

https://dspace.org/

The use of machine learning in search for new physics at the ATLAS and applications to model COVID-19

Mathaha, Thuso Stephen

2024
https://hdl.handle.net/10539/37413

Downloaded from a DSpace repository. https://dspace.org/

The use of Machine Learning in search for new physics at the ATLAS and applications to model COVID-19

Thuso Stephen Mathaha

Supervisor:

Professor Bruce Mellado



In partial fulfillment of the requirements for the degree of Master of Science

in the School of Physics and Institute for Collider Particle Physics

University of the Witwatersrand, Johannesburg

12 June 2023

Declaration

I, Thuso Stephen Mathaha, hereby certify that the work presented in this dissertation is my own original work, and that any sources used in the research have been properly acknowledged and cited in accordance with the guidelines of the Faculty of Science at the University of the Witwatersrand, Johannesburg.



Thuso Stephen Mathaha

12 June 2023

Acknowledgements

I would like to express my deepest gratitude to my supervisor, Prof. Bruce Mellado, for his guidance, support, and encouragement throughout my research journey. His valuable insights, constructive feedback and unwavering commitment to excellence has been instrumental in shaping the direction of my work and helping me achieve my academic goals.

I am also indebted to the institute members in the ICPP who have provided me with the necessary knowledge, resources, and opportunities to pursue my research interests. I would like to thank Dr. Mukesh Kumar, Dr. Xifeng Ruan, Dr. Salah-Eddine Dahbi, Dr. Edward Nkadimeng and Joshua Choma for their insightful discussions, helpful comments, and suggestions that have greatly enhanced the quality of my thesis.

I am also grateful to my friends and family members who have provided me with emotional support and encouragement during the ups and downs of the thesis writing process. In particular, I would like to thank Mr and Mrs Mathaha, Tshiliso Mathaha, and Nkosiphendule Njara for their unwavering support and encouragement throughout my academic journey.

Last but not least, I would like to express my sincere appreciation to all the participants who generously gave their time and expertise to take part

in my study. Without their contributions, this research would not have been possible. I would like to express my gratitude to the National Research Foundation (NRF) for their financial support.

List of Publications

Journals:

1. Alavinejad, M., Mellado, B., Asgary, A., Mbada, M., **Mathaha, T.**, Lieberman, B., ... & Kong, J. D. (2022). Management of hospital beds and ventilators in the Gauteng province, South Africa, during the COVID-19 pandemic. PLOS Global Public Health, 2(11), e0001113.

2. **Mathaha, T.**, Mafu, M., Mabikwa, O. V., Ndenda, J., Hillhouse, G., & Mellado, B. (2022). Leveraging artificial intelligence to optimize COVID-19 robust spread and vaccination roll-out strategies in Southern Africa. Frontiers in Artificial Intelligence, 5.

3. Tan, Y. R., Agrawal, A., Matsoso, M. P., Katz, R., Davis, S. L., Winkler, A. S., ..., **Mathaha, T.** ,...& Yap, P. (2022). A call for citizen science in pandemic preparedness and response: beyond data collection. BMJ global health, 7(6), e009389.

4. Alavinejad, M., Mellado, B., Asgary, A., Mbada, M., **Mathaha, T.**, Lieberman, B., ... & Kong, J. D. (2022). Management of Healthcare Resources in the Gauteng Province, South Africa, During the COVID-19 Pandemic. South Africa, During the COVID-19 Pandemic (March 3, 2022).

5. Mellado, B., Wu, J., Kong, J. D., Bragazzi, N. L., Asgary, A., Kawonga, M., ..., **Mathaha, T.** ,... & Orbinski, J. (2021). Leveraging artificial

intelligence and big data to optimize COVID-19 clinical public health and vaccination roll-out strategies in Africa. International Journal of Environmental Research and Public Health, 18(15), 7890.

6. Stevenson, F., Hayasi, K., Bragazzi, N. L., Kong, J. D., Asgary, A., Lieberman, B., ..., **Mathaha, T.** ... & Wu, J. (2021). Development of an early alert system for an additional wave of covid-19 cases using a recurrent neural network with long short-term memory. International Journal of Environmental Research and Public Health, 18(14), 7376

Conferences:

1. **Thuso Mathaha**, Abhaya Kumar Swain, Bruce Mellado, A study of top quark pair production in association with a high energy photon at the LHC, in The Proceedings of SAIP2022, the 66th Annual Conference of the South African Institute of Physics, edited by Prof. Aletta Prinsloo, (UJ), pp. 200 - 205. ISBN: 978-0-6397-4426-1. Available online at http://events.saip.org.za,

2. **Thuso Mathaha**, Abhaya Kumar Swain, Mukesh Kumar, Xifeng Ruan, Bruce Mellado, Understanding two same-sign and three leptons with b-jets in four top quark events at the LHC, in The Proceedings of SAIP2021, the 65th Annual Conference of the South African Institute of Physics, edited by Prof. Aletta Prinsloo, (UJ), pp. 96 - 101. ISBN: 978-0-620-97693-0. Available online at http://events.saip.org.za

Abstract

In this thesis, the production of a pair of top quarks in association with a heavy pseudo-scalar (A) is examined. The heavy pseudoscalar subsequently decays into another pair of top quarks, resulting in a final state of four top quarks ($t\bar{t}A \rightarrow t\bar{t}t\bar{t}$). The ATLAS public paper Ref. [1] provided the analytical framework for this study, which aimed to investigate the four top quarks production in the multilepton final state. The study focuses on final states with two same-sign leptons of different flavours (e.g. $e^{\pm}, \mu^{\pm}$) or at most three isolated leptons (muons and electrons) without any charge requirement, as well as jets. The analysis employs a multivariate discriminant that uses twelve discriminating kinematic variables to separate the signal from the background in an effort to understand the differences between the SM and BSM production mechanisms of four top quarks. The machine learning techniques deployed for the multivariate algorithm were transferred to tackle the COVID-19 pandemic. The COVID-19 pandemic has caused significant health, social, and economic damage worldwide, with many developed countries vaccinating their citizens while African nations relied on clinical public health (CPH) strategies. Recent studies in Botswana and South Africa found age, gender, hypertension and diabetes were significant factors in disease severity,

with the elderly population aged $\geq$ 60 years and those with major COVID-19 comorbidities recommended for vaccination. AI was also used to optimize vaccination roll-out strategies, targeting population groups needing insufficient vaccines.

Contents

Declaration i

Acknowledgements ii

List of Publications iv

Abstract vi

List of Figures xii

List of Tables xix

List of Abbreviations xx

List of Symbols xxiii

1 Introduction 1

1.1 Overview 1

1.2 The Standard Model of Particle Physics 4

1.2.1 The elementary particles 4

1.2.2 Production of Higgs boson 9

1.2.3 Decay modes of the Higgs boson 10

1.3 The multilepton anomalies at the LHC 11

1.3.1 Madala hypothesis and the multi-lepton final states . . . 14
1.3.2 The model . . . 15

2 The Large Hadron Collider and the ATLAS Detector 19
2.1 The Large Hadron Collider . . . 19
2.2 The ATLAS Detector . . . 21
2.2.1 The Inner Detector . . . 24
2.2.2 The ATLAS Calorimeter System . . . 26
2.2.3 The Muon Spectrometer . . . 28
2.2.4 Forward Detectors . . . 31
2.2.5 Magnets . . . 32
2.2.6 Trigger and Data Acquisition system of the ATLAS detector . . . 34

3 Monte Carlo samples 37
3.1 Data Set . . . 37
3.2 Monte Carlo simulated samples . . . 39

4 Event Reconstruction and Selection 43
4.1 Leptons . . . 43
4.1.1 Electrons . . . 44
4.1.2 Muons . . . 44
4.2 Jets . . . 45
4.3 Missing Transverse Momentum . . . 46
4.4 Preselection . . . 46
4.4.1 Signal region . . . 47

5 The production of four top quarks in SM and BSM **49**

5.1 SM four top quark 49

5.2 The multilepton channel 53

5.3 BSM four top quark 55

5.3.1 Heavy pseudo-scalar in the mass range 400-600 GeV . . . 56

5.4 Discriminatory variables 59

6 Machine Learning analysis **69**

6.1 Overview 69

6.2 Supervised Learning 71

6.3 Applications of machine learning in particle physics and COVID-19 72

6.4 Deep Neural Networks 75

6.4.1 Exploration of Deep Neural Network Architecture 76

6.4.2 Response distribution of DNN 78

7 Conclusions **84**

Appendices **86**

A Applications of ML techniques to model COVID-19 **87**

A.1 Introduction 87

A.2 Covid-19 Data 88

A.3 Leveraging ML techniques used in physics to model COVID-19 vaccine strategies 92

A.3.1 DNN classification results 93

A.4 Transferring COVID-19 vaccines strategies to Botswana 100

A.4.1 COVID-19 in Botswana 100
A.4.2 Age structures between South Africa (Gauteng) and Botswana 101
A.4.3 Comorbidities in Botswana 102
A.4.4 Comparative study between Botswana and South Africa 104
A.4.5 Mortality rates between South Africa and Botswana . . . 105

Bibliography **109**

List of Figures

1.1 The Standard Model categorizes fermions into two groups: leptons and quarks. These subatomic particles are further divided into three distinct generations, each made up of one charged lepton, two quarks and one neutral lepton, along with their corresponding antiparticles. Although fermions within the same generation possess similar characteristics, they possess varying masses. Additionally, the electric charge of these particles is expressed as fractions of the proton charge [18]. 6

1.2 The potential of Higgs as a function of ϕ in Re(ϕ) - Im(ϕ) plane for $\mu^2 < 0$ [21]. 9

1.3 Feyman diagrams at leading order showing the production of Higgs mechanism: (a) gluon-gluon fusion, (b) vector boson fusion, (c) Higgs strahlung, (d) top fusion [22]. 10

1.4 Feynman diagrams contributing to the dominant (a) production mode of H (ggF) and (b) decay mode ($H \rightarrow Sh$). The coupling parameters β_g and λ_{HhS} are defined in equation 1.9 and equation 1.15 [16]. 12

1.5 The decay modes of Higgs boson between 100 GeV and 200 GeV [27]. 13

1.6 Invariant mass distribution of ATLAS data (black dots) and SM background (red histogram) compared with the BSM signal (blue histogram) in events with two oppositely charged leptons and at least 1 b-tagged jet are required [16]. 17
1.7 The profile likelihood ratios for each of the individual fit considered here [17]. 18

2.1 The LHC particle accelerator. Protons are accelerated by multiple accelerators prior to injection into the LHC [31]. 20
2.2 The layout of the ATLAS detector [32]. 22
2.3 The ATLAS inner detector together with its sub-detectors [43]. . 26
2.4 ATLAS calorimeter system [44]. 27
2.5 Muon spectrometer showing its components [49]. 30
2.6 A schematic diagram showing the positions of the forward detectors at the ATLAS detectors. 32
2.7 The ATLAS magnet system with the central solenoid and the three toroids [53]. 33
2.8 An overview of the trigger system used by the ATLAS detector [55]. 35

4.1 Distrubtions of the number of jets after applying preselection requirements. The distribution on the left (right) is for $2SS\ell$ (3ℓ). 48

5.1 The Feynman diagram for the production of four top quarks at the leading order in QCD. 50
5.2 The decay of the top quark hadronically and leptonically [89]. . 51

5.3 The Feynman diagram for the production of four top quarks in association with heavy-pseudoscalar (A) decaying to top quark pair. 55

5.4 The plot of the branching ratios ($A \to xy$) against the mass of the heavy pseudoscalar m_A. it is shown that the heavy pseudoscalar A decays predominantly to ZH if its mass is above 600 GeV and when it is below 600 GeV $t\bar{t}$ predominates [90]. . . 56

5.5 Distributions of the kinematic variables in the $2SS\ell$ channel at mass of the pseudo-scalar $m_A = 400$ GeV after preselection. We represent the signal and background distributions in blue and orange-colored histograms. 61

5.6 Distributions of the kinematic variables in the 3ℓ channel at mass of the pseudo-scalar $m_A = 400$ GeV after preselection. We represent the signal and background distributions in blue and orange-colored histograms. 62

5.7 Distributions of the kinematic variables in the $2SS\ell$ channel at mass of the pseudo-scalar $m_A = 500$ GeV after preselection. We represent the signal and background distributions in blue and orange-colored histograms. 63

5.8 Distributions of the kinematic variables in the 3ℓ channel at mass of the pseudo-scalar $m_A = 500$ GeV after preselection. We represent the signal and background distributions in blue and orange-colored histograms. 64

5.9 Distributions of the kinematic variables in the $2SS\ell$ channel at mass of the pseudo-scalar $m_A = 600\,\text{GeV}$ after preselection. We represent the signal and background distributions in blue and orange-colored histograms. 65

5.10 Distributions of the kinematic variables in the 3ℓ channel at mass of the pseudo-scalar $m_A = 600\,\text{GeV}$ after preselection. We represent the signal and background distributions in blue and orange-colored histograms. 66

5.11 Distributions of the kinematic variables for $t\bar{t}A \to t\bar{t}ZH$ in the 2SSℓ channel at mass of the pseudo-scalar $m_A = 500\,\text{GeV}$ after preselection. ZH can further decay into either $bb\tau\tau$ (orange histogram), $WWbb$ (green histogram), $WW\tau\tau$ (red histogram) and $ZZbb$ (purple histogram). 67

5.12 Distributions of the kinematic variables for $t\bar{t}A \to t\bar{t}ZH$ in the 3ℓ channel at mass of the pseudo-scalar $m_A = 500\,\text{GeV}$ after preselection. ZH can further decay into either $bb\tau\tau$ (orange histogram), $WWbb$ (green histogram), $WW\tau\tau$ (red histogram) and $ZZbb$ (purple histogram). 68

6.1 The DNN output distributions for the three mass ranges for $2SS\ell$ channel. The top is for m_A=400 GeV, the middle is for m_A=500 GeV and the bottom is for m_A=600 GeV. 81

6.2 The DNN output distributions for the three mass ranges for 3ℓ channel. The top is for m_A=400 GeV, the middle is for m_A=500 GeV and the bottom is for m_A=600 GeV. 82

6.3 The ROC curves obtained from the DNN models for signal and background. The graphs correspond to $2SS\ell$ (top) and 3ℓ (bottom) channels. 83

A.1 The distribution plots of Age, Gender (sex), Ethnic group and the total sum of pre-existing comorbidities upon arrival at the hospital ward. 90

A.2 The distributions of comorbidities that are partof input variables are used to train the DNN. The blue histogram represents severe category, while the orange histogram represents general ward category. The peak at zero denotes patients who do not have a comorbidity, while the peak at one denotes patients with comorbidity. 91

A.3 The response distributions of the DNN after training (top left) and testing (top right). The bottom distributions show the comparison between traning and testing data for general ward (bottom left) and severe (bottom right) class. 94

A.4 The DNN response distributions for prediction on the data sample that represents the population at large (Gauteng population). The distribution on the left shows training data and the one on the right shows testing data. 95

A.5 The five classes distributions for Age. These distributions illustrate the level of severity in 20% intervals. 96

A.6 The five classes distributions for gender. These distributions illustrate the level of severity in 20% intervals. 97

A.7 The five classes distributions for Diabetes. These distributions illustrate the level of severity in 20% intervals. 98
A.8 The five classes distributions for hypertension. These distributions illustrate the level of severity in 20% intervals. 99
A.9 A ROC curve showing the possible outcome of illness reduction when using this model to roll-out vaccines. 99
A.10 Age structures of the populations of Botswana 102
A.11 Age structures of the population's of Gauteng (South Africa) . . 102
A.12 Age structures of the population's ratio of Botswana and Gauteng - South Africa 103
A.13 Crude prevalence of raised blood pressure for males and female trends in Botswana and South Africa from 1975 to 2015. 104
A.14 Crude prevalence of raised blood pressure for male and female ratios between Botswana and South Africa from 1975 to 2015. . 104
A.15 The distributions of the input variables are used to train the DNN. The blue histogram represents Botswana data, while the orange histogram represents South African data. For histograms related to comorbidities, the peak at zero denotes patients who do not have a comorbidity, while the peak at one denotes patients with comorbidity. 106
A.16 The DNN output distributions for Mortality behind COVID-19. The distribution on the left represents the mortality of Botswana, and the one on the right represents the mortality of Gauteng (South Africa). 107

A.17 Five classes were obtained from DNN output distributions. The distributions include Age, Diabetes and Hypertension. The first (last) row is the most (least) severe target group which is Class 1 (Class 5). 108

List of Tables

1.1 A summary of the SM+BSM fit results for each measurement considered in this article, as well as the combined result. 16

5.1 Summary of the $t\bar{t}t\bar{t}$ possible finale states. 52

5.2 Excesses in the search for (pseudo-)scalars in the 400-600 GeV range are summarized. The experiment's local significances are given. When available, global significance is shown in round brackets. The mass given in brackets for the final states of $ZH \to \ell\ell b\bar{b}$ corresponds to H. 58

List of Abbreviations

ALFA	**A**bsolute **L**uminosity **F**or **A**TLAS
ALICE	**A** **L**arge **I**on **C**ollider **E**xperiment
ATLAS	**A** **T**oroidal **LHC** Apparatu**S**
CERN	European Organization for Nuclear Research
CMS	**C**ompact **M**uon **S**olenoid
CSC	**C**athode **S**trip **C**hambers
EM	**E**vent **F**ilter
EM	**E**lectro **M**agnetic
EMB	**E**lectro **M**agnetic **B**arrel
EMEC	**E**lectro **M**agnetic **E**nd **C**ap
ECT	**E**nd **C**ap **T**oroid
FCAL	**F**orward **C**alorimeter
HECAL	**H**adronic **E**nd Cap **C**alorimeter
HL-LHC	**H**igh **L**uminosity-**LHC**
HLT	**H**igh-**L**evel **T**rigger
ID	**I**nner **D**etector
ICPP	**I**nstitute for **C**ollider **P**article **P**hysics
LAr	**L**iquid **A**rgon
LINAC 2	**Lin**ear **AC**celerator 2

LEP	**L**arge **E**lectron **P**ositron collider
LHC	**L**arge **H**adron Collider
LHCb	**L**arge **H**adron Collider beauty
LB	**L**ong **B**arrel
LO	**L**eading **O**rder
LUCID	**LU**minosity **C**erenkov **I**ntegrating **D**etector
MBTS	**M**inimum **B**ias **T**rigger **S**cintillators
MC	**M**onte **C**arlo
MDTC	**M**onitored **D**rift **T**ube **C**hambers
MS	**M**uon **S**pectrometer
NLO	**N**ext to **L**eading **O**rder
PS	**P**roton **S**ynchroton
QCD	**Q**uantum **C**hromo **D**ynamics
QED	**Q**uantum **E**lectro **D**ynamics
RF	**R**adio **F**requency
RPC	**R**esistive **P**late **C**hambers
ROC	**R**eceiver **O**perating **C**haracteristic curve
SCT	**S**emi **C**onductor **T**racker
SM	**S**tandard **M**odel
SPS	**S**uper **P**roton **S**ynchrotron
SR	**S**ignal **R**egion
TDAQ	**T**rigger **D**ata **A**c**Q**uisition
TGC	**T**hin **G**ap **C**hambers
TileCal	**T**ile **C**alorimeter
TRT	**T**ransition **R**adiation **T**racker

WLS	**W**avelength **S**hifting
ZDC	**Z**ero-**D**egree **C**alorimeter
2HDM	**T**wo **H**iggs **D**oublet **M**odel

List of Symbols

Global notations

η	Pseudo-rapidity
γ	Gamma photons
λ	Lambda
ϕ	denotes the coordinates
δ	Delta
θ, ϕ	incoming direction expressed in term of spherical coordinates
τ	Tau
x,y	spatial coordinates
p_T	Transverse momentum (Projection of the total momentum in the plane)
E_T	Transverse Energy
Q	Electric charge
u	Up quark
d	Down quark
c	Charm quark
s	Strange quark

t	Top quark
b	Bottom quark
e^-	Electron
μ	Muon

Chapter 1

Introduction

1.1 Overview

The two multi-purpose experiments of the Large Hadron Collider (LHC) ATLAS [2] and CMS [3] reported the discovery of the Standard Model (SM) higgs boson (h) in 2012 [4, 5]. This opened a new chapter in particle physics to further explore whether the witnessed boson is part of the spectrum of SM Higgs or it is an uncommon boson. Discrepancies between data collected at the LHC and Monte Carlo (MC) simulations have been observed in multiple SM results published by ATLAS and CMS collaborations related to multi-lepton production. This displays a lack of understanding of the available LHC data in the field of multi-lepton production. Measurements of the properties of this 125 GeV boson indicate that it is suitable with those predicted by the SM [6, 7]. Excesses in multi-lepton final states may be well examined by the Madala hypothesis [8, 9], which will be explained in greater detail in Section 1.3. According to this hypothesis, additional scalars H and S were introduced in a manner that H decays to SS, Sh, which in return decays to a vast range of final states that are dominated by jets and leptons.

