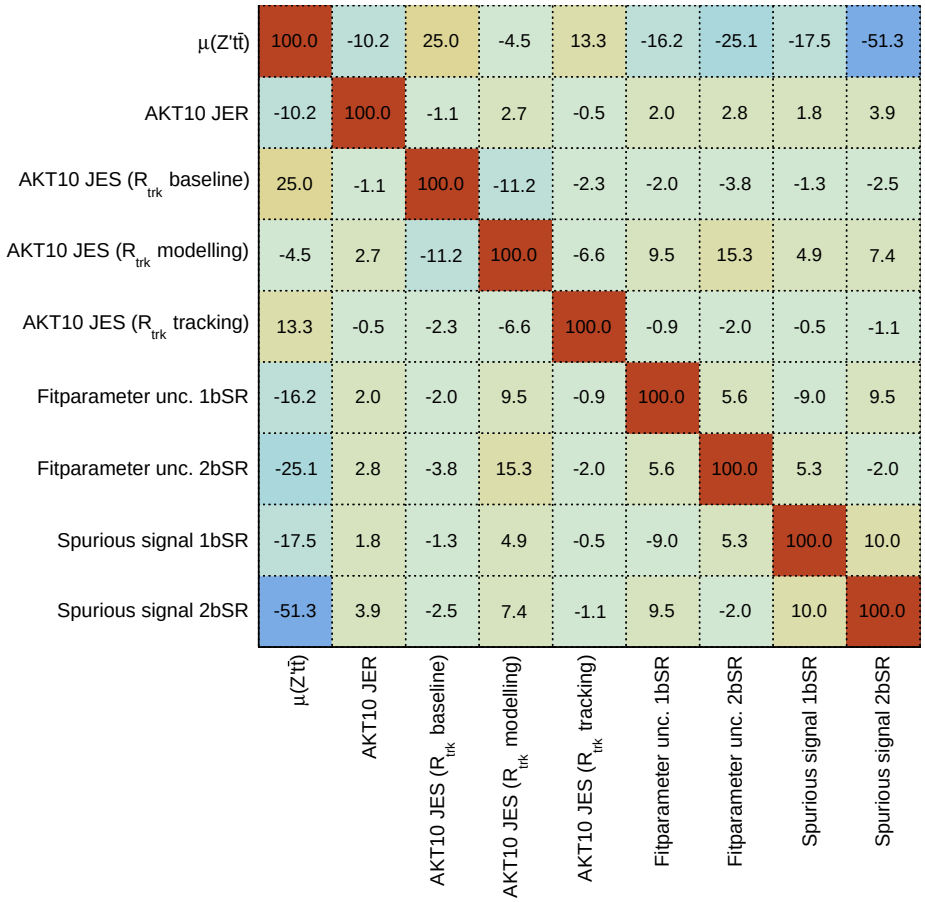


Figure 7.30: Post-fit $(s+b)$ correlation matrix of the systematic NPs and signal strength for Z' with a mass of 4 TeV. The background uncertainties have 15 – 45% correlation with the POI. Amongst signal systematics, JES have up to 25% correlation with the parameter of interest. Only variables with a correlation > 0.1 are shown.

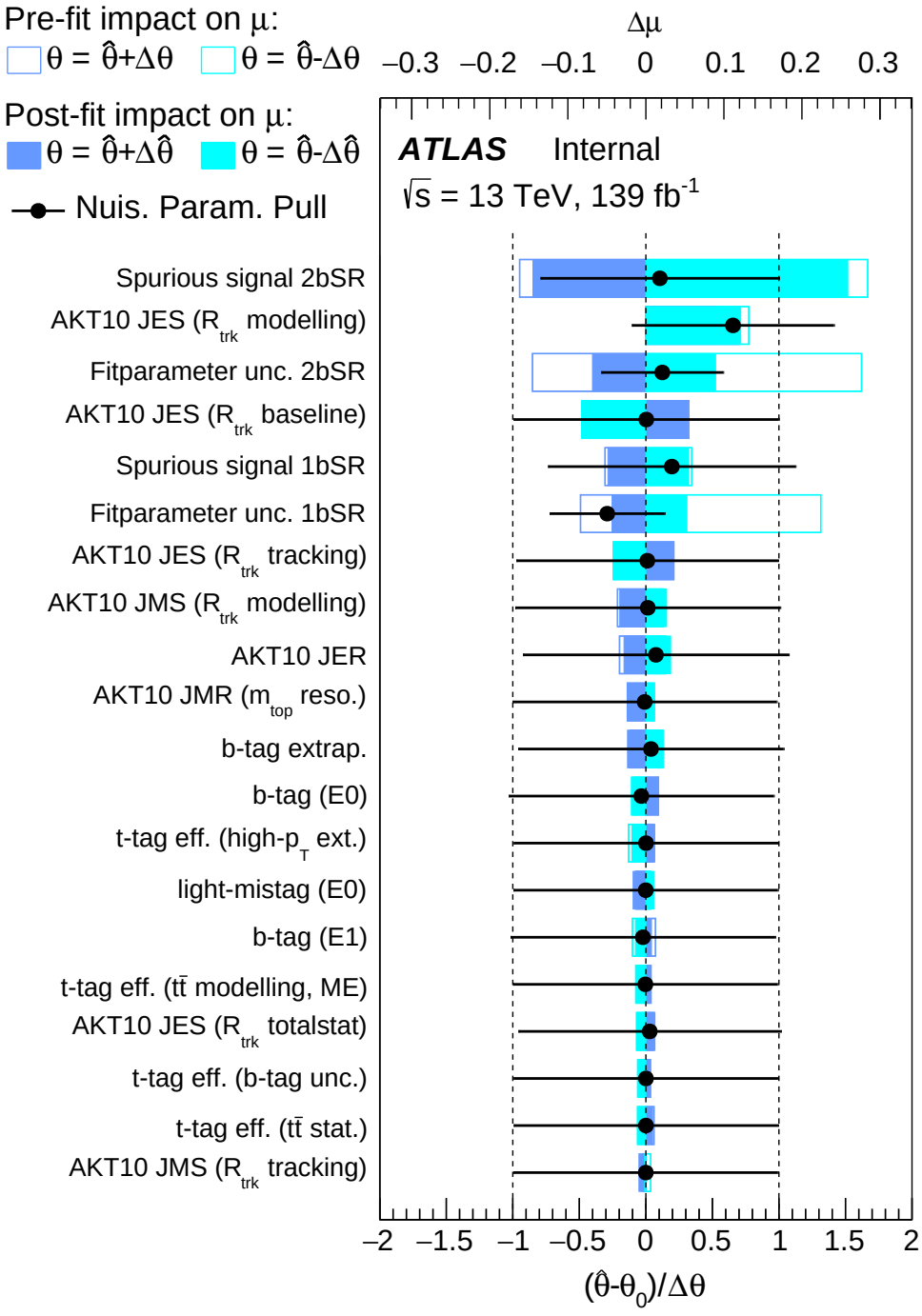


Figure 7.31: Systematics ranking plot for Z' mass of 4 TeV.

7.12.3 BumpHunter search results

The model independent BumpHunter algorithm, described in section 6.8, is used in this analysis to search for resonant signals. The algorithm searches for excesses on the $m_{t\bar{t}}$ spectrum. The minimum search window is set to two bins and the maximum search window is half of the spectrum for this search. The BumpHunter search results in the two signal regions are shown in Figure 7.24.

The BH local significance is almost always less than 2σ. The two vertical lines extending between the upper and lower panel represent the most significant deviation interval. The BH global p-values are 0.45 and 0.56 in the SR1b and SR2b, respectively. These correspond to around 0.2σ. So, it can be concluded no significant deviation is observed in data.

7.12.4 Model dependent search results

The model-dependent hypothesis test for significance (p_0) scan is based on the *profile likelihood ratio* method. The likelihood is built using the binned histograms. Z'_{TC2} signals with $\Gamma/m = 1.2\%$ are used in the search phase. The search is performed for signal pole mass between 1.75 TeV and 5.0 TeV. The local significances are calculated by combining the two signal regions in the likelihood. The significance scan result shows data is consistent with the Standard Model (background) hypothesis, and the local significance is always less than 1.56σ. The local p_0-values are shown in Figure 7.32 as a function of Z'_{TC2} pole mass. All the local p_0-values and the corresponding significances are summarized in Table B.1.

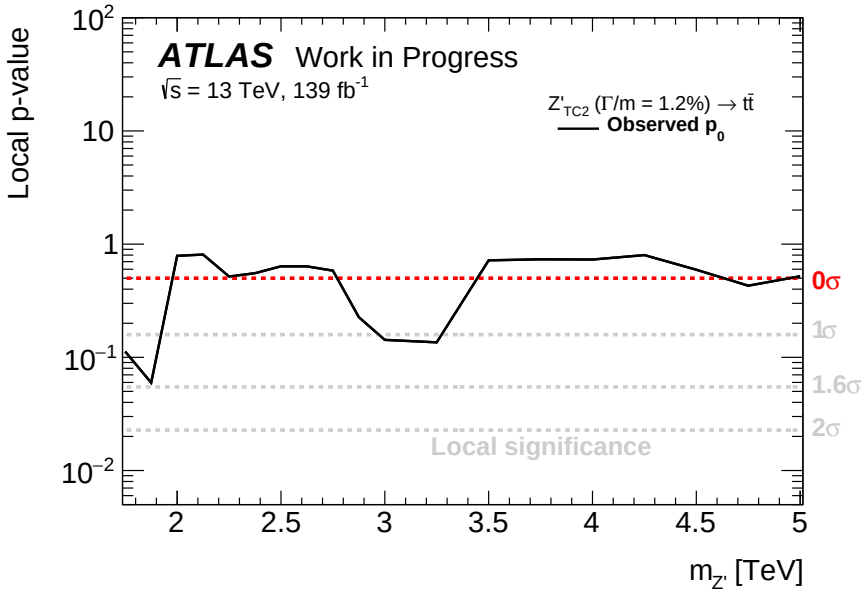


Figure 7.32: Local p_0-value and significances are calculated using profile likelihood ratio-based test statistic with Z'_{TC2} ($\Gamma/m = 1.2\%$) signals between 1.75 and 5.0 TeV. The search result shows data is consistent with the background hypothesis as local significance values are always less than 1.6σ in the search region.

7.13 Limit setting

In absence of new signal, 95% CL limit on cross-section (σ) × branching ratio (BR) of resonant new physics processes is set with frequentist CL_s method (subsection 5.3.6) using asymptotic formula described in subsection 5.3.5. The signal strength μ values are scanned within a specified range, and the CL_s is calculated for each μ. The 95% CL upper limit on signal strength, μ_{up} is the largest positive value of μ which gives $CL_s \leq 0.05$. μ_{up} is then converted to the upper limit for $\sigma\times$ BR by multiplying with the signal cross-section. The expected limit and the $\pm1\sigma$ and $\pm2\sigma$ uncertainty bands are calculated using the Asimov dataset using the asymptotic approximation without generating pseudo-experiments. The background and signal systematic uncertainties summarized in section 7.11 are included in the limit calculation. Statistical uncertainties are also included by adding bin-by-bin normalization factors (the γ factors). Similar to the p_0 scan, the limits are calculated by combining the two signal regions in the likelihood. All signal systematic uncertainties are assumed to be fully correlated across the two signal regions whereas the background uncertainties are treated as fully uncorrelated in the two signal regions. Expected and observed 95% CL upper limits are shown in Figure 7.33. The cross-section limits are calculated using the Z'_{TC2} with width $\Gamma/m = 1.2\%$ for masses from $m_{Z'_{TC2}} = 1.75$ TeV to $m_{Z'_{TC2}} = 5$ TeV. The first point is at 1.75 TeV to make sure the full signal peak is included in the mass spectrum, which starts at 1.4 TeV. Limits are calculated for the following simulated signal models (bold) and with interpolated signal models

- **1750**, 1875, **2000**, 2125, **2250**, 2375, **2500**, 2625, **2750**, 2875, **3000**, 3250, 3500, 3750, **4000**, 4250, 4500, 4750 and **5000** GeV.

Since the detector resolutions are much larger than the signal width, the calculated cross-section × branching ratio limits also hold for other narrow width signal models where $\Gamma/m = 1\%$ and 3%. The lower limits on Z' mass are calculated from the intersection of the observed limit and theory cross-section lines. Z'_{TC2} masses with $\Gamma/m = 1.2\%$ are excluded up to 4.1 TeV at 95% CL. For other Z'_{TC2} models with $\Gamma/m = 1\%$ and 3%, the Z'_{TC2} masses are excluded at 95% CL up to 3.9 and 4.7 TeV, respectively.

The mass exclusion limits on different Z'_{TC2} models are summarized in Table 7.17. Cross-section times branching ratio limits are calculated separately for the two signal regions to compare their sensitivity. It is seen that most of the sensitivity comes from the 2b signal region. Throughout this analysis, we used the asymptotic approximation of the test statistic distribution. The limits are also calculated from pseudo-experiments to check the validity of the asymptotic approximation. The observed $\sigma\times$ BR limits from the asymptotic approximation are found to be stronger than those from the pseudo-experiments by at most 20% at mass above 4 TeV. Since the difference was small, the limits calculated using the asymptotic approximation are used.

7.13.1 Comparison to previous results

The limits obtained with full Run-2 data are compared with the 36.1 fb^{-1} Run-2 results [229], where the analysis strategy was very different. The previous analysis used a two-variable top-tagger where the top quarks were identified based on jet mass (m^{comb}) and τ_{32}. It was shown in Figure 4.12 that the DNN top-tagger performs much better compared to the two-variable taggers. The background-rejection improved by a factor of almost 4 at 80% signal efficiency [197]. The b-tagging algorithm used in ATLAS also improved. The currently used DL1 algorithm performs better than the previously used MV2 algorithm, as was shown in Figure 4.10. The background

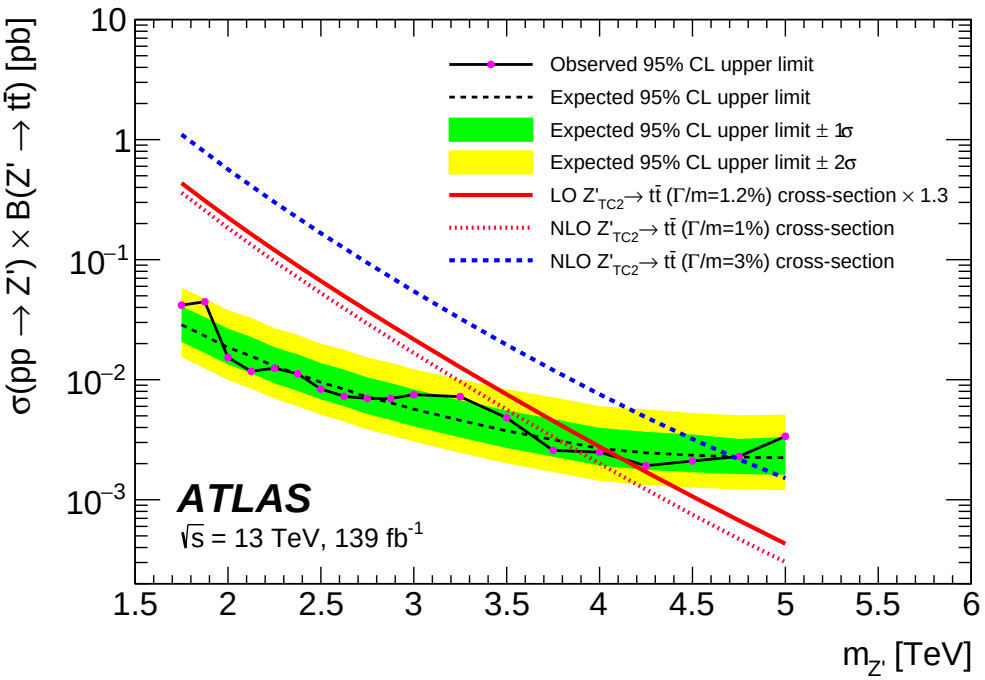


Figure 7.33: Observed and expected 95% CL upper limits on the cross-section times branching fraction of the $Z' \to t\bar{t}$ as a function of the Z' mass. The NLO theory cross-sections times branching fraction for the Z'_{TC2} with $\Gamma/m = 1\%$ and 3% are shown by the dotted and dashed lines, respectively. Shown by the solid line is the LO theory cross-section times branching fraction scaled by a factor 1.3 for the Z'_{TC2} with $\Gamma/m = 1.2\%$ [228].

Table 7.17: Summary of the expected and observed mass exclusion limits of Z' signals with $\Gamma = 1.2\%, 1\%, 3\%$. For the Z' with $\Gamma/m = 1\%$, 1.2% and 3%, the Z' masses are excluded at 95% CL up to 3.9, 4.1 and 4.7 TeV, respectively.

	Signal	**Mass mass upper limit**	
		Expected [GeV]	Observed [GeV]
Z'	$\Gamma/m = 1.2\%$	4027	4141
	$\Gamma/m = 1\%$	3801	3903
	$\Gamma/m = 3\%$	4724	4726

modeling strategy has been changed completely for this analysis. The total background is estimated from the data, whereas in the past, the $t\bar{t}$ was modeled using simulation, including all the associated $t\bar{t}$ modeling uncertainties.

To compare the performance, the current analysis is repeated for only the 36.1 fb^{-1} dataset. The cross-section × BR limits improved by a factor of 2 over the previously published results due to the analysis improvements. By the end of Run-2, the luminosity increased roughly by a factor of 4, which improves the limits by another factor of 2. These comparisons are shown in Figure 7.34. The mass limit on the Z'_{TC2} model improves by almost 1 TeV.

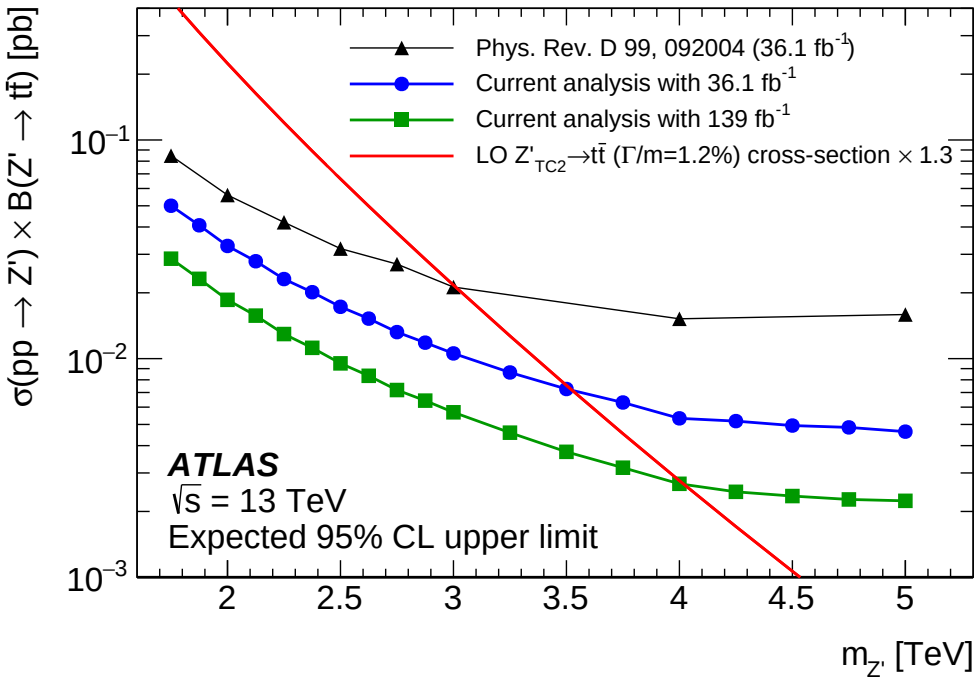


Figure 7.34: Comparison of expected upper limits on the cross-section times branching fraction of $Z'_{\mathrm{TC2}} \to t\bar{t}$ as a function of the Z' mass. The triangle markers show the expected limits from the previous analysis using 36.1 fb^{-1} data from 2015-2016. The circle markers show the expected limits obtained from that 36.1 fb^{-1} data by using the methods of current analysis. The square markers show the current expected limits obtained using 139 fb^{-1} of the full Run-2 data between 2015 and 2018. The NLO theory prediction, obtained by multiplying the LO cross-sections by a factor 1.3, for the Z'_{TC2} with $\Gamma/m = 1.2\%$ is shown by the solid line [228].

7.14 Inference with dark matter mediator model

Statistical inference is also done with the dark matter mediator models. In dark matter mediator models, normal matter interacts with dark matter, χ, via a high mass mediator particle. The search considers both vector mediators, Z'_V, and axial-vector mediators, Z'_A, proposed in the simplified models by the LHC dark matter working group [118, 119]. The five parameters of this model are the Z' coupling to dark matter particle (g_χ), quarks (g_q), leptons (g_l), and mediator Z' mass, dark matter mass, m_χ. The mediator mass is varied between 500 GeV and 5 TeV and the coupling parameters are set to $g_q = 0.25$, $g_\chi = 1$, $g_l = 0$ and $m_\chi = 10$ GeV. The two-dimensional limits in third-generation quark coupling (g_q) and mediator mass ($m_{Z'}$) are calculated using the full Run-2 data. Figure 7.35 shows the upper limit on the coupling to quarks, g_q, as a function of the axial-vector mediator mass, $m_{Z'_A}$. 95 % CL limits on third generation quark coupling lie in the range [0.1, 0.23] for mediator masses between 100 GeV and 200 GeV.

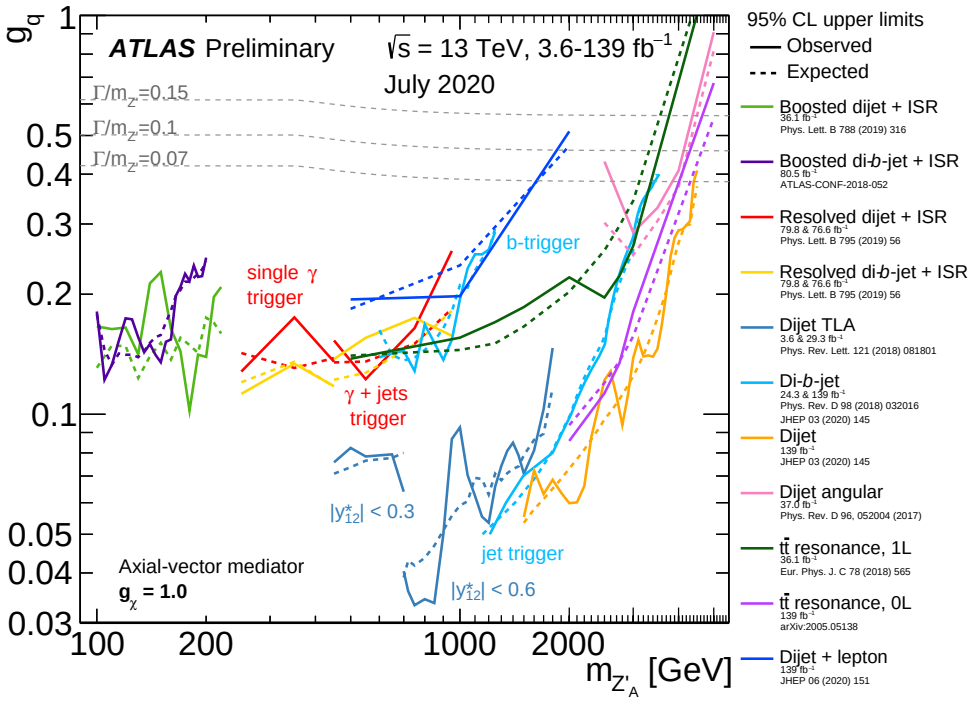


Figure 7.35: 95% CL upper limit on the quark coupling, g_q, as a function mediator mass, $m_{Z'_A}$. The full Run-2 all-hadronic $t\bar{t}$ results are shown in the purple. The quark coupling is excluded up to 0.1 for mediator masses of 200 GeV. Figure taken from [247]

7.15 Conclusion

A $t\bar{t}$ resonance search in the boosted all-hadronic channel was performed using the full Run-2 data collected during the years 2015-2018. Since no new signal is found, upper limits are placed on narrow width signal cross-section × BR at 95% CL. With an improved analysis strategy and more data, the limits improved by almost a factor of four. The current mass limit on narrow width (1.2%) Z'_{TC2} is 4.1 TeV. Although the results improved significantly, there is still room for further improvement. The jet energy scale uncertainties and the top-tagging uncertainties have large impacts on the sensitivity. With an improved jet calibration, the uncertainties could already be reduced by 50%. Jet mass resolution for high p_T jets improves if we include the track information [248]. Several other developments on jet reconstruction are currently ongoing, and new jet reconstruction approaches are being studied in ATLAS. The QCD background rejection is expected to improve by almost a factor of 1.5 after training the top-tagging algorithms on the new jet collections. Apart from these developments, signal sensitivity might be further improved using multivariate techniques. Recent developments in likelihood-free inference methods could be used in this analysis to improve the results [249].

CHAPTER

EIGHT

Search for high mass resonances in the semi-leptonic top-antitop final sate

This chapter summarizes the search for $t\bar{t}$ resonances in the semi-leptonic final state, where one of the top-quarks decays leptonically, and the other one decays hadronically. In the all-hadronic analysis, described in chapter 7, only the high momenta top-quarks were explored, whereas, in this search, the full momentum range of the decaying top-quarks is covered. Based on the decay topology of the top-quark, two categories are defined: the *resolved* final state where the momentum of the decaying top-quark is low, typically below 300 GeV, and the *boosted* final state where the momentum of the top-quark is higher. The motivation behind this search is the same as in the all-hadronic channel. This search is complementary to the all-hadronic $t\bar{t}$ resonance search in that the datasets are orthogonal, and here lower masses are also studied. The lower mass region will help us to probe signals with much weaker coupling than the signal models. This search is performed with the full Run-2 proton-proton collision data (139 fb^{-1}) collected during 2015-2018. This analysis is looking for enhanced $t\bar{t}$ production via a heavy boson. The search results are interpreted in terms of a heavy boson, Z', using several benchmark BSM models described in section 1.5. A similar search was carried out with partial Run-2 ATLAS data (36.1 fb^{-1}), and no trace of a new signal was found [250]. Z'_{TC2} with 1% relative width was excluded below 3 TeV by the data. The current analysis is not finalized, so the data distribution of the main discriminant invariant mass of the $t\bar{t}$ system is still blinded. Hence, the statistical results shown in the chapter are expected results only and evaluated using the Asimov dataset.

This chapter is organized as follows. An overview of this analysis is presented in section 8.1. The semi-leptonic top decay topology is described using schematic diagrams in section 8.2. Almost all the objects used in this analysis were already described in chapter 7, so only the main differences are highlighted in this chapter. Different techniques used to estimate the SM background processes are summarized in section 8.3. The information of the simulated signal models are provided in section 8.4. The object definitions and event selections are summarized in section 8.5 and section 8.6, respectively. A different non-conventional method is used in this search to remove overlap between lepton and jet responses, summarized in section 8.7. All the major SM background processes are modelled with simulations, but a data-driven method is used to model the multijet processes. This data-driven method is described in section 8.8. Different reconstruc-

tion strategies are used for top quarks with small and large momentum, discussed in section 8.9. Some control plots are presented in section 8.10 to show the data-background comparison. Finally, some preliminary statistical results are presented in section 8.11.

8.1 Analysis overview

The semi-leptonic top decay has the second-largest $t\bar{t}$ branching ratio and constitutes roughly 30% (considering only electron and muon final states) of all the $t\bar{t}$ decays. The presence of a charged lepton in this final state provides an efficient way of selecting events, using single electron and muon triggers. As a result, the sensitivity of this analysis is generally comparable to the all-hadronic search, despite the lower branching ratio. Requiring at least one charged lepton to be present also suppresses the background contributions from QCD multijet processes. So the QCD multijet background is much lower than in the all-hadronic channel. In the SRs of the all-hadronic $t\bar{t}$ resonance, the QCD multijet process contributes around 20-80% of the total background, whereas in this search, it contributes maximum up 10% of the total background. The analysis searches for a narrow resonance in the $t\bar{t}$ invariant mass spectrum, $m_{t\bar{t}}$. $m_{t\bar{t}}$ is computed from the momentum of the reconstructed tops. The top-quarks are reconstructed using different techniques in the resolved and boosted regions. In the boosted channel, a Deep Neural Network (DNN) based top-tagging method was applied to the large radius jets to identify the hadronic tops. Whereas the tops are reconstructed using a chi-squared algorithm in the resolved region using only the small radius jets. A deep-learning-based b-tagging method is used here to identify the b-jets coming from top decay. b-tagging helps to suppress non-$t\bar{t}$ backgrounds further. The event selection is optimized for narrow width signals. A Feynman diagram of the signal process ($Z' \to t\bar{t} \to qqbl\nu_l b$) is shown in Figure 8.1. The selected events are categorized in different signal

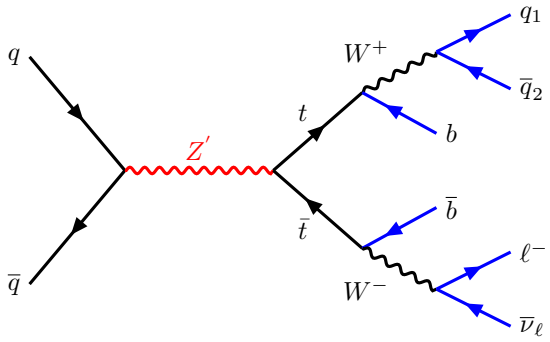


Figure 8.1: Feynman diagram of the Z' signal process where one of the top quarks decay hadronically and the other one decays leptonically.

regions based on the number of b-tagged jets in an event. All the SM background processes, other than QCD multijet, are modeled with Monte Carlo simulations. The QCD multijet background is estimated using a data-driven method, different from the one used in the all-hadronic analysis (described in section B.1 of Appendix B). Unlike the all-hadronic search, the total background is a combination of simulated and data-driven components. Different data-driven methods were tried to estimate the total background, but we could not find a reliable method that fits the data and also can extract a signal from data. The main difficulty of estimating the background directly from data comes from the shape of the $m_{t\bar{t}}$ spectra. It does not have a monotonically decreasing trend

in the search regions. Initially, a model-independent statistical search will be carried out using a bump-hunting technique. A model-dependent statistical search will then be performed using a combined profile-likelihood test-statistic. In the absence of new physics, an upper limit on the signal cross-section × branching ratio with 95% CL will be set. The lower mass limit on the Z'_{TC2} signal with 1.2% relative width is expected to be around 3.6 TeV, which is ∼ 700 GeV more than the previous expected limit [250].

8.2 Top quark decay topology

A boosted hadronic top gets reconstructed as a large-R jet as already described in section 7.2. The lepton and the b-jet coming from the leptonically decaying top get reconstructed as different objects in the resolved regime. But in the boosted region, the lepton and the b-jet are very close to each other. So the reconstruction of these two objects becomes difficult. It becomes even more complicated when the final state lepton is an electron where the calorimeter energy clusters are used to reconstruct the electrons and the jets. So, jets and electrons overlapping (in space) introduces the chance of either double counting the objects or losing both the objects. This situation is demonstrated in the schematic diagram of Figure 8.2.

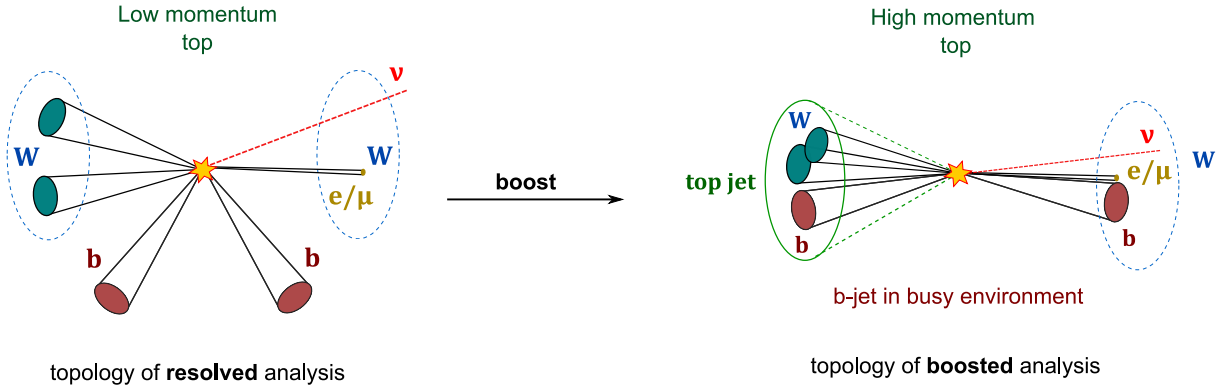
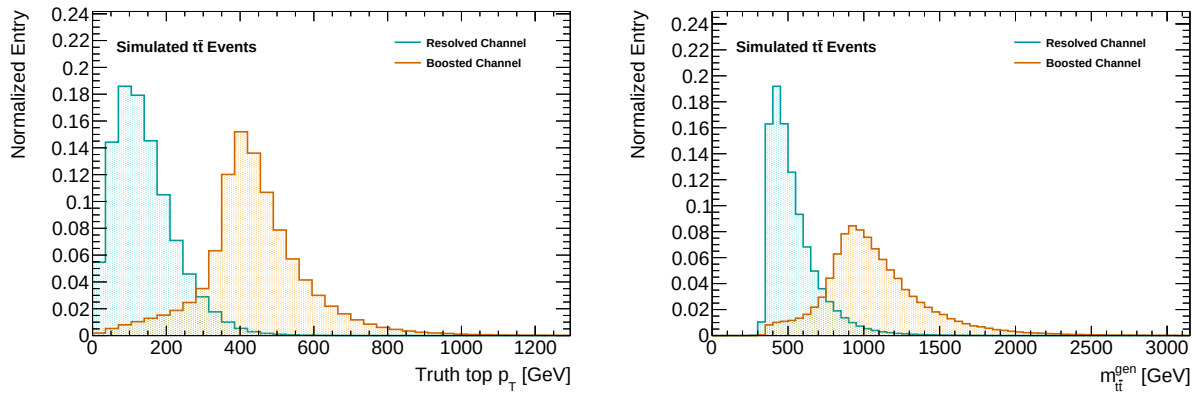


Figure 8.2: Schematic diagram showing the single-lepton $t\bar{t}$ final state in the resolved and boosted regimes.

Figure 8.3a shows the p_T distribution of top and anti-top quarks at the generator level (parton) before final-state radiation. The p_T of the top partons are mostly above 300 GeV in the boosted region whereas they are under 300 GeV in the resolved channel. Figure 8.3b shows the $m_{t\bar{t}}^{\mathrm{gen}}$ distribution, where $m_{t\bar{t}}^{\mathrm{gen}}$ is calculated from the momenta of parton before final-state radiation. The $m_{t\bar{t}}^{\mathrm{gen}}$ is mostly above 600 GeV in the boosted region and less than 600 GeV in the resolved region.

8.3 Background processes

In this search, almost all the background processes are estimated using Monte Carlo simulation. The major irreducible background contribution comes from SM $t\bar{t}$ process via s- and t-channel production. Since we are looking for events with $t\bar{t}$ final state, it is impossible to reduce the SM $t\bar{t}$ contribution significantly. In the boosted regime where the invariant mass of the $t\bar{t}$ system is very high, close to the TeV range, the SM $t\bar{t}$ background becomes smaller. The QCD contributions



(a) (b)

Figure 8.3: Distributions of (a) the p_T of top and anti-top quarks and (b) $m_{t\bar{t}}^{\text{gen}}$ distribution at the generator (parton) level before final-state radiation in the resolved and boosted region. The distributions are normalized to unit area to compare the shape the distributions.

mostly come from the events with at least one fake lepton, where a jet gets reconstructed as a lepton. This background is estimated using a data-driven method. Some other SM processes pass our event selection and constitute the reducible component of the background. Those processes are W+jets, Z+jets, single-top, diboson, and associate production of $t\bar{t}$ with a vector boson ($t\bar{t}V$).

$t\bar{t}$ samples are generated using the same the next-to-leading-order (NLO) Powheg generator along with Pythia 8 with the same settings described in subsection 7.3.1. The single top production is also generated using Powheg and Pythia 8 is used with NNPDF3.0 NLO PDF set for showering and hadronization with A14 tune.

W+jets and Z+jets events are generated using the Sherpa 2.2.1 generator where matrix element calculation includes up to two partons at NLO in QCD, and four partons at the LO using the Comix and OpenLoops matrix element generator. The NNPDF3.0 NLO PDF set is used while generating the events. Generated events are normalized to the NNLO cross-section.

Diboson (WW, WZ, ZZ) processes are generated using the Sherpa 2.2.1 generator. The final states have four charged leptons (4ℓ), three charged lepton and a neutrino ($3\ell + \nu$), two charged leptons and two neutrinos (2ℓ+2ν) or one charged lepton and one neutrino ($\ell\nu q\bar{q}'$). The NNPDF3.0 NLO PDF set is used with the dedicated parton shower tuning developed by the Sherpa authors.

The same simulated multijet events as used in the all-hadronic $t\bar{t}$ analysis (subsection 7.3.2) are used for some initial studies. Since the simulated multijet process does not describe the data well, the multijet background is estimated from data using a *matrix method*, a simplified version of section 6.4.1. It is briefly described in section 8.8.

8.4 Signal modeling

Many different signal models are considered in this search. The resonance widths vary from very narrow (1% of the mass) to 30% of the $t\bar{t}$ invariant mass. The signal processes are generated with Monte Carlo simulation. Like the all-hadronic analysis, signal interference with SM $t\bar{t}$ is not considered for the Z' signal model due to its negligible effect. The primary benchmark signal model is the Top-color assisted-technicolor (TC2) model. The narrow width (1.2% width) Z'_{TC2} events

are the same described in subsection 7.4.1. Spin-1 color-octet Kaluza-Klein gluon, g_{KK}, spin-2 color-singlet Kaluza-Klein graviton, G_{KK} are also used as benchmark signal models. The $g_{KK} \to t\bar{t}$ process is simulated with Pythia 8.165 at LO with NNPDF2.3 LO PDF set and the A14 set of tuned parameters. The spin-2 color-singlet signal is modeled using MadGraph5_aMC@NLO NNPDF2.3 LO PDF set, with parton showering performed by Pythia 8.165 with the A14 set of tuned parameters.

The final results will also be interpreted using dark matter mediator models described in subsection 1.5.4. The s-channel mediators with vector, $Z'_{DM,vec}$, and axial-vector, $Z'_{DM,ax}$, coupling are considered in this analysis. This process is simulated using MadGraph5_aMC@NLO NNPDF23LO PDF set, where the parton showering is done with Pythia 8. This process is simulated only at truth level for some representative signal mass points. To save computing resources, the reconstructed level distributions are obtained by rewighting the Z'_{TC2} events. The Heavy Vector Triplet model described in subsection 7.4.2 will be used to combine these results with those from other searches. The modeling of the signal processes are summarized in Table 8.1.

Table 8.1: Simulation details of the signal processes used in the semi-leptonic $t\bar{t}$ resonance search.

Process	Generator	Order	Corrected	Parton Shower	PDF
$Z'_{SSM} \to t\bar{t}$	Pythia 8.186 [210]	LO	NNLO	Pythia 8.186 [210]	NNPDF23LO [220]
$g_{KK} \to t\bar{t}$	Pythia 8.165 [210]	LO	LO	Pythia 8.165 [210]	NNPDF23LO [220]
$G_{KK} \to t\bar{t}$	MG5_aMC@NLO v2_6_0 [240]	LO	LO	Pythia 8.165 [210]	NNPDF23LO [220]
$Z'_{DM} \to t\bar{t}$	MG5_aMC@NLO v2_6_0 [240]	LO	LO	Pythia 8.186 [210]	NNPDF23LO [220]
$Z'_{HVT} \to t\bar{t}$	MG5_aMC@NLO v2_6_0 [240]	LO	LO	Pythia 8.186 [210]	NNPDF23LO [220]

8.5 Object definitions

The set of relevant objects for this analysis is very similar to what is used in the $t\bar{t}$ resonance search in the all-hadronic final state described in section 7.5. One big difference is that the leptons are important in this analysis, and they are used for triggering and selecting events. The overlap removal procedure between the leptons and the jets also becomes crucial for this search. Along with leptons and jets, the missing transverse energy (MET) is used in this search to reconstruct the neutrino momentum. Since all other objects were already defined in section 7.5, they are not repeated here. This section only includes the MET definition.

Missing transverse energy

In the leptonic top-quark decay, we expect one neutrino in the final state. The ATLAS detector cannot detect the neutrinos, so their momenta are estimated using the MET information. A summary of MET reconstruction was given in section 4.12, and the same procedure is followed here. A *tight* operating point is used where forward jets are required to have $p_T > 30$ GeV.

8.6 Event selection

This search considers two different top-quark decay topologies. So two different sets of criteria are defined to select the events. A few requirements are common in both cases. All the events

are expected to have at least one primary vertex. The primary vertex is defined as the vertex with highest $\sum p_{T,\text{track}}^2$, where $p_{T,\text{track}}$ is the p_T of the associated tracks. The events are cleaned by removing pre-identified noise bursts and corrupted data. The events in the data are required to pass a good run list of the corresponding data-taking period. Selected data events need to pass either a single-electron or a single-muon trigger. Similar offline trigger implementations are used to select simulated events. The exact trigger names are listed in Table 8.2. All the events are required to have exactly one reconstructed lepton, either an electron or a muon with $p_T > 30$ GeV. Since we expect to see MET coming from the neutrino, events with MET > 20 GeV are selected to suppress the multijet and other non-$t\bar{t}$ events. In addition to the MET cut, a combination MET and W boson transverse mass, MET + $m_{T,W} > 60$ GeV is required[15]. This cut also helps to reduce backgrounds.

Table 8.2: List of the single-lepton high Level triggers (HLT) used in this analysis. All the triggers are used together with an 'OR' condition.

Year	lepton-channel	Trigger
2015	electron	e24_lhmedium_L1EM20VH, e60_lhmedium, e120_lhloose
	muon	mu20_iloose_L1MU15, mu50
2016-18	electron	e26_lhtight_nod0_ivarloose, e60_lhmedium_nod0, e140_lhloose_nod0
	muon	mu26_ivarmedium, mu50

In order to reduce the QCD background coming from the light-flavored jets, at least one b-tagged jet is required in the selected events. Either the calorimeter jets with $R = 0.4$ or the variable radius track jets can be used for b-tagging. Both the options are studied, and the performance is similar. And it is decided that the calorimeter jets will be used for b-tagging. The b-tagging is done on the using deep learning algorithm (DL1) with 77% efficiency working point.

8.6.1 Boosted selection

In the boosted final state, we target top-quarks with high momenta. As a result, the top-quark decay products will be very collimated along the direction of the top-quark. So a boosted hadronically decaying top-quark forms a single large-R jet in the calorimeter. Hence, the events in the boosted region are required to have at least one top-tagged large-R jet. The top-tagging is done with a DNN top-tagger using 80% efficiency working point. On the leptonic side, the lepton is, in general, close to a high p_T jet coming from the b-quark decay. This is ensured by requiring ΔR (lepton, R=0.4 jet) < 2.0. If multiple jets satisfy these criteria, we select the highest p_T jet and call it the *selected jet*. In the signal events, the leptonic and the hadronic top-quark are almost back-to-back in the angular planes. To ensure that, we further require

- ΔR (R=1.0 jet, lepton) > 1.0
- ΔR (R=1.0 jet, selected R=0.4 jet) > 1.0

8.6.2 Resolved selection

In the resolved top-quark decay topology, we require at least four $R = 0.4$ jets in the event, at least one of which should be b-tagged. Although four jets are expected at leading order, we often see

[15] $m_{T,W}$ is the transverse mass of the W boson.

more than four jets in the event due to initial and final state radiation making it more challenging to construct the $t\bar{t}$ system.

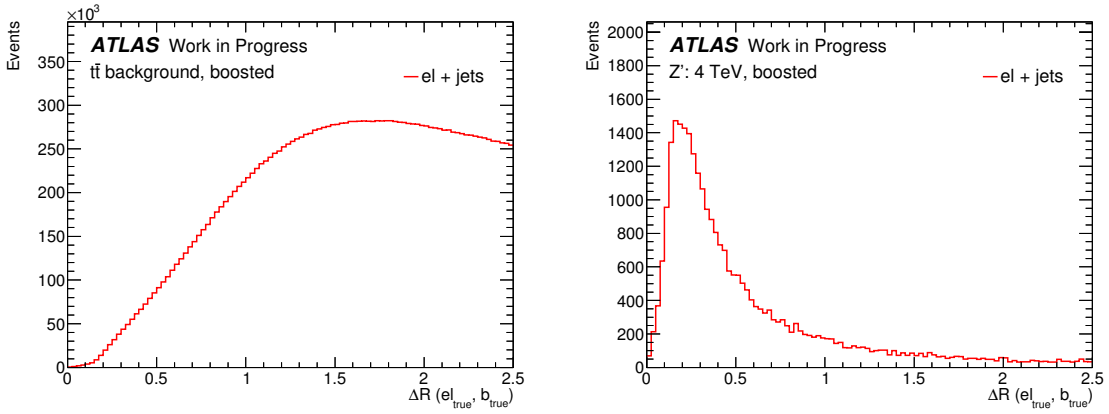
A summary of the event selection is shown in Table 8.3.

Table 8.3: Summary of the event selection criteria used in the semi-leptonic $t\bar{t}$ resonance search.

Selection	**Criteria**
Common Selection	
GRL	Data events in the Good Run List (2015-2018)
Single lepton trigger	Separate el and μ triggers: Table 8.2
Event cleaning	Remove calorimeter noise, LAr, SCT errors, noise burst
Exactly one lepton	== 1 e or μ with $p_T > 30$ GeV
MET cut	MET > 20 GeV
MET + W transverse mass	MET + $m_{T,W} > 60$ GeV
b-tagging	≥ 1 b-tagged jet
Boosted Selection	
Large-R jet	≥ 1 large-R jet, $p_T > 300$ GeV
top-tagging	≥ 1 with DNN top-tagger (80% WP)
Close-to-lepton jet	≥ 1 jet with ΔR (lepton, R=0.4 jet) < 2.0 $\rightarrow$ **selected jet**
Back-to-back	ΔR (R=1.0 jet, lepton) > 1.0
Back-to-back	ΔR (R=1.0 jet, selected R=0.4 jet) > 1.0
Resolved Selection	
At least four jets	≥ 4 jets, $p_T > 30$ GeV

8.7 Overlap removal

Overlap removal (OR) is a procedure of removing overlaps between the different physics objects. It protects us from double counting the same detector responses, being reconstructed by different algorithms. In this search, the OR between the jets and the leptons is important. In the resolved topology, the lepton responses are usually outside the jet radius ($R = 0.4$ in this case). But in the boosted topology, the leptons can be very close to a jet. It is not a big issue in the muon channel since the muons are reconstructed from the MS and ID tracks, whereas the jets are reconstructed using calorimeter clusters and ID tracks. So, even if a muon and a jet are close and $\Delta R < 0.4$, it is possible to reconstruct them both reliably. The situation gets complicated in the electron channel because both electrons and jets are built using calorimeter energy clusters and ID tracks. So, there is a high chance of double-counting. To avoid the double-counting of the calorimeter energy clusters, it is standard to require $\Delta R(\text{el, jet}) > 0.4$ in ATLAS to clearly separate them at the cost of efficiency. But in our signal events, we expect the electron and the jet coming from top-quark decay to be very close to each other. One such situation is demonstrated in Figure 8.4, where ΔR at the generator level partons between the electron and the b-quark is shown for SM $t\bar{t}$ and a Z' signal process with $m_{Z'} = 4$ TeV. This distribution shows that the distance between the electron and the b-quark is smaller than 0.4 in most signal events. In the case of SM $t\bar{t}$, the separation is almost always larger than 0.4. Figure 8.5 shows this angular separation, $\Delta R(el_{\text{true}}, b_{\text{true}})$, as a function of the electron p_T. There is a significant number of high p_T electrons with $\Delta R(el_{\text{true}}, b_{\text{true}}) < 0.4$. So,

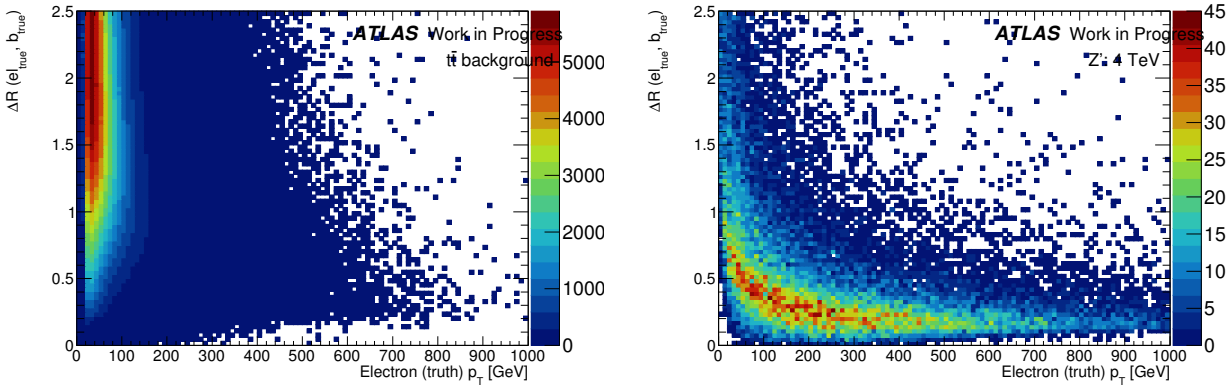


(a) SM $t\bar{t}$ process

(b) BSM Z', $m_{Z'} = 4$ TeV

Figure 8.4: Angular distance, ΔR, between the electron and b-quark at the parton level in the boosted electron channel for (a) SM $t\bar{t}$ and (b) Z' signal of $m_{Z'} = 4$ TeV.

the standard method of removing jet and electron overlap is far from optimal. To overcome this issue, and increase signal acceptance, a different electron-jet OR method, *electron-in-jet subtraction method* is developed. This method will be discussed in the next section.



(a) SM $t\bar{t}$ process

(b) BSM Z', $m_{Z'} = 4$ TeV

Figure 8.5: $\Delta R(el_{\text{true}}, b_{\text{true}})$ as a function of truth electron p_T for (a) SM $t\bar{t}$ and (b) 4 TeV Z'.

In the muon channel, the OR is done in two steps. First, isolated muons are selected, then all jets with $\Delta R < 0.2$ or with fewer than three tracks are removed. In the next step, a sliding ΔR cut is used where the threshold is p_T-dependent. As the p_T of the muon increases, the threshold value becomes very small, allowing a jet close to a high-p_T muon. An overlap removal between an electron and a muon is also applied. In general, an electron is removed if they have a shared ID track.

Table 8.4: Muon-jet and electron-muon overlap removal definition used in the semi-leptonic $t\bar{t}$ resonance analysis.

Reject	Against	Criteria
electron	muon	shared ID track
jet	muon	NumTrack < 3 and $\Delta R < 0.2$
muon	jet	$\Delta R < \min(0.4, 0.04 + 10\ \text{GeV}/p_T^{\mu})$

8.7.1 Electron-in-jet subtraction method

This method is a two-step procedure like in the muon channel. First, an isolated well-reconstructed electron is selected. If there are small-R jets close to the electron with $\Delta R(el, \text{jet}) < 0.4$, then the electron 4-momentum is subtracted from the jet. The resulting jets are defined as electron-subtracted jets. A p_T dependent threshold is defined to check if an electron-subtracted jet has high enough p_T to be still considered as a jet. If the p_T of an electron-subtracted jet is less than the threshold value, it is assumed that the jet is built with the electron clusters, and the jet is not a real jet. So the jet is removed from the list of jets in that event. If an electron-subtracted jet has p_T more than the threshold value, the ΔR between the electron-subtracted jet and the electron is recalculated, and the following steps are done:

- If $\Delta R(el, \text{electron-subtracted jet}) > 0.2$
  - Keep both electron and the electron-subtracted jet. Subsequently use the electron-subtracted jet as a proxy of the actual jet.
- If $\Delta R(el, \text{electron-subtracted jet}) < 0.2$
  - It is assumed that the electron is coming from a B-hadron decay. Then electron is removed from the event and it's p_T is added back to the el-subtracted jet p_T.

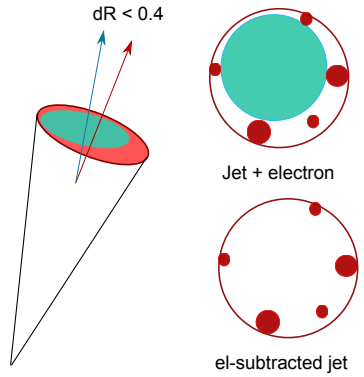
This situation is shown schematically in Figure 8.6, where the electron energy deposits are shown in cyan and the jet energy clusters are shown in dark red. An electron-subtracted jet that does not pass the p_T threshold is shown in Figure 8.6a and the other case where it passes the p_T threshold is shown in Figure 8.6b.

8.8 Data-driven multijet background

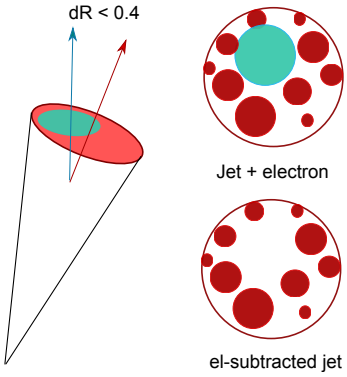
Multijet processes do not contain any prompt lepton in the final state, but the jets can sometimes get reconstructed as a lepton. The matrix method uses the prompt lepton identification efficiency (*real efficiency*), r, and the non-prompt lepton misidentification probability (*fake efficiency*), f. These quantities are defined using *tight* and *loose* lepton definitions. Tight leptons are the leptons passing the nominal analysis selection of identification and isolation criteria. In the loose definition, the identification criteria are loosened, and the isolation requirement is not applied. The tight and loose lepton definitions are summarized in Table 8.5.

The number of events with leptons satisfying the loose definition, N_L is defined as

$$N_L = N_{\text{prompt}} + N_{\text{non-prompt}}\ , \tag{8.1}$$



(a) Subtracted jet fails p_T threshold.



(b) Subtracted jet passes p_T threshold.

Figure 8.6: Schematic diagram demonstrating the electron-in-jets subtraction method. The diagram shows an electron-subtracted jets in the case (a) where it fails the p_T threshold and in case (b) where it passes the p_T threshold. The electron contribution is shown in the cyan color and the energy clusters contributing to the jet are shown in dark red color.

Table 8.5: Definitions of the leptons used for the matrix method QCD estimation.

Lepton definition	**Electron**	**Muon**
Loose	MediumLH and No isolation	Loose Quality and No isolation
Tight	TightLH and TightTrackOnly	Medium Quality and TightTrackOnly_VarRad

where N_{prompt} ($N_{\text{non-prompt}}$) is the number of events with prompt (non-prompt) leptons. These quantities can be measured directly from data. But we can count the number of events with with tight leptons, N_T, from the experimental data. Then N_T can be written as

$$N_T = r \times N_{\text{prompt}} + f \times N_{\text{non-prompt}} \tag{8.2}$$

where the real efficiency, r, and the fake efficiency, f are defined as

$$r = \frac{N^{\text{tight}}_{\text{prompt}}}{N^{\text{loose}}_{\text{prompt}}}, \qquad f = \frac{N^{\text{tight}}_{\text{non-prompt}}}{N^{\text{loose}}_{\text{non-prompt}}}. \tag{8.3}$$

Let's define the number of anti-tight leptons which pass the loose selection but does not pass the tight selection as $N_A = N_L - N_T$. Now, with these definitions, we can write the following matrix equation

$$\begin{pmatrix} N_T \\ N_A \end{pmatrix} = \begin{pmatrix} r & f \\ 1-r & 1-f \end{pmatrix} \begin{pmatrix} N_{\text{prompt}} \\ N_{\text{non-prompt}} \end{pmatrix} \tag{8.4}$$

We can invert the matrix equation to write the number of prompt and non-prompt leptons in terms of reconstructible quantities, N_T and N_A as

$$\begin{pmatrix} N_{\text{prompt}} \\ N_{\text{non-prompt}} \end{pmatrix} = \frac{1}{(r-f)} \begin{pmatrix} 1-f & -f \\ (r-1) & r \end{pmatrix} \begin{pmatrix} N_T \\ N_A \end{pmatrix}. \tag{8.5}$$

Hence, the number of fake leptons in the signal region coming from the multijet process can be written as

$$N_T^{\text{Multijet}} = f \times N_{\text{non-prompt}} = \frac{(r-1)f}{r-f} N_T + \frac{rf}{r-f} N_A \,, \tag{8.6}$$

where N_T is the number of leptons passing tight selection and N_A is the number of leptons passing loose selection but failing tight selection. The weights are applied as event weight.