In order to account for several features in Run 1 data, Refs. [10, 11] introduced the effective model and scalars H and S. These features include distortions in the transverse momentum spectrum of the Higgs boson, along with increased activity in associated jets, elevated rates of leptons in association with b-tagged jets during the search for the associated production of the Higgs boson with top quarks, and results from the search for double Higgs boson and weak boson production. The impact on some signal strengths of the newly discovered Higgs boson has been assessed in Ref. [12].

The two-Higgs doublet models (2HDM+S) [13] are simple extensions of the SM that require an additional Higgs-doublet in the model, resulting in a scalar spectrum populated with two CP-even (h, H), one CP-odd (A) and charged $(H\pm)$ scalar bosons. Various facets related to the theory, phenomenology, and constraints of 2HDM+S have been explored in the literature [13, 14]. However, a 2HDM alone cannot account for the aforementioned features in the data, as pointed out in Refs. [15, 11, 8]. To address this, a scalar singlet S is introduced in the 2HDM+S model, which can also explore scenarios with dark matter. Ref. [11] discusses that the 2HDM+S model would lead to anomalous production of multiple leptons, which has been compared to data [16]. Large discrepancies between the data and SM Monte Carlos have been observed, which cannot be resolved with the current understanding of theoretical systematics. Ref. [17] studied these features with additional data, indicating that they have become more pronounced with more data.

In this paper, we build on the phenomenology presented in Ref. [16, 11]. Our first objective is to determine the parameter space of the 2HDM+S model that can account for the observed features in the data discussed in Ref. [16]. Additionally, we examine the consequences of selecting this parameter space for the charged scalar and the heavy pseudo-scalar, with a focus on the CP-odd scalar in the 2HDM+S model.

The four top-quark production which is associated with a high jet multiplicity and leptons in its final states is one of the relevant processes that can be studied in detail with the Madala hypothesis. We focus on investigating the associated production of top quark pair with a heavy pseudoscalar $t\bar{t}A$, and its decay into $A \to t\bar{t}, ZH$, leading to four top quarks in the final state when only considering $A \to t\bar{t}$, while considering $A \to ZH$ we are able to consider the decay modes $H \to SS, Sh, hh$. This analysis results in several noteworthy final states featuring leptons and b-tagged jets. We analyze the four top quarks in the context of the 2HDM+S model, considering the scalar boson properties discussed in Refs. [10, 11, 16], where the masses of H and S are approximately 270 GeV and 150 GeV, respectively. The structure includes events with more than two b- tagged jets, which can be interpreted within a 2HDM as coming from a bottom-quark induced production of the CP-odd boson at intermediate values of tanβ.

The mass of A will be examined in range 400 GeV to 600 GeV, where $A \to t\bar{t}, ZH$ are the leading decays [16]. The SM and the anatomy of multi-lepton

anomalies are presented in section 1.2. An overview of the LHC and a description of the ATLAS detector are provided in Chapter 2. Chapter 3 elaborates on the monte carlo simulation of the heavy pseudoscalar A process involved in this work. The physics objects and trigger choices used for this study are described in Chapter 4. The four top quark background is presented in Chapter 5 and the conclusion of our study is given in Chapter 7.

1.2 The Standard Model of Particle Physics

The world of particle physics is best described by the SM theory as it provides a description of the interaction of elementary particles. It was postulated during the 1960s and it is a relativistic quantum field theory that combines three of the four fundamental interactions; strong, weak and electromagnetic forces in gauge symmetries. It consist of twelve fermions and four gauge bosons: $W^{\pm}$ and Z bosons, gluon (g) and photon (γ). The SM forms electroweak gauge theory by combining electromagnetic and weak forces.

1.2.1 The elementary particles

The behaviour and interactions of the elementary particles that make up matter are governed by the SM. Their properties are classified according to their spin which is associated with their internal angular momentum. Depending on their spin correlation, these particles are separated into two classes, namely fermions and bosons. Fermions are particles with half-integer spin and they adhere to Fermi-Dirac statistics, while bosons are particles with integer spin and they adhere to Bose-Einstein statistics. According to the principle of

Pauli exclusion, more than one boson is allowed to occupy the same quantum state, while this is not possible for fermions. Therefore fermions are considered as particles that make up matter and bosons mediate their interaction. Fermions exist in two categories: leptons and quarks. Leptons interact with electromagnetic and weak forces while quarks interact with weak, strong and electromagnetic forces. Leptons are further divided based on their charge, with neutral leptons and charged leptons. The charged leptons are electron (e^-), muon (μ^-) and tau (τ^-) whereas neutral leptons are electron neutrino (ν_e), muon neutrino (ν_μ) and tau neutrino (ν_τ). Although these leptons have the same charge ($Q = -1$), they have different masses with the electron with the smallest mass which is 200 times less than the muon while tau is 17 times heavier than the muon. The dense mass of muon makes it accelerate less compared to the electron in the presence of electric fields and when it is transmitted through a material, it produces lower electromagnetic irradiation, thus penetrating remarkably. The tau lepton which is the heaviest lepton decays promptly into either a muon or electron and their associated neutrinos in about 2.9×10^{-13} s. The masses of the neutrinos are assumed to be zero in SM and they do not decay and they barely interact with matter. There are six types of quarks which are separated into three groups of generations, with the first generation containing the up quark (u) and the down quark (d), the second containing the charm quark (c) and strange quark (s) and the third generation contains the top quark (t) and the bottom quark (b). Quarks are also charge sensitive and they are categorized according to their charge, with the up, charm and top quarks having a charge of ($Q = +2/3$) while the down, strange and bottom quarks have a charge of ($Q = -1/3$). The other property

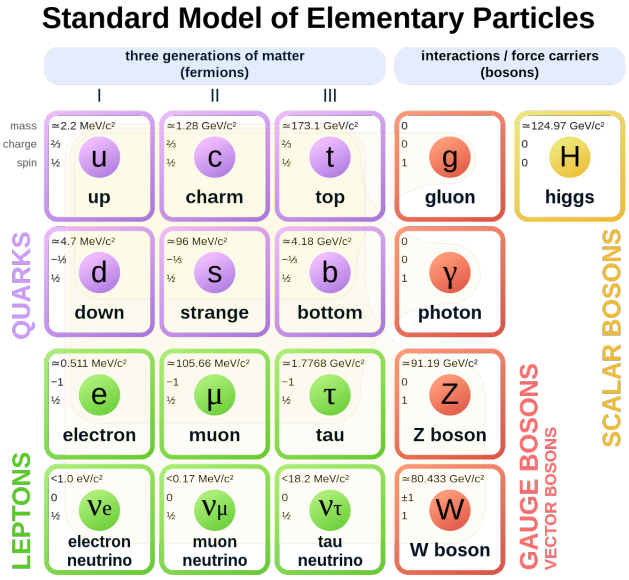


FIGURE 1.1: The Standard Model categorizes fermions into two groups: leptons and quarks. These subatomic particles are further divided into three distinct generations, each made up of one charged lepton, two quarks and one neutral lepton, along with their corresponding antiparticles. Although fermions within the same generation possess similar characteristics, they possess varying masses. Additionally, the electric charge of these particles is expressed as fractions of the proton charge [18].

carried by quarks is their colour change when they interact with strong forces which makes them distinguishable. Quarks are allowed to carry either red, blue or green charge and antiquarks carry anticolour charges. There is yet to be an explanation of the three repetitions of the fermion families (generations). The quarks of the first generation are found to make up most of the mesons including pions and nucleons that are responsible for nuclear binding, while quarks from the second and third generations are constituents of short-lived heavier particles. The particles from the last two generations decay rapidly through the weak force to the particles of the first generation. Thus to be precise, stable matter is made up of particles from the first generation. Quarks

exist as hadrons, which are combinations of integer electrical charge particles while leptons exist as isolated (free) particles in nature.

The interactions of elementary particles in the SM are governed by the exchange of spin-1 gauge bosons. The massless gluon mediates the strong force between quarks. The photon, another particle with zero mass that does not have an electric charge and an infinitely wide range of interaction, mediates the electromagnetic force and it acts between electrically charged particles. This interaction is driven by the electromagnetism theory of quantum electrodynamics (QED). The weak force is mediated by Z and $W^{\pm}$ bosons and it affects charged and neutral particles. These bosons are required to be massless for local gauge invariance, and they gain mass by spontaneous symmetry breaking. The masses and charges of the particles in the SM are summarized in figure 1.1.

The grand unification of the weak and electromagnetic interactions of leptons under the gauge group of $SU(2)_L \times U(1)_Y$ is implied by the coupling of the $W^{\pm}$ and Z bosons with the γ. The subscript L refers to the left-handed and Y refers to the weak hypercharge given by $Y = Q - T^3$, where T^3 is the third component of the weak operator and Q is the electric charge operator. The theory of strong forces upon quarks which are mediated by gluons is known as quantum chromodynamics (QCD). It is well described by the gauge group which includes a local $SU(3)_C$ symmetry, where the subscript refers to the colour charge since gluons couple with colour-charged particles. The interactions between gluons emitted by these colour charges restrict the existence of single (individual) quarks using the colour confinement process. Thus quarks

and gluons cannot exist as free particles.

The Higgs boson, a neutral scalar particle that results from the Brout-Englert-Higgs mechanism used to generate mass to the fermions, $W^{\pm}$ and Z bosons, is the last particle predicted by the SM. This scalar field was hypothesized in 1964 [5, 4] and was discovered by the ATLAS and CMS collaborations [19, 20] in 2012 at CERN. As of yet, the characteristics of this particle have not deviated from those expected for the SM Higgs boson. The Lagrangian of the SM with fermion and scalar fields is defined in Equation 1.1:

$$\mathcal{L}_{SM} = \mathcal{L}_H + \mathcal{L}_G + \mathcal{L}_F + \mathcal{L}_Y, \tag{1.1}$$

where the components in Equation 1.1 are as follows:

$$\mathcal{L}_H = |D_\mu \phi|^2 - V(\phi), \tag{1.2}$$

$$\mathcal{L}_G = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu}, \tag{1.3}$$

$$\mathcal{L}_F = i\bar{\psi} D\psi + h.c., \tag{1.4}$$

$$\mathcal{L}_Y = \bar{\psi}^i y^{ij} \psi^j \phi + h.c., \tag{1.5}$$

where $\mathcal{L}_H$ represents couplings to the Higgs boson, $\mathcal{L}_G$ represents the gauge fields' kinetic energies, $\mathcal{L}_F$ represents couplings to fermions and $\mathcal{L}_Y$ represents the Yukawa couplings. Equation 1.1 describes the SM Lagrangian which is invariant under gauge symmetry $U(1)_Y \times SU(2)_L \times SU(3)_C$.

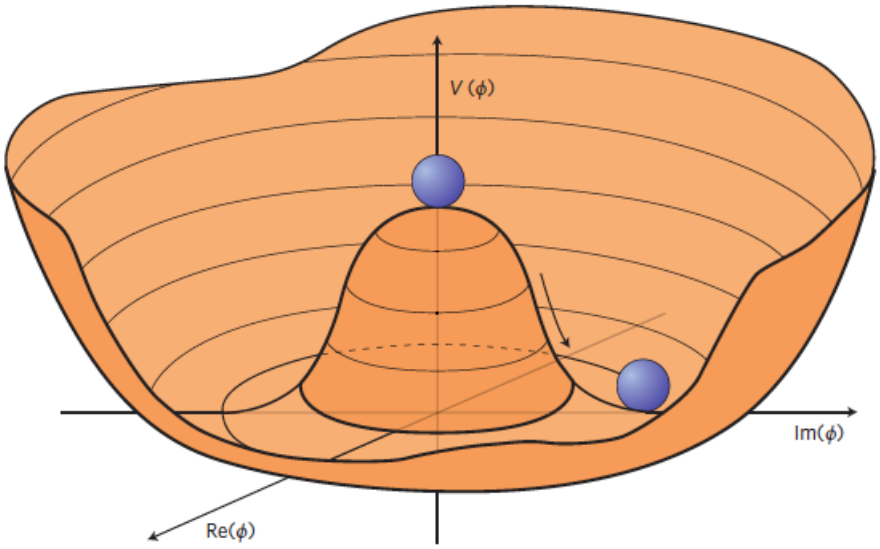


FIGURE 1.2: The potential of Higgs as a function of ϕ in Re(ϕ) - Im(ϕ) plane for $\mu^2 < 0$ [21].

1.2.2 Production of Higgs boson

Various mechanisms produce the SM Higgs boson at the LHC at CERN. Depending on the cross sections, the four main productions are shown in figure 1.3. Gluon-gluon fusion is regarded as the main Higgs boson production because of its cross-section is large and it is mediated by top quark as shown in figure 1.3a. The second largest cross-section of the production mode of the Higgs boson is the vector boson fusion (VBF) displayed in figure 1.3b. Figure 1.3c displays the production mode of Higgs boson associated with Z boson $q\overline{q} \rightarrow ZH(ZH)$ and Higgs boson associated with W boson $q\overline{q} \rightarrow WH(WH)$ is shown in figure 1.3d.

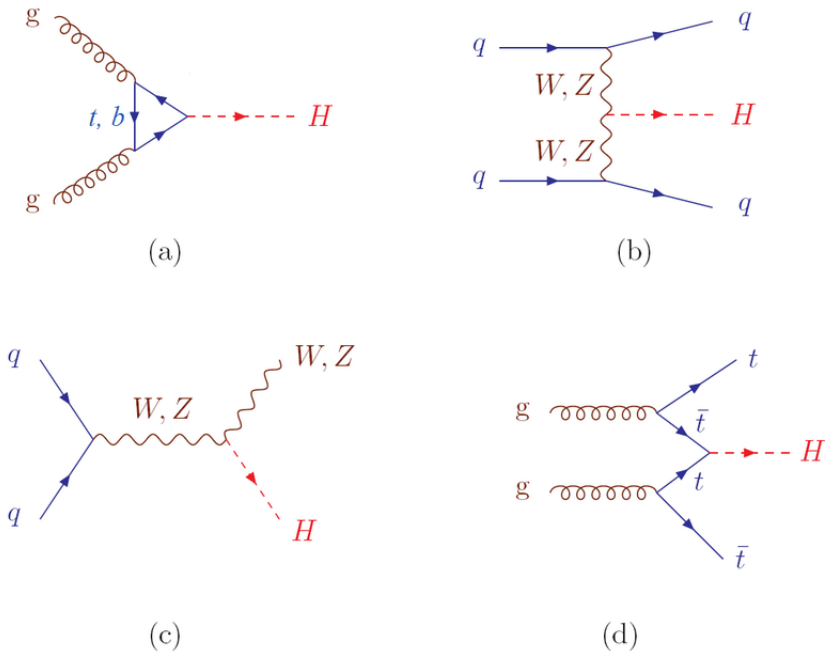


FIGURE 1.3: Feyman diagrams at leading order showing the production of Higgs mechanism: (a) gluon-gluon fusion, (b) vector boson fusion, (c) Higgs strahlung, (d) top fusion [22].

1.2.3 Decay modes of the Higgs boson

The leading decay mode of the Higgs boson which occurs 57.1% of the time is the fermionic decay $h \to b\bar{b}$. The decay of $h \to \tau^+\tau^-$ and $h \to c\bar{c}$ occur 6% and 3% of the time, while the off-shell decay of Higgs to W/Z boson, with higgs $h \to WW^{(*)}$ and $h \to ZZ^{(*)}$ occur 22.1% and 2.9% of the time. The effects of the W and Z off-shell lead to $h \to ZZ^{(*)} \to 4l$ final state as one of the discovery modes of SM Higgs boson. A loop of virtual heavy quarks can lead the Higgs boson to decay to a pair of gluons gg and photons $\gamma\gamma$. The photon pair $\gamma\gamma$ decay is mediated by top quark and the W boson loops which occur 0.22% of the time and the gluon pair is mediated by loops of the bottom and top quarks which occurs 8.5% of the time.

1.3 The multilepton anomalies at the LHC

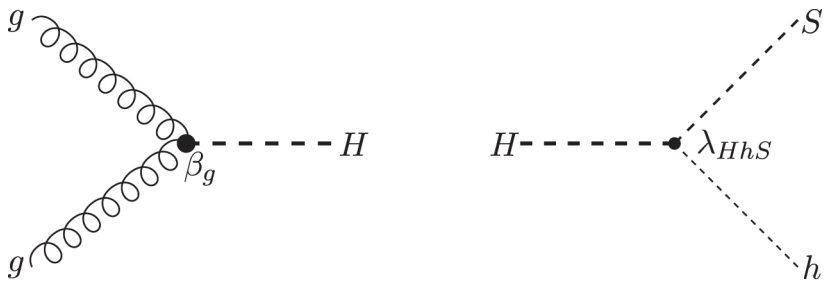
A new era in particle physics has been ushered in by the detection of a Higgs boson (h) [23, 24, 25, 26] at the Large Hadron Collider (LHC) by both ATLAS [19] and CMS [20] experiments. Data collected during Run 1 of the LHC displayed a number of features that led to the formulation of the Madala hypothesis which predicts the existence of two new scalar bosons, H and S. Furthermore, ATLAS and CMS experiments have observed excesses in multilepton data that can be explained by the Madala hypothesis as postulated in Ref. [17]. The Madala hypothesis considered in this work is the extension of the two Higgs doublet model with a singlet scalar S to form a 2HDM+S model with S and H interacting with the SM h. In this model, the singlet scalar is assumed to be a Higgs-like scalar in order to reduce the number of free parameters while allowing S to obtain the branching ratios of the Higgs boson as seen in figure 1.5. The decay of H to singlet scalar and the Higgs boson ($H \to Sh$) is assumed to have a branching ratio of 100%. Figure 1.4 presents the production and decay modes of H. The total effective Lagrangian for a simplified model which includes H and S is given by:

$$\mathcal{L} = \mathcal{L}_{SM} + \mathcal{L}_{BSM}, \tag{1.6}$$

where all the new interactions and states hold in the $\mathcal{L}_{BSM}$ term and it is given by:

$$\mathcal{L}_{BSM} = \mathcal{L}_{H} + \mathcal{L}_{S}. \tag{1.7}$$

The terms in the Lagrangian shown in equation 1.7 are defined as follows:



(a) (b)

FIGURE 1.4: Feynman diagrams contributing to the dominant (a) production mode of H (ggF) and (b) decay mode ($H \to Sh$). The coupling parameters β_g and λ_{HhS} are defined in equation 1.9 and equation 1.15 [16].

$$\mathcal{L}_H = \frac{1}{2}\partial_\mu H \partial^\mu H - \frac{1}{2}m_H^2 HH + \mathcal{L}_G + \mathcal{L}_Y, \tag{1.8}$$

$$\mathcal{L}_G = -\frac{1}{4}\beta_g k_{Hgg}^{SM} G_{\mu v} G^{\mu v} H + \beta_V k_{hVV}^{SM} V_\mu V^\mu H, \tag{1.9}$$

$$\mathcal{L}_Y = -\frac{1}{\sqrt{2}}[y_{ttH}\bar{t}tH + y_{bbH}\bar{b}bH]. \tag{1.10}$$

The Lagrangian for singlet scalar can be written as:

$$\mathcal{L}_S = \mathcal{L}_K + \mathcal{L}_{SVV} + \mathcal{L}_{Sf\bar{f}} + \mathcal{L}_{hHS}, \tag{1.11}$$

where the terms in the Lagrangian shown in equation 1.11 are defined as follows:

$$\mathcal{L}_K = -\frac{1}{2}\partial_\mu S \partial^\mu S - \frac{1}{2}m_S^2 SS, \tag{1.12}$$

$$\begin{aligned}\mathcal{L}_{SVV} = &\frac{1}{4}k_{Sgg}\frac{\alpha_S}{12\pi v}SG^{a\mu v}G^a_{\mu v} + \frac{1}{4}k_{S\gamma\gamma}\frac{\alpha}{\pi v}SF^{\mu v}F_{\mu v} + \frac{1}{4}k_{SZZ}\frac{\alpha}{\pi v}SZ^{\mu v}Z_{\mu v} \\ &+ \frac{1}{4}k_{SZ\gamma}\frac{\alpha}{\pi v}SZ^{\mu v}F_{\mu v} + \frac{1}{4}k_{SWW}\frac{\alpha}{\pi v}SW^{+\mu v}W^-_{\mu v},\end{aligned} \tag{1.13}$$

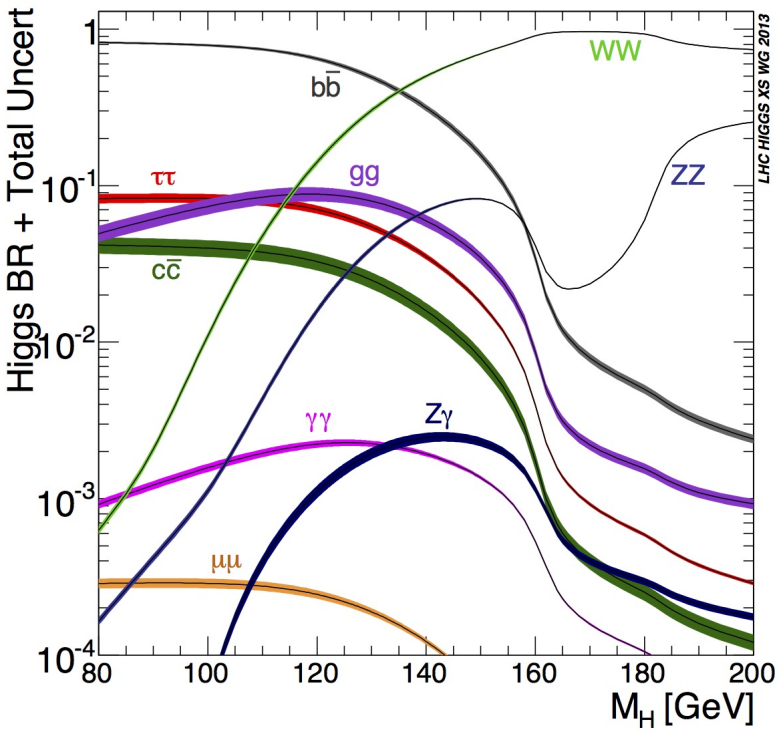


FIGURE 1.5: The decay modes of Higgs boson between 100 GeV and 200 GeV [27].

$$\mathcal{L}_{Sf\bar{f}} = -\sum_{f} K_{Sf}\frac{m_f}{v}S\bar{f}f, \tag{1.14}$$

$$\mathcal{L}_{HhS} = -\frac{1}{2}v[\lambda_{hhS}hhS + \lambda_{HHS}HHS + \lambda_{HSS}HSS + \lambda_{HhS}HhS]. \tag{1.15}$$

The Z and W are represented by V symbol and $W^{+}_{\mu v} = D_{\mu}W^{\pm}_{v} - D_{v}W^{\pm}_{v}$, where $D_{v}W^{\pm}_{\mu} = [\partial_{\mu} \pm ieA_{\mu}]W^{\pm}_{v}$. The first term in equation 1.9 represents the effective coupling of gluons to the H boson, where k^{SM}_{Hgg} is the SM coupling defined as $k^{SM}_{Hgg} = \alpha_s/3\pi v$. The production rate of $gg \rightarrow H$ is regulated by the parameter β_g. The effective coupling strength of gluons to the H boson is defined by the term $\beta_g k^{SM}_{Hgg}$ while the second term in equation 1.9 denotes the effective coupling of VVH. The branching ratio of $H \rightarrow VV$ is supervised by the term $k^{SM}_{hVV} = m^2_V/v$. The SM top Yukawa coupling is represented by the first term in equation 1.10 while the second term represents the bottom Yukawa coupling. The $H \rightarrow b\bar{b}$ decay is catered for by the bottom Yukawa coupling y_{bbH} however because the production of bottom quarks via ggF is insignificant, this decay is negligible. On the other hand, the $H \rightarrow Sh$ decay produces a variety of possible final states that consist of leptons, jets, missing transverse momentum (E^{mis}_T) and photons. In this work, we take advantage of these possible final states and focus on two same-sign leptons and three leptons channels arising from the production of four top quarks in association with a heavy pseudo-scalar (A), where we would evaluate three mass points of the pseudoscalar.

1.3.1 Madala hypothesis and the multi-lepton final states

Figure 1.6 illustrates the dilepton invariant mass distribution, which compares ATLAS data, SM background and BSM signal after requiring events with

two leptons with opposite charge of different flavours and a minimum of one b-tagged jet [16]. The invariant mass distribution between 0 and 100 Gev experiences some discrepancy which can be better explained by the $H \to Sh$ signal, where the mass of H and S are set to $m_H = 270\,\text{GeV}$ and $m_S = 150\,\text{GeV}$.

The BSM model considered here displayed multiple issues that deserved attention. The BSM model is capable of shearing significant light with regards to the magnitude of the excesses in data using just one degree of freedom β_g^2. Simultaneous fits were made for all measurements considered in table 1.1 and their respective profile likelihood ratios are shown in figure 1.7. The best fit is a combination of all measurements with $\beta_g^2 = 2.92 \pm 0.35$ with a Correspondence significance value of 8.04σ. This significance value illustrates how the simplified BSM model considered here can describe the discrepancies in multi-lepton final states at the LHC. It is believed that since this simplified model can provide a vast number of final states, it is only reasonable to examine one of the rare processes that have been predicted by the SM, which is the production of four top quarks in the dilepton channel due to recent success of this model being able to explain excesses in the multi-lepton sector.

1.3.2 The model

To interpret the findings reported by ATLAS [28], we have studied a 2HDM model extended with a real singlet scalar, Φ_S, referred to as the 2HDM+S model [11, 29]. The same notation as in Ref. [11] is used in this study. The

Selection	Best-Fit β_g^2	Significance (σ)
ATLAS Run 1 SS ll + b-jets	6.51±2.99	2.37
ATLAS Run 1 OS $e\mu$ + b-jets	4.06±1.37	2.99
ATLAS Run 2 SS ll + b-jets	2.22±1.19	2.01
CMS Run 2 SS ll + b-jet	1.41±0.80	1.75
CMS Run 2 OS $e\mu$	2.79±0.52	5.45
ATLAS Run 2 OS $e\mu$ + b-jets	5.42±1.28	4.06
CMS Run 2 lll + E_T^{miss}	9.70±3.88	2.36
ATLAS Run 2 lll + E_T^{miss}	9.05±3.35	2.52
Combination	2.92±0.35	8.04

TABLE 1.1: A summary of the SM+BSM fit results for each measurement considered in this article, as well as the combined result.

potential of the 2HDM+S model is expressed as:

$$\begin{aligned} V(\Phi_1,\Phi_2,\Phi_S) = {} & m_{11}^2|\Phi_1|^2 + m_{22}^2|\Phi_2|^2 - m_{12}^2(\Phi_1^\dagger\Phi_2 + h.c.) \\ & + \frac{\lambda_1}{2}(\Phi_1^\dagger\Phi_1)^2 + \frac{\lambda_2}{2}(\Phi_2^\dagger\Phi_2)^2 + \lambda_3(\Phi_1^\dagger\Phi_1)(\Phi_2^\dagger\Phi_2) \\ & + \lambda_4(\Phi_1^\dagger\Phi_2)(\Phi_2^\dagger\Phi_1) + \frac{\lambda_5}{2}[(\Phi_1^\dagger\Phi_2)^2 + h.c.] \\ & + \frac{1}{2}m_S^2\Phi_S^2 + \frac{\lambda_6}{8}\Phi_S^4 + \frac{\lambda_7}{2}(\Phi_1^\dagger\Phi_1)\Phi_S^2 + \frac{\lambda_8}{2}(\Phi_2^\dagger\Phi_2)\Phi_S^2. \end{aligned} \tag{1.16}$$

Here the Φ_1 and Φ_2 are the $SU(2)_L$ Higgs doublets and Φ_S is the singlet field. The first three lines correspond to the potential for the 2HDM and the last line is the contribution from the singlet field. In order to avoid the tree-level FCNC's, all the quarks of a given charge must couple to a single Higgs doublet, which can be done by imposing a Z_2 symmetry and it is softly broken by the m_{12}^2 term. Moreover, if we extend the Z_2 symmetry to the Yukawa sector it guarantees the absence of FCNC at tree level. Here, we consider a case where the real singlet field acquires a vacuum expectation value (vev) with a Z_2 symmetry. In other words, if this symmetry is respected then the singlet

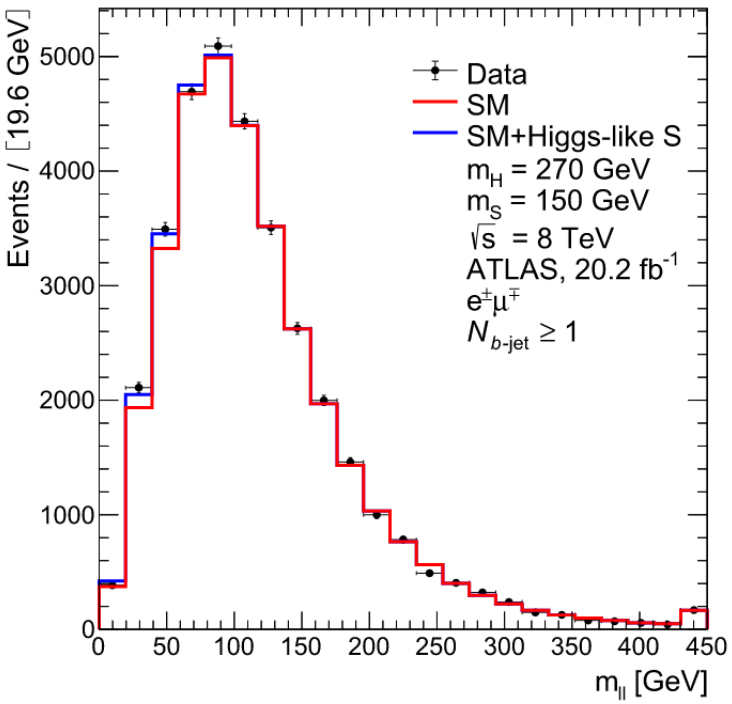


FIGURE 1.6: Invariant mass distribution of ATLAS data (black dots) and SM background (red histogram) compared with the BSM signal (blue histogram) in events with two oppositely charged leptons and at least 1 b-tagged jet are required [16].

scalar becomes a viable dark matter candidate.

In this study, we set $m_{12}^2 \neq 0$ in the 2HDM+S potential, Eqn. 1.16, which corresponds to a soft breaking of the Z_2 symmetry. Since we do not consider explicit CP violation, we assume λ_i to be real. Minimizing the potential with the three Higgs fields and assuming the *vevs* of the fields $\Phi_1 \to v_1/\sqrt{2}$, $\Phi_2 \to v_2/\sqrt{2}$ and $\Phi_S \to v_S$, the three minimization conditions are:

$$\frac{\partial V}{\partial[v_1, v_2, v_S]} = 0. \tag{1.17}$$

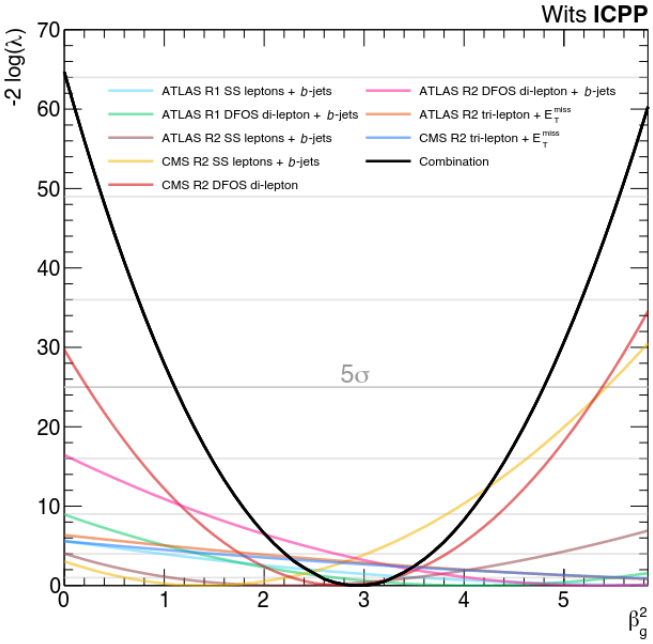


FIGURE 1.7: The profile likelihood ratios for each of the individual fit considered here [17].

With these conditions on the fields are:

$$m_{11}^2 = -\tfrac{1}{2}(v_1^2\lambda_1 + v_2^2\lambda_{345} + v_S^2\lambda_7) + \tfrac{v_2}{v_1}m_{12}^2, \tag{1.18}$$

$$m_{22}^2 = -\tfrac{1}{2}(v_2^2\lambda_2 + v_1^2\lambda_{345} + v_S^2\lambda_8) + \tfrac{v_2}{v_1}m_{12}^2, \tag{1.19}$$

$$m_S^2 = -\tfrac{1}{2}(v_1^2\lambda_7 + v_2^2\lambda_8 + v_S^2\lambda_6), \tag{1.20}$$

where $\lambda_{345} = \lambda_3 + \lambda_4 + \lambda_5$.

Chapter 2

The Large Hadron Collider and the ATLAS Detector

2.1 The Large Hadron Collider

The LHC [30] is a double-ring superconducting hadron accelerator and particle collider built by CERN. It is stationed beneath the borders of France and Swiss near Geneva 100 meters underground and consists of a 27 km ring of superconducting magnets. In order to bypass a new threshold of discovery, a large amount of energy is needed while not forgetting the importance of luminosity as well. The primary design of the LHC was to collide proton beams at higher center-of-mass energies and luminosity. Its center-of-mass energy and luminosity make it the largest particle accelerator to be ever built in the world.

The energies of these particles are increased by accelerators that were once the leading accelerators of their time that are in series before reaching the LHC, as shown in figure 2.1. The ionization of hydrogen atoms provides protons through the electric field in a duoplasmatron. Then they are directed to

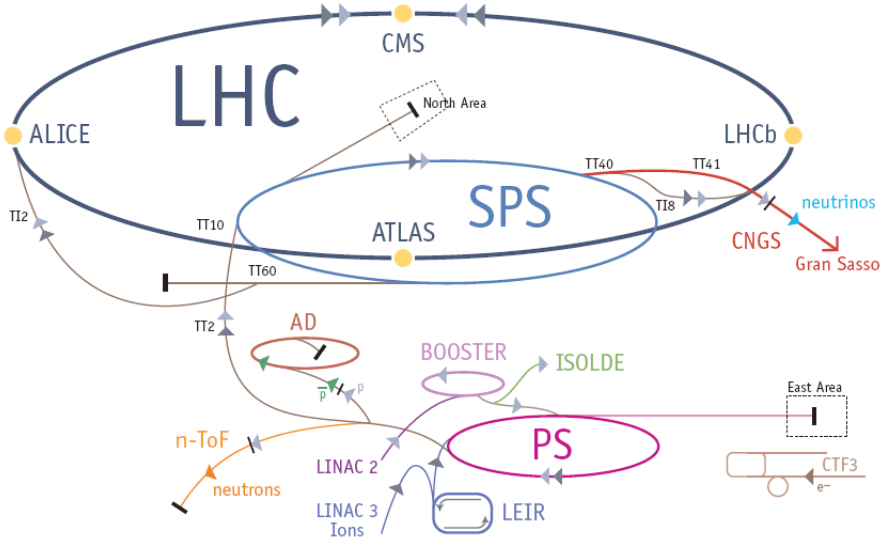


FIGURE 2.1: The LHC particle accelerator. Protons are accelerated by multiple accelerators prior to injection into the LHC [31].

a linear accelerator called LINAC 2 which accelerates the protons to 50 MeV. Then the protons are injected into the PS (proton synchrotron) Booster which is the first circular accelerator of the chain, which accelerates the energy of the protons to 1.4 GeV. Afterwards, the beam of protons is supplied to PS which will increase the energy to 25 GeV. The beam of the proton is then delivered from the PS to the Super PS (SPS), where its energy is increased to 450 GeV before being injected into the LHC in both clockwise and counterclockwise directions.

The injection of protons from SPS is performed in bunches. The bunches are separated into patterns of 72 bunches which are split by twelve empty buckets, forming a train of bunches. The beams are captured, accelerated and stored by the LHC's radio frequency (RF) cavities, which operate at approximately 400 MHz. Because of this, there are approximately 35640 locations in the RF

field that have energies that are favourable for the bunches to be positioned on the beam. Within the LHC the beam collides around the accelerator ring at four interaction points that represent the location of the particle detectors. These detectors include A Toroidal LHC ApparatuS (ATLAS), the Compact Muon Solenoid (CMS), the Large Hadron Collider (LHCb) and A Large Ion Collider Experiment (ALICE). In this study we have been working with the ATLAS, thus we will present more detail about it in the following section.