Real and fake rate

The real rate is determined using the simulated $t\bar{t}$ events, where we expect to have one prompt lepton. To estimate the fake rate, a background control region, enriched with QCD-like events, is defined. The control region is defined by inverting the MET and MET+$m_{T,W}$ cut. The fake rate, f, is measured from data in the QCD-enriched control region. The fakes rates for electrons vary from 10% to 85%, with the largest value occurring in the region with highest lepton p_T. This behavior is observed because the track-based lepton isolation becomes very loose for high p_T electrons. The fake rates in the muon channel vary from 4% to 74% depending on muon p_T.

Finally these rates are used to calculate the number of multijet events in the signal regions. The multijet background is roughly 10% of the total background in the boosted region and 4% in the resolved region.

8.9 Event reconstruction

Following the event selection, an observable $m_{t\bar{t}}$ is constructed using the objects in the event to approximate the invariant mass of the $t\bar{t}$ system. The reconstruction of the $t\bar{t}$ system is more involved than in the all-hadronic analysis. The selected events in the boosted and the resolved region have different objects. So, the $m_{t\bar{t}}$ is constructed differently in these two regions. In this section, the reconstruction methods are summarized. Neutrinos are reconstructed using the same technique in the boosted and resolved regions. Following the neutrinos, the reconstruction of top-quarks is described in this section.

8.9.1 Neutrino reconstruction

Neutrinos do not interact with the ATLAS detector. So, we assume the transverse momentum of a neutrino, p_T^{ν}, is same as the missing transverse energy, E_T^{miss}. We have access to the azimuthal angle, ϕ, of the MET. It gives us the ability to reconstruct the x and y component of the neutrino momentum, leaving the z-component of the neutrino vector undetermined. To determine the neutrino vector z-component, we use momentum conservation

$$p_W = p_l + p_\nu \,, \tag{8.7}$$

where p_l is the lepton 4-vector and p_ν is the neutrino 4-vector. We also assume the mass of a neutrino to be zero, $m_\nu = 0$. Which implies $p_\nu^2 = 0$. We further assume that the lepton and the neutrino come from a on-shell W boson decay and impose the W mass constraint on the lepton-neutrino system. We get the following equation from the momentum conservation and on-shell W decay approximation

$$m_W^2 = (p_l + p_\nu)^2 = m_l^2 + 2p_l \cdot p_\nu \,. \tag{8.8}$$

Solving this equation we get the neutrino momentum z-component, $p_{z,\nu}$ in terms of the other known quantities. We find

$$p_{z,\nu} = \frac{4c_1 p_{z,l} \pm \sqrt{(4c_1 p_{z,l})^2 - 4 \cdot 4\,(E_l^2 - p_{z,l}^2) \cdot 4\,(p_{T,\nu}^2 E_l^2 - c_1^2)}}{2 \cdot 4\,(E_l^2 - p_{z,l}^2)} \tag{8.9}$$

where

$$c_1 = m_W^2 - m_l^2 + 2(p_{x,l}p_{x,\nu} + p_{y,l}p_{y,\nu})\,.$$

If no real solution is found, it is assumed that complex solutions appear due to mis-measurement of E_T^{miss}. Hence, the E_T^{miss} vector is rescaled and rotated minimally until the real solution is found. If there are two real solutions for $p_{z,\nu}$, the solution with smaller absolute value is used.

8.9.2 Boosted region

In the boosted region, it is relatively simple to reconstruct the $t\bar{t}$ system. The top-tagged large-R jet is used as the hadronic top. The leptonic top is reconstructed using the lepton, the highest-p_T close-by R=0.4 jet and the neutrino calculated from MET.

8.9.3 Resolved region

In the resolved region, the reconstruction becomes more complicated. Selected events usually have more than four jets, as a result, it becomes difficult to decide which jet should get associated with the leptonic top and which three jets should be used to reconstruct the hadronic top. A Chi-squared algorithm was developed to find the best combination of jets forming the leptonic and hadronic top candidates. The χ^2 is defined using the 4-momenta of the lepton, neutrino, and all the R=0.4 jets of an event.

$$\chi^2 = \left[\frac{m_{jj} - m_{W_h}}{\sigma_{W_h}}\right]^2 + \left[\frac{m_{jjb} - m_{jj} - m_{t_h - W_h}}{\sigma_{t_h - W_h}}\right]^2 + \left[\frac{m_{b\ell\nu} - m_{t_\ell}}{\sigma_{t_\ell}}\right]^2 + \left[\frac{(p_{T,jjb} - p_{T,b\ell\nu}) - (p_{T,t_h} - p_{T,t_\ell})}{\sigma_{p_{T,t_h} - p_{T,t_\ell}}}\right]^2 \tag{8.10}$$

The first term is a constraint on the dijet mass to from the hadronic W boson. The second term corresponds to the hadronic top-quark. The quantities m_{jj} and m_{jjb} are highly correlated, so the contribution of the hadronic W boson is subtracted to decouple the second term from the first term. The third term takes care of the leptonic top-quark. The combined mass of the lepton, neutrino and a jet is constrained to match to the leptonically decaying top-quark mass. The last term is used to establish the transverse momentum balance of the two top-quarks. It constrains the p_T of the hadronic and leptonic top-quark to be similar as expected in a resonance. The expected values of the parameters m_{W_h}, $m_{t_h - W_h}$, m_{t_ℓ}, $p_{T,jjb} - p_{T,b\ell\nu}$ as well as their uncertainties σ_{W_h}, $\sigma_{t_h - W_h}$, σ_{t_ℓ}, and $\sigma_{p_{T,t_h} - p_{T,t_\ell}}$ are obtained from the simulated Z' events by matching the reconstructed objects to the truth partons. Poorly reconstructed events are discarded by requiring with $\log_{10}\chi^2 < 0.9$.

This χ^2-method of top reconstruction was used in the previous $t\bar{t}$ resonance searches done in ATLAS. Although the performance of this algorithm is very good, new methods based on neural networks are being explored. Two main options are currently studied.

- **Classification method:** This method tries to classify the reconstructed jets coming from the W decay, and the two b-jets. Once the jets are uniquely identified, leptonic top and hadronic top can be reconstructed by combining these jets, lepton and neutrino vectors. For this study, a classifier is being trained on the simulated samples where the jets are labeled based on the truth parton-level information.

- **Regression method:** In this method the goal is to learn the kinematic variables of the truth level top-quark, W-boson and b-quark. So, a deep neural network is trained to learn the kinematics as well as the deconvolution of the detector effects. In this method, the network is trained with simulated $t\bar{t}$ process and truth parton-level quantities are used as target variables.

Neither of the these methods have been sufficiently studied to be included in this dissertation. So, the chi-squared method is the nominal method.

8.9.4 $t\bar{t}$ invariant mass

Once the leptonic top, t_{lep}, and the hadronic top, t_{had}, are reconstructed, the invariant mass of the $t\bar{t}$ system is built by adding the two 4-vectors as

$$m^2_{t\bar{t}} = (p_{t_{\text{lep}}} + p_{t_{\text{had}}})^2. \tag{8.11}$$

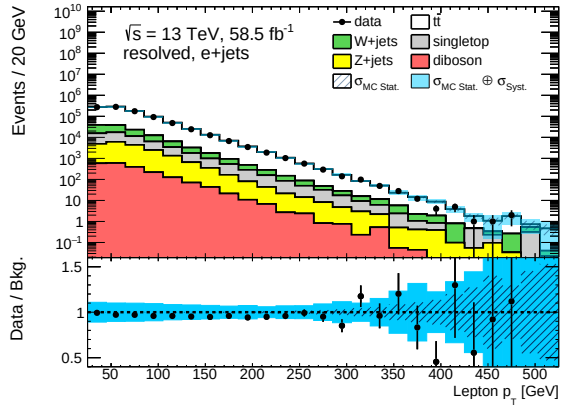
$p_{t_{\text{lep}}}$ and $p_{t_{\text{had}}}$ are the 4-vectors of the leptonic and hadronic top respectively.

8.10 Kinematic distributions

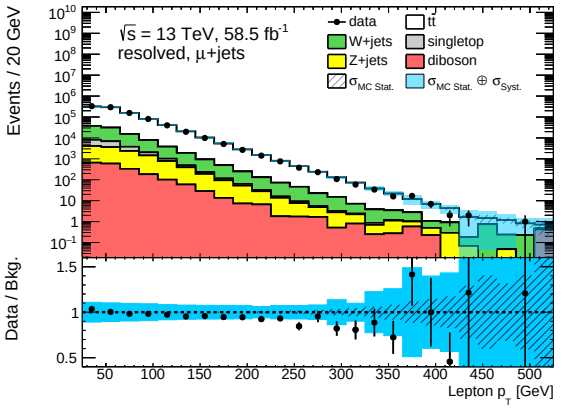
This analysis is still on-going, so we are still blinded in the signal regions, and we are not allowed to look at the $m_{t\bar{t}}$ distributions. Instead, we are allowed to look at the data distributions of other kinematic variables related to the leptons and the jets. Some of these control plots are shown in this section, using only the 2018 data. The NNLO correction on the $t\bar{t}$ background is not applied here, since it is not yet finalized. Also, not all systematic uncertainties are included in the systematics band though the dominant jet-related uncertainties are shown. A subset of uncertainties related to the MET, electron, and muon are not included. The data-driven estimates of QCD contributions in the signal regions are not fully finalized yet, so that background contribution is also not added in the control plots.

8.10.1 Resolved region

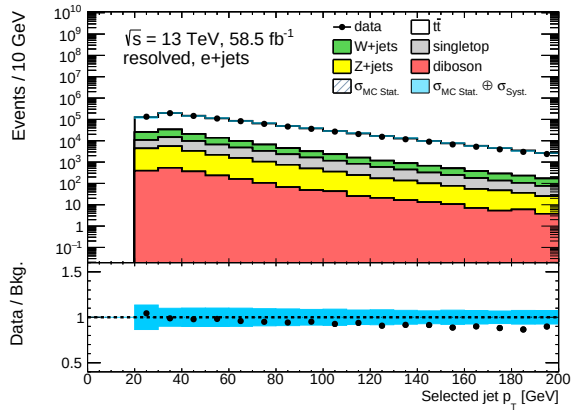
Although $m_{t\bar{t}}$ is the main discriminant, the other kinematic variables of the lepton, jet, and MET help us understand the general trend between data and simulation. Figure 8.7 shows the p_T distributions of (top) the lepton and (middle) the jet (closest to lepton), and the (bottom) E_T^{miss} distribution in the electron and muon final states. The lepton p_T distributions demonstrate a good agreement between data and background. While the overall agreement in the p_T distribution of the jet closest to the lepton is good, there is a minor slope in the ratio probably coming from the $t\bar{t}$ background modelling. There is a small slope in the ratio of E_T^{miss} distribution probably also coming from the $t\bar{t}$ background modelling. The modelling issue is expected to become better after applying the NNLO QCD and NLO electroweak corrections, hence, the data-background agreement is expected to improve after applying the $t\bar{t}$ corrections. Nonetheless, overall the data agrees quite well with the background in the resolved region. Although the QCD background is missing and some of the corrections are not applied yet, it is expected that the ratio between data and background is still within the uncertainty band.



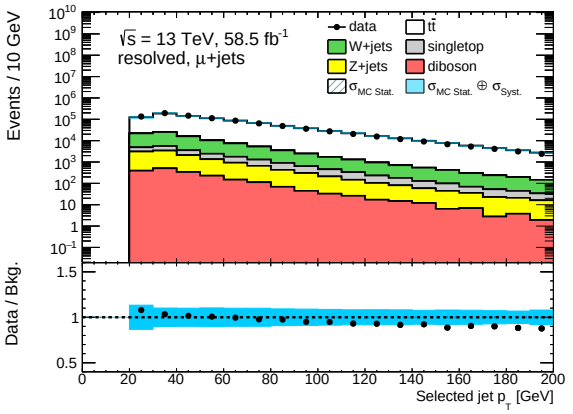
(a) Resolved Electron channel



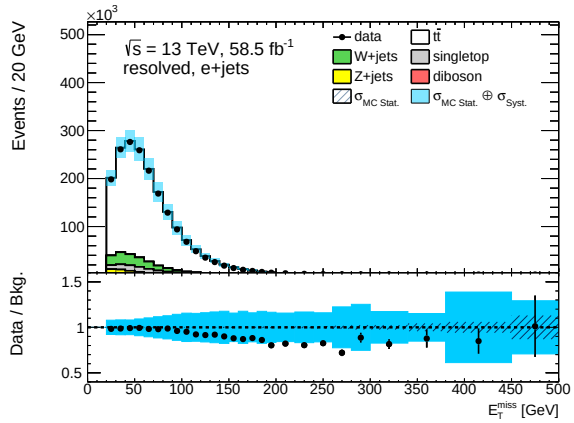
(b) Resolved Muon channel



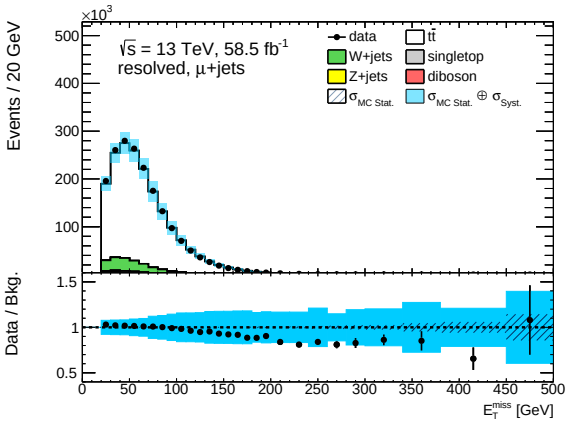
(c) Resolved Electron channel



(d) Resolved Muon channel



(e) Resolved Electron channel

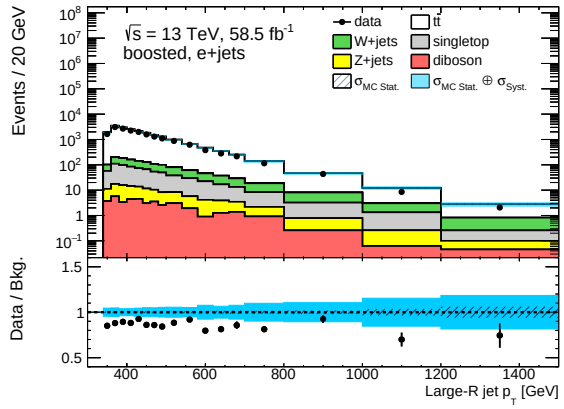


(f) Resolved Muon channel

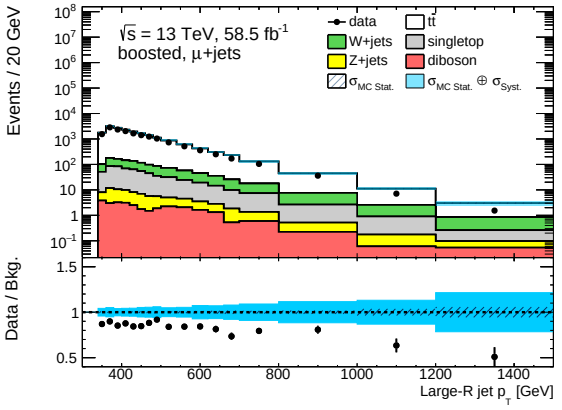
Figure 8.7: The p_T distributions of the (top) lepton, (middle) the jet (closest to lepton), and (bottom) E_T^{miss} distributions in the (a, c, e) boosted electron and (b, d, f) boosted muon channel. The ratio of data and background is mostly within the systematic band.

8.10.2 Boosted region

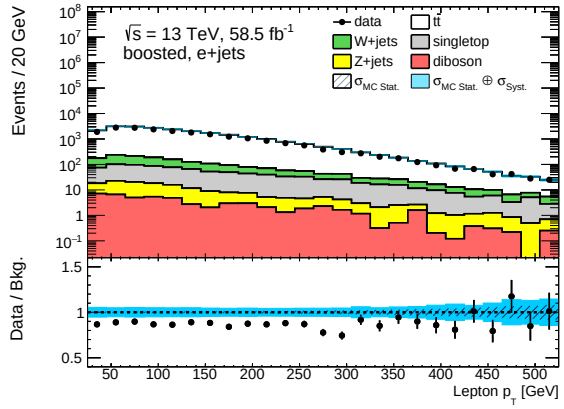
In this section, large-R jet p_T, lepton p_T, and jet (closest to lepton) p_T are shown for the electron and muon channels. Along with the uncertainties missing in the resolved channel, the top-tagging uncertainties are not included in the boosted region. The overall agreement between data and background is not very good. Although there are no clear slopes or trends, there is an overall deficit of $\sim 10-20\%$ in data. A similar deficit was observed while analyzing $2015-2016$ data [250]. This is also seen in the differential $t\bar{t}$ cross-section measurements and is thought to come from the poor modelling of the top p_T spectrum in our current generators, which becomes large in this channel as we are searching high momentum top quarks. Figure 8.8 shows the large-R jet p_T in the top, lepton p_T in the middle and small-R jet (close to the lepton) p_T distributions in both the channels. There is a small slope in the large-R jet p_T distribution in the higher p_T region. The other distributions look consistent.



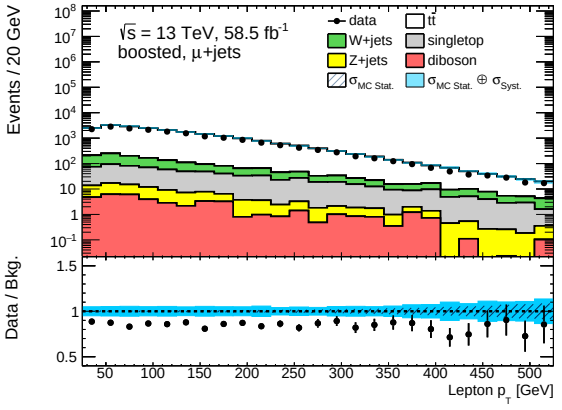
(a) Boosted Electron channel



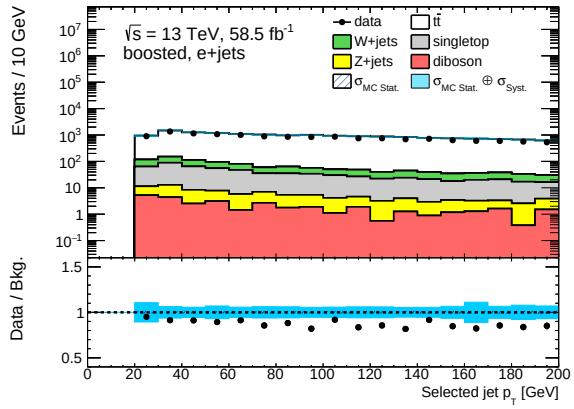
(b) Boosted Muon channel



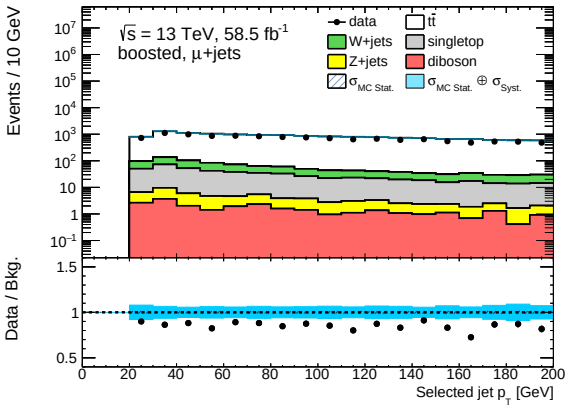
(c) Boosted Electron channel



(d) Boosted Muon channel



(e) Boosted Electron channel



(f) Boosted Muon channel

Figure 8.8: The p_T distribution of the (top) large-R jet p_T, (middle) lepton p_T, and (bottom) close-to-lepton small-R jet in the boosted (a), (c), (e) electron and (b), (d), (f) muon channel.

8.11 Statistical analysis

The statistical analysis is very similar to what used in the all-hadronic analysis in three steps as described in section 7.12. The model-independent search is not set up yet, so those studies are not included in this dissertation. The model-dependent search will be done after unblinding the data in the signal region, and those results are also not included in this section. Only the initial expected cross-section limits are shown in this section. The statistical treatment for calculating limits corresponds to the description in section 7.12. The same statistical tools and methods are used for this analysis. The likelihood function used in this search is the same as the one described in Equation 5.19.

Results using only the Z'_{TC2} model are included in this dissertation. In the future, the other models, described in section 8.4 will be included. For these studies, four signal regions are used: boosted electron (be), boosted muon (bmu), resolved electron (re), and resolved muon (rmu). All the systematics uncertainties are assumed to be fully-correlated across the four signal regions. It is seen in the all-hadronic $t\bar{t}$ resonance search that having multiple signal regions categorized by the number of b-tagged jets improves the limits slightly. In the future, these four signal regions will be subdivided into categories based on the number of b-tagged jets to improve the sensitivity. In the absence of data, all these studies are done with Asimov data. Two sets of results are shown here:

- Background-only (b-only) fit: The Asimov data is created from the nominal background spectrum,
- Signal + Background (s+b) fit: The Asimov data is created either using the signal + background spectrum or background spectrum depending on hypothesized signal strength (μ) value.

Most of the experimental uncertainties are considered in this study, but the theoretical uncertainties like $t\bar{t}$ modelling uncertainties are not included yet. They will be added to the statistical analysis in the future. The effect of each systematic variation gets decomposed into shape effect and normalization effect. The normalization components are always kept in this analysis, but the shape components are dropped if the overall effect is less than 0.5%.

8.11.1 b-only profile likelihood Asimov fit

The Asimov data is created using the simulated background spectrum. In the background-only fit, the signal strength is not included in the likelihood. Two free-floating normalization factors, $k(t\bar{t}^r)$, $k(t\bar{t}^b)$ are used for the overall $t\bar{t}$ background normalization in the resolved and boosted channel, respectively. The normalization factors are included in the likelihood to take care of the normalization difference between data and simulation. The best-fit values of the normalization factors are 1.0 as expected (Figure 8.9), but there are expected to take different values in the presence of data. We can see that we expect to be able to constrain these $t\bar{t}$ normalization factors to 2-3%. The best-fit values of the other nuisance parameters (NP) are shown in Figure C.1. Only some of the large jet related systematics are constrained in the fit.

8.11.2 s+b profile likelihood Asimov fit

In this case, also, the Asimov data is created using the background spectrum estimated from the MC simulation. Along with the normalization factors ($k(t\bar{t}^r)$,$k(t\bar{t}^b)$) the signal strength (μ) is used

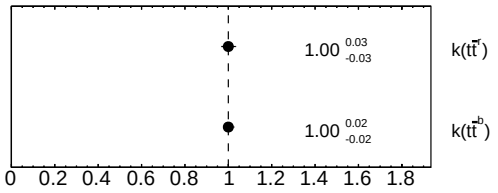


Figure 8.9: The best-fit value of the normalization factors in the b-only fit.

as a free-floating parameter. The best-fit values of μ and the two normalization factors are shown in Figure 8.10, where the 4 TeV Z'_{TC2} model is used as a signal. Since the Asimov data is created from the background spectrum, the best-fit μ value $\hat{\mu}$ is 0 as expected. The best-fit values of the two normalization factors are 1 with similar uncertainties as seen in the b-only fit. Figure 8.11 shows the nuisance parameter ranking plot, where the NPs are ranked according to their expected impact on μ. Only the top 20 NPs are shown in this plot. It is seen that $k(t\bar{t}^b)$ has almost 10% impact on the μ value. It is also worth noting that some of the γ-factors that represents the bin-by-bin statistical uncertainties are highly ranked. In the future, it has to be studied if any sort of smoothing of the nominal spectrum improves these effects. All the NPs constraints look consistent at this point and none of the NPs are constrained more than 40%. The post-fit pulls of all the nuisance parameters are shown in Appendix C.

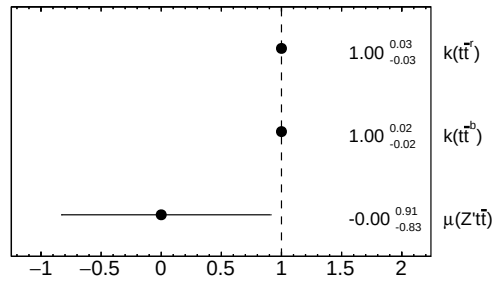


Figure 8.10: The best-fit value of the signal strength and the normalization factors in the $s+b$ fit. The uncertainty bar shows the combined effect of the statistical and systematic uncertainties.

8.11.3 Limit setting

95% CL limit on the cross-section times branching fraction of resonant new physics processes is set with the Frequentist CL_s method (subsection 5.3.6) using asymptotic formula (described in subsection 5.3.5) in the same fashion as in the all-hadronic analysis (section 7.13). The 95% upper limit on the signal strength is converted to the upper limit for cross-section times branching fraction by multiplying by expected signal cross-section times branching fraction. The expected limits, and the $\pm1\sigma$ and $\pm2\sigma$ uncertainty bands are calculated making use of the Asimov dataset using the asymptotic formula.

Expected 95% CL upper limits are calculated for two cases, with all the systematic and statistical uncertainties and with the statistical uncertainty only. Z'_{TC2} model with $\Gamma/m = 1.2\%$ is used

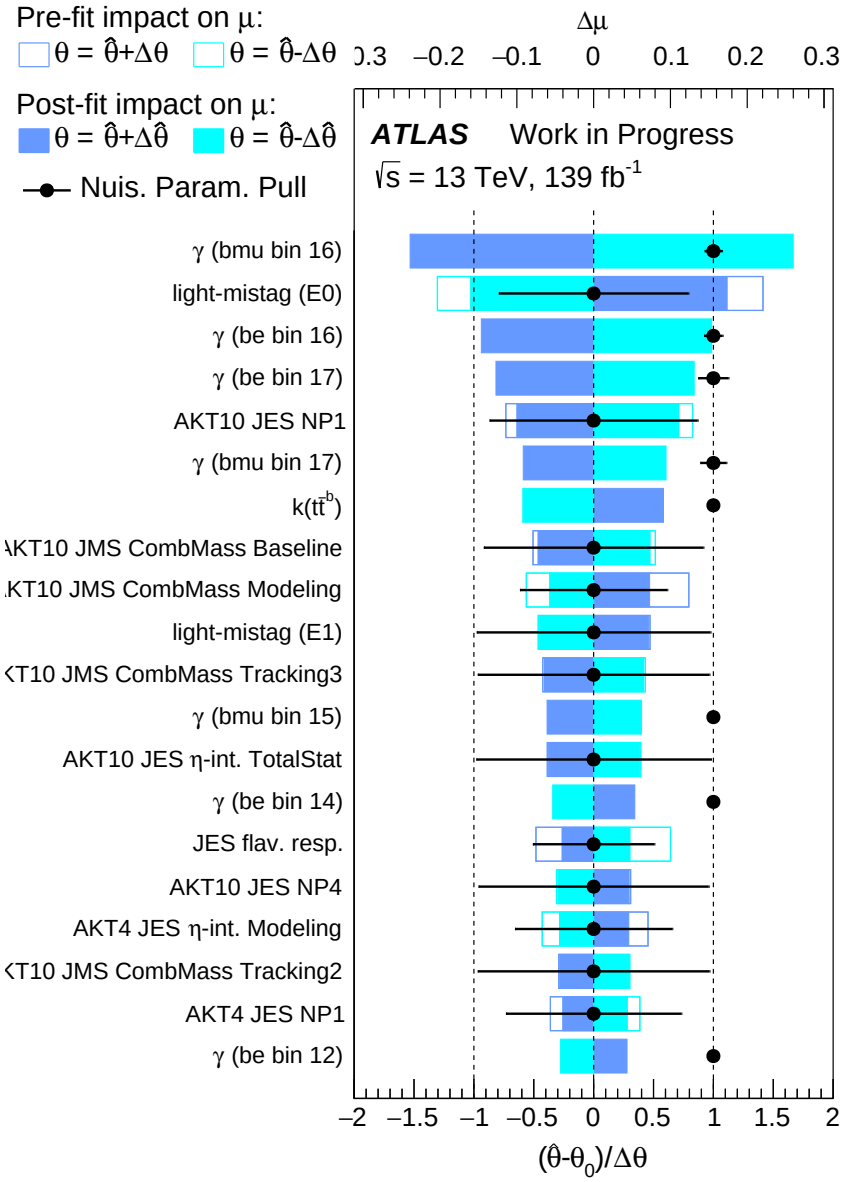


Figure 8.11: Systematics ranking plot where the 4 TeV Z'_{TC2} signal is used. Some of the gamma-factors (bin-by-bin statistical uncertainties) along with some of the larger components of jet and b-tagging systematics have large impact on μ.

as the signal model. Limits are calculated starting from Z'_{TC2} 500 GeV to 5 TeV. The limits are calculated for the following points using MC signal simulations

- 500, 750, 1000, 1250, 1500, 1750, 2000, 2250, 2500, 2750, 3000, 4000, 5000 GeV.