2.2 The ATLAS Detector

The ATLAS experiment is one of the two general-purpose detectors built at the LHC, designed to affirm the measurements of the SM and to discover new physics. Presently it has more than 5500 scientists who are from over 245 institutions in 42 countries, thus making it one of the largest collaborations globally. It has a diameter of 25 meters and it is 44 meters long with a weight of approximately 7000 tonnes. The requirements of its design were influenced by important signals for the possible detection of new physics , as well as the need to sustain the hard radiation environment.

The ATLAS detector is made up of a sequence of sub-detector systems with cylindrical symmetry arranged around the interaction point where the LHC proton beams collide. From the interaction point to the outside of the detector, these sub-detectors are identified as follows; the inner detector (ID) which is used to identify charged particles followed by the calorimeters which enclose the ID, which is used to measure the energies of the particles going through

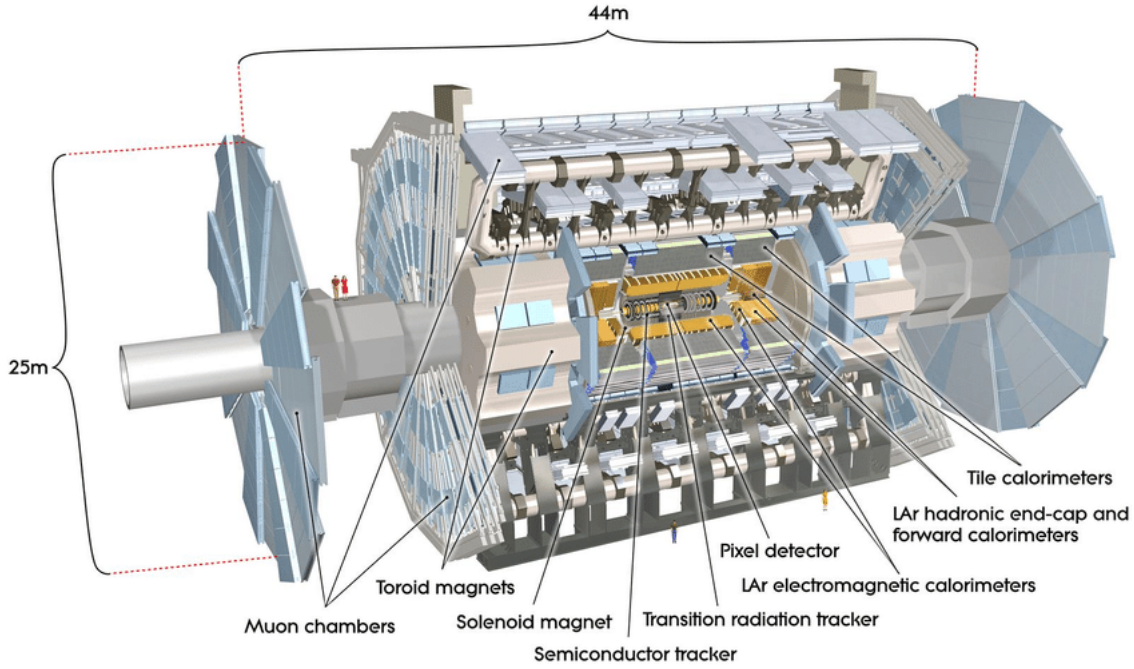


FIGURE 2.2: The layout of the ATLAS detector [32].

them. The last outermost sub-detector is the muon spectrometer (MS) which identifies and measures the momenta of the muons. A system of magnets is used to deflect the trajectory of the charged particles for momentum measurement. A trigger system has also been assembled to determine which events are to be measured in a short period of time. The detector is designed to be completely sealed as possible in order to be able to detect all particles produced in proton-proton collisions except for neutrinos. The muons can be accounted for by measuring the momentum imbalance amongst detected particles. The full layout of the detector is displayed in figure 2.2.

The Cartesian coordinate system used by the ATLAS detector is right-handed

coming from the origin of the detector at the interaction point of the proton-proton collisions. The positive z-axis is defined by the counter-clockwise beam direction, while the $x-y$ plane is perpendicular to the z-axis. The positive x-axis points to the center of the LHC ring from the interaction point of the detector and the positive y-axis points towards the surface of the earth, upwards from the interaction point. The detector is divided into two parts by the symmetric in the positive and negative z, with side "A" associated with $z > 0$ while side "C" is associated with $z < 0$. The detector has two endcaps in the transversal planes to the beam axis, both of them are referred to as Endcap-A and Endcap-C.

An unknown fraction of the momentum of the proton is carried by partons in proton-proton collisions. Because a significantly large portion of the longitudinal momentum is lost in remains that manage to escape through the beam pipe, the overall energy of the collision is unknown. According to the geometry of the detector, the cylindrical coordinate system can be defined by:

$$r = \sqrt{x^2 + y^2}, \tag{2.1}$$

$$\phi = arctan(\frac{y}{x}), \tag{2.2}$$

$$\theta = arctan(\frac{r}{z}). \tag{2.3}$$

The radial dimension represented by equation 2.1 measures the distance from the beam. The polar angle represented by equation 2.3 is the angle between the transverse plane and the beam axis. Equation 2.2 represents the angle measured around the beam axis. Because of the remains lost in the beam axis,

the cylindrical and polar coordinates are substituted by the Lorentz invariant quantities. As a result, θ is rarely used and has been replaced by the rapidity which is defined as:

$$y = \frac{1}{2} ln[(E + p_Z)/(E - p_Z)], \tag{2.4}$$

where p_Z represents the momentum along the direction of the beam and E is the energy of the particle. It can also be replaced by pseudorapidity, which is defined as the rapidity in the relativistic approximation,

$$\eta = -ln(tan(\frac{\theta}{2})). \tag{2.5}$$

The $x-y$ plane which is transverse to the direction of the beam is often referred to $\eta - \phi$ angle space, where the surfaces of the cylindrical detectors appear as flat sheets and the ΔR cone is defined as:

$$\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}. \tag{2.6}$$

2.2.1 The Inner Detector

The closest component of the ATLAS detector to the beam axis is the ID [33, 34], which is designed to identify and track charged particles going through it. It also provides exquisite momentum resolution and both primary and secondary vertex measurements for the tracks of charged particles with $p_T >$ 500 MeV within the region of $|\eta| < 2.5$. Particles below 500 MeV cannot be reconstructed since they do not fully cross the ID. It is contained within a 3512 mm long cylinder envelope, with an outer radius of 1150 mm. It is absorbed in a magnetic field of 2 T produced by a superconducting solenoid that

is required to bend charged particles for momentum measurements.

During proton-proton collisions, a large number of particles are formed at the interaction point creating enormous track density within the detector, thus a fine granularity is required in this part of the detector. This is achieved by means of three great precision sub-detectors within the ID, namely the Pixel detector [35, 36] commonly referred to as "Pixel", is the closest part of the detector to the interaction point and is made up of silicon pixel sensors. It is followed by the SemiConductor Tracker (SCT) [37, 38, 39] that is covering the pixel detector and lastly the Transition Radiation Tracker (TRT) [40, 41, 42]. The configuration of ID is shown in figure 2.3. The Pixel and SCT detectors use semiconductor technology to provide hits on track, while small drift tubes are being used by the TRT to provide semi-continuous tracking out to large radius.

Each sub-detector is divided into a barrel region and an endcap region. The barrels region is arranged in concentric cylinders around the beam axis whilst the endcap region has sensors mounted on disks that are perpendicular to the beam axis and are used to expand the coverage of the detector in η without increasing the surface of the detector. The material used to build the sub-detector system in the ID reduces the quality of the energy measured. Therefore these detectors are designed with a structure that reduces the amount of matter that exits the detector while also supporting the high radiation conditions set for the detectors closer to the beam axis.

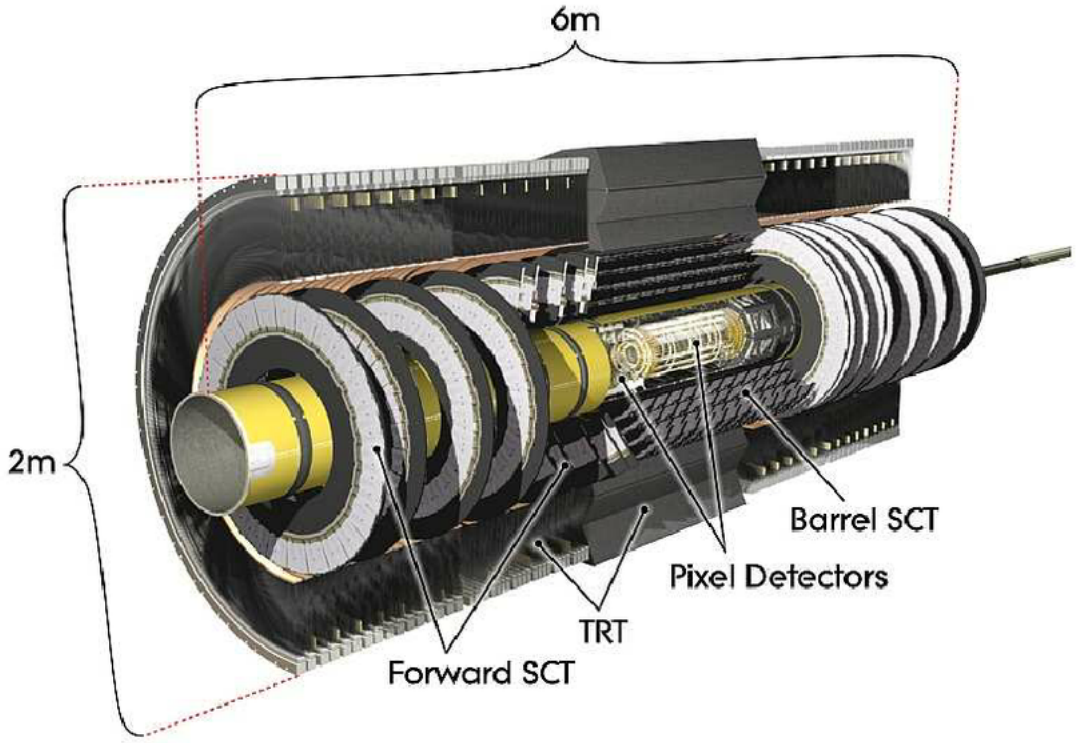


FIGURE 2.3: The ATLAS inner detector together with its sub-detectors [43].

2.2.2 The ATLAS Calorimeter System

Outside the solenoid magnet that surrounds the ID is the ATLAS calorimeter [45, 46, 47] system. It is made up of multiple non-compensating thirteen sampling calorimeters that make use of various technologies to analyze the various physics processes and to withstand the hard radiation environment. This system is designed to measure the energies of hadrons, photons and electrons in the pseudorapidity range of up to $|\eta| < 4.9$ around the beam axis. The particles emerging from the interaction point interact with the calorimeter via multiple processes that can be traced from the shower of particles produced

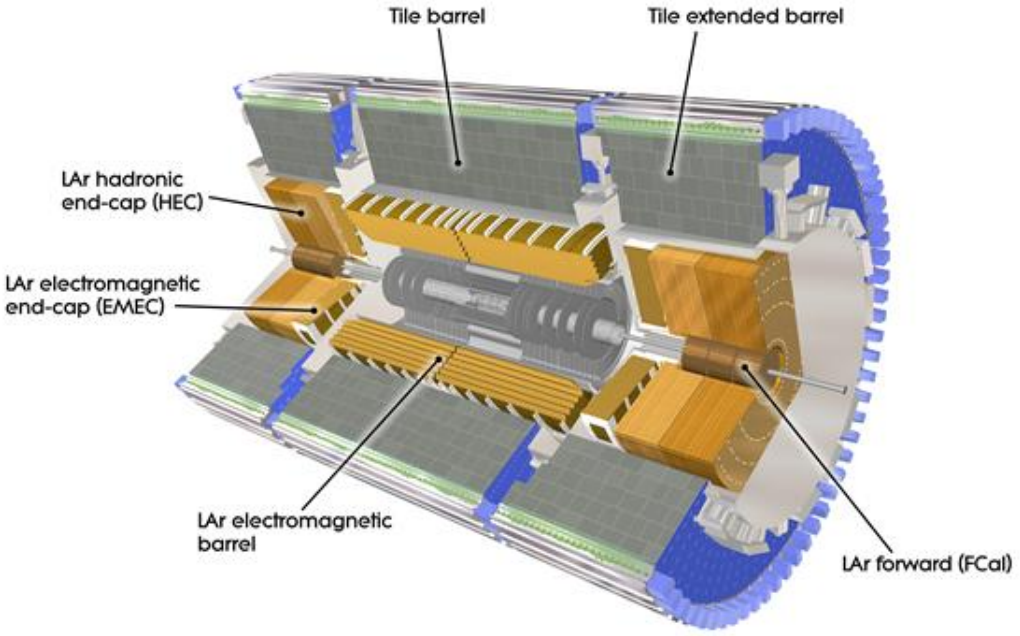


FIGURE 2.4: ATLAS calorimeter system [44].

in the active medium. In order to contain the particles that meet the energy range of the calorimeter, the absorbing medium has to be a dense material that stops particles before they reach the muon spectrometer.

The calorimeter system provides measurements of both electromagnetic and hadronic energy. Therefore, ATLAS calorimeters are classified as electromagnetic or hadronic calorimeters based on their design. The electromagnetic (EM) calorimeter is required to have high granularity, in order to make accurate measurements of electrons and photons. It also helps to determine the energy of hadrons in jets. The ATLAS calorimeter is made up of four components: in the pseudorapidity region of $|\eta|$ <3.2 we have the Liquid Argon (LAr) EM calorimeter, in the region of $|\eta| < 1.7$ we have the Tile hadronic

calorimeter (TileCal), the LAr hadronic end-cap calorimeter in the region of $1.5 < |\eta| < 3.2$ and in the region of $3.1 < |\eta| < 4.9$. The energies of the photons and electrons are measured by the LAr EM calorimeter while the energies of the hardons are measured by the TileCal and LAr hadronic end-cap. The overall ATLAS calorimeter system is shown in figure 2.4, where each calorimeter system is placed.

The calorimeters closest to the interaction point are contained in a barrel and endcaps. The LAr EM barrel calorimeter (EMB) is located in the barrel, while each of the two endcaps contains a LAr EM endcap (EMEC). A LAr hadronic endcap (HEC) is located at the rear of the EMEC and a forward calorimeter (FCal) covers the region adjacent to the beam axis. The active medium in all of these calorimeters is liquid argon. The argon showers release electrons, which are later collected and recorded. Because of its radiation resistance, linear behaviour and response stability, liquid argon was selected as the active medium of the ATLAS calorimeter.

2.2.3 The Muon Spectrometer

The outermost part of the ATLAS detector is the Muon Spectrometer (MS) [48]. It is designed to detect charged particles that made it through the end-cap and barrel calorimeters and measure their momentum in the pseudorapidity region of $|\eta| < 2.7$. The role of the MS is to allow the identification of muons and the measurement of muon tracks in large superconducting air-core toroid magnets with bending powers between 1 to 7.5 Tm is used to determine the magnetic

deflection of muons. The large barrel toroid provides the magnetic field required to direct the particle trajectories in the pseudorapidity range of up to $|\eta| < 1.4$, while the smaller magnets in the endcaps placed in both extremes of the barrel toroid bend muon tracks in the pseudorapidity region $1.6 < |\eta| < 2.7$. The magnetic deflection in the transition region $1.4 < |\eta| < 1.6$ is provided by a combination of the barrel and endcap fields. This magnet configuration provides a magnetic field that is perpendicular to the trajectories of the muons, thus reducing the deterioration of resolution due to multiple scattering.

The MS is composed of two end-caps and barrel regions. The chambers of the barrel are configured as three cylinders aligned with the direction of the beam, placed at three distinct radius distances of about 5 m, 7.5 m and 10 m from the axis of the beam while covering the pseudorapidity range up to $|\eta| < 1$. The endcap chambers are arranged in four disks placed at $|z| \approx$ 7.4 m, 10.8 m, 14 m and 21.5 m from the point of interaction, aligned with the beam axis covering the pseudorapidity range of $1 < |\eta| < 2.7$. Except in the region where $\eta \approx 0$, where open space has been created in the chamber coverage to allow for the passage of cables and services for the inner detector, central solenoid, and calorimeters, all of the chambers together offer almost complete coverage.

The chambers in the detector are separated into two categories: small and large. The small chambers are located near the barrel toroid coils, while the large ones are situated between them. The barrel chambers and a part of the endcap chambers are mounted on the barrel toroid structure. The remaining endcap chambers are mounted on the experimental hall's service structures

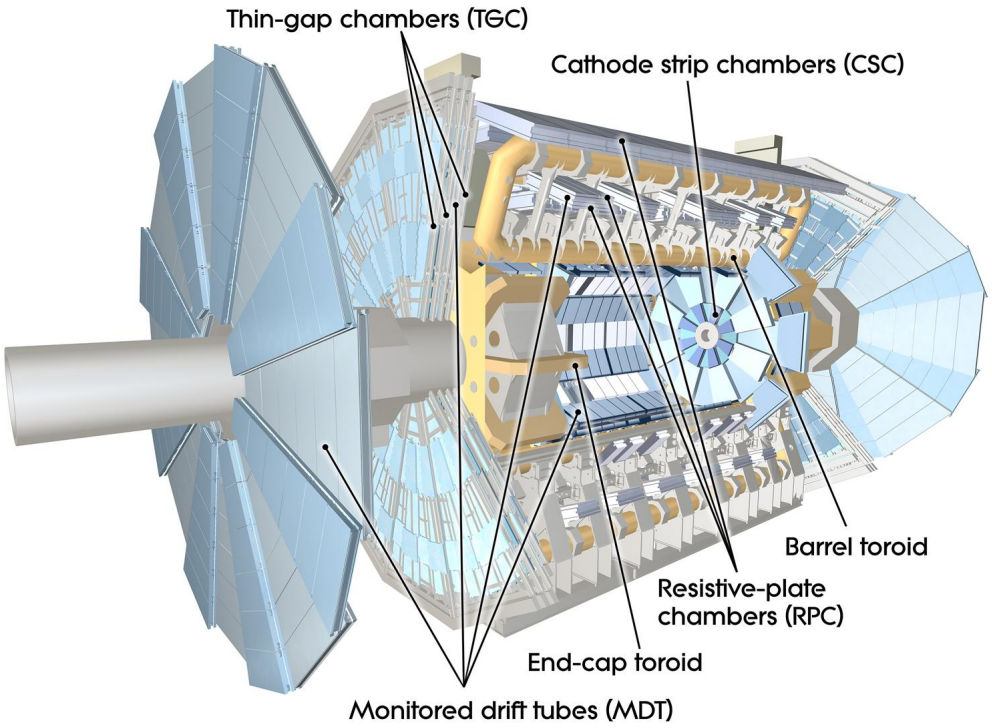


FIGURE 2.5: Muon spectrometer showing its components [49].

and special support frames. The four types of chambers used in the detector are Monitored Drift Tube Chambers (MDT), Cathode Strip Chambers (CSC), Resistive Plate Chambers (RPC), and Thin Gap Chambers (TGC). The first two types are used for precise measurement of muon tracks, while the other two types are used for the trigger and data acquisition system. All of these types of chambers are illustrated in figure 2.5.

2.2.4 Forward Detectors

The ATLAS detector also includes other multiple forward detector systems [50] which are; Minimum Bias Trigger Scintillators (MBTS), Zero-Degree Calorimeter (ZDC), Absolute Luminosity For ATLAS (ALFA) Roman pots and LUminosity measurement using Cerenkov Integrating Detector (LUCID). These forward detector systems provide additional physics measurements and monitoring purposes when used together.

The MBTS system consists of thirty-two scintillation paddles, each 2 cm thick, placed on both sides of the interaction point in two disks. These paddles are mounted on the inner side of the endcap calorimeter, located at a distance of $|z| = 3.65$ m from the collision point and cover a pseudorapidity range of $2.09 < |\eta| < 3.84$. The LUCID detector is a Cerenkov detector located at $|z| = 17$ m from the interaction point and 10 cm radially from the beam axis, covering the pseudorapidity range of $5.61 < |\eta| < 5.93$. LUCID detects the scattering of protons to measure the integrated luminosity and it also tracks the amount of instantaneous luminosity produced at any given moment, as well as the conditions of the beams. The ZDC system detects photons and neutrons with pseudorapidity $|\eta| < 8.3$ in proton-proton and heavy-ion collisions [51]. The ALFA system is composed of four stations using Raman pot [52], with two on each side of the collision point located 4 m apart. It is situated at a distance of $|z| = 240$ m from the collision point and is situated about 1-1.5 mm away from the LHC beams. These forward detector systems are placed just outside the detector on both sides as seen in figure 2.6.

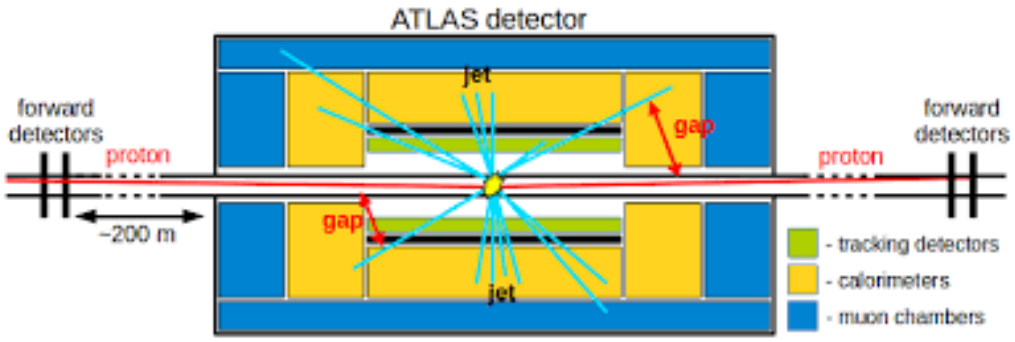


FIGURE 2.6: A schematic diagram showing the positions of the forward detectors at the ATLAS detectors.

2.2.5 Magnets

The ATLAS detector's magnet system generates a magnetic field that bends the trajectories of charged particles for momentum measurement, which is determined by the radius of the tracks' curvature within the detector. This system, measuring 26 m in length and 22 m in diameter, creates a magnetic field within a volume of approximately 12000 m using a combination of four large superconducting magnets. These magnets are cooled using helium, which is forced through tubes attached to the windings' casing, and kept at a temperature of around 4.5 K. The central solenoid, consisting of four magnets, is aligned with the beam axis and generates an axial magnetic field for the ID. The ID is enclosed by three large toroidal magnets, which create toroidal magnetic fields for the MS, as illustrated in figure 2.7.

The central solenoid is engineered to generate a 2 T magnetic field along the beam axis for the ID, with the highest magnetic field reaching 2.6 T. The hadronic calorimeter's steel returns the magnetic field flux. Because the solenoid magnet is located in front of the calorimeter system, it must have a

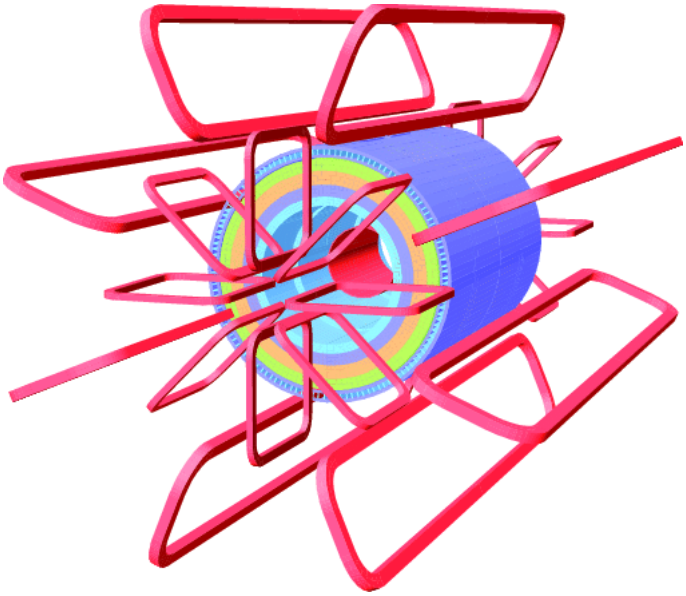


FIGURE 2.7: The ATLAS magnet system with the central solenoid and the three toroids [53].

minimal impact on the energy measurement in the calorimeters. As a result, it was designed to have a low material thickness as possible. The toroidal magnet system is made up of one barrel magnet and two end-cap magnets. In the central and endcap regions, the toroids are designed to generate toroidal magnetic fields of 0.5 T and 1 T respectively for the MS and it is situated just outside of the calorimeters but within the muon system. The toroids in the MS system provide a magnetic field that bends the tracks of particles in specific ranges of pseudorapidity. The barrel toroid has a bending power in the range $|\eta| < 1.4$, while the endcap toroids have bending power in the range $1.6 < |\eta| < 2.7$. The overlap of the magnetic fields generated by the two magnets results in a lower bending power in the transition region of $1.4 < |\eta| < 1.6$.

2.2.6 Trigger and Data Acquisition system of the ATLAS detector

The LHC studies rare phenomena that have a low expected cross-section, making high luminosity necessary to produce these events. With an initial rate of 40 MHz per bunch crossing and an average of 23 interactions per crossing, the proton-proton interaction rate increases to 1 GHz at the target luminosity of $10^{34}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$. However, most of these events are not of high importance for the processes studied because they would contain very low p_T particles which lead to minimum bias events. Due to a large number of events from proton-proton collisions, event selection is vital because of the limited permanent storage in the configuration of the design which is about 200 Hz. To ensure that a sufficient number of events can be stored in the available storage, a large number of events of approximately a factor of 10^6 should be rejected and this process is done by the ATLAS trigger system [54].

While rejecting unusable events and recording interesting events in a limited time frame, the TDAQ system is designed to use simple criteria to make this decision. It is composed of three levels of selecting events which are Level-1 (L1), Level-2 (L2) and event filter (EF). The high-level trigger (HLT) is formed by the combination of L2 and EF. Each trigger level refines the previous level's decisions by applying more selection criteria sequentially. The Level 1 (L1) trigger is executed using specialized, high-speed electronics boards. It makes event selections based on limited data from the calorimeters and muon detector. The HLT on the other hand is entirely implemented in software. ATLAS detector. [52]It makes use of information from the calorimeters, muon

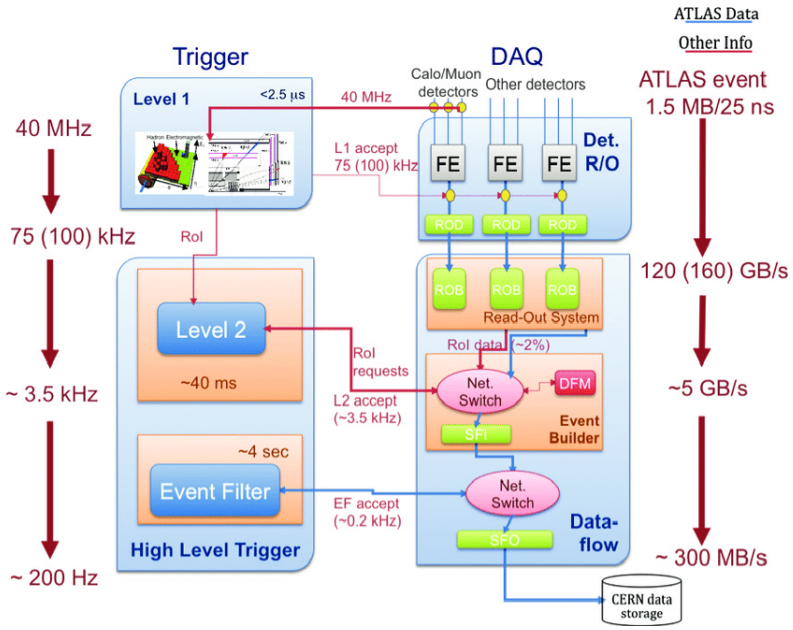


FIGURE 2.8: An overview of the trigger system used by the ATLAS detector [55].

detector and ID.

The L1 trigger is responsible for the first stage of event selection, which involves identifying specific patterns based on the reduced-granularity data from the calorimeters and muon detector. The L2 trigger is the next stage of the trigger system and the first in the ATLAS TDAQ system to receive data from the ID. It makes use of all granularity and precision information associated with the region of interest defined at the L1 trigger. Just like L1, event selection at L2 uses inclusive criteria such as high p_T objects above the defined thresholds to refine the selection made at L1. The EF is in charge of the final stage of online event selection. It receives events chosen by the L2 and processes them using full detector granularity, with offline analysis procedures on fully-built events. It employs more complex algorithms than L2 to reduce event rates to around

200 Hz.

Chapter 3

Monte Carlo samples

The analysis of data is a critical component of research in many fields, as it enables researchers to draw meaningful conclusions and make informed decisions based on the results. However, in the course of my research, I did not personally conduct the analysis. Rather, I relied on the results presented in a publicly available ATLAS paper. This paper provided a detailed description of the methodology used to analyze the simulated data, as well as the results obtained from the analysis. It's important to note that the analysis was conducted on simulated data and not real data.

3.1 Data Set

The ATLAS detector collected the data similar to one generated using monte carlo (MC) in this analysis between 2015 and 2018 at $\sqrt{s} = 13\,\text{TeV}$. The considered events are for all the detector subsystems that were operational. The data set corresponds to a $139\,\text{fb}^{-1}$ integrated luminosity [56, 57]. Single-lepton or dilepton triggers were used to collect events. The selection of events with

a single lepton is done through the use of triggers, which have specific criteria for identification and transverse momentum (p_T). These triggers can have either low threshold requirements for transverse momentum and an isolation requirement or a looser identification criterion and higher threshold with no isolation requirement. The lowest threshold for transverse momentum used in single-lepton triggers varies between 20 to 26 GeV, depending on the lepton flavour and the data collection period [58, 59]. The p_T thresholds used in the dilepton triggers ranged from 8 to 24 GeV depending on the lepton flavour and the time period of data collection. Dilepton triggers are used to select events containing leptons without the need for isolation; they are used to validate the fake/non-prompt background estimation. At least one vertex must be reconstructed from at least two ID tracks with transverse momenta greater than 0.4 GeV. The thresholds for p_T used in dilepton triggers vary between 8 and 24 GeV, depending on the type of lepton and the time period of data collection. These triggers are used to select events that contain leptons without requiring isolation and are employed to validate the estimation of fake/non-prompt background. To be considered, an event must have at least one vertex reconstructed from at least two ID tracks with transverse momenta greater than 0.4 GeV. The primary vertex for each event is determined as the vertex with the highest sum of the squares of the transverse momentum for all associated ID tracks [60].

Electron candidates are identified by energy deposits in the EM calorimeter [61] that are associated with tracks in the ID. The pseudorapidity of the calorimeter energy cluster must be $|\eta_{cluster}| < 2.47$, excluding the area where

the barrel and endcap calorimeters overlap. Muon candidates are reconstructed by combining tracks from the ID and tracks from the MS [62] and must have $|\eta| < 2.5$. Both the electron and muon candidates must have p_T greater than 28 GeV.

3.2 Monte Carlo simulated samples

To model signals and background processes, MC event generators are employed. These generators take into account multiple proton-proton interactions per bunch crossing and the detector's response to collisions from bunch crossings before or after the one containing the hard interaction. The simulated samples are reweighted to match the distribution of the number of proton-proton collisions in the data. The detector response is modelled using a combination of Geant4 [63] simulation for the inner tracker and muons and a fast simulation [64] for shower development in the calorimeter. The same algorithms used for data reconstruction are applied to these simulated samples. To account for small differences in efficiencies, momentum scales, and momentum resolutions between data and simulation, scale factors and smearing techniques are used to correct the simulated response.

The SM expectation is used to model the production of $t\bar{t}t\bar{t}$ events. The nominal sample used to model the $t\bar{t}t\bar{t}$ signal was generated with the MadGraph5 aMC@NLO v2.7.3 [65] generator, which generates matrix elements (ME) at NLO in the strong coupling constant α_S, with the NNPDF3.0nlo [66] PDF set. Pythia 8.2 [67] with the A14 set of tuned parameters [68] and the NNPDF2.3lo

PDF set was used to simulate the parton shower, fragmentation and underlying event. The mass of the top qaurk (m_t) in the simulated samples is set to 173 GeV.

Top quark pair production with additional jets ($t\bar{t}$+jets) is the most important background process, accounting for 90-95% of the background. Single top quark processes, including t-channel and tW production, are the next most important. These are followed by Z+jets and W+jets electroweak processes, which only consider the leptonic decay modes of the bosons. The following rare processes, $t\bar{t}+Z, W, H$, involve the production of a top quark-antiquark pair as well as a Z, W, or h boson. Finally, we consider $t\bar{t}$ production in conjunction with dibosons, $t\bar{t}WW$, $t\bar{t}WZ$, $t\bar{t}ZZ$, $t\bar{t}WH$, $t\bar{t}ZW$, $t\bar{t}HH$, and triple top quark production ($t\bar{t}t$+jets and $t\bar{t}tW$), which we refer to as $t\bar{t}$XY. In the simulation, $t\bar{t}$+Z and $t\bar{t}+H$ are grouped together based on their signature resemblance and comparability of production rates to the $t\bar{t}t\bar{t}$ signal, while $t\bar{t}+W$ and $t\bar{t}$XY are grouped together.

The Sherpa 2.2.1 [69] generator with the NNPDF3.0nlo PDF set was used to generate the $t\bar{t}W$ simulated events. The calculation of the matrix element was performed using the Comix [69] and OpenLoops [70] libraries, which are capable of including one additional parton in the NLO QCD calculation. The five-flavour scheme, including both c and b-quarks, was also used to consider two partons at LO QCD. The resulting calculation was merged with the Sherpa parton shower [71] using the MEPS@NLO [72] method and with a merging scale of 30 GeV. The simulated $t\bar{t}W$ sample is normalized to the

$601\,\text{fb}^{-1}$ cross-section computed at NLO in QCD with the leading NLO electroweak corrections [73]. Using MadGraph5 aMC@NLO v2.7.3 generator, with the same PDF used in the nominal sample, an alternative $t\bar{t}$W sample was generated at NLO in QCD with no additional partons. Using the A14 tune and the NNPDF2.3lo PDF set, the events were interfaced with Pythia 8.2.

The production of $t\bar{t}$ and single-top-quark events were modelled using the Powheg-Box [74] v2 generator at NLO in QCD with the NNPDF 3.0 nlo PDF set. The diagram removal technique [75] was used to remove the overlap between the $t\bar{t}$ and tW final states. The $t\bar{t}$ and single-top-quark simulated samples are normalized to the cross-sections calculated in QCD at next-to-next-to-leading order (NNLO), including the resummation of NNLL soft-gluon terms [76].

The production of $t\bar{t}H$ events was modelled using the Powheg-Box v2 generator at NLO in QCD and the five-flavour scheme, with the NNPDF 3.0 nlo PDF set. The simulated sample is normalized to the NLO cross-section in QCD with the leading NLO electroweak corrections [73]. An alternative sample generated with the same settings as the MadGraph5 aMC@NLO v2.7.3 generator is used to assess the uncertainty in $t\bar{t}$H modelling caused by the generator choice.

The WH and ZH processes were generated with Pythia 8.2 generator, A14 tune and NNPDF 2.3 lo PDF set. Their theoretical cross-sections were calculated at NNLO in QCD and NLO electroweak accuracies. Sherpa 2.2.1 generator was used to model the production of diboson (VV) and triboson (VVV)

events, using NNPDF 3.0 nnlo PDF set. For the simulation of Z+jets and W+jets production, the Sherpa 2.2.2 generator with the same PDF set was used. The theoretical cross-sections calculated at NLO QCD were used to normalize the VV and VVV samples, and the Z+jets and W+jets backgrounds are used to normalize the NNLO cross-sections [77].

The signal and background samples were modelled through Geant4 [63], which simulated the ATLAS detector geometry and response [34], followed by reconstruction using the same software used for actual collider data. Some of the alternative samples used to assess systematic uncertainties were instead subjected to a fast detector simulation using parameterised showers in the calorimeters [78]. Corrections were made to the simulated events so that the selection efficiencies, energy scales and energy resolutions of the physics objects matched those determined from data control samples.

Chapter 4

Event Reconstruction and Selection

The physical objects used in the analyses must be reconstructed from the data obtained by the detector. This data is derived primarily from hits and energy depositions left by collision particles in the detector. After the proton-proton collision, the vertex with the largest sum of squared transverse momentum (p_T^2) of tracks with $p_T > 0.4\,\text{GeV}$ is identified as the primary vertex. The physical objects, such as electrons, muons, and jet candidates are then reconstructed, with each having to meet specific minimum requirements as outlined in subsequent sections.

4.1 Leptons

Candidates for events must have at least one vertex with at least two associated tracks with p_T greater than 400 MeV. The leptons that we will focus on in this study are electrons and muons.

4.1.1 Electrons

Electron candidates are reconstructed from energy deposits associated with ID tracks in the EM calorimeter. Electron candidates must meet specific requirements, such as having a calorimeter energy cluster with a pseudo rapidity of $|\eta| < 2.47$, excluding the area between the barrel and endcap calorimeters with a pseudorapidity $1.37 < |\eta| < 1.52$, and having a transverse momentum $p_T > 25\,\text{GeV}$. This exclusion is due to the fact that the transition region contains a large amount of material and the electromagnetic objects are measured with lower resolution and efficiency, thus we will not use these objects in this analysis.

Electrons must be identified using the 'Tight' likelihood-based criterion and isolated through a set of criteria based on the calorimeter's topological clusters and the surrounding ID tracks. The transverse impact parameter ($|d_0|$) divided by its uncertainty ($\sigma(d_0)$) must be less than 5. The longitudinal impact parameter ($|z_0 sin(\theta)|$) must be less than 0.5 mm.