The expected limits on Z'_{TC2} signal cross-section times branching fraction are shown in Figure 8.12. The 95% CL expected limit on Z'_{TC2} ($\Gamma/m = 1.2\%$) mass is around 3.6 TeV whereas the stat-only limit is around 3.8 TeV.

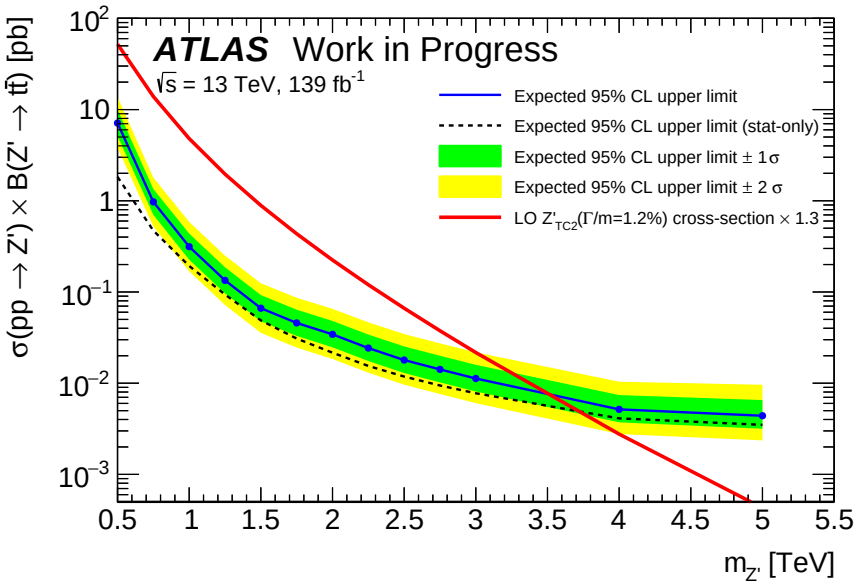


Figure 8.12: The expected cross-section 95% CL upper limits on Z'_{TC2} signals.

8.12 Conclusion

This chapter described a $t\bar{t}$ resonance search in the single lepton final state, where one top-quark decays leptonically, and the other top-quark decays hadronically. In this search, we are looking for a narrow resonance using the full Run-2 data (139 fb^1) collected by the ATLAS detector during the year 2015-2018. Several benchmark signal models such as narrow width Z'_{TC2}, Kaluza-Klein graviton, Kaluza-Klein gluons will be used to interpret the results. The s-channel dark-matter mediator models will also be used, where the new resonance can be interpreted as a dark matter mediator. The analysis involves several physics objects due to the complicated $t\bar{t}$ final state. Both the boosted and the resolved top-decay topologies are studied in this search. The main discriminant is the invariant mass of the $t\bar{t}$ system, where the signal is expected to show up as a bump. The primary strategy is to perform a model-independent statistical search on the $t\bar{t}$ mass spectrum. In the absence of statistically significant deviations, an upper limit (at 95% CL) on the cross-section times branching ratio will be placed on the tested signal model. A similar search was carried out with partial Run-2 data (36.1 fb^1), and no evidence of the new signal was found [250]. A narrow width Z'_{TC2} (1.2% width) boson of mass lower than 3.2 TeV was excluded at 95% CL. Following the previous analysis, several improvements are done in this full Run-2 search. As a result, the expected cross-section limits improved by almost a factor of 1.5. The expected mass limit (1.2% width Z'_{TC2}) including only the statistical uncertainties improved from 3 TeV to 3.8 TeV based on the current optimization. However, the expected sensitivity is less than that of the all-hadronic search in which the expected stat only mass limit was 4.2 TeV. But in the future, a better sensitivity can be achieved by combining these two channels along with the dilepton $t\bar{t}$ final state. This chapter presented the expected sensitivity only on the Z'_{TC2} model; all the other models will be used in the future. The results will also be interpreted with the dark matter mediator models after the completion of this analysis.

CHAPTER

NINE

Conclusions and outlook

The standard model (SM) of particle physics describes the ordinary matter particles and their interactions. Besides its successes, the SM cannot explain some of the experimental observations like particle origin of dark matter, or the origin of neutrino mass. To explain these observations and address the other problems of the SM, several theories beyond the SM (BSM) evolved over the last few decades. Most of these theories predict new particles, which can be related to new forces of nature. The discovery of such new particles would reshape our understanding of the universe.

This dissertation presented experimental searches for new heavy particles that decay into two SM particles. These searches used LHC proton-proton collision data at $\sqrt{s} = 13$ TeV collected by the ATLAS experiment. The studies presented here particularly look for the resonant production of these new BSM bosons. These bosons can come from additional $U(1)$ symmetries of the BSM models. An additional $U(1)$ symmetry can either come from breaking the symmetry of a BSM model with a larger summery group, or we can ignore the exact symmetry breaking mechanism and consider a minimal extension of the SM with this additional $U(1)$ symmetry. In the dilepton resonance search, we considered both of these cases to interpret the results. Whereas in the $t\bar{t}$ resonance searches, we considered benchmark models with alternative electroweak symmetry breaking that gives additional new gauge bosons. These search results can also be interpreted in the context of s-channel dark matter mediator models with minimal extension to the SM. Two lepton final states give us the cleanest signature in the detector, so it was one of the obvious places to look for new physics at the beginning of the Run-2; with the increase of collision center-of-mass energy in Run-2, these searches became exciting. The dilepton resonances search was particularly exciting in 2016 because the data set increased almost ten times from the previous results published with the data collected in 2015. A total of 36.1 fb^{-1} of data, collected during 2015 and 2016, was used in the dilepton resonance search presented in this dissertation. There was no sign of the resonant signal in the two-electron or two-muon final states. In the absence of a new signal, we set an upper limit on the signal cross-section times branching ratio at 95% CL. The cross-section limits were also interpreted as lower limits on the mass of the BSM particle. Z'_{SSM} (3% width) bosons are excluded with masses below 4.3 TeV in the di-electron channel. There was almost 1 TeV improvement in the mass limit from the previous results published using data in 2015 only. Following 2016, the analysis sensitivity with more data collected during 2017 and 2018 was not expected to improve significantly. So I wanted to move on and explore BSM scenarios where the

new particles couple preferentially with the third generation quarks; I chose the $t\bar{t}$ final state to search for high mass top-philic resonances.

The dissertation described two searches for resonant $t\bar{t}$ production in the boosted all-hadronic and semi-leptonic final states. These searches use 139 fb^{-1} data collected during 2015 to 2018. The sensitivities of these two channels were found to be similar in the past. Hence I studied both channels. The SM background contribution in the boosted all-hadronic $t\bar{t}$ resonance search was estimated by fitting an analytic function to the $t\bar{t}$ invariant mass spectrum in the data. We performed a model-independent search for resonant signals and did not observe any evidence of a new signal. Hence an upper limit on the signal cross-section times branching ratio was placed at 95% CL. New top-tagging, b-tagging, and background modeling strategies significantly improved the sensitivity in this analysis. As a result, the signal cross-section limits improved by a factor of four. The lower limit on Z'_{TC2} mass improved by 1 TeV, and Z'_{TC2} bosons were excluded up to masses 4 TeV. A similar search in the semi-leptonic final state is currently on-going. Unlike the all-hadronic search, both boosted and resolved final states are studied in the semi-leptonic $t\bar{t}$ resonance search. The event reconstruction and background estimation strategies are very different from the boosted all-hadronic $t\bar{t}$ resonance search. All the major SM backgrounds are modeled using Monte Carlo simulation, whereas the QCD multijet processes are estimated using a data-driven method. Since the analysis is still blinded, the dissertation presented the expected sensitivities with the full Run-2 data. Based on the current optimization, the expected limit on Z'_{TC2} mass is 3.6 TeV. Although the signal cross-section limits are slightly weaker than the boosted all-hadronic final state, a combination of these two channels will improve the sensitivity. Soon we want to combine results from all the $t\bar{t}$ final states to improve the overall sensitivity. The $t\bar{t}$ resonance search results will also be combined with the other resonance searches in ATLAS using the benchmark Heavy Vector Triplet model. With our $t\bar{t}$ resonance search results, we will be able to probe the coupling of the BSM boson to the third generation quarks. The results from the boosted all-hadronic final state are also interpreted in the context of the dark matter mediator models; the same will be done for the semi-leptonic final state.

Following Run-2, the next run of the LHC will begin in early 2022. The collected data is expected to be roughly double the amount of current data by the end of Run-3 in 2024. This increase in luminosity alone is not going to improve the sensitivities of these searches significantly. So, we need to apply novel techniques to extract even more of the information from data. In the searches presented in this dissertation, we used one discriminant to identify the signals. A multivariate approach with likelihood-free inference could help us regain some of the information lost while building a summary test statistics from the discriminant. It could be a potential direction for the Run-3 analyses. Furthermore, the $t\bar{t}$ resonance searches could benefit from the improvement in jet reconstruction in ATLAS. The identification of the top-jets is also expected to improve if we use optimized jet definitions in the top-tagging algorithms. The searches described in this dissertation focus mostly on the narrow width resonances. In the absence of any hint of such signals, searches for broad excesses are also becoming important. Such a search already exists in the dilepton final state, but there is no such search in $t\bar{t}$ final states. Another interesting model-independent way of looking for new physics is to parameterize cross-section measurement results with Effective Field Theory (EFT) operators. There are several technical difficulties of using EFT approaches that are currently being worked on by both experimentalists and theorists, but it could be a potential new direction for finding hints for new physics.

Besides doing physics searches, it is also necessary to study the reconstruction of the objects

used in physics analysis. Lately, the usage of several machine learning (ML) techniques increased a lot in various steps in object reconstruction. One such application was discussed in this dissertation. It allowed me to build an experience of using ML-based algorithms. It was valuable since this is something I also want to pursue further in future research. We studied the tracking associated with the high p_T jets. These jets are used in several physics searches including the $t\bar{t}$ resonance searches described in this dissertation. In the core of high p_T jets, the charged particle tracks are very close to each other. They are often so close that the average separation is comparable to the granularity of individual sensors and their charge clusters in the pixel detector get merged. Currently, ATLAS uses an ML algorithm based on neural networks to split these merge charge clusters. The dissertation discussed a new algorithm based on Mixture Density Networks (MDN) developed to do the same job more efficiently and more accurately. MDN used fewer steps compared to the current algorithm, hence their use will speed up the process. The performance of the MDN-based algorithm is either comparable to or better than the current algorithm. It is expected that MDN will improve the impact parameter resolution and b-tagging performance. The impact of using MDN in tracking is currently being studied in the context of Run-3. Even though MDN performance is good, there is still room for improving MDN by optimizing the mixture model and changing the training input variables. Following the end of Run-3, the High Luminosity LHC (HL-LHC) will begin in 2026, and the ATLAS inner detector will be replaced with new detectors. During HL-LHC runs, the average pileup will increase roughly by a factor of four, and track reconstruction will be even more challenging. MDN will serve as a good starting point for future algorithms necessary in the tracking during HL-LHC.

Bibliography

[1] D. Griffiths. *Introduction to elementary particles*. 2008. ISBN: 978-3-527-40601-2 (Cited on page 1).

[2] L. Kelvin. "Contact Electricity and Electrolysis According to Father Boscovich". *Nature* **56**:1439 (1897), 84–85. ISSN: 1476-4687. DOI: 10.1038/056084a0 (Cited on page 1).

[3] J. J. Thomson. "XL. Cathode rays". *Philosophical Magazine Letters* **44**:269 (1897), 293–316. DOI: 10.1080/09500830701306165 (Cited on page 1).

[4] I. Falconer. "Corpuscles, Electrons and Cathode Rays: J.J. Thomson and the 'Discovery of the Electron'". *The British Journal for the History of Science* **20**:3 (1987), 241–276. DOI: 10.1017/S0007087400023955 (Cited on page 1).

[5] J. G. O'Hara. "George Johnstone Stoney, F. R. S. and the concept of the electron". *Notes and Records of the Royal Society of London* **29**:2 (1975), 265–276. DOI: 10.1098/rsnr.1975.0018 (Cited on page 1).

[6] E. Rutherford. "LXXIX. The scattering of α and β particles by matter and the structure of the atom". *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science* **21**:125 (1911), 669–688. DOI: 10.1080/14786440508637080 (Cited on page 2).

[7] N. Bohr. "I. On the constitution of atoms and molecules". *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science* **26**:151 (1913), 1–25. DOI: 10.1080/14786441308634955 (Cited on page 2).

[8] J. Chadwick. "Possible Existence of a Neutron". *Nature* **129**:3252 (1932), 312. DOI: 10.1038/129312a0 (Cited on page 2).

[9] A. Einstein. "Über einen die Erzeugung und Verwandlung des Lichtes betreffenden heuristischen Gesichtspunkt". *Annalen der Physik* **322**:6 (1905), 132–148. DOI: 10.1002/andp.19053220607 (Cited on page 2).

[10] A. H. Compton. "A Quantum Theory of the Scattering of X-rays by Light Elements". *Phys. Rev.* **21**:5 (1923), 483–502. DOI: 10.1103/PhysRev.21.483 (Cited on page 2).

[11] H. Yukawa. "On the Interaction of Elementary Particles. I". *Nippon Sugaku-Buturigakkwai Kizi Dai 3 Ki* **17**: (1935), 48–57. DOI: 10.11429/ppmsj1919.17.0_48 (Cited on page 2).

[12] S. H. Neddermeyer and C. D. Anderson. "Note on the Nature of Cosmic-Ray Particles". *Phys. Rev.* **51**:10 (1937), 884–886. DOI: 10.1103/PhysRev.51.884 (Cited on page 2).

[13] C. Lattes et al. "Processes Involving Charged Mesons". *Nature* **159**:4047 (1947), 694–697. ISSN: 1476-4687. DOI: 10.1038/159694a0 (Cited on page 3).

[14] W. Heisenberg. "Über den Bau der Atomkerne. I". *Zeitschrift für Physik* **77**:1 (1932), 1–11. ISSN: 0044-3328. DOI: 10.1007/BF01342433 (Cited on page 3).

[15] E. Wigner. "On the Consequences of the Symmetry of the Nuclear Hamiltonian on the Spectroscopy of Nuclei". *Phys. Rev.* **51**:2 (1937), 106–119. DOI: 10.1103/PhysRev.51.106 (Cited on page 3).

[16] P. A. M. Dirac and R. H. Fowler. "The quantum theory of the electron". *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character* **117**:778 (1928), 610–624. DOI: 10.1098/rspa.1928.0023 (Cited on page 3).

[17] C. D. Anderson. "THE APPARENT EXISTENCE OF EASILY DEFLECTABLE POSITIVES". *Science* **76**:1967 (1932), 238–239. ISSN: 0036-8075. DOI: 10.1126/science.76.1967.238 (Cited on page 3).

[18] P. M. S. Blackett, G. P. S. Occhialini, and E. Rutherford. "Some photographs of the tracks of penetrating radiation". *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character* **139**:839 (1933), 699–726. DOI: 10.1098/rspa.1933.0048 (Cited on page 3).

[19] R. P. Feynman. "Space-Time Approach to Non-Relativistic Quantum Mechanics". *Rev. Mod. Phys.* **20**:2 (1948), 367–387. DOI: 10.1103/RevModPhys.20.367 (Cited on page 3).

[20] O. Chamberlain et al. "Observation of Antiprotons". *Phys. Rev.* **100**:3 (1955), 947–950. DOI: 10.1103/PhysRev.100.947 (Cited on page 3).

[21] B. Cork et al. "Antineutrons Produced from Antiprotons in Charge-Exchange Collisions". *Phys. Rev.* **104**:4 (1956), 1193–1197. DOI: 10.1103/PhysRev.104.1193 (Cited on page 3).

[22] E. Fermi. "Tentativo di una Teoria Dei Raggi β". *Il Nuovo Cimento (1924-1942)* **11**:1 (2008), 1. ISSN: 1827-6121. DOI: 10.1007/BF02959820 (Cited on page 4).

[23] L. M. Brown. "The idea of the neutrino". *Physics Today* **31**:9 (1978), 23 (Cited on page 4).

[24] C. L. Cowan et al. "Detection of the Free Neutrino: a Confirmation". *Science* **124**:3212 (1956), 103–104. ISSN: 0036-8075. DOI: 10.1126/science.124.3212.103 (Cited on page 4).

[25] G. Danby et al. "Observation of High-Energy Neutrino Reactions and the Existence of Two Kinds of Neutrinos". *Phys. Rev. Lett.* **9**:1 (1962), 36–44. DOI: 10.1103/PhysRevLett.9.36 (Cited on page 4).

[26] M. L. Perl et al. "Evidence for Anomalous Lepton Production in $e^+ - e^-$ Annihilation". *Phys. Rev. Lett.* **35**:22 (1975), 1489–1492. DOI: 10.1103/PhysRevLett.35.1489 (Cited on page 4).

[27] DONUT Collaboration. "Observation of tau neutrino interactions". *Phys. Lett. B* **504**:3 (2001), 218–224. ISSN: 0370-2693. DOI: 10.1016/S0370-2693(01)00307-0 (Cited on page 4).

[28] B. Pontecorvo. "Mesonium and antimesonium". *Soviet Phys. JETP* **6**: (1958), 429. URL: https://www.osti.gov/biblio/4344536 (Cited on page 4).

[29] Particle Data Group. "Review of Particle Physics". *Progress of Theoretical and Experimental Physics* **2020**:8 (2020). 083C01. ISSN: 2050-3911. DOI: 10.1093/ptep/ptaa104 (Cited on pages 4, 8, 9, 16, 18).

[30] G. Rochester and C. Butler. "Evidence for the Existence of New Unstable Elementary Particles". *Nature* **160**:4077 (1947), 855–857. ISSN: 1476-4687. DOI: 10.1038/160855a0 (Cited on page 4).

[31] A. J. Seriff et al. "Cloud-Chamber Observations of the New Unstable Cosmic-Ray Particles". *Phys. Rev.* **78**:3 (1950), 290–291. DOI: 10.1103/PhysRev.78.290 (Cited on page 4).

[32] T. Nakano and K. Nishijima. "Charge Independence for V-particles*". *Progress of Theoretical Physics* **10**:5 (1953), 581–582. ISSN: 0033-068X. DOI: 10.1143/PTP.10.581 (Cited on page 4).

[33] K. Nishijima. "Charge Independence Theory of V Particles*". *Progress of Theoretical Physics* **13**:3 (1955), 285–304. ISSN: 0033-068X. DOI: 10.1143/PTP.13.285 (Cited on page 4).

[34] M. Gell-Mann. "The interpretation of the new particles as displaced charge multiplets". *Il Nuovo Cimento (1955-1965)* **4**:2 (1956), 848–866. ISSN: 1827-6121. DOI: 10.1007/BF02748000 (Cited on page 4).

[35] S. Sakata. "On a Composite Model for the New Particles". *Progress of Theoretical Physics* **16**:6 (1956), 686–688. ISSN: 0033-068X. DOI: 10.1143/PTP.16.686 (Cited on page 5).

[36] M. Ikeda, S. Ogawa, and Y. Ohnuki. "A Possibile Symmetry in Sakata's Model for Bosons-Baryons System". *Progress of Theoretical Physics* **22**:5 (1959), 715–724. ISSN: 0033-068X. DOI: 10.1143/PTP.22.715 (Cited on page 5).

[37] M. Gell-Mann. "The Eightfold Way: A Theory of strong interaction symmetry" (1961). DOI: 10.2172/4008239 (Cited on page 5).

[38] Y. Ne'eman. "Derivation of strong interactions from a gauge invariance". *Nucl. Phys.* **26**:2 (1961), 222–229. ISSN: 0029-5582. DOI: 10.1016/0029-5582(61)90134-1 (Cited on page 5).

[39] V. E. Barnes et al. "Observation of a Hyperon with Strangeness Minus Three". *Phys. Rev. Lett.* **12**:8 (1964), 204–206. DOI: 10.1103/PhysRevLett.12.204 (Cited on page 5).

[40] M. Gell-Mann. "A schematic model of baryons and mesons". *Phys. Lett.* **8**:3 (1964), 214–215. ISSN: 0031-9163. DOI: 10.1016/S0031-9163(64)92001-3 (Cited on page 5).

[41] G. Zweig. "An SU(3) model for strong interaction symmetry and its breaking. Version 1" (1964) (Cited on page 5).

[42] G. Zweig. "An SU(3) model for strong interaction symmetry and its breaking. Version 2". *Developments in the quark theory of hadrons: A Reprint Collection. 1964-1978*. Ed. by D. Lichtenberg and S. P. Rosen. 1964, 22–101 (Cited on page 5).

[43] LHCb Collaboration. "Observation of $J/\psi p$ Resonances Consistent with Pentaquark States in $\Lambda_b^0 \to J/\psi K^- p$ Decays". *Phys. Rev. Lett.* **115**:7 (2015), 072001. DOI: 10.1103/PhysRevLett.115.072001 (Cited on page 6).

[44] Belle Collaboration. "Study of $e^+e^- \to \pi^+\pi^- J/\psi$ and Observation of a Charged Charmoniumlike State at Belle". *Phys. Rev. Lett.* **110**:25 (2013), 252002. DOI: 10.1103/PhysRevLett.110.252002 (Cited on page 6).

[45] BESIII Collaboration. "Observation of a Charged Charmoniumlike Structure in $e^+e^- \to \pi^+\pi^- J/\psi$ at $\sqrt{s}$=4.26 GeV". *Phys. Rev. Lett.* **110**:25 (2013), 252001. DOI: 10.1103/PhysRevLett.110.252001 (Cited on page 6).

[46] LHCb Collaboration. "Observation of the Resonant Character of the $Z(4430)^-$ State". *Phys. Rev. Lett.* **112**:22 (2014), 222002. DOI: 10.1103/PhysRevLett.112.222002 (Cited on page 6).

[47] Y. Nambu. "The confinement of quarks". *Scientific American* **235**:5 (1976), 48–63 (Cited on page 6).

[48] K. A. Johnson. "The bag model of quark confinement". *Scientific American* **241**:1 (1979), 112–121 (Cited on page 6).

[49] C. Rebbi. "The lattice theory of quark confinement". *Scientific American* **248**:2 (1983), 54–65 (Cited on page 6).

[50] M. Breidenbach et al. "Observed Behavior of Highly Inelastic Electron-Proton Scattering". *Phys. Rev. Lett.* **23**:16 (1969), 935–939. DOI: 10.1103/PhysRevLett.23.935 (Cited on page 6).

[51] E. D. Bloom et al. "High-Energy Inelastic $e-p$ Scattering at 6° and 10°". *Phys. Rev. Lett.* **23**:16 (1969), 930–934. DOI: 10.1103/PhysRevLett.23.930 (Cited on page 6).

[52] O. W. Greenberg. "Spin and Unitary-Spin Independence in a Paraquark Model of Baryons and Mesons". *Phys. Rev. Lett.* **13**:20 (1964), 598–602. DOI: 10.1103/PhysRevLett.13.598 (Cited on page 6).

[53] C. N. Yang and R. L. Mills. "Conservation of Isotopic Spin and Isotopic Gauge Invariance". *Phys. Rev.* **96**:1 (1954), 191–195. DOI: 10.1103/PhysRev.96.191 (Cited on page 6).

[54] B. S. DeWitt. "Theory of Radiative Corrections for Non-Abelian Gauge Fields". *Phys. Rev. Lett.* **12**:26 (1964), 742–746. DOI: 10.1103/PhysRevLett.12.742 (Cited on page 6).

[55] B. S. DeWitt. "Quantum Theory of Gravity. I. The Canonical Theory". *Phys. Rev.* **160**:5 (1967), 1113–1148. DOI: 10.1103/PhysRev.160.1113 (Cited on page 6).

[56] L. Faddeev and V. Popov. "Feynman Diagrams for the Yang-Mills Field". *Phys. Lett. B* **25**:1 (1967). Ed. by J.-P. Hsu and D. Fine, 29–30. DOI: 10.1016/0370-2693(67)90067-6 (Cited on page 6).

[57] G. 't Hooft. "Renormalization of Massless Yang-Mills Fields". *Nucl. Phys. B* **33**:1 (1971), 173–199. DOI: 10.1016/0550-3213(71)90395-6 (Cited on page 6).

[58] G. 't Hooft. "Renormalizable Lagrangians for Massive Yang-Mills Fields". *Nucl. Phys. B* **35**:1 (1971). Ed. by J. Taylor, 167–188. DOI: 10.1016/0550-3213(71)90139-8 (Cited on page 6).

[59] H. Fritzsch, M. Gell-Mann, and H. Leutwyler. "Advantages of the color octet gluon picture". *Phys. Lett. B* **47**:4 (1973), 365–368. ISSN: 0370-2693. DOI: 10.1016/0370-2693(73)90625-4 (Cited on page 6).

[60] TASSO Collaboration. "Evidence for planar events in e^+e^- annihilation at high energies". *Phys. Lett. B* **86**:2 (1979), 243–249. ISSN: 0370-2693. DOI: 10.1016/0370-2693(79)90830-X (Cited on page 6).

[61] MARK-J Collaboration. "Discovery of Three-Jet Events and a Test of Quantum Chromodynamics at PETRA". *Phys. Rev. Lett.* **43**:12 (1979), 830–833. DOI: 10.1103/PhysRevLett.43.830 (Cited on page 6).

[62] PLUTO Collaboration. "Evidence for gluon bremsstrahlung in e^+e^- annihilations at high energies". *Phys. Lett. B* **86**:3 (1979), 418–425. ISSN: 0370-2693. DOI: 10.1016/0370-2693(79)90869-4 (Cited on page 6).

[63] JADE Collaboration. "Observation of planar three-jet events in e^+e^- annihilation and evidence for gluon bremsstrahlung". *Phys. Lett. B* **91**:1 (1980), 142–147. ISSN: 0370-2693. DOI: 10.1016/0370-2693(80)90680-2 (Cited on page 6).

[64] B. J. Bjørken and S. L. Glashow. "Elementary particles and SU(4)". *Phys. Lett.* **11**:3 (1964), 255–257. ISSN: 0031-9163. DOI: 10.1016/0031-9163(64)90433-0 (Cited on page 7).

[65] S. L. Glashow, J. Iliopoulos, and L. Maiani. "Weak Interactions with Lepton-Hadron Symmetry". *Phys. Rev. D* **2**:7 (1970), 1285–1292. DOI: 10.1103/PhysRevD.2.1285 (Cited on page 7).

[66] J. J. Aubert et al. "Experimental Observation of a Heavy Particle *J*". *Phys. Rev. Lett.* **33**:23 (1974), 1404–1406. DOI: 10.1103/PhysRevLett.33.1404 (Cited on page 7).

[67] E598 Collaboration, Augustin, J. -E. et. al. "Discovery of a Narrow Resonance in e^+e^- Annihilation". *Phys. Rev. Lett.* **33**:23 (1974), 1406–1408. DOI: 10.1103/PhysRevLett.33.1406 (Cited on page 7).

[68] J. H. Christenson et al. "Evidence for the 2π Decay of the K_2^0 Meson". *Phys. Rev. Lett.* **13**:4 (1964), 138–140. DOI: 10.1103/PhysRevLett.13.138 (Cited on page 7).