4.1.2 Muons

Muons are reconstructed by combining tracks from the ID and tracks from the MS [79]. To be considered valid, the muon candidates must have a pseudorapidity of less than 2.5 and a transverse momentum of at least 25 GeV. Muons must meet the 'Medium' cut-based identification criterion, which includes requirements on the number of ID and MS hits as well as compatibility between momentum measurements in the ID and MS. Muons must also meet the isolation requirement based on the properties of the ID tracks surrounding

the reconstructed muon. The muon candidates must have a $|d_0|/\sigma(d_0)$ value that is less than three and a $|z_0 sin(\theta)|$ value that is less than 0.5mm. These requirements are similar to those for electrons.

4.2 Jets

Jets are reconstructed using the anti-k_t algorithm [80] from the energy deposits in the calorimeters as topological clusters [81] and then calibrated as described in Ref.[82]. To be considered, the jets must have a transverse momentum of $p_T > 25\,\text{GeV}$ and be located in the pseudorapidity range of $|\eta| < 2.5$. To mitigate the effect of additional proton-proton collisions in the same or nearby bunch crossing, known as pile-up, jets with $p_T < 120\,\text{GeV}$ and $|\eta| < 2.4$ are only considered when they meet a requirement based on the response of a multivariate classifier known as the jet-vertex-tagger [83].

A set of selection criteria are applied to eliminate events containing jets originating from non-collision sources or detector noise [84]. To identify jets originating from b-hadrons, the MV2c10 multivariate algorithm [85] is used. A b-tagged jet is defined as one that meets a specified set of criteria, resulting in an average efficiency of 77% for identifying b-quark jets in simulated $t\bar{t}$ events while achieving rejection factors of 110 and 4 against light-quark or gluon jets and c-quark jets, respectively.

4.3 Missing Transverse Momentum

The missing transverse momentum, referred to as E_T^{mis}, in an event is calculated as the negative sum of the transverse momentum of all reconstructed and calibrated objects in the event, including the momentum of unmatched ID tracks associated with the primary vertex [86]. The tracks considered must have a transverse momentum greater than 0.4 MeV and be connected to the hard scatter vertex, according to strict criteria on longitudinal and transverse impact parameters. The coverage of the E_T^{mis} pseudorapidity is limited to $|\eta| < 2.5$, which is less than calorimeter coverage of $|\eta| < 4.9$.

4.4 Preselection

The event selection criteria are based on the standard procedures utilized by ATLAS for top quark physics studies. The primary purpose of these criteria is to eliminate events that are not relevant to the signal under study and only retain events that correspond to the desired final state. The selected events must have either one same-sign lepton pair or three isolated leptons without any charge specification. Additionally, at least one of the reconstructed leptons must correspond to a lepton that activated the trigger in each event. To reduce the charge misassignment background caused by low-mass resonances and Z-boson decay, events with two same-sign electrons must have an invariant mass $m_{ee} > 15\,\text{GeV}$ and $|m_{ee}$ - $91\,\text{GeV}| > 10\,\text{GeV}$. To minimize interference from the decay of Z-bosons in events with three or more leptons, all opposite-sign, same-flavor lepton pairs must satisfy $|m_{ll}$ - $91\,\text{GeV}| > 10\,\text{GeV}$.

The events from $t\bar{t}t\bar{t}$ production are selected based on their high number of light-flavor jets and b-tagged jets and their high overall activity. The sum of the transverse momentum of the isolated leptons and jets in the event, referred to as H_T, is used to measure this activity. The SR requires at least six jets, at least two b-tagged jets and $H_T > 500\,\text{GeV}$.

4.4.1 Signal region

The signal region (SR) is determined by a phase space with high signal sensitivity and low background contamination. The production of top-quark pairs in association with additional jets and/or bosons dominates the SR's background composition. The SR requires two leptons of the same charge with any flavour ($e^{\pm}\mu^{\pm}$,$e^{\pm}e^{\pm}$,$\mu^{\pm}\mu^{\pm}$) or three isolated leptons without charge requirement. The leptons considered here need to satisfy $p_T > 25\,\text{GeV}$ in both channels. The jets required need to satisfy $p_T > 25$ and $|\eta| < 2.5$. To maximize signal sensitivity, the jet p_T requirements are optimized. Figure 4.1 shows the jet multiplicity in the first row and the b-tagged jet multiplicity in the second row after applying the preselection mentioned in section 4.4. From the distributions of the number of jets, we could see that the SM $t\bar{t}t\bar{t}$ in the first two bins ($N_{jets} = 6,\ 7$) is slightly dominating over the two signal samples ($t\bar{t}t\bar{t}, t\bar{t}ZH$). This explains the reason why these channels require a large jet activity where the signal is so sensitive. Beyond the first two bins, the signal appears quite well as you can see the signals are high over SM $t\bar{t}t\bar{t}$. The b-jets requirement also places a significant role in displaying the signal. From the distributions of the number of b-jets in both channels in the first bin,

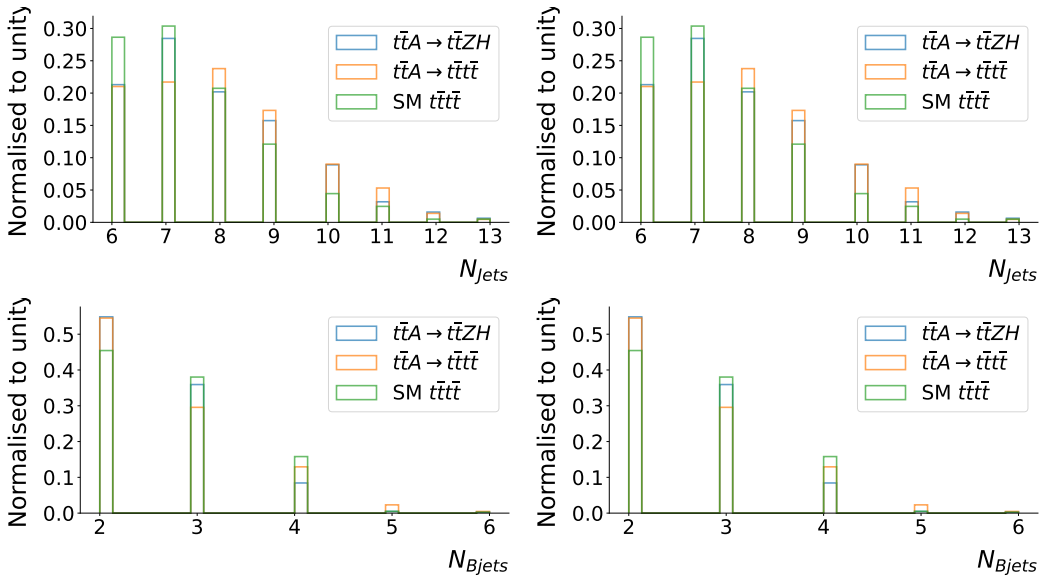


FIGURE 4.1: Distrubtions of the number of jets after applying preselection requirements. The distribution on the left (right) is for $2SS\ell$ (3ℓ).

the signal is visible enough over the background. This shows that the signal regional chosen is sufficient to perform our analysis.

Chapter 5

The production of four top quarks in SM and BSM

5.1 SM four top quark

Four-top production is a rare type of SM process that occurs during proton-proton collisions through gluon fusion or less frequently, the annihilation of a quark and antiquark. Gluons can mediate this process in QCD, photons in electromagnetism, Z bosons in weak forces, or a Higgs boson, as depicted in figure 5.1. The cross-section of the $t\bar{t}t\bar{t}$ production is a sensitive indicator of the CP properties and strength of the Yukawa couplings of the top quark to the Higgs boson, as it can be generated through an off-shell Higgs boson as described in Refs. [87, 88]. The SM predicts a cross-section of 12 fb for this process, which is of the order of 10^3 smaller relative to the cross-section of the top quark pair ($t\bar{t}$) which is predicted to be around 832 pb [76].

In four-top production, each of the four top quarks decays into a W boson and a b-quark as shown in figure 5.2. The b-quark forms a hadron, which may

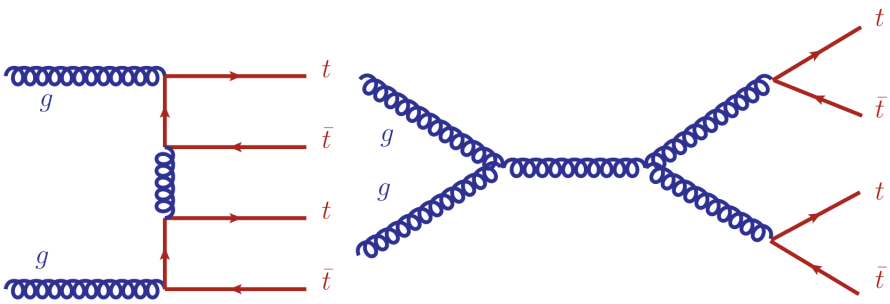


FIGURE 5.1: The Feynman diagram for the production of four top quarks at the leading order in QCD.

appear as a b-jet in the detector. The W boson further decays either into a quark-antiquark pair that forms hadrons, which appear as jets in the detector or into a lepton and its associated neutrino. The neutrino passes through the detector and can only be inferred from missing energy (E_T^{mis}). Leptons that are muons or electrons create tracks or showers, respectively, but taus pose a challenge as they may also decay hadronically and are difficult to reconstruct. Thus, in this study when we refer to leptons, we refer to electrons and muons only.

Based on how all four top quark decay, there are a variety of possible final states with varying lepton multiplicities, as outlined in table 5.1. The four-top analysis aims to identify a specific final state pattern in comparison to backgrounds. Each channel has a unique probability of occurring and predicts a specific number of events. These channels also have backgrounds that closely resemble the signal and must be distinguished from it. In order for a $t\bar{t}t\bar{t}$ signal to be detected, it must have a sufficient number of events that stand out statistically when compared to the background. These states are separated into two analysis channels, determined by the main sources of background: the

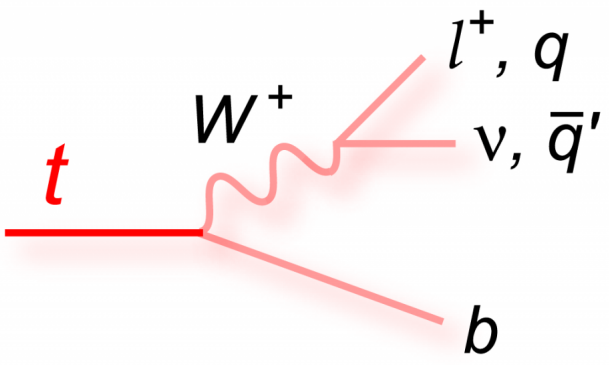


FIGURE 5.2: The decay of the top quark hadronically and leptonically [89].

SSML channel (comprising of same-sign dilepton and multilepton events) and the 1L / 2LOS channel (composed of single lepton and opposite-sign dilepton events).

In the SM, several processes can create a similar outcome to the production of four top quarks. Specifically, when the final state includes two leptons with the same charge or three leptons, the probability of SM processes producing this channel is very low due to this kind of event being rare in the SM. The main source of these types of events is from instrumental backgrounds, where objects are incorrectly identified or reconstructed.

Background from top-quark pair production ($t\bar{t}$) is similar to the signal due to its resemblance to it. However, the signature of $t\bar{t}$ is different as it contains half the number of expected top quarks. Discrimination is complicated by reconstruction efficiencies, the potential radiation of extra jets or overlap with pileup. The most sensitive four-top channels are those that reduce this

Finale State	**Bakground**	**Analysis channel**
All hadronic	31 %	
Single lepton	42 %	
Opposite-sign dilepton	14 %	1L / 2LOS
Same-sign dilepton	7 %	
≥ 3 leptons	5 %	SSML

TABLE 5.1: Summary of the $t\bar{t}t\bar{t}$ possible finale states.

background, namely signatures where two or more of the four tops, of which two are (anti)tops, decay leptonically into at least two leptons of the same sign (so two (anti) electrons, two (anti) muons, or an (anti) muon-(anti) electron pair). This eliminates $t\bar{t}$ as a background except in the case of lepton sign mis-reconstruction, making such "same-sign dilepton and multilepton" final states the most sensitive despite not having the largest branching ratios.

In this study, we focus on events that have either two isolated leptons with the same charge (2SSℓ) or at least three isolated leptons (3ℓ). Despite their relatively low branching fractions, these channels provide promising results due to their clean nature and small background from the SM. The addition of a Higgs-doublet to the SM introduces a scalar spectrum with two CP-even bosons (h, H), one CP-odd boson (A) and charged scalar bosons ($H^{\pm}$), which offers opportunities to study the scalar spectrum. Our aim is to examine the CP-odd scalar in the 2HDM+S model by investigating the production of A in association with two top quarks and its decay into $t\bar{t}$ ($A \to t\bar{t}$), using the 2SSℓ and 3ℓ channels.

5.2 The multilepton channel

In this analysis channel, events that contain two leptons with the same charge or at least three leptons are selected. Under these conditions, four top quarks still produce many jets in the final state, with a minimum of six jets expected, many of which are likely to originate from b-quarks, referred to as b-jets. To account for this, at least six jets are required in the selected events and two of these jets must be identified as b-jets by the b-tagging algorithm. The high lepton multiplicity in the final state also leads to a significant number of events with non-prompt or fake leptons. To reduce this contribution, especially for objects with lower p_T, an additional cut of $H_T > 500$ GeV is applied, where H_T is the sum of the p_T of all leptons and jets in the final state.

These selections define the signal region for four top quarks, and the main backgrounds come from $t\bar{t}W$ production ($\sim$ 37%) and instrumental backgrounds ($\sim$ 25%). Accurately estimating these two backgrounds is essential for a precise measurement of the four top quark process. Instrumental backgrounds are further divided into different classes depending on their origin. The most significant sources include non-prompt leptons from heavy-flavor hadron decays, electron pairs from photon conversion or Drell-Yann decays, and charge misidentification. The background from charge misidentification comes from events with pairs of opposite-sign leptons that are reconstructed as same-sign pairs. This background is determined by measuring the charge misidentification probability.

The charge misidentification probability is calculated by comparing the number of events that have same-sign pairs of leptons and the number of events that have opposite-sign dilepton pairs that reconstruct the Z-boson resonance. These computed values are then used to adjust events in the opposite-sign dilepton region to match the same-sign one. The other backgrounds are estimated by utilizing a template method, where the normalization is measured in dedicated control regions and the shapes are assumed to be similar to those from MC simulations.

The template method used to estimate the other backgrounds involves measuring the normalization in dedicated control regions and using Monte Carlo simulations to determine the shape of the distribution. This method allows for a more precise measurement of the backgrounds, as it accounts for variations in the shapes of distributions due to different kinematic properties or acceptance. Additionally, the control regions are designed to be as similar as possible to the signal region, but with a higher expected background contribution. This allows for more accurate measurement of background normalization. Furthermore, the template method also allows for a better understanding of systematic uncertainties by comparing the shapes of the backgrounds in data and simulation. It is worth noting that background estimation is crucial for the interpretation of the results and the significance of any observed signal. Therefore, a robust estimation of the background is essential for a precise measurement of the four top quark process.

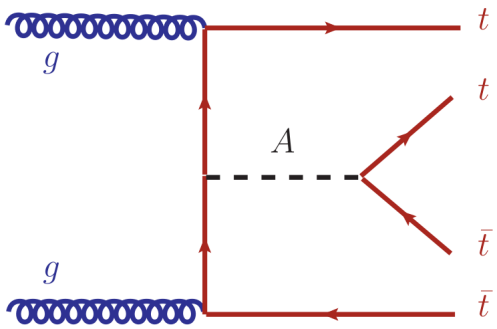


FIGURE 5.3: The Feynman diagram for the production of four top quarks in association with heavy-pseudoscalar (A) decaying to top quark pair.

5.3 BSM four top quark

The alignment limit in physics offers a useful framework for understanding the production of heavy Higgs bosons. In this limit, with a small value of tan β, the couplings between HW^+W^- and $b\bar{b}H(A)$ are reduced, making the main contributions to $H(A)$ production come from the $t\bar{t}H(A)$ vertex. This results in a range of production processes involving $t\bar{t}$ pairs, which can be generated from the SM top production processes with the addition of a heavy scalar radiated from the internal top quark as seen in figure 5.3.

Many BSM models have been predicted to enhance the cross-section of the production of for top-quarks due to the top quark having a large coupling to the SM higgs. One of these models is the production of a heavy pseudo-scalar boson in association with a top-quark pair ($t\bar{t}$) using the type-II 2HDM.

A number of anomalies related to the multilepton final states have been explained using a 2HDM+S model, where the mass of S and H are set to $m_S = 150\,\text{GeV}$ and $m_H = 270\,\text{GeV}$. Using the same model, we consider the

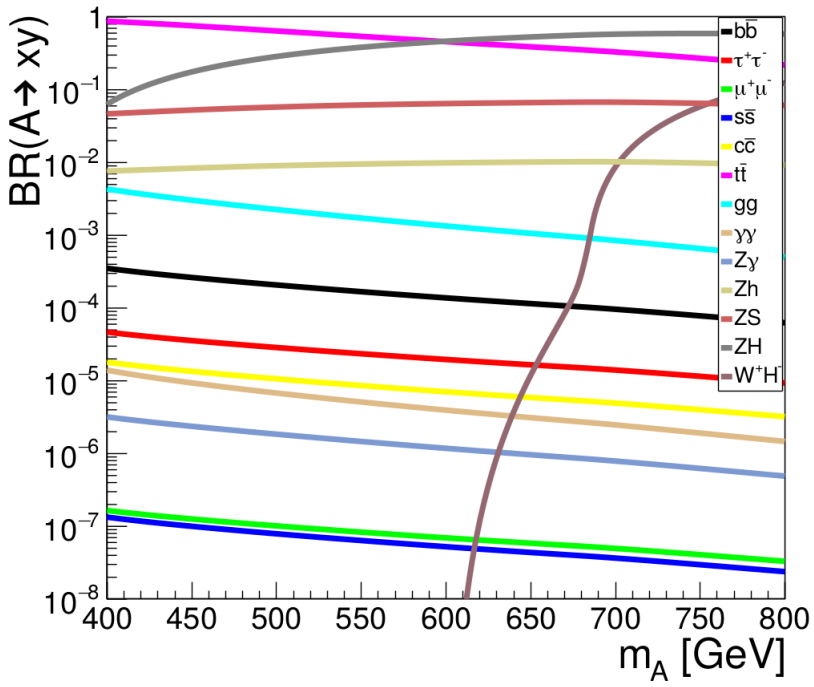


FIGURE 5.4: The plot of the branching ratios ($A \to xy$) against the mass of the heavy pseudoscalar m_A. it is shown that the heavy pseudoscalar A decays predominantly to ZH if its mass is above 600 GeV and when it is below 600 GeV $t\bar{t}$ predominates [90].

assocated production $t\bar{t}A$, with $A \to t\bar{t}$, leading to four top quarks as well as $A \to ZH$. The mass of the heavy pseudo-scalar will be examined between 400 GeV and 600 GeV, where the two leading decays are ZH and $t\bar{t}$ as seen in figure 5.4.

5.3.1 Heavy pseudo-scalar in the mass range 400-600 GeV

The situation regarding heavy resonance searches remains inconclusive, as a number of analyses have not been reported by both the ATLAS and CMS experiments, as well as the entire Run 2 data sets. In some important areas,

results are only available with partial Run 2 data sets and from a single experiment. Nonetheless, a number of minor excesses in the mass range 400-600 GeV have emerged in recent years. These are summarized in table 5.2, which also reports the mass for which the significance of the deviation occurs.

The CMS experiment reported an excess in the search for $t\bar{t}$ resonances with $35.9\,\text{fb}^{-1}$ of integrated luminosity at $\sqrt{s} = 13\,\text{TeV}$ [91]. The maximum excess occurs at 400 GeV, with a 3.5σ (1.9σ) local (global) significance. The ATLAS experiment reported a search with $20.3\,\text{fb}^{-1}$ at $\sqrt{s} = 8\,\text{TeV}$, with significantly less sensitivity [92], where no results are available with Run 2 data. It has been observed that this excess appears at the threshold for $t\bar{t}$ production. Weak corrections are significant and may contribute to the excess [93].

The ATLAS experiment reported with $139\,\text{fb}^{-1}$ of integrated luminosity results for the search for $\tau^+\tau^-$ resonances with at least one τ-lepton decaying into hadrons [94]. In order to isolate the $bbH(A)$ production mechanism, the search is conducted comprehensively but also in conjunction with b-quarks. The inclusive and b-quark associated searches yield excesses of 2.2σ and 2.7σ at 400 GeV, respectively. The CMS experiment yields no results, despite the fact that the data set corresponds to $35.9\,\text{fb}^{-1}$ of integrated luminosity [95].

The ATLAS experiment reported an excess in the search for a pseudoscalar $A \to Zh$, where $Z \to \ell^+\ell^-, \nu\nu$ where $\ell = e, \mu$, and $h \to b\bar{b}$ [100]. This includes the search in association with b-quark. An excess is seen at 440 GeV with a local (global) significance of 3.6σ (2.4σ). The corresponding search by the

Final state	Mass (GeV)	Significance (σ)	Ref.
$t\bar{t}$	400	3.5 (1.9)	[91]
$\tau^+\tau^-(+b)$	400	2.2-2.7	[94]
$ZH \to \ell\ell b\bar{b}$	610 (290)	3.1 (1.3)	[96]
$ZH \to \ell\ell b\bar{b}$	630 (270)	2.5	[97]
$ZH \to \ell\ell b\bar{b}$	630 (160)	3.9 (1.3)	[97]
$b\bar{b}H \to b\bar{b}b\bar{b}$	640	2	[98, 99]

TABLE 5.2: Excesses in the search for (pseudo-)scalars in the 400-600 GeV range are summarized. The experiment's local significances are given. When available, global significance is shown in round brackets. The mass given in brackets for the final states of $ZH \to \ell\ell b\bar{b}$ corresponds to H.

CMS experiment sees a significantly smaller excess of 1-2σ at 400 GeV for the inclusive analysis and the search in association with b-quarks, respectively. No excess is seen at 440 GeV, thus diminishing the excess observed by ATLAS.

Another group of excesses includes structures just above 600 GeV, which are mostly caused by $A \to ZH \to \ell\ell b\bar{b}$ decays, where $\ell = e, \mu$. Searches are carried out in a two-dimensional plane (m_A, m_H), where m_A and m_H represent the masses of the pseudo-scalar and heavy scalar, respectively. Using 139 fb^{-1} of integrated luminosity [96], the ATLAS experiment reported an excess at $(610, 290)$ GeV with a local (global) significance of 3.1σ (1.3σ). The CMS experiment reported on the same search, but with a lower integrated luminosity of 35.9 fb^{-1}. The largest excess is reported for $(630, 160)$ GeV with a local (global) significance of 3.9σ (1.3σ), where ATLAS sees no excess [97]. CMS does, however, detect a local excess of 2.5σ at $(630, 270)$ GeV, which is consistent with the ATLAS results. CMS does, however, detect a local excess of 2.5σ at $(630, 270)$ GeV, which is consistent with the ATLAS results. ATLAS

finds a small local excess with $\approx 2\sigma$ around 640 GeV using the $b\bar{b}H \to b\bar{b}b\bar{b}$ final state with only 27.8 fb^{-1} integrated luminosity.

In light of this, it is worthwhile to investigate the possible compatibility of a pseudoscalar state in the mass range 400-600,GeV with the multi-lepton final states described above within the context of a 2HDM+S model.

5.4 Discriminatory variables

The discriminant variables were identified as the leading jet transverse momentum, $p_T^{j_0}$, the second leading jet transverse momentum, $p_T^{j_1}$, the sixth leading jet transverse momentum, $p_T^{j_5}$, total number of b-tagged jets, N_b, the leading b-tagged jet transverse momentum, $p_T^{b_0}$, the leading lepton transverse momentum, $p_T^{\ell_0}$, the missing transverse energy, E_T^{mis}, the scalar sum of transverse momenta over all leptons and jets excluding the leading jet, H_T, the minimum distance defined as $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$ between two leptons out of all possible pairs, $\Delta R_{min}^{\ell\ell}$, the sum of the distances between two leptons for all possible pairs, $\Delta R_S^{\ell\ell}$, the maximum distance between a b-tagged jet and a lepton among all possible pairs $\Delta R_{max}^{b\ell}$, and the minimum distance between a jet and a b-tagged jet among all possible pairs ΔR_{min}^{bj}. The distributions for the process $t\bar{t}A \to t\bar{t}t\bar{t}$ in the $2SS\ell$ channel are shown in figure 5.5,5.7 and 5.9 for the three masses of the heavy pseudoscalar. The distributions for the process $t\bar{t}A \to t\bar{t}t\bar{t}$ in 3ℓ channel are shown in figure 5.6, 5.8 and 5.10 for the three masses of the heavy pseudoscalar. The distribution of the process $t\bar{t}A \to t\bar{t}ZH$ in the $2SS\ell$ channel are shown in figure 5.11 and for the 3ℓ channel are shown

in figure 5.12 for a mass of the heavy pseudoscalar at $m_A = 500\,\text{GeV}$. For other masses the distributions are assumed to show similar characteristics to the distributions shown in figure 5.11 and 5.12. These kinematic variables will be used to train a classification model. It will classify the SM four top production from the BSM four top production using a deep neural network algorithm.

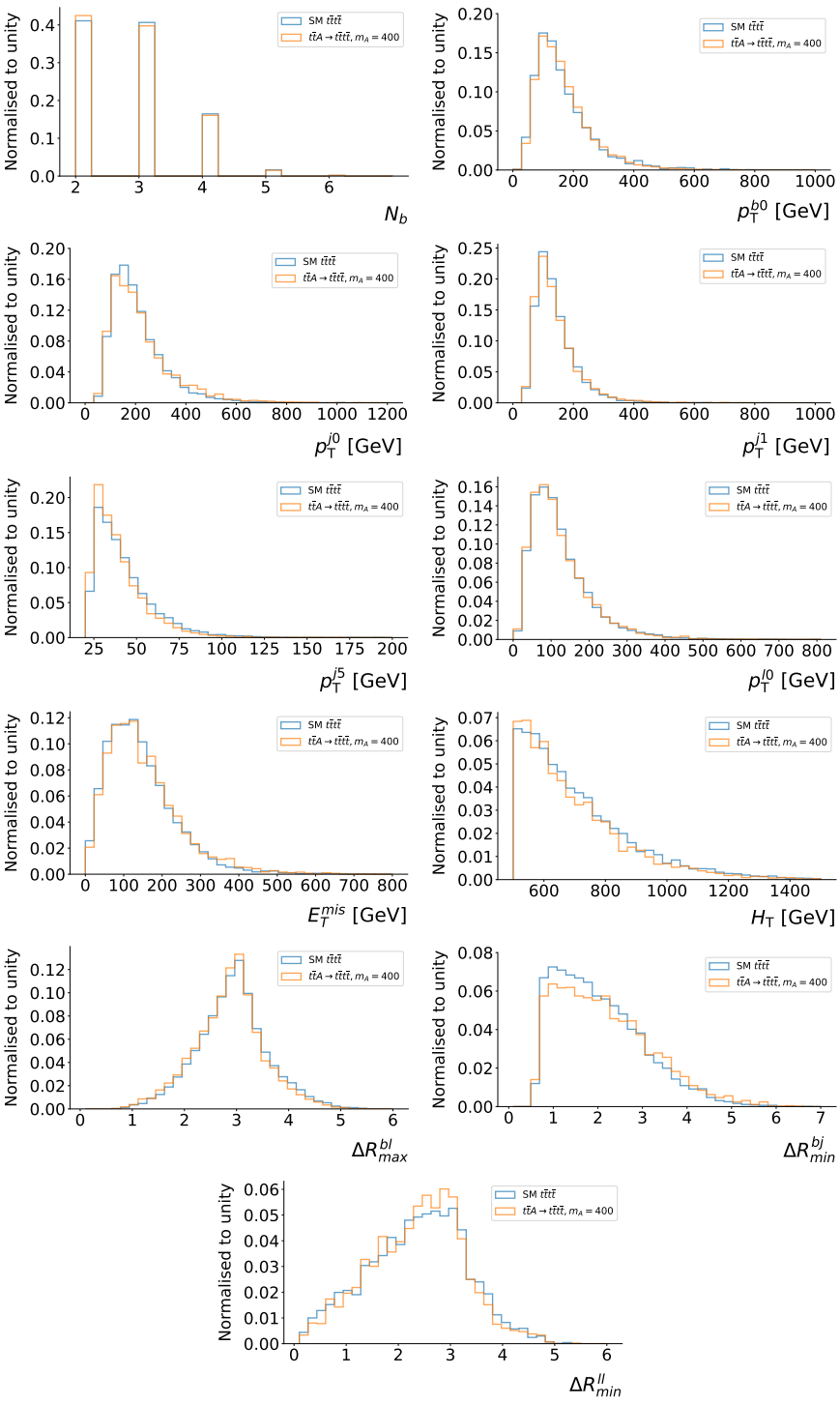


FIGURE 5.5: Distributions of the kinematic variables in the $2SS\ell$ channel at mass of the pseudo-scalar $m_A = 400\,\text{GeV}$ after preselection. We represent the signal and background distributions in blue and orange-colored histograms.

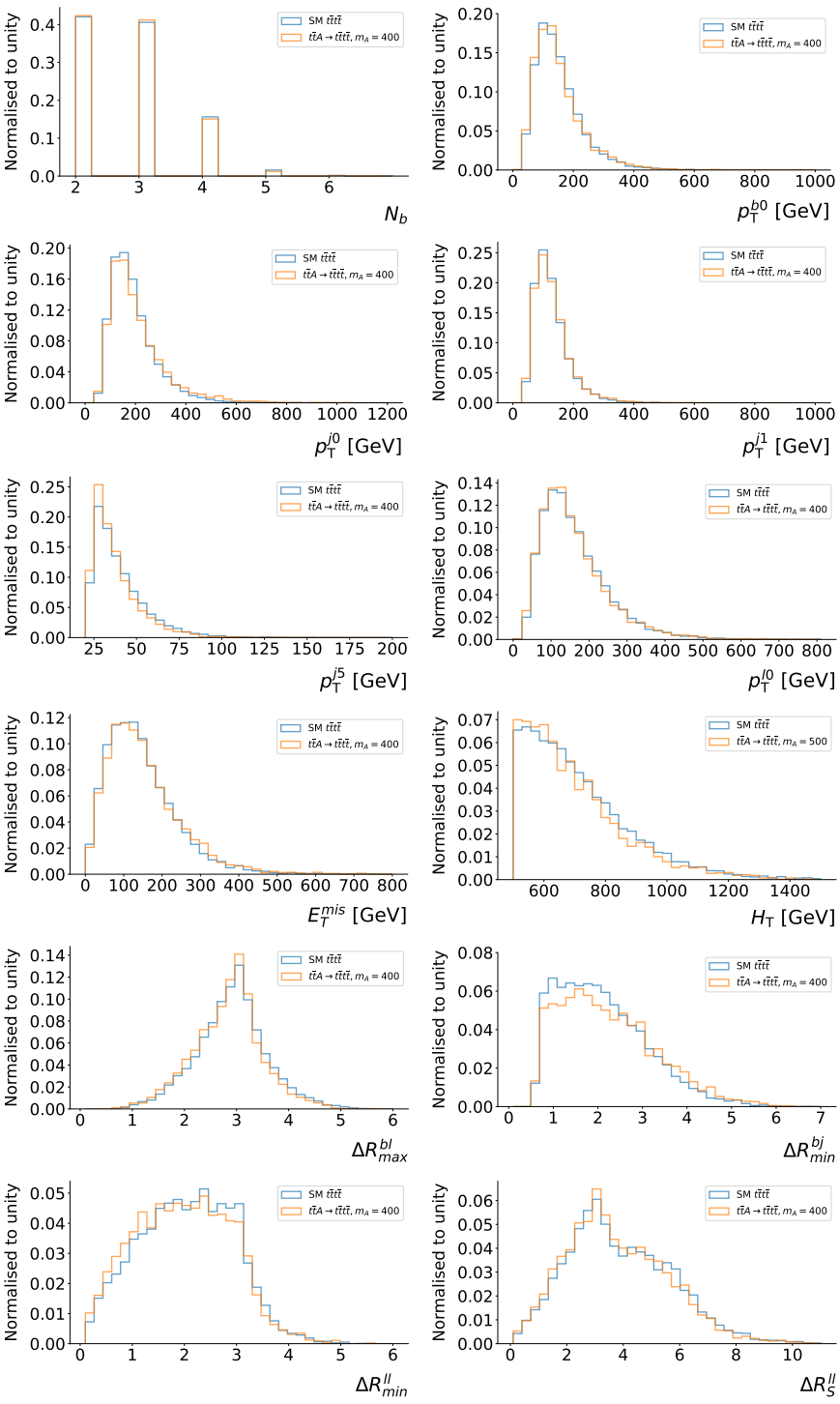


FIGURE 5.6: Distributions of the kinematic variables in the 3ℓ channel at mass of the pseudo-scalar $m_A = 400\,\text{GeV}$ after preselection. We represent the signal and background distributions in blue and orange-colored histograms.

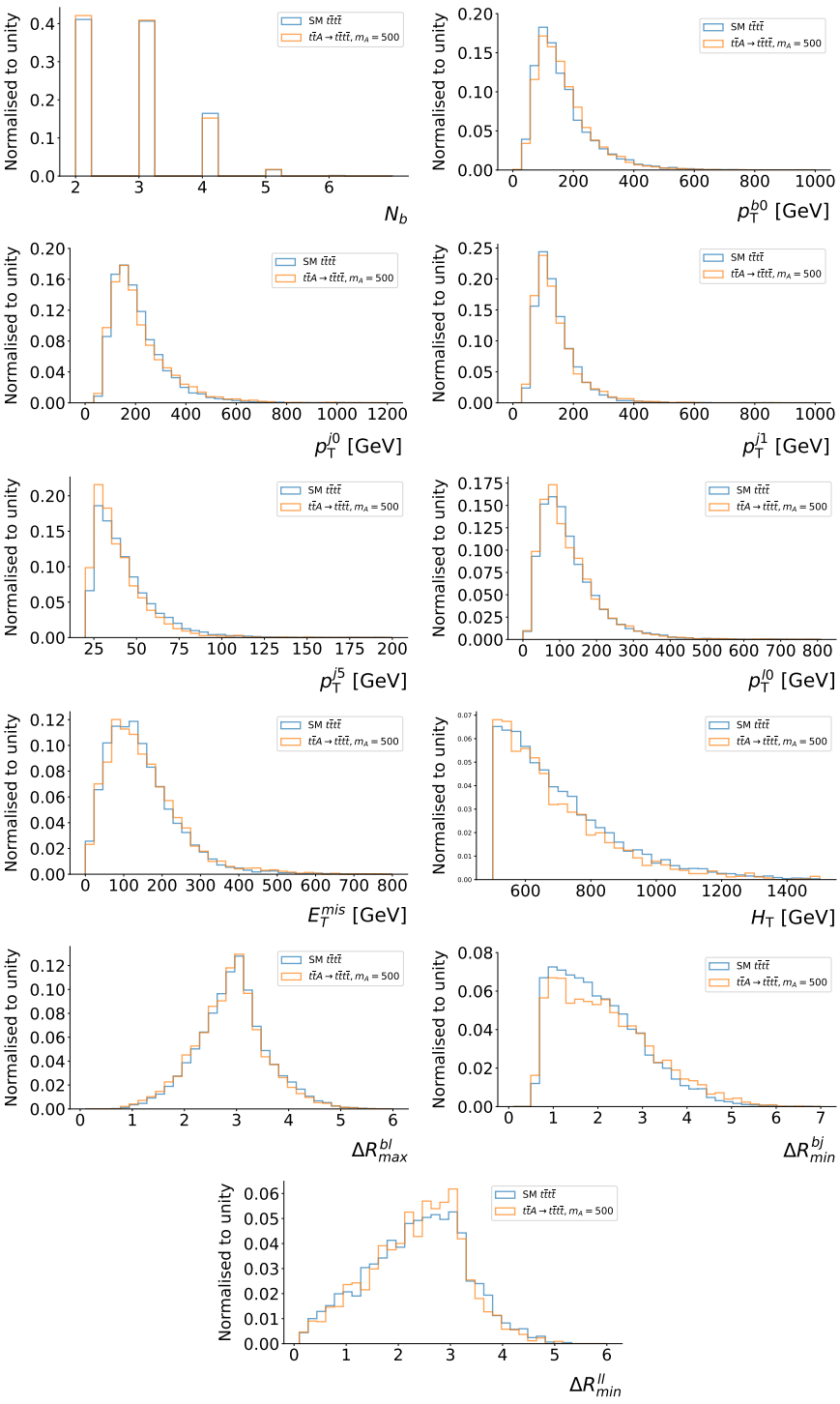


FIGURE 5.7: Distributions of the kinematic variables in the $2SS\ell$ channel at mass of the pseudo-scalar $m_A = 500\,\text{GeV}$ after preselection. We represent the signal and background distributions in blue and orange-colored histograms.