[69] M. Kobayashi and T. Maskawa. "CP-Violation in the Renormalizable Theory of Weak Interaction". *Progress of Theoretical Physics* **49**:2 (1973), 652–657. ISSN: 0033-068X. DOI: 10.1143/PTP.49.652 (Cited on page 7).

[70] S. W. Herb et al. "Observation of a Dimuon Resonance at 9.5 GeV in 400-GeV Proton-Nucleus Collisions". *Phys. Rev. Lett.* **39**:5 (1977), 252–255. DOI: 10.1103/PhysRevLett.39.252 (Cited on page 7).

[71] CDF Collaboration. *Phys. Rev. Lett.* **74**:14 (1995), 2626–2631. DOI: 10.1103/PhysRevLett.74.2626 (Cited on page 7).

[72] D0 Collaboration. *Phys. Rev. Lett.* **74**:14 (1995), 2632–2637. DOI: 10.1103/PhysRevLett.74.2632 (Cited on page 7).

[73] T. D. Lee and C. N. Yang. "Parity Nonconservation and a Two-Component Theory of the Neutrino". *Phys. Rev.* **105**:5 (1957), 1671–1675. DOI: 10.1103/PhysRev.105.1671 (Cited on page 7).

[74] T. D. Lee and C. N. Yang. "Question of Parity Conservation in Weak Interactions". *Phys. Rev.* **104**:1 (1956), 254–258. DOI: 10.1103/PhysRev.104.254 (Cited on page 7).

[75] C. S. Wu et al. "Experimental Test of Parity Conservation in Beta Decay". *Phys. Rev.* **105**:4 (1957), 1413–1415. DOI: 10.1103/PhysRev.105.1413 (Cited on page 7).

[76] R. P. Feynman and M. Gell-Mann. "Theory of the Fermi Interaction". *Phys. Rev.* **109**:1 (1958), 193–198. DOI: 10.1103/PhysRev.109.193 (Cited on page 7).

[77] E. C. G. Sudarshan and R. E. Marshak. "Chirality Invariance and the Universal Fermi Interaction". *Phys. Rev.* **109**:5 (1958), 1860–1862. DOI: 10.1103/PhysRev.109.1860.2 (Cited on page 7).

[78] S. L. Glashow. "Partial-symmetries of weak interactions". *Nucl. Phys.* **22**:4 (1961), 579–588. ISSN: 0029-5582. DOI: 10.1016/0029-5582(61)90469-2 (Cited on page 7).

[79] A. Salam. "Weak and Electromagnetic Interactions". *Conf. Proc. C* **680519**: (1968), 367–377. DOI: 10.1142/9789812795915_0034 (Cited on page 7).

[80] S. Weinberg. "A Model of Leptons". *Phys. Rev. Lett.* **19**:21 (1967), 1264–1266. DOI: 10.1103/PhysRevLett.19.1264 (Cited on page 7).

[81] UA1 Collaboration. "Experimental Observation of Isolated Large Transverse Energy Electrons with Associated Missing Energy at *sqrts* = 540 GeV". *Phys. Lett. B* **122**:1 (1983), 103–116. DOI: 10.1016/0370-2693(83)91177-2 (Cited on page 7).

[82] UA2 Collaboration. "Observation of Single Isolated Electrons of High Transverse Momentum in Events with Missing Transverse Energy at the CERN anti-p p Collider". *Phys. Lett. B* **122**:5 (1983), 476–485. DOI: 10.1016/0370-2693(83)91605-2 (Cited on page 7).

[83] UA1 Collaboration. "Experimental Observation of Lepton Pairs of Invariant Mass Around 95-GeV/c^2 at the CERN SPS Collider". *Phys. Lett. B* **126**:5 (1983), 398–410. DOI: 10.1016/0370-2693(83)90188-0 (Cited on page 7).

[84] UA2 Collaboration. "Evidence for $Z^0 \to e^+e^-$ at the CERN $\bar{p}p$ Collider". *Phys. Lett. B* **129**:1 (1983), 130–140. DOI: 10.1016/0370-2693(83)90744-X (Cited on page 7).

[85] T. Y. Cao. "References". *Conceptual Developments of 20th Century Field Theories*. 2nd ed. Cambridge University Press, 2019, 376–421 (Cited on page 8).

[86] H. E. Logan. "TASI 2013 lectures on Higgs physics within and beyond the Standard Model" (2014). arXiv: 1406.1786 [hep-ph] (Cited on page 8).

[87] M. Thomson. *Modern particle physics*. New York: Cambridge University Press, 2013. ISBN: 978-1-107-03426-6 (Cited on page 8).

[88] P. Langacker. *The standard model and beyond*. 2010. ISBN: 978-1-4200-7906-7 (Cited on page 8).

[89] R. Feynman. "The behavior of hadron collisions at extreme energies". *Conf. Proc. C* **690905**: (1969), 237–258 (Cited on page 9).

[90] J. D. Bjorken and E. A. Paschos. "Inelastic Electron-Proton and γ-Proton Scattering and the Structure of the Nucleon". *Phys. Rev.* **185**:5 (1969), 1975–1982. DOI: 10.1103/PhysRev.185.1975 (Cited on page 9).

[91] G. Altarelli and G. Parisi. "Asymptotic Freedom in Parton Language". *Nucl. Phys. B* **126**:2 (1977), 298–318. DOI: 10.1016/0550-3213(77)90384-4 (Cited on page 10).

[92] Y. L. Dokshitzer. "Calculation of the Structure Functions for Deep Inelastic Scattering and e+ e- Annihilation by Perturbation Theory in Quantum Chromodynamics." *Sov. Phys. JETP* **46**: (1977), 641–653 (Cited on page 10).

[93] V. Gribov and L. Lipatov. "Deep inelastic e p scattering in perturbation theory". *Sov. J. Nucl. Phys.* **15**: (1972), 438–450 (Cited on page 10).

[94] F. Englert and R. Brout. "Broken Symmetry and the Mass of Gauge Vector Mesons". *Phys. Rev. Lett.* **13**:9 (1964), 321–323. DOI: 10.1103/PhysRevLett.13.321 (Cited on page 12).

[95] P. W. Higgs. "Broken Symmetries and the Masses of Gauge Bosons". *Phys. Rev. Lett.* **13**:16 (1964), 508–509. DOI: 10.1103/PhysRevLett.13.508 (Cited on page 12).

[96] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble. "Global Conservation Laws and Massless Particles". *Phys. Rev. Lett.* **13**:20 (1964), 585–587. DOI: 10.1103/PhysRevLett.13.585 (Cited on page 12).

[97] ATLAS Collaboration. "Combined measurements of Higgs boson production and decay using up to 80 fb^{-1} of proton-proton collision data at $\sqrt{s} = 13$ TeV collected with the ATLAS experiment". *Phys. Rev. D* **101**:1 (2020), 012002. DOI: 10.1103/PhysRevD.101.012002 (Cited on page 15).

[98] M. Czakon, P. Fiedler, and A. Mitov. "Total Top-Quark Pair-Production Cross Section at Hadron Colliders Through $\mathcal{O}(\alpha_S^4)$". *Phys. Rev. Lett.* **110**:25 (2013), 252004. DOI: 10.1103/PhysRevLett.110.252004 (Cited on pages 16, 115).

[99] LHC Top Working Group. *NNLO+NNLL top-quark-pair cross sections*. https://twiki.cern.ch/twiki/bin/view/LHCPhysics/TtbarNNLO. Accessed: 2020-09-30. 2015 (Cited on page 17).

[100] LHC Top Working Group. *Standalone Summary Plots of LHCTopWG*. https://twiki.cern.ch/twiki/bin/view/LHCPhysics/LHCTopWGSummaryPlots. Accessed: 2020-09-30. 2019 (Cited on page 17).

[101] D0 Collaboration. *Top Pair Branching Fractions*. https://www-d0.fnal.gov/Run2Physics/top/top_public_web_pages/top_feynman_diagrams.html. Accessed: 2020-09-30. 2011 (Cited on page 18).

[102] ATLAS Collaboration. *Summary plots, Standard Model Results*. https://twiki.cern.ch/twiki/bin/view/AtlasPublic/StandardModelPublicResultsRetired. Accessed: 2020-09-30. 2018 (Cited on pages 19, 20).

[103] ATLAS Collaboration. "Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC". *Phys. Lett. B* **716**:1 (2012), 1–29. DOI: 10.1016/j.physletb.2012.08.020 (Cited on page 20).

[104] CMS Collaboration. "Observation of a New Boson at a Mass of 125 GeV with the CMS Experiment at the LHC". *Phys. Lett. B* **716**:1 (2012), 30–61. DOI: 10.1016/j.physletb.2012.08.021 (Cited on page 20).

[105] D. Buttazzo et al. "Investigating the near-criticality of the Higgs boson". *Journal of High Energy Physics* **2013**:12 (2013), 89. ISSN: 1029-8479. DOI: 10.1007/JHEP12(2013)089 (Cited on page 21).

[106] D. E. Morrissey, T. Plehn, and T. M. Tait. "Physics searches at the LHC". *Phys. Rept.* **515**:1 (2012), 1–113. DOI: 10.1016/j.physrep.2012.02.007 (Cited on page 21).

[107] Y. Gershtein et al. "Working Group Report: New Particles, Forces, and Dimensions". *Community Summer Study 2013: Snowmass on the Mississippi*. 2013. arXiv: 1311.0299 [hep-ex] (Cited on page 21).

[108] P. Langacker. "The physics of heavy Z' gauge bosons". *Reviews of Modern Physics* **81**:3 (2009), 1199 (Cited on page 22).

[109] V. Barger, W. Keung, and E. Ma. "Sequential W and Z bosons". *Phys. Lett. B* **94**:3 (1980), 377–380. ISSN: 0370-2693. DOI: 10.1016/0370-2693(80)90900-4 (Cited on page 22).

[110] E. Farhi and L. Susskind. "Technicolour". *Phys. Rept.* **74**:3 (1981), 277–321. ISSN: 0370-1573. DOI: https://doi.org/10.1016/0370-1573(81)90173-3 (Cited on page 23).

[111] C. T. Hill and E. H. Simmons. "Strong dynamics and electroweak symmetry breaking". *Phys. Rept.* **381**:4 (2003). [Erratum: Phys.Rept. 390, 553–554 (2004)], 235–402. ISSN: 0370-1573. DOI: https://doi.org/10.1016/S0370-1573(03)00140-6 (Cited on page 23).

[112] S. Weinberg. "Implications of dynamical symmetry breaking: An addendum". *Phys. Rev. D* **19**:4 (1979). [Addendum: Phys.Rev.D 19, 1277–1280 (1979)], 1277–1280. DOI: 10.1103/PhysRevD.19.1277 (Cited on page 23).

[113] L. Susskind. "Dynamics of Spontaneous Symmetry Breaking in the Weinberg-Salam Theory". *Phys. Rev. D* **20**:10 (1979), 2619–2625. DOI: 10.1103/PhysRevD.20.2619 (Cited on page 23).

[114] R. M. Harris, C. T. Hill, and S. J. Parke. "Cross-Section for Topcolor Z'_t Decaying to $t\bar{t}$" (1999). arXiv: hep-ph/9911288 (Cited on page 23).

[115] C. T. Hill. "Topcolor: top quark condensation in a gauge extension of the standard model". *Phys. Lett. B* **266**:3 (1991), 419–424. ISSN: 0370-2693. DOI: https://doi.org/10.1016/0370-2693(91)91061-Y (Cited on pages 23, 116).

[116] C. T. Hill. "Topcolor assisted technicolor". *Phys. Lett. B* **345**:4 (1995), 483–489. ISSN: 0370-2693. DOI: https://doi.org/10.1016/0370-2693(94)01660-5 (Cited on pages 23, 116).

[117] R. M. Harris and S. Jain. "Cross Sections for Leptophobic Topcolor Z' Decaying to Top-Antitop". *Eur. Phys. J. C* **72**:7 (2012), 2072. DOI: 10.1140/epjc/s10052-012-2072-4 (Cited on pages 23, 116).

[118] D. Abercrombie et al. "Dark Matter Benchmark Models for Early LHC Run-2 Searches: Report of the ATLAS/CMS Dark Matter Forum". *Phys. Dark Univ.* **27**: (2020). Ed. by A. Boveia et al., 100371. DOI: 10.1016/j.dark.2019.100371 (Cited on pages 24, 158).

[119] G. Busoni et al. "Recommendations on presenting LHC searches for missing transverse energy signals using simplified *s*-channel models of dark matter". *Phys. Dark Univ.* **27**: (2020). Ed. by A. Boveia et al., 100365. DOI: 10.1016/j.dark.2019.100365 (Cited on pages 24, 158).

[120] A. Albert et al. "Recommendations of the LHC Dark Matter Working Group: Comparing LHC searches for dark matter mediators in visible and invisible decay channels and calculations of the thermal relic density". *Phys. Dark Univ.* **26**: (2019), 100377. DOI: 10.1016/j.dark.2019.100377 (Cited on page 24).

[121] F. Marcastel. "CERN's Accelerator Complex" (2013). General Photo. URL: https://cds.cern.ch/record/1621583 (Cited on page 26).

[122] L. Evans and P. Bryant. "LHC Machine". *JINST* **3**:08 (2008), S08001–S08001. DOI: 10.1088/1748-0221/3/08/s08001 (Cited on page 25).

[123] S. Myers. "The LEP collider, from design to approval and commissioning" (1991). DOI: 10.5170/CERN-1991-008 (Cited on page 25).

[124] ATLAS Collaboration. "The ATLAS Experiment at the CERN Large Hadron Collider". *JINST* **3**: (2008), S08003. DOI: 10.1088/1748-0221/3/08/S08003 (Cited on pages 25, 28, 30, 31, 33, 35).

[125] CMS Collaboration. "The CMS experiment at the CERN LHC". *JINST* **3**:08 (2008), S08004–S08004. DOI: 10.1088/1748-0221/3/08/s08004 (Cited on page 25).

[126] LHCb Collaboration. "The LHCb Detector at the LHC". *JINST* **3**:08 (2008), S08005–S08005. DOI: 10.1088/1748-0221/3/08/s08005 (Cited on page 25).

[127] ALICE Collaboration. "The ALICE experiment at the CERN LHC". *JINST* **3**:08 (2008), S08002–S08002. DOI: 10.1088/1748-0221/3/08/s08002 (Cited on page 25).

[128] R. Bailey and P. Collier. *Standard Filling Schemes for Various LHC Operation Modes*. Tech. rep. LHC-PROJECT-NOTE-323. Geneva: CERN, 2003. URL: https://cds.cern.ch/record/691782 (Cited on page 27).

[129] B. G. Taylor. "Timing distribution at the LHC" (2002). DOI: 10.5170/CERN-2002-003.63 (Cited on page 27).

[130] ATLAS Collaboration. "Luminosity determination in *pp* collisions at $$\sqrt{s}$$s= 8 TeV using the ATLAS detector at the LHC". *Eur. Phys. J. C* **76**:12 (2016), 653. ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-016-4466-1 (Cited on page 28).

[131] ATLAS Collaboration. "Luminosity determination in pp collisions at $\sqrt{s} = 13$ TeV using the ATLAS detector at the LHC". ATLAS-CONF-2019-021 (2019). URL: https://cds.cern.ch/record/2677054 (Cited on pages 28, 144).

[132] ATLAS Collaboration. *Luminosity Public Results Run-2*. https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResultsRun2. Accessed: 2020-09-30. 2019 (Cited on page 29).

[133] *ATLAS magnet system: Technical Design Report, 1*. Technical Design Report ATLAS. Geneva: CERN, 1997. URL: https://cds.cern.ch/record/338080 (Cited on page 30).

[134] A. Yamamoto et al. "Progress in ATLAS central solenoid magnet". *IEEE Transactions on Applied Superconductivity* **10**:1 (2000), 353–356 (Cited on page 30).

[135] ATLAS Collaboration. "The ATLAS Inner Detector commissioning and calibration". *Eur. Phys. J. C* **70**:3 (2010), 787–821. ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-010-1366-7 (Cited on page 31).

[136] ATLAS Collaboration. "ATLAS pixel detector electronics and sensors". *JINST* **3**:07 (2008), P07007. DOI: 10.1088/1748-0221/3/07/p07007 (Cited on pages 31, 39).

[137] ATLAS Collaboration. *ATLAS Insertable B-Layer Technical Design Report Addendum*. Tech. rep. CERN-LHCC-2012-009. ATLAS-TDR-19-ADD-1. Addendum to CERN-LHCC-2010-013, ATLAS-TDR-019. 2012. URL: https://cds.cern.ch/record/1451888 (Cited on page 31).

[138] ATLAS Collaboration. *ATLAS Insertable B-Layer Technical Design Report*. Tech. rep. CERN-LHCC-2010-013. ATLAS-TDR-19. 2010. URL: https://cds.cern.ch/record/1291633 (Cited on pages 31, 39).

[139] ATLAS SCT Collaboration. "The silicon microstrip sensors of the ATLAS semiconductor tracker". *Nucl. Instrum. Meth. Phys. Res. A* **578**:1 (2007), 98–118. ISSN: 0168-9002. DOI: 10.1016/j.nima.2007.04.157 (Cited on page 32).

[140] ATLAS TRT Collaboration. "The ATLAS Transition Radiation Tracker (TRT) proportional drift tube: Design and performance". *JINST* **3**:2 (2008), P02013. DOI: 10.1088/1748-0221/3/02/P02013 (Cited on page 32).

[141] ATLAS Collaboration. *ATLAS liquid-argon calorimeter: Technical Design Report*. Technical Design Report ATLAS. Geneva: CERN, 1996. URL: https://cds.cern.ch/record/331061 (Cited on pages 33, 34).

[142] C. Grupen and I. Buvat, eds. *Handbook of particle detection and imaging, vol. 1 and vol.2*. Berlin, Germany: Springer, 2012. ISBN: 978-3-642-13271-1. DOI: 10.1007/978-3-642-13271-1 (Cited on page 33).

[143] ATLAS Collaboration. *ATLAS tile calorimeter: Technical Design Report*. Technical Design Report ATLAS. Geneva: CERN, 1996. URL: https://cds.cern.ch/record/331062 (Cited on page 34).

[144] A. Henriques. "The ATLAS tile calorimeter". *4th International Conference on Advancements in Nuclear Instrumentation Measurement Methods and their Applications*. IEEE Nucl.Sci.Symp.Conf.Rec. 2015, 7465554. DOI: 10.1109/ANIMMA.2015.7465554 (Cited on page 34).

[145] ATLAS Collaboration. "Muon reconstruction performance of the ATLAS detector in proton-proton collision data at $\sqrt{s}$ =13 TeV". *Eur. Phys. J. C* **76**:5 (2016), 292. DOI: 10.1140/epjc/s10052-016-4120-y (Cited on pages 34, 56–58).

[146] ATLAS Collaboration. "Performance of the ATLAS muon triggers in Run 2". *JINST* **15**:09 (2020), P09015. DOI: 10.1088/1748-0221/15/09/p09015 (Cited on page 36).

[147] ATLAS Collaboration. "Operation of the ATLAS trigger system in Run 2" (2020). arXiv: 2007.12539 [physics.ins-det] (Cited on pages 36, 37).

[148] H. Bertelsen et al. "Operation of the upgraded ATLAS Central Trigger Processor during the LHC Run 2". *JINST* **11**:02 (2016), C02020–C02020. DOI: 10.1088/1748-0221/11/02/c02020 (Cited on page 37).

[149] ATLAS Collaboration. *Athena*. Version 22.0.1. 2019. DOI: 10.5281/zenodo.2641997 (Cited on page 37).

[150] G. Barrand et al. "GAUDI - A software architecture and framework for building HEP data processing applications". *Computer Physics Communications* **140**:1 (2001). CHEP2000, 45–55. ISSN: 0010-4655. DOI: 10.1016/S0010-4655(01)00254-5 (Cited on page 37).

[151] ATLAS Collaboration. "ATLAS data quality operations and performance for 2015–2018 data-taking". *JINST* **15**:04 (2020), P04003. DOI: 10.1088/1748-0221/15/04/P04003 (Cited on page 37).

[152] S. Agostinelli et al. "GEANT4–a simulation toolkit". *Nucl. Instrum. Meth. A* **506**:3 (2003), 250–303. DOI: 10.1016/S0168-9002(03)01368-8 (Cited on page 38).

[153] J. Allison et al. "Geant4 developments and applications". *IEEE Trans. Nucl. Sci.* **53**:1 (2006), 270. DOI: 10.1109/TNS.2006.869826 (Cited on page 38).

[154] ATLAS Collaboration. "The ATLAS Simulation Infrastructure". *Eur. Phys. J. C* **70**:3 (2010), 823–874. DOI: 10.1140/epjc/s10052-010-1429-9 (Cited on page 38).

[155] ATLAS Collaboration. "The simulation principle and performance of the ATLAS fast calorimeter simulation FastCaloSim". ATL-PHYS-PUB-2010-013 (2010). URL: https://cds.cern.ch/record/1300517 (Cited on page 38).

[156] J. S. and. "The new ATLAS Fast Calorimeter Simulation". *Journal of Physics: Conference Series* **898**: (2017), 042006. DOI: 10.1088/1742-6596/898/4/042006 (Cited on page 38).

[157] ATLAS collaboration. "A neural network clustering algorithm for the ATLAS silicon pixel detector". *JINST* **9**:09 (2014), P09009–P09009. DOI: 10.1088/1748-0221/9/09/p09009 (Cited on pages 39, 40).

[158] A. Rosenfeld and J. L. Pfaltz. "Sequential Operations in Digital Picture Processing". *J. ACM* **13**:4 (1966), 471–494. ISSN: 0004-5411. DOI: 10.1145/321356.321357 (Cited on page 40).

[159] ATLAS Collaboration. *Training and validation of the ATLAS pixel clustering neural networks*. Tech. rep. CERN, 2018. URL: https://cds.cern.ch/record/2309474 (Cited on pages 40, 41).

[160] C. M. Bishop. "Mixture density networks". Aston University, 1994. ISBN: NCRG/94/004 (Cited on pages 41, 42).

[161] G. J. McLachlan and K. E. Basford. *Mixture models: Inference and applications to clustering*. Vol. 84. Marcel Dekker, 1988 (Cited on page 42).

[162] J. S. Bridle. "Probabilistic interpretation of feedforward classification network outputs, with relationships to statistical pattern recognition". *Neurocomputing*. Springer, 1990, 227–236. DOI: 10.1007/978-3-642-76153-9_28 (Cited on page 42).

[163] R. A. Jacobs et al. "Adaptive mixtures of local experts". *Neural computation* **3**:1 (1991), 79–87. DOI: 10.1162/neco.1991.3.1.79 (Cited on page 42).

[164] C. M. Bishop. *Pattern recognition and machine learning*. springer, 2006 (Cited on page 42).

[165] F. Chollet et al. *Keras*. 2015. URL: https://keras.io (Cited on page 43).

[166] J. Bergstra et al. "Theano: Deep learning on gpus with python" (2011) (Cited on page 43).

[167] D. P. Kingma and J. Ba. "Adam: A method for stochastic optimization" (2014). arXiv: 1412.6980 (Cited on page 43).

[168] I. Goodfellow, Y. Bengio, and A. Courville. *Deep Learning*. MIT Press, 2016. URL: http://www.deeplearningbook.org (Cited on page 46).

[169] A. Krogh and J. A. Hertz. "A simple weight decay can improve generalization". *Advances in neural information processing systems*. 1992, 950–957 (Cited on page 46).

[170] M. B. Christopher. *Pattern recognition and Machine Learning*. Springer-Verlag New York, 2016 (Cited on page 46).

[171] S. Ioffe and C. Szegedy. "Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift". *Proceedings of the 32nd International Conference on Machine Learning*. Ed. by F. Bach and D. Blei. Vol. 37. Proceedings of Machine Learning Research. Lille, France: PMLR, 2015, 448–456 (Cited on page 46).

[172] J. Pequenao and P. Schaffner. "How ATLAS detects particles: diagram of particle paths in the detector" (2013). URL: https://cds.cern.ch/record/1505342 (Cited on page 50).

[173] T. Cornelissen et al. "The new ATLAS track reconstruction (NEWT)". *Journal of Physics: Conference Series* **119**:3 (2008), 032014. DOI: 10.1088/1742-6596/119/3/032014 (Cited on page 50).

[174] ATLAS Collaboration. "Performance of the ATLAS Track Reconstruction Algorithms in Dense Environments in LHC Run 2". *Eur. Phys. J. C* **77**:10 (2017), 673. DOI: 10.1140/epjc/s10052-017-5225-7 (Cited on pages 50, 51).

[175] ATLAS Collaboration. "Reconstruction of primary vertices at the ATLAS experiment in Run 1 proton–proton collisions at the LHC". *Eur. Phys. J. C* **77**:5 (2017), 332. DOI: 10.1140/epjc/s10052-017-4887-5 (Cited on page 51).

[176] G. Piacquadio, K. Prokofiev, and A. Wildauer. "Primary vertex reconstruction in the ATLAS experiment at LHC". *J. of Phys.: Conf. Ser.* **119**:3 (2008), 032033. DOI: 10.1088/1742-6596/119/3/032033 (Cited on page 51).

[177] W. Waltenberger, R. Frühwirth, and P. Vanlaer. "Adaptive vertex fitting". *Journal of Physics G: Nuclear and Particle Physics* **34**:12 (2007), N343–N356. DOI: 10.1088/0954-3899/34/12/n01 (Cited on page 51).

[178] W. Lampl et al. *Calorimeter Clustering Algorithms: Description and Performance*. Tech. rep. ATL-LARG-PUB-2008-002. ATL-COM-LARG-2008-003. Geneva: CERN, 2008. URL: https://cds.cern.ch/record/1099735 (Cited on page 52).

[179] ATLAS Collaboration. "Topological cell clustering in the ATLAS calorimeters and its performance in LHC Run 1". *Eur. Phys. J. C* **77**:7 (2017), 490. DOI: 10.1140/epjc/s10052-017-5004-5 (Cited on pages 52, 60, 63).

[180] ATLAS Collaboration. "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. C* **79**:8 (2019), 639. arXiv: 1902.04655 [hep-ex]. URL: https://doi.org/10.1140/epjc/s10052-019-7140-6 (Cited on pages 54, 55).

[181] ATLAS Collaboration. "Electron and photon performance measurements with the ATLAS detector using the 2015–2017 LHC proton-proton collision data". *JINST* **14**:12 (2019), P12006–P12006. DOI: 10.1088/1748-0221/14/12/p12006 (Cited on pages 53, 55).

[182] M. Cacciari, G. P. Salam, and G. Soyez. "The anti-k_t jet clustering algorithm". *JHEP* **2008**:04 (2008), 063. DOI: 10.1088/1126-6708/2008/04/063 (Cited on page 59).

[183] ATLAS Collaboration. "Jet energy scale measurements and their systematic uncertainties in proton-proton collisions at $\sqrt{s}$ = 13 TeV with the ATLAS detector". *Phys. Rev. D* **96**:7 (2017), 072002. DOI: 10.1103/PhysRevD.96.072002 (Cited on pages 60, 61, 199–201).

[184] ATLAS Collaboration. "Performance of pile-up mitigation techniques for jets in pp collisions at $\sqrt{s}$ = 8 TeV using the ATLAS detector". *Eur. Phys. J. C* **76**:11 (2016), 581. DOI: 10.1140/epjc/s10052-016-4395-z (Cited on page 61).

[185] ATLAS Collaboration. "Jet reconstruction and performance using particle flow with the ATLAS Detector". *Eur. Phys. J. C* **77**:7 (2017), 466. DOI: 10.1140/epjc/s10052-017-5031-2 (Cited on pages 62, 202).

[186] ATLAS Collaboration. *Variable Radius, Exclusive-k_T, and Center-of-Mass Subjet Reconstruction for Higgs($\to b\bar{b}$) Tagging in ATLAS*. Tech. rep. ATL-PHYS-PUB-2017-010. Geneva: CERN, 2017. URL: http://cds.cern.ch/record/2268678 (Cited on page 62).

[187] D. Krohn, J. Thaler, and L.-T. Wang. "Jet Trimming". *JHEP* **02**: (2010), 084. DOI: 10.1007/JHEP02(2010)084 (Cited on page 63).

[188] S. D. Ellis and D. E. Soper. "Successive combination jet algorithm for hadron collisions". *Phys. Rev. D* **48**: (1993), 3160–3166. DOI: 10.1103/PhysRevD.48.3160 (Cited on page 63).

[189] M. Cacciari, G. P. Salam, and G. Soyez. "The Catchment Area of Jets". *JHEP* **04**: (2008), 005. DOI: 10.1088/1126-6708/2008/04/005 (Cited on page 63).

[190] ATLAS Collaboration. "In situ calibration of large-radius jet energy and mass in 13 TeV proton–proton collisions with the ATLAS detector". *Eur. Phys. J. C* **79**:2 (2019), 135. DOI: 10.1140/epjc/s10052-019-6632-8 (Cited on pages 63, 64, 202).

[191] ATLAS Collaboration. "Performance of jet substructure techniques for large-R jets in proton-proton collisions at $\sqrt{s}$ = 7 TeV using the ATLAS detector". *JHEP* **2013**:9 (2013), 076. DOI: 10.1007/JHEP09(2013)076 (Cited on pages 64, 144).

[192] W. Commons. *File:B-tagging diagram.png — Wikimedia Commons, the free media repository*. [Online; accessed 9-July-2020]. 2016. URL: https://commons.wikimedia.org/w/index.php?title=File:B-tagging_diagram.png&oldid=216090609 (Cited on page 65).

[193] ATLAS Collaboration. "ATLAS b-jet identification performance and efficiency measurement with $t\bar{t}$ events in pp collisions at $\sqrt{s}$ = 13 TeV". *Eur. Phys. J. C* **79**:11 (2019), 970. DOI: 10.1140/epjc/s10052-019-7450-8 (Cited on pages 66, 68).

[194] ATLAS Collaboration. *Optimisation and performance studies of the ATLAS b-tagging algorithms for the 2017-18 LHC run*. Tech. rep. ATL-PHYS-PUB-2017-013. Geneva: CERN, 2017. URL: https://cds.cern.ch/record/2273281 (Cited on pages 66, 67).

[195] ATLAS Collaboration. *Secondary vertex finding for jet flavour identification with the ATLAS detector*. Tech. rep. ATL-PHYS-PUB-2017-011. Geneva: CERN, 2017. URL: https://cds.cern.ch/record/2270366 (Cited on page 66).

[196] ATLAS Collaboration. *Topological b-hadron decay reconstruction and identification of b-jets with the JetFitter package in the ATLAS experiment at the LHC*. Tech. rep. ATL-PHYS-PUB-2018-025. Geneva: CERN, 2018. URL: https://cds.cern.ch/record/2645405 (Cited on page 67).

[197] ATLAS Collaboration. "Performance of top-quark and W-boson tagging with ATLAS in Run 2 of the LHC". *Eur. Phys. J. C* **79**:5 (2019), 375. DOI: 10.1140/epjc/s10052-019-6847-8 (Cited on pages 70, 71, 156).

[198] ATLAS Collaboration. "Performance of missing transverse momentum reconstruction with the ATLAS detector using proton-proton collisions at $\sqrt{s}$ = 13 TeV". *Eur. Phys. J. C* **78**:11 (2018), 903. DOI: 10.1140/epjc/s10052-018-6288-9 (Cited on page 72).

[199] F. James and M. Roos. "Minuit: A System for Function Minimization and Analysis of the Parameter Errors and Correlations". *Comput. Phys. Commun.* **10**:6 (1975), 343–367. DOI: 10.1016/0010-4655(75)90039-9 (Cited on page 78).

[200] J. Neyman, E. S. Pearson, and K. Pearson. "IX. On the problem of the most efficient tests of statistical hypotheses". *Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character* **231**:694-706 (1933), 289–337. DOI: 10.1098/rsta.1933.0009 (Cited on page 78).

[201] S. S. Wilks. "The Large-Sample Distribution of the Likelihood Ratio for Testing Composite Hypotheses". *Ann. Math. Statist.* **9**:1 (1938), 60–62. DOI: 10.1214/aoms/1177732360 (Cited on pages 79, 82).

[202] G. Cowan et al. "Asymptotic formulae for likelihood-based tests of new physics". *Eur. Phys. J. C* **71**:2 (2011). [Erratum: Eur.Phys.J.C 73, 2501 (2013)], 1554. DOI: 10.1140/epjc/s10052-011-1554-0 (Cited on pages 80, 81, 84).

[203] A. Wald. "Tests of Statistical Hypotheses Concerning Several Parameters When the Number of Observations is Large". *Transactions of the American Mathematical Society* **54**:3 (1943), 426–482. ISSN: 00029947. URL: http://www.jstor.org/stable/1990256 (Cited on page 81).

[204] A. L. Read. "Presentation of search results: The CL(s) technique". *J. Phys. G* **28**: (2002). Ed. by M. Whalley and L. Lyons, 2693–2704. DOI: 10.1088/0954-3899/28/10/313 (Cited on pages 84, 151).

[205] G. Choudalakis. "On hypothesis testing, trials factor, hypertests and the BumpHunter". *Proceedings, PHYSTAT 2011 Workshop on Statistical Issues Related to Discovery Claims in Search Experiments and Unfolding, CERN,Geneva, Switzerland 17-20 January 2011*. 2011. arXiv: 1101.0390 [physics.data-an] (Cited on page 88).

[206] ATLAS Collaboration. "Search for new high-mass phenomena in the dilepton final state using 36 fb^{-1} of proton-proton collision data at $\sqrt{s}$ = 13 TeV with the ATLAS detector". *JHEP* **2017**:10 (2017), 182. DOI: 10.1007/JHEP10(2017)182 (Cited on pages 93, 102, 110–112).

[207] ATLAS Collaboration. "Search for high-mass new phenomena in the dilepton final state using proton-proton collisions at $\sqrt{s}$ = 13 TeV with the ATLAS detector". *Phys. Lett. B* **761**: (2016), 372–392. DOI: 10.1016/j.physletb.2016.08.055 (Cited on page 93).

[208] S. Alioli et al. "A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX". *J. High Energy Phys.* **2010**:6 (2010), 043. DOI: 10.1007/JHEP06(2010)043 (Cited on pages 97, 98, 115, 116).

[209] H.-L. Lai et al. "New parton distributions for collider physics". *Phys. Rev. D* **82**:7 (2010), 074024. DOI: 10.1103/PhysRevD.82.074024 (Cited on pages 97, 98).

[210] T. Sjostrand, S. Mrenna, and P. Z. Skands. "A Brief Introduction to PYTHIA 8.1". *Comput. Phys. Commun.* **178**:11 (2008), 852–867. DOI: 10.1016/j.cpc.2008.01.036 (Cited on pages 97, 98, 101, 116, 118, 164).

[211] ATLAS Collaboration. "Measurement of the Z/γ^* boson transverse momentum distribution in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector". *JHEP* **2014**:9 (2014), 55. DOI: 10.1007/JHEP09(2014)145 (Cited on page 97).

[212] S. Dulat et al. "The CT14 Global Analysis of Quantum Chromodynamics" (2015). arXiv: 1506.07443 [hep-ph] (Cited on page 97).

[213] C. Anastasiou et al. "High precision QCD at hadron colliders: Electroweak gauge boson rapidity distributions at NNLO". *Phys. Rev. D* **69**:9 (2004), 094008. DOI: 10.1103/PhysRevD.69.094008 (Cited on pages 97, 104).

[214] S. G. Bondarenko and A. A. Sapronov. "NLO EW and QCD proton-proton cross section calculations with mcsanc-v1.01". *Comput. Phys. Commun.* **184**:10 (2013), 2343–2350. DOI: 10.1016/j.cpc.2013.05.010 (Cited on page 97).

[215] A. D. Martin et al. "Parton distributions incorporating QED contributions". *Eur. Phys. J. C* **39**:2 (2005), 155–161. DOI: 10.1140/epjc/s2004-02088-7 (Cited on page 97).

[216] T. Gleisberg et al. "Event generation with SHERPA 1.1". *JHEP* **2009**:02 (2009), 007. DOI: 10.1088/1126-6708/2009/02/007 (Cited on pages 97, 98).

[217] Skands, Peter Zeiler. "Tuning Monte Carlo Generators: The Perugia Tunes". *Phys. Rev. D* **82**:7 (2010), 074018. DOI: 10.1103/PhysRevD.82.074018 (Cited on page 97).

[218] M. Czakon and A. Mitov. "Top++: A Program for the Calculation of the Top-Pair Cross-Section at Hadron Colliders". *Comput. Phys. Commun.* **185**: (2014), 2930. DOI: 10.1016/j.cpc.2014.06.021 (Cited on pages 98, 115).

[219] T. Sjostrand, S. Mrenna, and P. Z. Skands. "PYTHIA 6.4 Physics and Manual". *JHEP* **2006**:05 (2006), 026. DOI: 10.1088/1126-6708/2006/05/026 (Cited on page 98).

[220] R. D. Ball et al. "Parton distributions with LHC data". *Nucl. Phys. B* **867**:2 (2013), 244–289. DOI: 10.1016/j.nuclphysb.2012.10.003 (Cited on pages 100, 101, 116, 118, 164).

[221] J. Gao and P. Nadolsky. "A meta-analysis of parton distribution functions". *JHEP* **2014**:7 (2014), 035. DOI: 10.1007/JHEP07(2014)035 (Cited on page 104).

[222] S. Dulat et al. "New parton distribution functions from a global analysis of quantum chromodynamics". *Phys. Rev. D* **93**:3 (2016), 033006. DOI: 10.1103/PhysRevD.93.033006 (Cited on page 104).

[223] J. Butterworth et al. "PDF4LHC recommendations for LHC Run II". *J. Phys. G* **43**:2 (2016), 023001. DOI: 10.1088/0954-3899/43/2/023001 (Cited on pages 104, 105).

[224] P. Motylinski et al. "Updates of PDFs for the 2nd LHC run". *Nucl. Part. Phys. Proc.* **273-275**: (2016). Ed. by M. Aguilar-Benitez et al., 2136–2141. DOI: 10.1016/j.nuclphysbps.2015.09.347 (Cited on page 104).

[225] R. D. Ball et al. "Parton distributions for the LHC run II". *JHEP* **2015**:4 (2015), 40. DOI: 10.1007/JHEP04(2015)040 (Cited on pages 104, 115, 116).

[226] A. Caldwell, D. Kollár, and K. Kröninger. "BAT – The Bayesian analysis toolkit". *Computer Physics Communications* **180**:11 (2009), 2197–2209. ISSN: 0010-4655. DOI: 10.1016/j.cpc.2009.06.026 (Cited on page 111).

[227] ATLAS Collaboration. "Search for high-mass dilepton resonances using 139 fb^{-1} of pp collision data collected at $\sqrt{s}$ =13 TeV with the ATLAS detector". *Phys. Lett. B* **796**: (2019), 68–87. DOI: 10.1016/j.physletb.2019.07.016 (Cited on page 112).

[228] ATLAS Collaboration. "Search for $t\bar{t}$ resonances in fully hadronic final states in pp collisions at $\sqrt{s}$ = 13 TeV with the ATLAS detector". *JHEP* **2020**:10 (2020), 061. DOI: 10.1007/JHEP10(2020)061 (Cited on pages 113, 126, 127, 143, 157, 158).

[229] ATLAS Collaboration. "Search for heavy particles decaying into a top-quark pair in the fully hadronic final state in pp collisions at $\sqrt{s}$ = 13 TeV with the ATLAS detector". *Phys. Rev. D* **99**:9 (2019), 092004. DOI: 10.1103/PhysRevD.99.092004 (Cited on pages 114, 117, 156).

[230] P. Nason. "A new method for combining NLO QCD with shower Monte Carlo algorithms". *JHEP* **2004**:11 (2004), 040. DOI: 10.1088/1126-6708/2004/11/040 (Cited on page 115).

[231] S. Frixione, P. Nason, and C. Oleari. "Matching NLO QCD computations with parton shower simulations: the POWHEG method". *JHEP* **2007**:11 (2007), 070. DOI: 10.1088/1126-6708/2007/11/070 (Cited on page 115).

[232] M. Cacciari et al. "Top-pair production at hadron colliders with next-to-next-to-leading logarithmic soft-gluon resummation". *Phys. Lett. B* **710**:4 (2012), 612–622. DOI: 10.1016/j.physletb.2012.03.013 (Cited on page 115).

[233] M. Beneke et al. "Hadronic top-quark pair production with NNLL threshold resummation". *Nucl. Phys. B* **855**:3 (2012), 695–741. DOI: 10.1016/j.nuclphysb.2011.10.021 (Cited on page 115).

[234] P. Baernreuther, M. Czakon, and A. Mitov. "Percent-Level-Precision Physics at the Tevatron: Next-to-Next-to-Leading Order QCD Corrections to $q\bar{q} \to t\bar{t} + X$". *Phys. Rev. Lett.* **109**:13 (2012), 132001. DOI: 10.1103/PhysRevLett.109.132001 (Cited on page 115).

[235] M. Czakon and A. Mitov. "NNLO corrections to top-pair production at hadron colliders: the all-fermionic scattering channels". *JHEP* **2012**:12 (2012), 54. DOI: 10.1007/JHEP12(2012)054 (Cited on page 115).

[236] M. Czakon and A. Mitov. "NNLO corrections to top pair production at hadron colliders: the quark-gluon reaction". *JHEP* **2013**:1 (2013), 80. DOI: 10.1007/JHEP01(2013)080 (Cited on page 115).

[237] ATLAS Collaboration. *Studies on top-quark Monte Carlo modelling for Top2016*. ATL-PHYS-PUB-2016-020. 2016. URL: https://cds.cern.ch/record/2216168 (Cited on page 115).

[238] J. Kühn, A. Scharf, and P. Uwer. "Weak interactions in top-quark pair production at hadron colliders: An update". *Phys. Rev. D* **91**:1 (2015), 014020. DOI: 10.1103/PhysRevD.91.014020 (Cited on page 115).

[239] Duccio Pappadopulo, Andrea Thamm, Riccardo Torre and Andrea Wulzer. "Heavy vector triplets: bridging theory and data". *JHEP* **2014**:9 (2014), 60. ISSN: 1029-8479. DOI: 10.1007/JHEP09(2014)060 (Cited on pages 116, 117).

[240] J. Alwall et al. "The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations". *JHEP* **2014**:7 (2014), 079. DOI: 10.1007/JHEP07(2014)079 (Cited on pages 117, 118, 164).

[241] ATLAS Collaboration. *Selection of jets produced in 13TeV proton-proton collisions with the ATLAS detector*. Tech. rep. ATLAS-CONF-2015-029. Geneva: CERN, 2015. URL: https://cds.cern.ch/record/2037702 (Cited on page 118).

[242] ATLAS Collaboration. "Search for new phenomena in dijet mass and angular distributions from pp collisions at $\sqrt{s}$ = 13 TeV with the ATLAS detector". *Phys. Lett. B* **754**: (2016), 302–322. DOI: 10.1016/j.physletb.2016.01.032 (Cited on pages 134, 135).

[243] ATLAS Collaboration. "Search for boosted dijet resonances decaying to two b-quarks and produced in association with a jet". ATL-COM-PHYS-2018-300 (2018). URL: https://cds.cern.ch/record/2310645 (Cited on page 135).

[244] K. Cranmer et al. *HistFactory: A tool for creating statistical models for use with RooFit and RooStats*. Tech. rep. CERN-OPEN-2012-016. New York: New York U., 2012. URL: https://cds.cern.ch/record/1456844 (Cited on page 151).

[245] W. Verkerke and D. Kirkby. *The RooFit toolkit for data modeling*. Tech. rep. physics/0306116. Stanford, CA: SLAC, 2003. URL: https://cds.cern.ch/record/622147 (Cited on page 151).

[246] L. Moneta et al. "The RooStats Project". *PoS* **ACAT2010**: (2010). Ed. by T. Speer et al., 057. DOI: 10.22323/1.093.0057 (Cited on page 151).

[247] *Dark matter summary plots for s-channel mediators*. Tech. rep. ATL-PHYS-PUB-2020-021. Geneva: CERN, 2020. URL: https://cds.cern.ch/record/2725266 (Cited on page 159).

[248] ATLAS Collaboration. "Jet mass reconstruction with the ATLAS Detector in early Run 2 data". ATLAS-CONF-2016-035 (2016). URL: https://cds.cern.ch/record/2200211 (Cited on page 159).

[249] J. Hollingsworth and D. Whiteson. "Resonance Searches with Machine Learned Likelihood Ratios" (2020). arXiv: 2002.04699 [hep-ph] (Cited on page 159).

[250] ATLAS Collaboration. "Search for heavy particles decaying into top-quark pairs using lepton-plus-jets events in proton–proton collisions at $\sqrt{s}$ = 13 TeV with the ATLAS detector". *Eur. Phys. J. C* **78**:7 (2018), 565. DOI: 10.1140/epjc/s10052-018-5995-6 (Cited on pages 160, 162, 174, 179).

[251] ATLAS Collaboration. "Measurements of $t\bar{t}$ differential cross-sections of highly boosted top quarks decaying to all-hadronic final states in pp collisions at $\sqrt{s}$ = 13 TeV using the ATLAS detector". *Phys. Rev. D* **98**:1 (2018), 012003. DOI: 10.1103/PhysRevD.98.012003 (Cited on page 203).

APPENDIX

A

Jet Calibration

In this appendix, some of the steps used to calibrate different types of jets in ATLAS is described. Calibration of EMTopo jets are described in section A.1. The calibration sequence for PFlow jets are shown in section A.2. MC based calibration of large-R jet is summarized in section A.3.

A.1 EMTopo jet calibration

This section describes the steps used to calibrate the EMTopo jets. A more detail description is available in [183].

Jet Origin Correction

The directions of the topoclusters in $\eta-\phi$ plane are defined as from the center of the ATLAS detector to the energy-weighted barycentre of the associated calorimeter cells. Thus the reconstructed jet momentum points to the center of the ATLAS detector. The origin correction recalculates the jet four-vectors such that it points to the hard-scatter primary vertex while keeping the jet energy constant. This correction improves the η resolution ($\eta_{\text{reco}} - \eta_{\text{truth}}$) from roughly 0.06 to around 0.045 at jet $p_T = 20$ GeV. The improvement is even better for jets with p_T higher than 200 GeV for which the resolution improves from 0.03 to below 0.006 [183].

Pileup Corrections

Pileup corrections are used to remove the additional energy deposited within the jet radius due to in-time and out-of-time pileup effects. First, an area-based correction is done by removing the per-event pileup contribution to the jet p_T. This effect is estimated from the median jet p_T density, ρ, in the $\eta - \phi$ plane. The pileup effect is assumed to be distributed roughly uniformly. So the resulting pileup effect on jets should be proportional to the jet area. Certainly, the pileup is not exactly uniform and that has to be considered. Hence after the area-based correction, there is still a residual dependence on the reconstructed jet p_T. This additional correction is parametrized in terms of the number of primary vertices (N_{PV}) and the number of average interactions, μ. This way of parametrization takes into account both in-time and out-of-time pileup effects. Typically

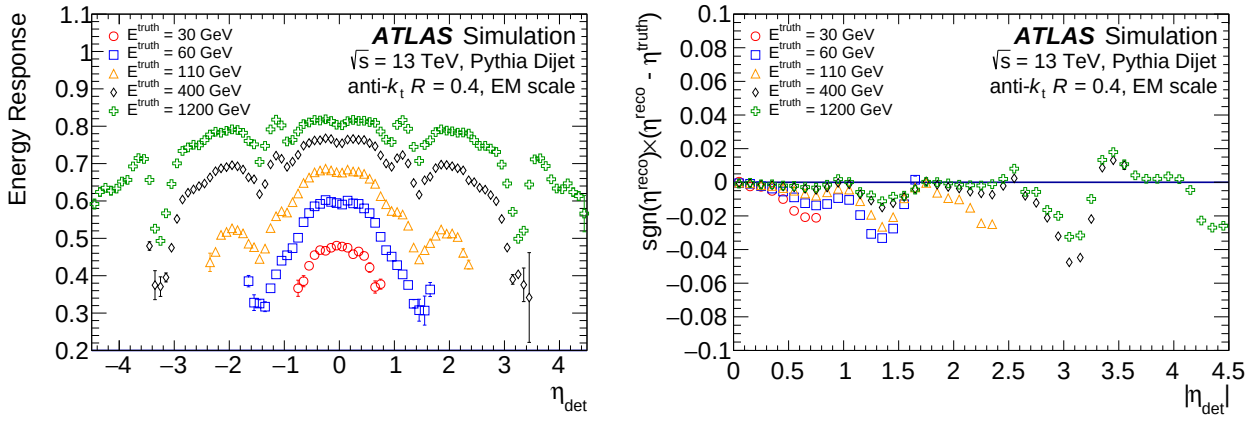
the jets in the central detector region gets corrected by ±1 GeV from these residual corrections. The pileup-corrected jet p_T can be written as

$$p_T^{\text{corr}} = p_T^{\text{reco}} - \underbrace{\rho \times A}_{\text{area-based}} \underbrace{- \alpha \times (N_{\text{PV}} - 1)}_{\text{in-time residual}} - \underbrace{\beta \times \mu}_{\text{out-of-time residual}} \tag{A.1}$$

where α and β parameters are derived from simulation and can be mathematically expresses at the slope of individual residual term, $\alpha = \partial p_T / \partial N_{PV}$ and $\beta = \partial p_T / \partial \mu$.

Jet Energy Scale and η correction

Next, the absolute jet energy scale calibration is applied to correct the reconstructed jet energy (EM-scale) to the truth particle jet energy. In this process, the reconstructed jet four-momentum is corrected to that of the associated truth jet using Pythia MC samples. This correction accounts for the energy loss due to dead material, out-of-cone radiations (radiations outside the jet radius), reconstruction inefficiencies, non-compensating calorimeter response, and also accounts for bias in jet η reconstruction. To derive this calibration, reconstructed jets are geometrically matched to truth jets within $\Delta R = 0.3$. Only isolated jets with no other calorimeter jet with $p_T > 7$ GeV within $\Delta R = 0.6$, and with only one associated truth jet with $p_T^{\text{truth}} > 7$ GeV within $\Delta R = 1.0$ are used to avoid ambiguity in truth matching. The inverse of the energy response, $E^{\text{reco}}/E^{\text{truth}}$, is used as the correction factor. Figure A.1a shows the average energy response as function of η_{det}. η_{det} is the η pointing from the geometric center of the detector and is useful to remove the ambiguity of the jet position. A bias is observed in the reconstructed jet η distribution as shown in Figure A.1b. The bias is larger in the calorimeter transition regions, barrel-endcap ($|\eta_{\text{det}}| \sim 1.4$) and endcap-forward ($|\eta_{\text{det}}| \sim 3.1$) transition regions where the calorimeter geometry and technology changes. In order to address this effect, a second correction factor is derived. The second correction factor is the difference between the reconstructed and the truth jet η, as shown in Figure A.1b. Jets after these two corrections have been applied, are referred to as EM+JES jets.



(a) (b)

Figure A.1: EM topo jet energy response and η correction factor as a function of detector pseudo-rapidity, η_{det}. (a) The average energy response as a function of η_{det} for jets for a truth energy of 30, 60, 110, 400, and 1200 GeV. (b) Signed difference between the truth and reconstructed jet η as a function of η_{det}. An η correction applied to address the bias. Figure taken from [183].

Global Sequence Calibration

After doing the previous calibrations, residual effects are observed on the transverse and longitudinal features of the jets. Global sequence calibration [183] is a series of corrections that accounts for the remaining residual effects. Such differences occur since the calorimeter response and the jet reconstruction algorithm are sensitive to fluctuations in jet particle composition and energy distribution within the jet. A quark-initiated jet often comprises of fewer hadrons, each with a higher fraction of the jet p_T and thus penetrates further into the calorimeter. Hence produces a higher calorimeter response on average. A gluon-initiated jet, on the other hand, contains more particles with relatively softer p_T, producing a lower calorimeter response and a wider transverse profile. The correction is derived using a set of five observables to flatten the jet energy response without changing the mean jet energy. It also uses muon track segments associated with the jet to reduce the tail of the response distribution caused by high p_T jets, that are not always fully-contained in the calorimeter system.

In-situ calibration: EMTopo jet

This is the last stage of the jet calibration, and it accounts for differences in the jet response between data and MC. The differences arise due to several imperfections in the detector description and response in simulation [183] as well as non-constant calorimeter performance during data taking. The η-intercalibration corrects the average response of forward jets ($0.8 < |\eta| < 4.5$) to that of central jets ($|\eta/| < 0.8$) using p_T balance in dijet events. Then Z/γ+jets events are used to balance the recoiled jets against a calibrated Z-boson or a γ in the central region. Finally, multijet events are used to calibrate a high-p_T jet against a system of well-calibrated low p_T jets, the so-called multijet balance (MJB). The Z/γ+jet and MJB are computed for central jets but also applied to the forward jets. A statistical combination of the Z/γ+jet and MJB is done to get a smooth calibration application over the full momentum range.

A.2 PFlow Energy scale correction

The steps are the same as used for EMTopo jets. The energy response ($E^{\text{reco}}/E^{\text{truth}}$) prior to the correction is shown in Figure A.2. The response is shown for the MC events where the distributions are binned in energy of the matched truth jet.

A.3 Large-R jet energy and mass scale calibration

The MC based calibration is very similar to small-R jets but it's a two-step process. First, the jet energy is corrected with the inverse of the jet energy response as discussed in section A.1. Figure A.3a shows the pre-calibrated jet energy response as a function of large-R jet η_{det}. The correction is roughly 5-10% and has an η dependent structure because of the detector geometry. Subsequently, a jet mass calibration is applied for the large-R jets. Since jet mass is more sensitive to soft, wide-angle contributions, cluster splitting, and merging and calorimeter geometry, it is important when jets are used in physics analysis. In order to do the jet mass calibration, the jet mass response, $m^{\text{reco}}/m^{\text{truth}}$, is defined in the same way as jet energy response. In this thesis, top quarks are used, as a result, the jet mass calibration to the top mass scale is very important. Figure A.3b shows the pre-calibrated mass response for truth top quark mass. The mass response

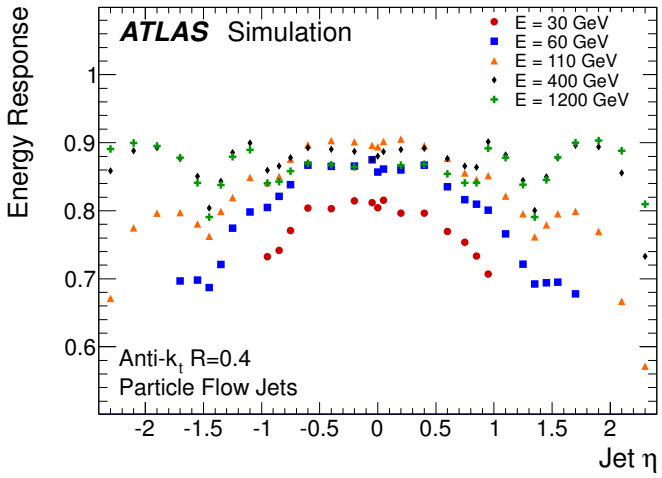
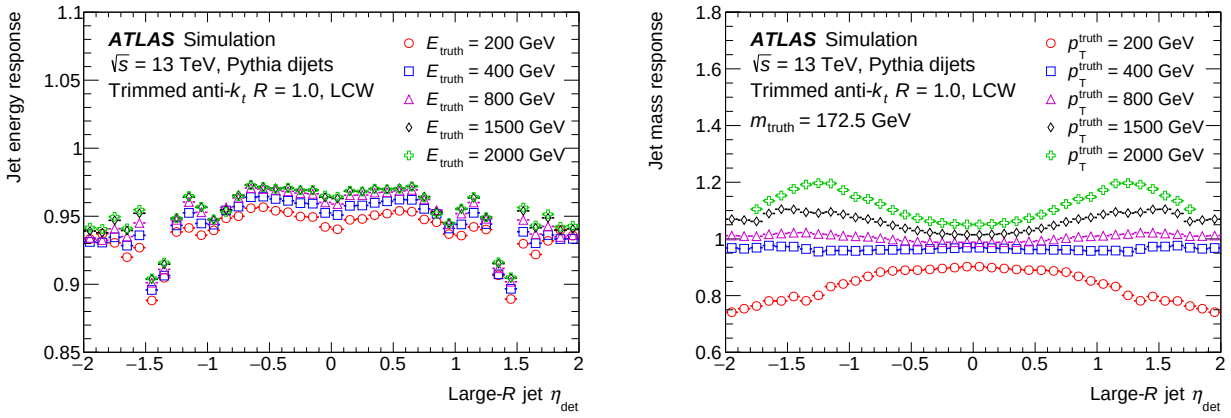


Figure A.2: The response, $E^{\text{reco}}/E^{\text{truth}}$, of PFlow jet is shown as function of jet η for the MC dijet events. The different colored distributions shown in this figure correspond to jet truth energy 30, 60, 110, 400, 1200 GeV. Figure taken from [185].

is close to unity for the jets with p_T between 200 and 800 GeV and becomes much higher or lower outside this range due to splitting or merging clusters.



(a) (b)

Figure A.3: (a) Large-R energy response as a function of detector pseudorapidity, η_{det} for several truth energy values between 200 GeV and 2 TeV. (b) Large-R mass response as a function of η_{det} for truth jet mass set to top quark mass (172.5 GeV). Figure taken from [190]

APPENDIX

B

Additional studies in boosted all-hadronic $t\bar{t}$ resonance search

This Appendix includes some additional studies and plots related to the boosted all-hadronic $t\bar{t}$ resonance search described in chapter 7. The data driven multijet estimation method is discussed in details in section B.1. Additional plots related to the statistical analysis of the boosted all-hadronic $t\bar{t}$ resonance search are shown in section B.3.

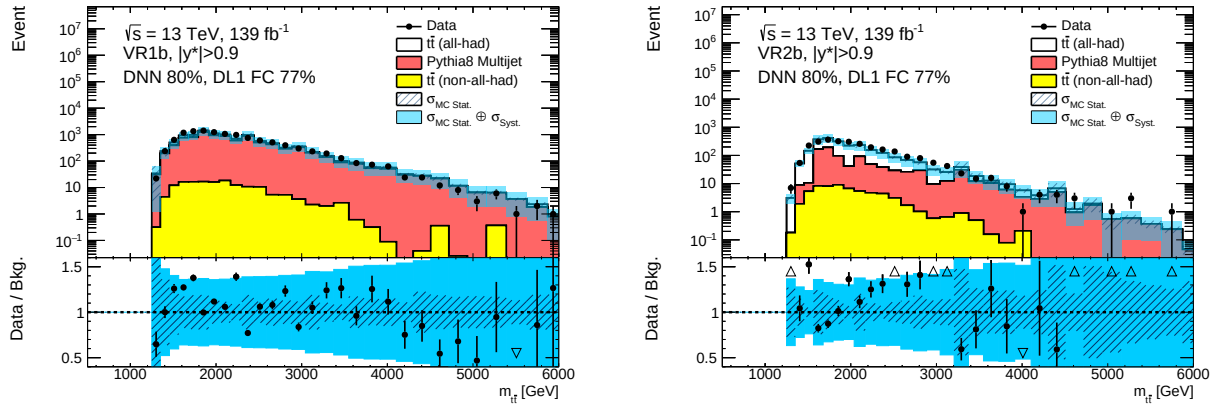
B.1 Data-driven estimate of SM multijet process

It is difficult to simulate multijet processes. The multijet simulation mentioned in subsection 7.3.2, is basically LO dijet simulation with additional radiation. The simulated multijet process does not describe the data well. Figure B.1 shows the comparison between data and simulation (Pythia 8 multijet + $t\bar{t}$) of the $m_{t\bar{t}}$ distribution in VR1b and VR2b. There is a significant difference between simulated background and data. The background only agrees with data within large top-tagging systematic uncertainties, mostly coming from the effect of the jet energy scale variations on the top-tagging scale factor. Also, the low MC statistics of the multijet sample in the low mass region leads to large statistical fluctuations there.

Therefore, to estimate the multijet background a data-driven approach based on extended "ABCD" background estimation [251] is used. In a standard ABCD method, the 2-dimensional phase space, spanned by two independent variables, is divided into four orthogonal regions, A, B, C, and D. It is assumed that the two directions corresponding to features of the events are independent and the contribution of the QCD background is not correlated. So, with that assumption, we can write the number of QCD events in region A in terms of the QCD counts in regions B, C, and D as

$$\frac{N_A^{\text{QCD}}}{N_B^{\text{QCD}}} = \frac{N_C^{\text{QCD}}}{N_D^{\text{QCD}}}.$$

Region A in the example can be though as a signal region and the other three regions are like QCD control regions.



(a) **VR1*b***

(b) **VR2*b***

Figure B.1: $m_{t\bar{t}}$ distributions with the simulated multijet background in the two validation regions: (a) VR1b and (b) VR2b. The multijet background is estimated using Pythia 8 multijet MC samples. Data only agrees with background within the large systematic band.

In this analysis, a total of 16 regions are defined based on top-tagging of the large-R jets and the number of associated b-tagged track jets. Table 7.12 shows the regions defined according to whether the leading and subleading large-R jets pass or fail the top- and/or b-tagging requirements. Previously defined SRs and CR are included in these 16 regions. The SRs defined in subsection 7.7.2 correspond to SR1$b_{tb\not{t}\not{b}}$, SR1$b_{t\not{b}tb}$ and SR2b_{tbtb}. The first (last) two subscripts correspond to the leading (subleading) large-R jet. These two subscripts indicate whether the jet is tagged (j) or not ($\not{j}$) where $j = t$ (b) corresponding to top-tagging (b-tagging). SR1b consists of the sum of the two regions, SR1$b_{tb\not{t}\not{b}}$ and SR1$b_{t\not{b}tb}$. In the SR2b signal region both large-R jets are top- as well as b-tagged. So, SR2b is denoted as SR2b_{tbtb} in the new notation. The previously defined CR0b in subsection 7.7.1 is here defined as the 'template region' TR$_{t\not{b}t\not{b}}$. It is dominated by multijet events with a purity of 97%. Hence TR$_{t\not{b}t\not{b}}$ is used to extract the shape of the multijet distribution. The multijet background contributions expected in the SRi ($i = 1b_{tb\not{t}\not{b}}$, $1b_{t\not{b}tb}$, $2b_{tbtb}$), $N_{\text{exp}}(\text{SR}i)$, is obtained by multiplying the observed events in the TR, $N(\text{TR}_{t\not{b}t\not{b}})$, by scale factors R for either one or both leading top-tagged jets to be b-tagged:

$$N_{\text{exp}}(\text{SR}i) = N(\text{TR}_{t\not{b}t\not{b}}) \times R(\text{TR}_{t\not{b}t\not{b}} \to \text{SR}i). \tag{B.1}$$

Here the N without subscript represents the observed number of events in the region given in the brackets.

SR1b is a combination of two regions, thus two scale factors need to be calculated. The two scale factors are

$$R(\text{TR}_{t\not{b}t\not{b}} \to \text{SR1}b_{t\not{b}tb}) = \kappa_{t_1 b_2} \times \frac{N(\text{A}_{t\not{b}tb})}{N(\text{B}_{t\not{b}t\not{b}})}, \tag{B.2}$$

$$R(\text{TR}_{t\not{b}t\not{b}} \to \text{SR1}b_{tb\not{t}\not{b}}) = \kappa_{t_2 b_1} \times \frac{N(\text{I}_{tb\not{t}\not{b}})}{N(\text{H}_{t\not{b}\not{t}\not{b}})}\,. \tag{B.3}$$

The $\kappa_{t_1 b_2}$ ($\kappa_{t_2 b_1}$) factor accounts for the correlation between the top-tagging of the leading (subleading) large-R jet and the b-tagging of the subleading (leading) large-R jet. The scale factors

from $\mathrm{TR}_{t\not{b}t\not{b}}$ to $\mathrm{SR2}b_{tbtb}$ have two values, R_1 and R_2, corresponding to the two possible paths: $\mathrm{TR}_{t\not{b}t\not{b}} \to \mathrm{SR1}b_{t\not{b}tb} \to \mathrm{SR2}b_{tbtb}$ and $\mathrm{TR}_{t\not{b}t\not{b}} \to \mathrm{SR1}b_{tbt\not{b}} \to \mathrm{SR2}b_{tbtb}$, respectively.

$$R_1(\mathrm{TR}_{t\not{b}t\not{b}} \to \mathrm{SR2}b_{tbtb}) = R(\mathrm{TR}_{t\not{b}t\not{b}} \to \mathrm{SR1}b_{t\not{b}tb}) \times \kappa_{t_1 b_1} \times \frac{N(\mathrm{E}_{\not{t}bt\not{b}})}{N(\mathrm{B}_{\not{t}\not{b}t\not{b}})}, \tag{B.4}$$

$$R_2(\mathrm{TR}_{t\not{b}t\not{b}} \to \mathrm{SR2}b_{tbtb}) = R(\mathrm{TR}_{t\not{b}t\not{b}} \to \mathrm{SR1}b_{tbt\not{b}}) \times \kappa_{t_2 b_2} \times \frac{N(\mathrm{G}_{t\not{b}\not{t}b})}{N(\mathrm{H}_{t\not{b}\not{t}\not{b}})}. \tag{B.5}$$

The two resulting $N_{\mathrm{exp}}(\mathrm{SR2}b_{tbtb})$ values agree within 20%, and the average is used as the final prediction for the region $\mathrm{SR2}b_{tbtb}$. The $\kappa_{t_1 b_2}$ ($\kappa_{t_2 b_1}$) factor is obtained as follows:

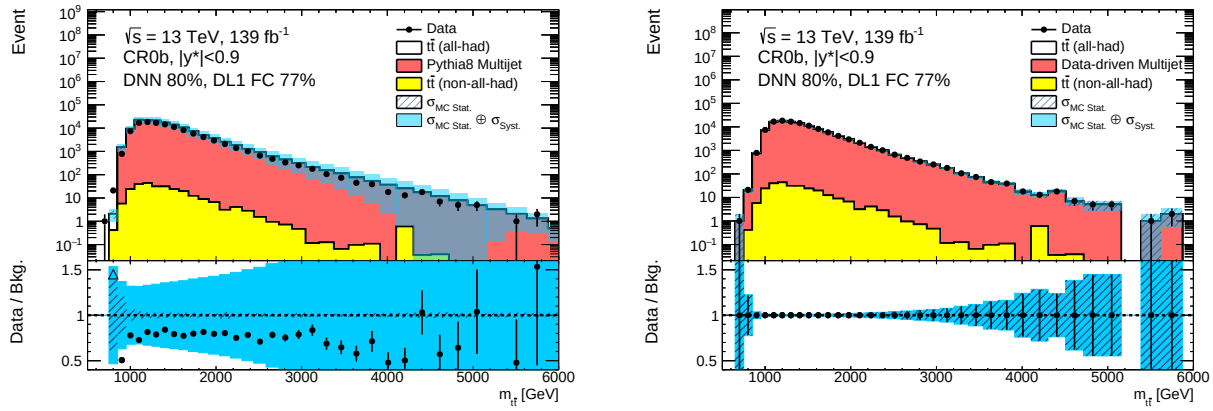
$$\kappa_{t_1 b_2} = \left(\frac{N(\mathrm{G}_{t\not{b}\not{t}b})}{N(\mathrm{H}_{t\not{b}\not{t}\not{b}})}\right) \Big/ \left(\frac{N(\mathrm{C}_{\not{t}\not{b}\not{t}b})}{N(\mathrm{D}_{\not{t}\not{b}\not{t}\not{b}})}\right), \qquad \kappa_{t_2 b_1} = \left(\frac{N(\mathrm{E}_{\not{t}bt\not{b}})}{N(\mathrm{B}_{\not{t}\not{b}t\not{b}})}\right) \Big/ \left(\frac{N(\mathrm{F}_{\not{t}b\not{t}\not{b}})}{N(\mathrm{D}_{\not{t}\not{b}\not{t}\not{b}})}\right). \tag{B.6}$$

Similarly, the $\kappa_{t_1 b_1}$ ($\kappa_{t_2 b_2}$) factor is defined as follows:

$$\kappa_{t_1 b_1} = \left(\frac{N(\mathrm{I}_{tb\not{t}\not{b}})}{N(\mathrm{H}_{t\not{b}\not{t}\not{b}})}\right) \Big/ \left(\frac{N(\mathrm{F}_{\not{t}b\not{t}\not{b}})}{N(\mathrm{D}_{\not{t}\not{b}\not{t}\not{b}})}\right), \qquad \kappa_{t_2 b_2} = \left(\frac{N(\mathrm{A}_{\not{t}\not{b}tb})}{N(\mathrm{B}_{\not{t}\not{b}t\not{b}})}\right) \Big/ \left(\frac{N(\mathrm{C}_{\not{t}\not{b}\not{t}b})}{N(\mathrm{D}_{\not{t}\not{b}\not{t}\not{b}})}\right). \tag{B.7}$$

These calculations in Equation B.1–Equation B.7 are performed bin-by-bin for the $m_{t\bar{t}}$ distributions in all relevant CRs.The values of the $\kappa_{t_1 b_2}$ and $\kappa_{t_2 b_1}$ factors range between 0.94 and 1.0 while those of $\kappa_{t_1 b_1}$ and $\kappa_{t_2 b_2}$ have larger values around 1.6–1.7 in the $m_{t\bar{t}}$ region of 1.4–6.0 TeV. This indicates only very small correlation between the top- and b-tagging of different jets but a relatively strong correlation between the probability of b-tagging and top-tagging a given jet.

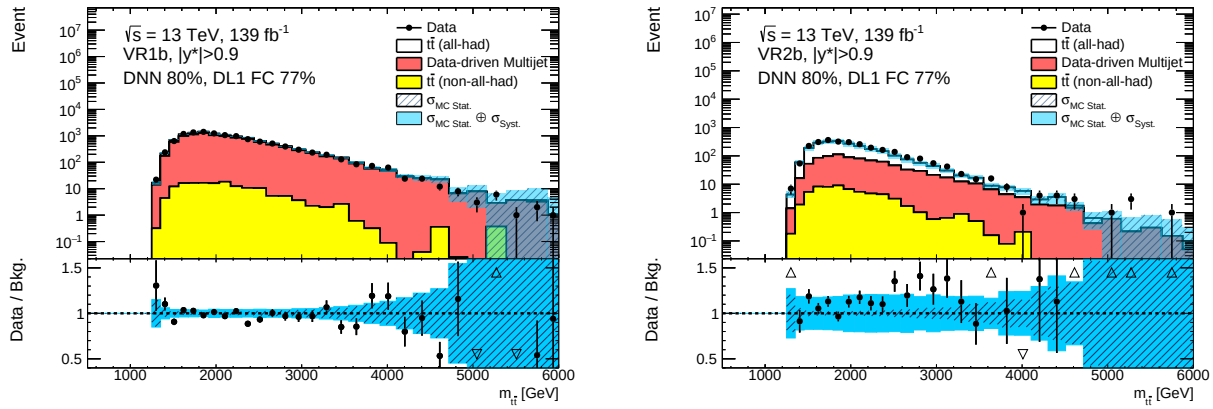
The $m_{t\bar{t}}$ distributions in the multijet control region (CR0b) are shown in Figure B.2. The data-driven multijet distributions are obtained by simply subtracting the $t\bar{t}$ contribution from the data. As a result the background agrees perfectly with data. Similar distributions with the data-driven multijet estimation in the validation regions VR1b and VR2b are shown in Figure B.3. It is seen that the agreement of data and background becomes much better in the validation regions. The agreement is better in VR1b and there is no clear trend observed in the ratio plot. The $m_{t\bar{t}}$ distributions in the signal regions are shown in Figure B.4. It can be seen that the data agrees well with the background.



(a) **CR0*b***: Pythia 8 multijet

(b) **CR0*b***: Data-driven multijet

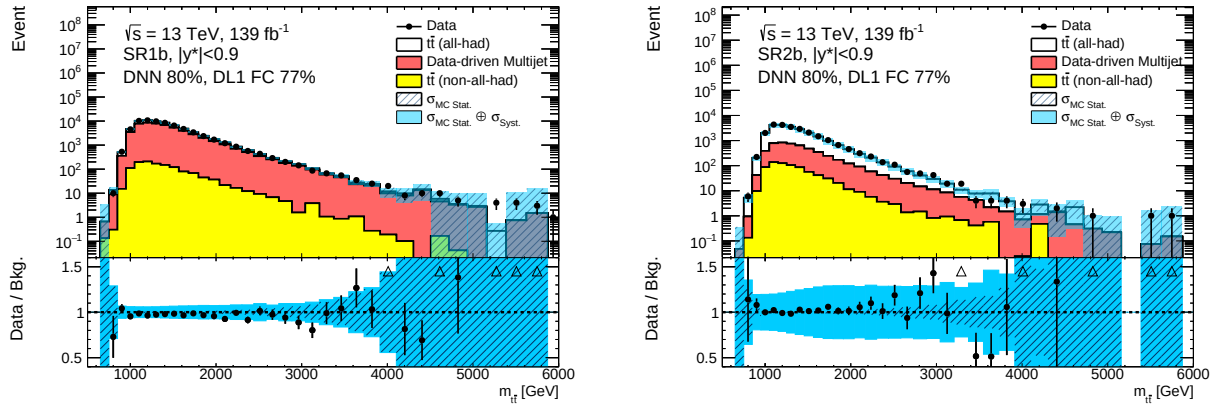
Figure B.2: $m_{t\bar{t}}$ distributions with (a) simulated (Pythia 8) and (b) data-driven multijet estimate are compared in the control region. The multijet estimation is done from the data after subtracting the $t\bar{t}$ contribution.



(a) **VR1*b***

(b) **VR2*b***

Figure B.3: $m_{t\bar{t}}$ distributions with data-driven multijet background in the two validation regions: (a) VR1b and (b) VR2b. The multijet background is estimated using data in the multijet extend ABCD control regions. Data agrees better with the background when using the data-driven multijet estimate.



(a) **SR1*b***

(b) **SR2*b***

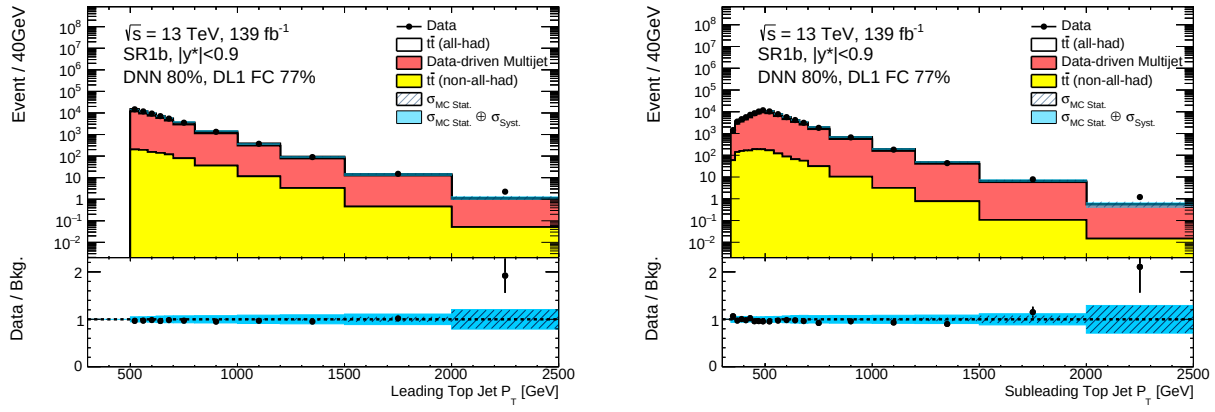
Figure B.4: $m_{t\bar{t}}$ distributions with the data-driven multijet background in the two signal regions: (a) SR1b and (b) SR2b. The multijet background is estimated using data in the extend ABCD control regions.

B.2 Control plots

The distributions of some of the kinematic variables are shown in this section. In these control plots, the background is estimated using MC simulated $t\bar{t}$ and data-driven multijet process. Although the background contribution to the final discriminant, $m_{t\bar{t}}$, is estimated from data directly the plots in this section give us some qualitative understanding of the background estimation (MC $t\bar{t}$+ data-driven Multijet).

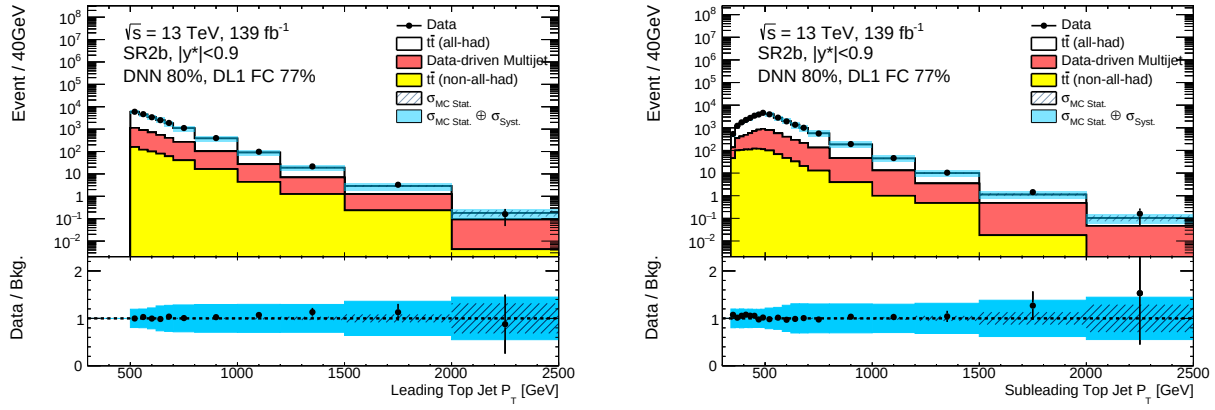
B.2.1 Signal region plots

The p_T distributions of the leading and the subleading large-R jets are well modeled and data agrees with background quite well as shown in Figure B.5 in the signal regions. The large-R jet mass on the other hand is not well modeled in both the signal regions. The modeling is relatively well in the 1b signal region but in the 2b signal region data does not agree with the background good as shown in Figure B.6. The systematic band does not include all the jet calibration systematic uncertainties. Similar behaviour is observer in the jet substructure variable τ_{32} distribution (Figure B.6c and Figure B.6d). The disagreement in the $SR2b$ indicates that it is likely coming from SM $t\bar{t}$ modeling.



(a) Leading large-R jet

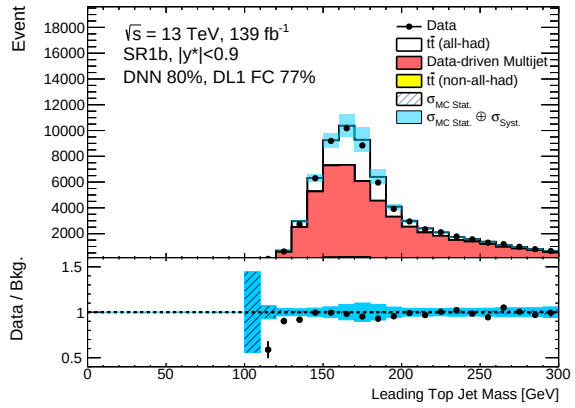
(b) Subleading large-R jet



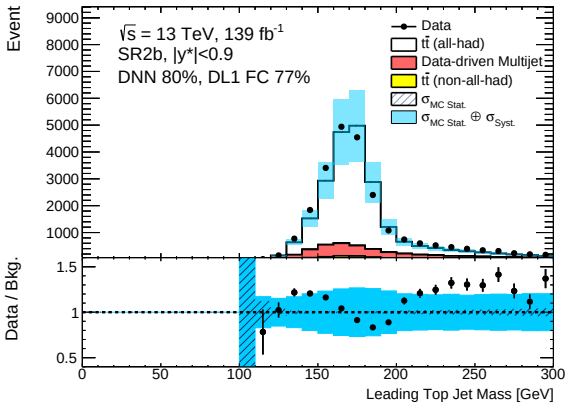
(c) Leading large-R jet

(d) Subleading large-R jet

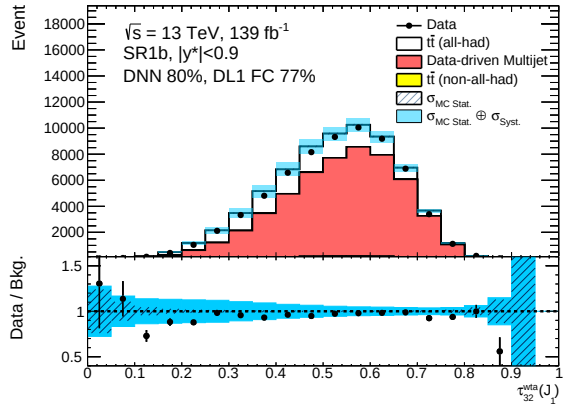
Figure B.5: Leading and subleading large-R jet p_T distribution in SR1b (top row) and SR2b (bottom row). All the jet uncertainties are not included in the systematic band. The small disagreement in SR1b is most likely coming from $t\bar{t}$ background.



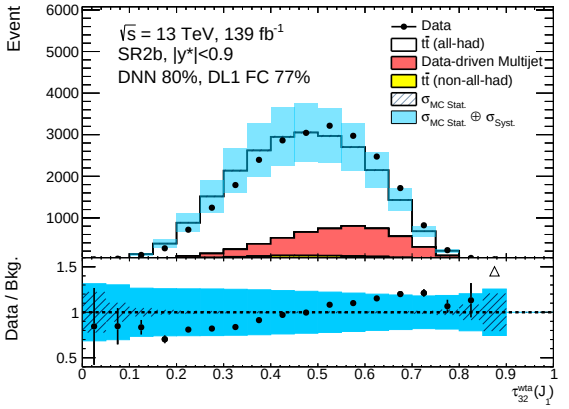
(a) **SR1*b***



(b) **SR2*b***



(c) **SR1*b***



(d) **SR2*b***

Figure B.6: Distributions of the leading large-R jet (a, b) mass and (c, d) τ_{32} in (a, c) SR1b and (b, d) SR2b. The distributions are not well modeled in SR2b due to $t\bar{t}$ mis-modeling. All the jet uncertainties are not included in the systematic band.

B.3 All-hadronic stat

This section contains additional plots related to the boosted all-hadronic $t\bar{t}$ resonance search. The fit results on the Asimov data are not included in chapter 7 and those are presented here. In addition, the section contains some of extra plots related related to b-only and $s+b$ fit using data.

B.3.1 b-only profile likelihood fit: Asimov

The Asimov data distribution is the same as the background distribution, but the bin error are adjusted according to Poisson error. In this is section the results of the background-only fit are discussed. The b-only fit is relatively simple and the likelihood only includes the four background uncertainties. Figure B.7 shows the best-fit NP pulls for the four background NPs. The fit parameter uncertainties get constrained up to $\sim 50-60\%$ in both signal regions.

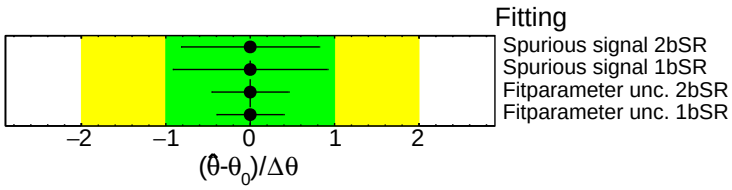


Figure B.7: The best-fit values of the NPs in the background-only likelihood fit on the Asimov data. The fit parameter uncertainties can be constrained to $\sim 50-60\%$ by the fit.

The correlation matrix is also calculated with the background systematics. Figure B.8 shows the NP correlations greater than 0.1. As a result, only two systematics out of four are shown in each signal region.

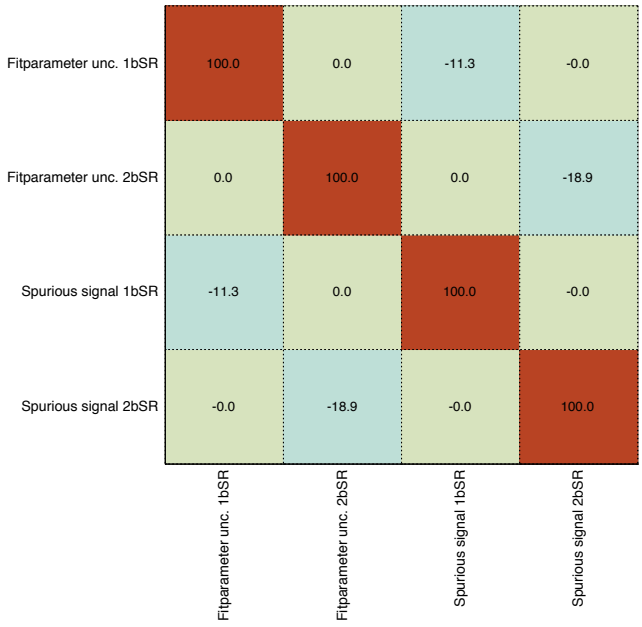


Figure B.8: Post-fit (ASIMOV) correlation matrix of the four background systematic NPs

B.3.2 s+b profile likelihood fit: Asimov

In this section the signal + background fit results are discussed where the 4 TeV Z'_{TC2} signal model is used. Since we are using Asimov data again, the best-fit value of the signal strength is expected to be zero. This time, all the signal systematics along with the background systematic uncertainties are included in the likelihood. None of the uncertainties are pruned, and they are symmetrized before including in the likelihood. The best-fit signal strength value can be seen in Figure B.9. In this picture, the error bars show the combined effect of the statistical and systematic uncertainties. Figure B.10 shows the best-fit signal strength value from a similar likelihood fit where the systematic uncertainties are not included.

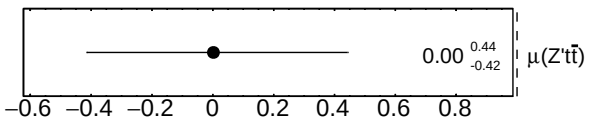


Figure B.9: The best-fit signal strength value where statistical and all the systematics uncertainties are included in the likelihood. The error bar shows the combination of the statistical and systematic uncertainties.

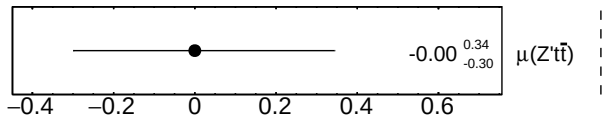


Figure B.10: The best-fit signal strength value where the systematic uncertainties are not included in the likelihood. The error bars represent only the statistical uncertainties.

The post-fit values of all other nuisance parameters are shown in Figure B.11. Since we are using Asimov data the signal uncertainties are not expected to affect the fit much. As a result, Figure B.11 shows that none of the signal NPs can be constrained. On the other hand, some background systematics can be constrained significantly. The fit-parameter uncertainties in both the signal regions can be constrained more than 50%. The spurious signal uncertainty in the 1b-tag signal region can be constrained approximately 10%.

The correlation matrix is also calculated for all the NPs and the signal strength. Figure B.12 shows the correlation matrix where only correlations above 0.1 are included. Most of the signal uncertainties have almost zero correlation. Hence, only the background uncertainties appear in this matrix.

To understand the effect of the NPs on the likelihood fit, ranking of the all NPs are studies based on their impact on the estimated (best-fit value) signal strength. Figure B.13 the top twenty systematics. The background systematics have effect on the fit, as a result they are ranked at the top. The impact of the four background systematics goes up to 15%. All the signal uncertainties have almost zero impact, as expected in the background-only Asimov data.

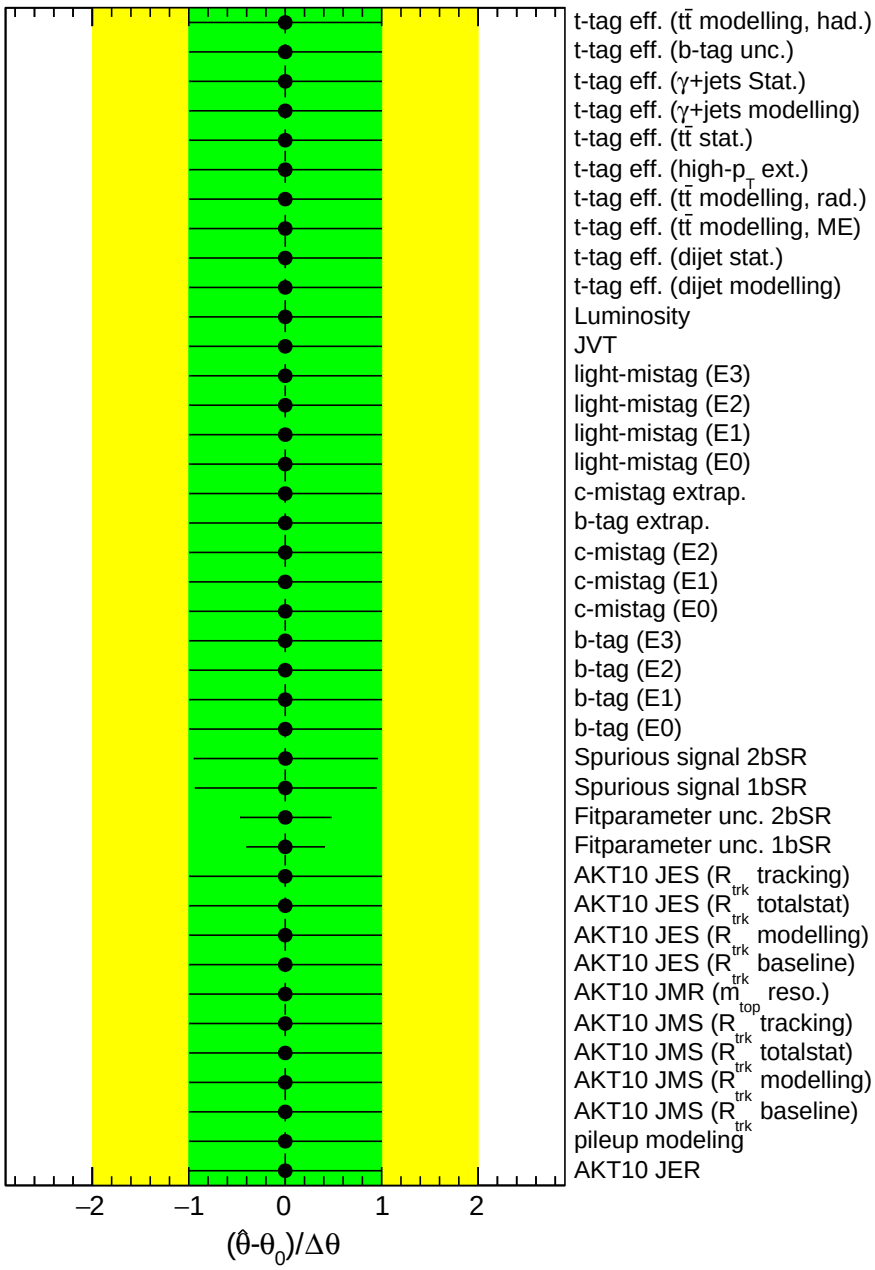


Figure B.11: Post-fit NP pulls where 4 TeV Z'_{TC2} signal model is used. Signal NPs cannot be constrained in absence of signal. Only the four background NPs can be constrained by the fit.

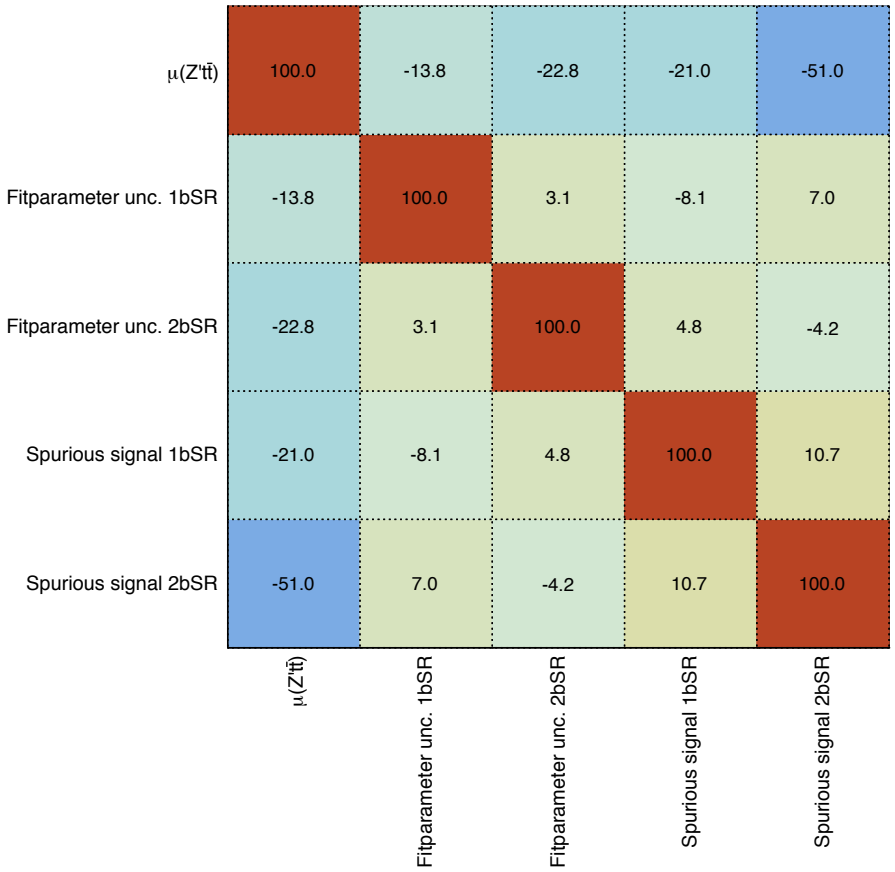


Figure B.12: Post-fit correlation matrix of the systematic NPs and signal strength for Z' mass 4 TeV. Most of the correlations are zero. Only the background uncertainties have strong correlation with the POI.

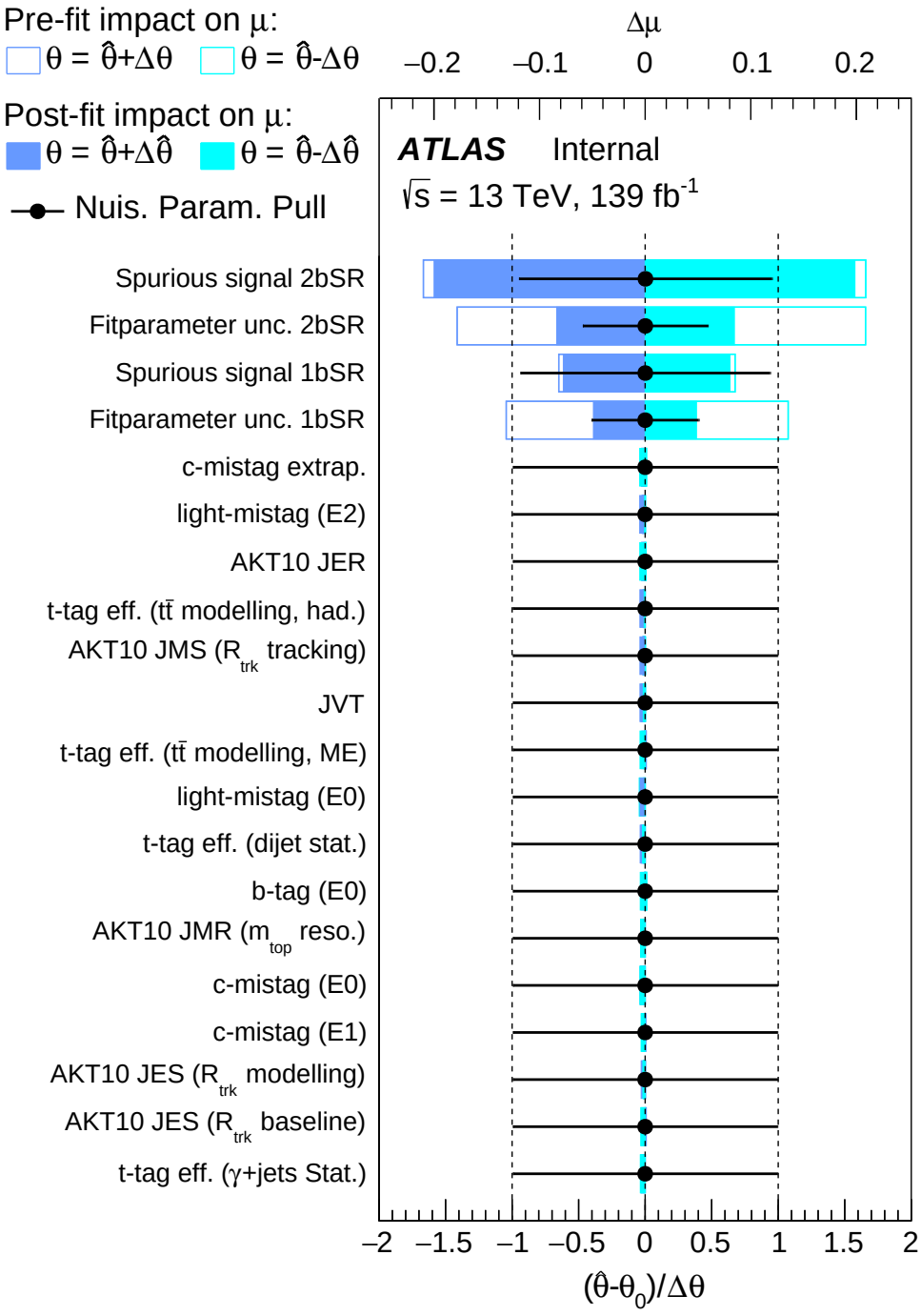


Figure B.13: Systematics ranking plot based on there impact on the signal strength for the 4 TeV Z'_{TC2} signal model. Only the four background uncertainties has significant impact. Since, the POI value is very close to zero and the signal systematics have very negligible impact.

B.3.3 b-only profile likelihood fit: Data

This section shows the NP correlation matrix obtained from the background-fit using data as discussed in subsection 7.12.1. Figure B.14 shows the correlation matrix where only correlations above 0.1 are included. There is some correlation between the fit parameter uncertainties and the spurious signal signal uncertainties.

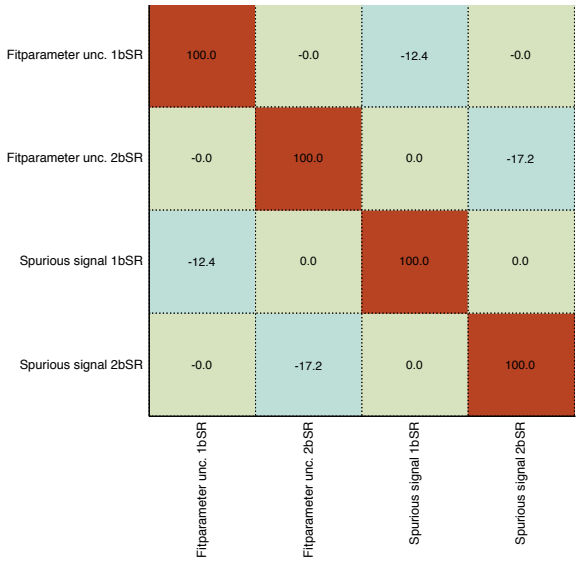


Figure B.14: Post-fit (b-only) correlation matrix of the background systematic NPs observed in data.

B.3.4 s+b-only profile likelihood fit: Data

This section includes the extra results of the signal + background fit discussed in subsection 7.12.1 where the 4 TeV Z'_{TC2} signal is used. The best-fit signal strength value is $-0.24^{+0.47}_{-0.33}$ as shown in Figure B.15. The error bars show the combined effect of the statistical and systematic uncertainties. Figure B.16 shows the best-fit signal strength value from a similar likelihood fit where the systematic uncertainties are not included. The best-fit value of signal strength is $-0.18^{+0.32}_{-0.20}$ which is close to the value obtained from the previous setup.

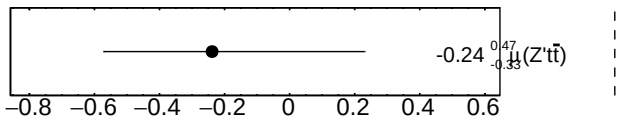


Figure B.15: The best-fit signal strength value observed on data in presence of statistical and all the systematic uncertainties. The error bars are combination of statistical and systematic uncertainties.

The post-fit nuisance parameter pulls are shown in Figure B.17. The fit-parameter uncertainties are constrained more than 50% by the data. In addition both of them are slightly pulled. The

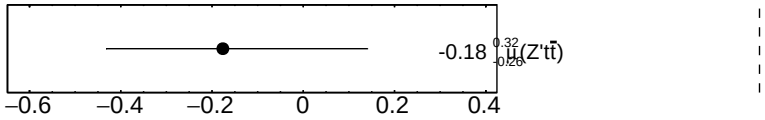


Figure B.16: The best-fit signal strength value observed on data in presence of only the statistical uncertainty. The error bars represent the statistical uncertainties only.

spurious signal uncertainty in 1b-tag signal region are constrained by 10% and slightly pulled. JES modelling uncertainty is the dominant signal systematics and as a result gets constrained as well as pulled by the maximum amount. It is pulled by $0.6 - 0.7\sigma$, other JES components are comparatively less affected in the fit.

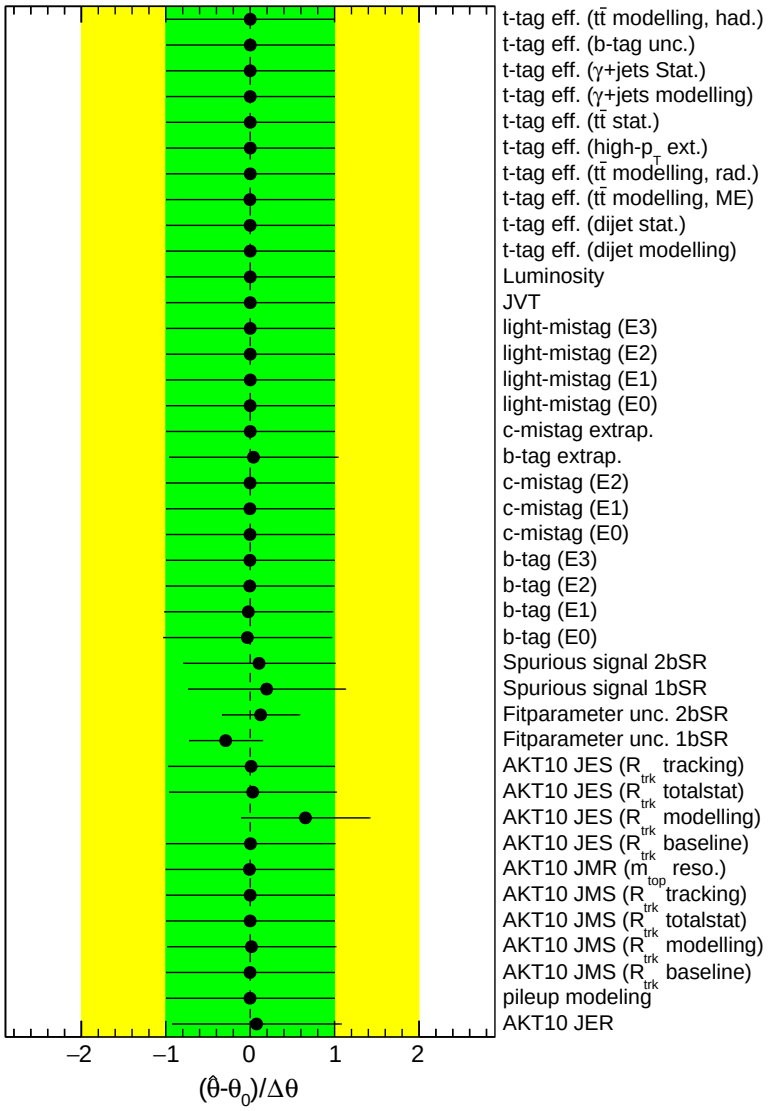


Figure B.17: Post-fit ($s+b$ fit) NP pulls where 4 TeV Z'_{TC2} signal model is used. Fit parameter uncertainties are contained more than 50% by the data. The JES modelling uncertainty gets constrained and pulled by the data.

B.4 Model-dependent search

This section contains a table (Table B.1 with all the local p-values of the profile likelihood-based significant scan for different signal mass point. Z'_{TC2} signals with $\Gamma = 1.2\%$ are used for this scan.

Table B.1: Summary of the significance scan of Z'_{TC2} signals with relative width $\Gamma = 1.2\%$. Local significance values are within 1.6σ in the search range. The minimum local p_0-value is found to be 0.06 (1.56σ) at Z'_{TC2} mass of 1.875 TeV.

Z' pole mass (TeV)	local p_0-value	local significance (σ)
1.75	0.112	1.2
1.875	0.060	1.6
2.0	0.788	-0.8
2.125	0.809	-0.9
2.25	0.517	-0.0
2.375	0.554	-0.1
2.5	0.637	-0.4
2.625	0.636	-0.3
2.75	0.583	-0.2
2.875	0.226	0.8
3.0	0.143	1.1
3.25	0.136	1.1
3.5	0.719	-0.6
3.75	0.735	-0.6
4.0	0.731	-0.6
4.25	0.800	-0.8
4.5	0.594	-0.2
4.75	0.429	0.2
5.0	0.520	-0.0

APPENDIX

C

Additional studies in semi-leptonic $t\bar{t}$ resonance search

In this Appendix, some additional plots related to the semi-leptonic $t\bar{t}$ resonance search described in chapter 8 are included. The post-fit nuisance parameter fits are shown in this Appendix.

C.1 Statistical analysis

In this section, post-fit the nuisance parameter plots are shown.

C.1.1 b-only profile likelihood Asimov fit

In the background-only fit, the signal strength is not included in the likelihood. All the other systematic uncertainties are included in the likelihood. The best-fit values of the all the nuisance parameters (NP) are shown in Figure C.1. Only some of the large jet related systematics are constrained in the fit.

C.1.2 s+b profile likelihood Asimov fit

The best-fit values of all the NPs are shown in Figure C.2. In the Asimov fit, they all take best-fit value of 0 and most of the NPs do not get contained in the s+b fit. Some of the large systematics are constrained by almost 50%. The NP constraints are very similar to that of the b-only fit.

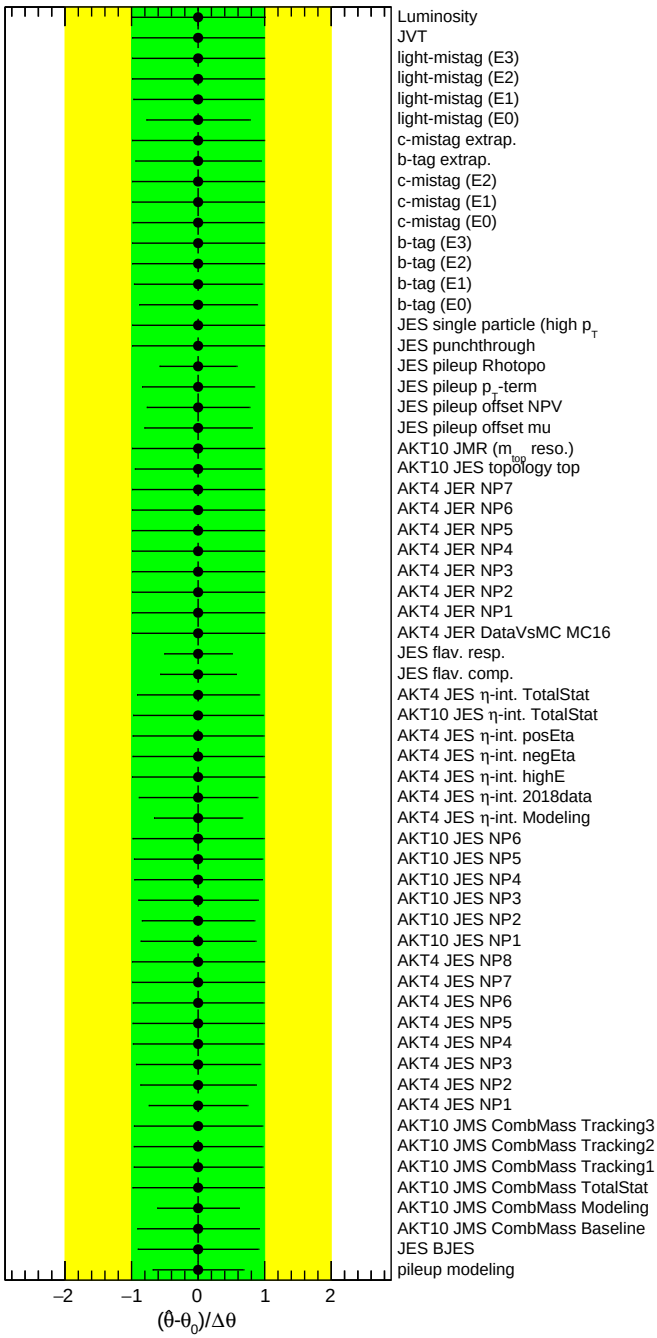


Figure C.1: The best-fit values of the NPs in the background-only likelihood fit. The best-fit values are all zero in the Asimov fit, and the uncertainty bars give a measure of constraints in the fit. Most of the NPs are very minimally constrained, since the absolute values of the error bars are close to 1. Some of the large jet systematics are constrained by roughly 50%.

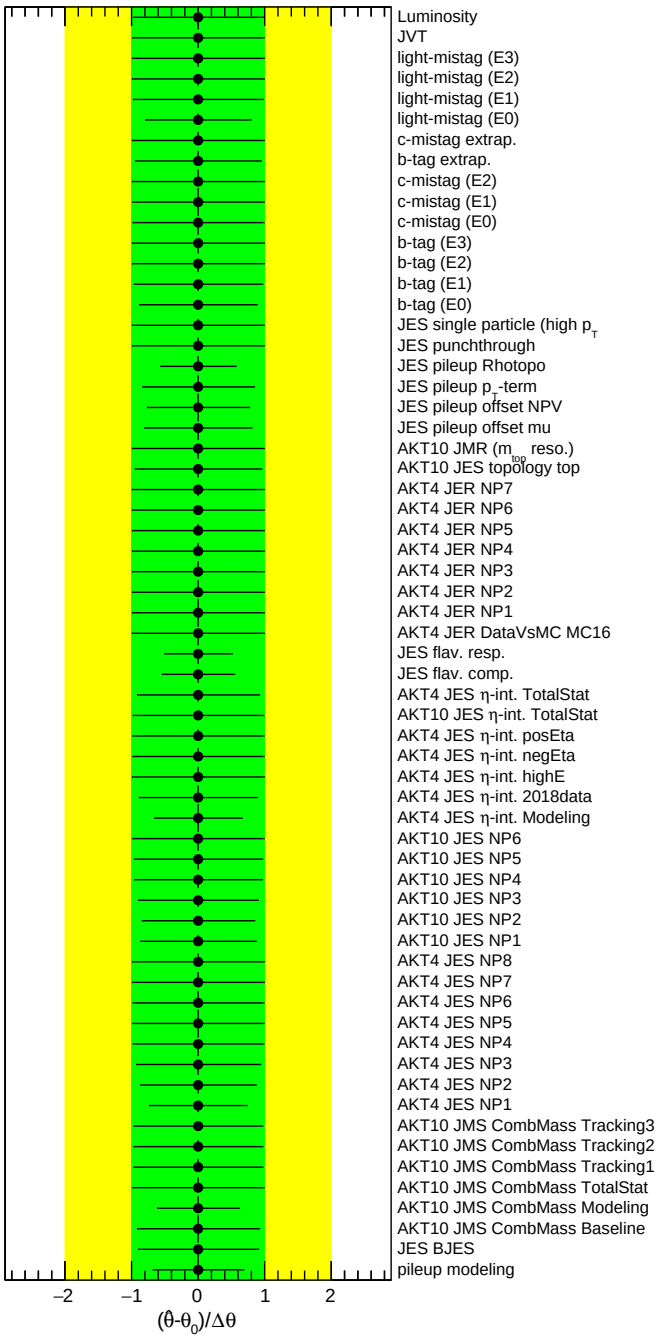


Figure C.2: The best-fit values of the NPs in the signal+background likelihood fit where 4 TeV Z'_{TC2} signal model is used. Here Z'_{TC2} with mass 4 TeV is used as the signal model. The best-fit values are all zero in the Asimov fit, and the uncertainty bars give a measure of constraints in the fit. Most of the NPs are very minimally constrained, since the absolute values of the error bars are close to 1. Some of the large jet systematics are constrained by roughly 50%.