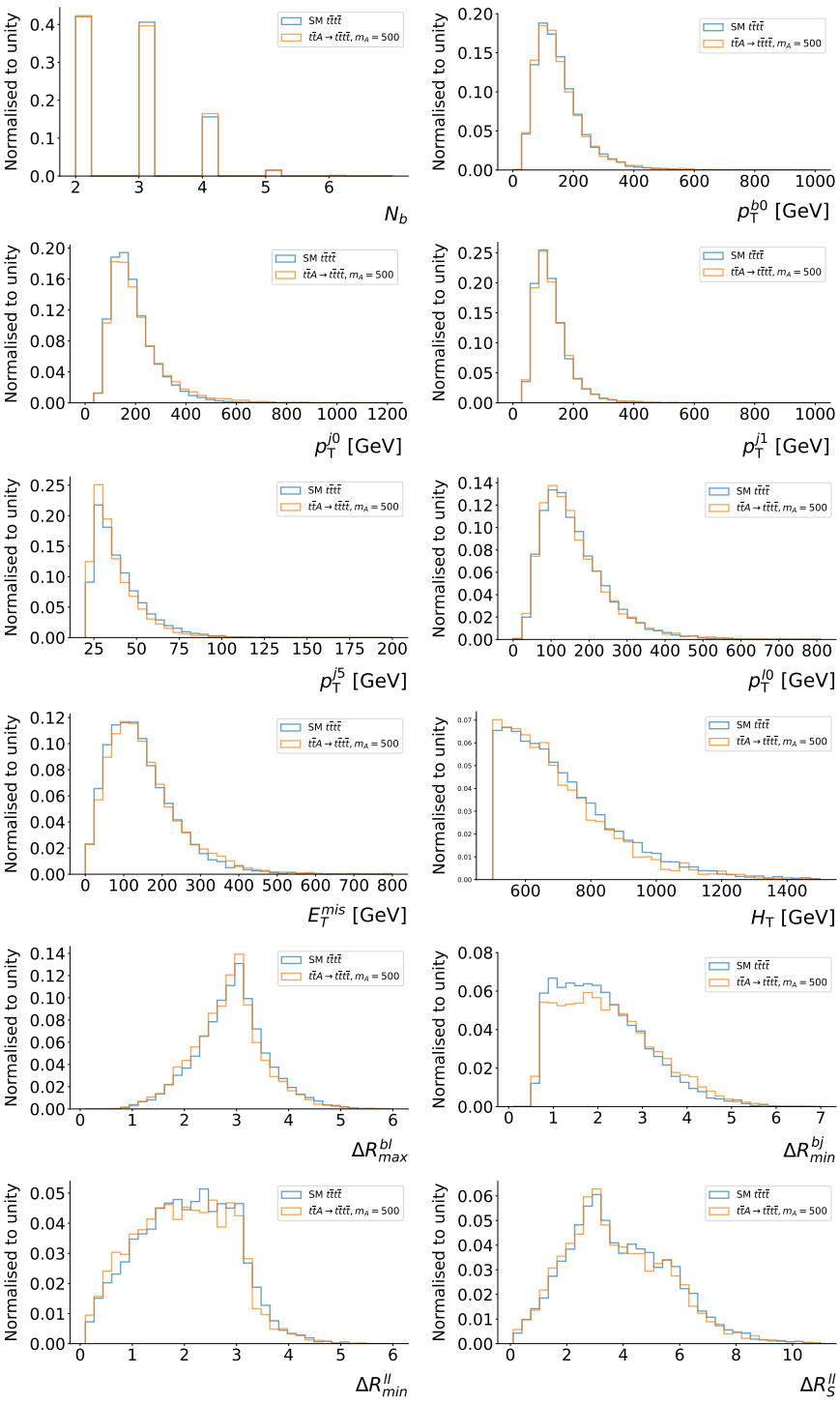


FIGURE 5.8: Distributions of the kinematic variables in the 3ℓ channel at mass of the pseudo-scalar $m_A = 500\,\text{GeV}$ after preselection. We represent the signal and background distributions in blue and orange-colored histograms.

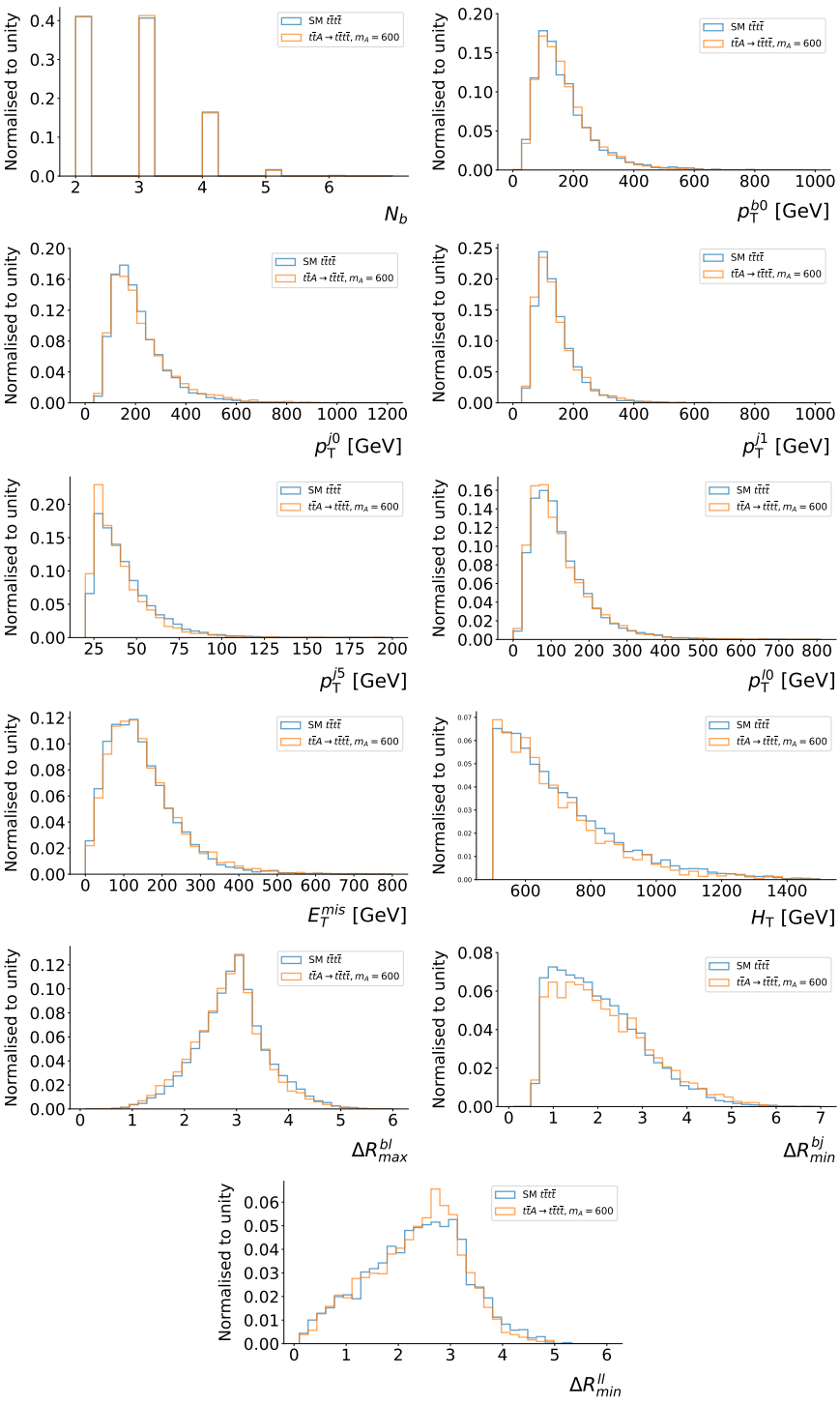


FIGURE 5.9: Distributions of the kinematic variables in the $2SS\ell$ channel at mass of the pseudo-scalar $m_A = 600\,\text{GeV}$ after preselection. We represent the signal and background distributions in blue and orange-colored histograms.

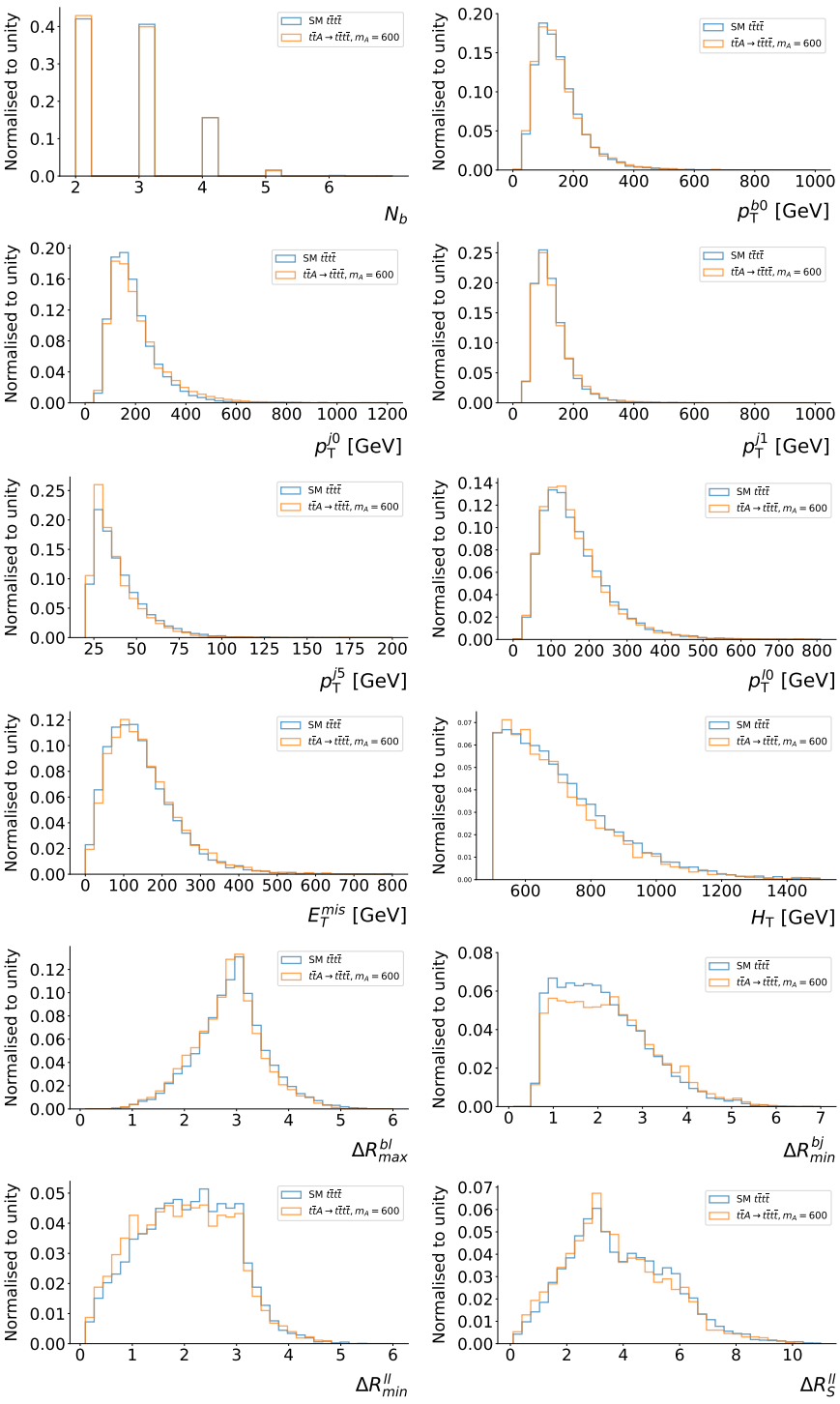


FIGURE 5.10: Distributions of the kinematic variables in the 3ℓ channel at mass of the pseudo-scalar $m_A = 600\,\text{GeV}$ after preselection. We represent the signal and background distributions in blue and orange-colored histograms.

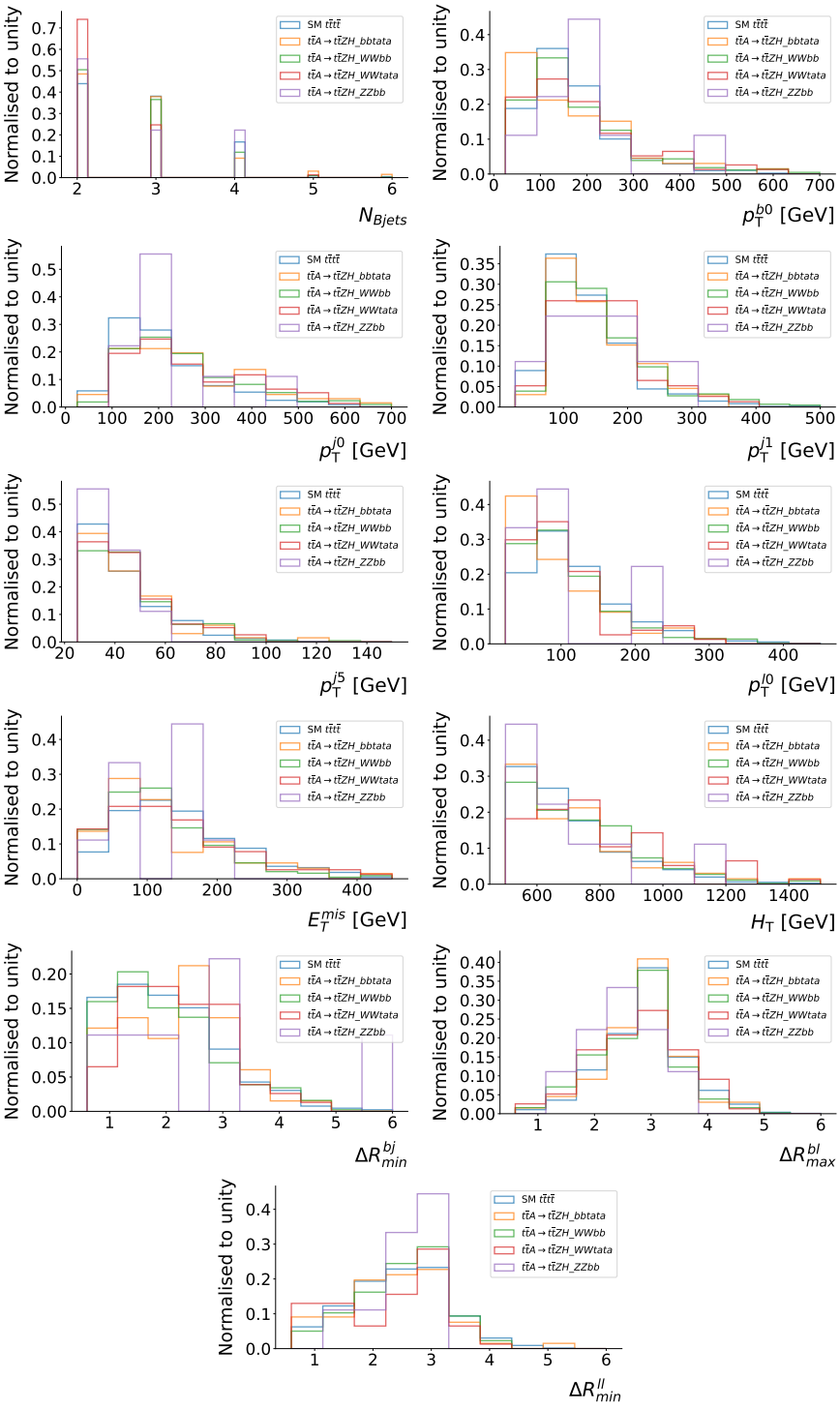


FIGURE 5.11: Distributions of the kinematic variables for $t\bar{t}A \to t\bar{t}ZH$ in the 2SSℓ channel at mass of the pseudoscalar $m_A = 500\,\text{GeV}$ after preselection. ZH can further decay into either $bb\tau\tau$ (orange histogram), $WWbb$ (green histogram), $WW\tau\tau$ (red histogram) and $ZZbb$ (purple histogram).

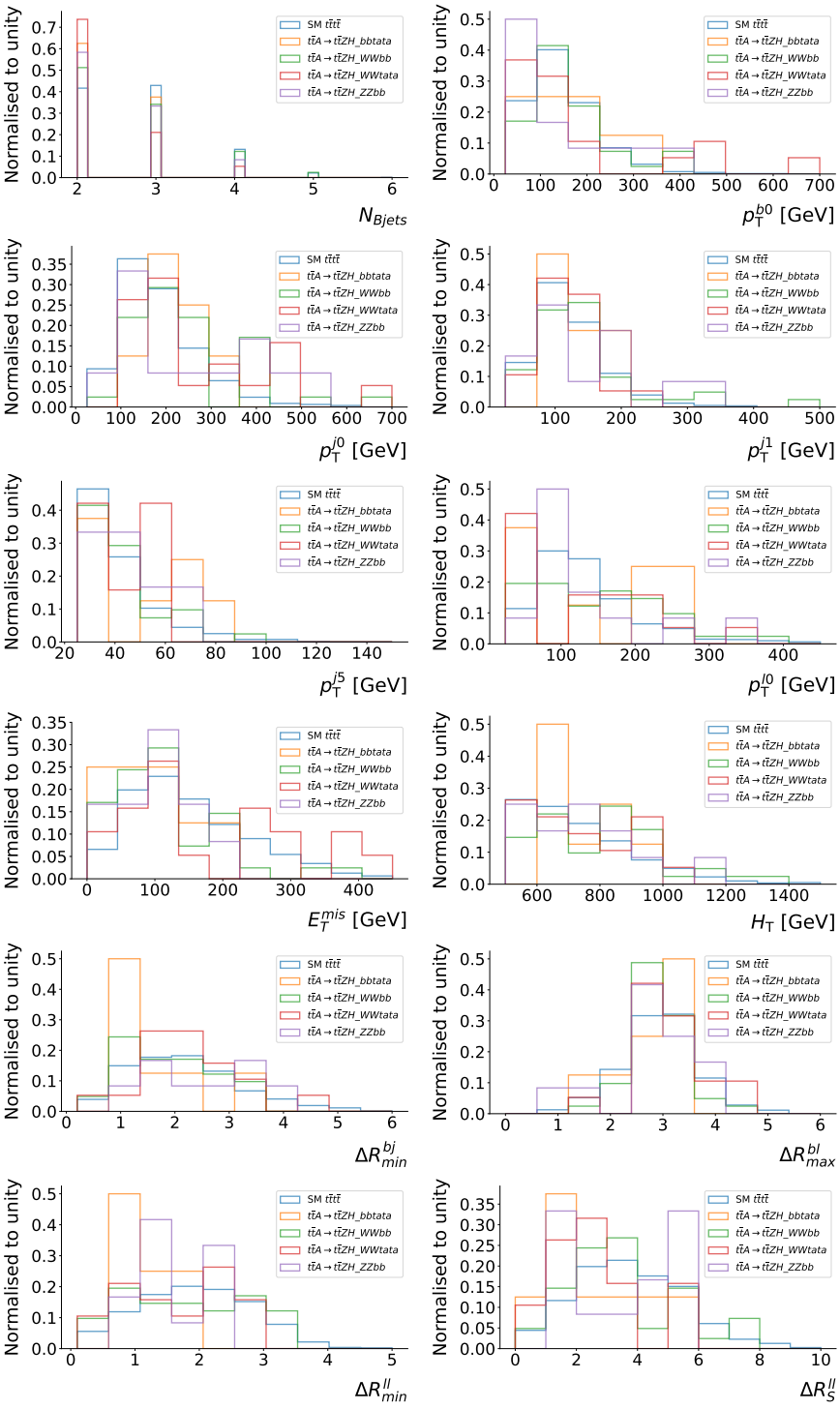


FIGURE 5.12: Distributions of the kinematic variables for $t\bar{t}A \to t\bar{t}ZH$ in the 3ℓ channel at mass of the pseudo-scalar $m_A = 500\,\text{GeV}$ after preselection. ZH can further decay into either $bb\tau\tau$ (orange histogram), $WWbb$ (green histogram), $WW\tau\tau$ (red histogram) and $ZZbb$ (purple histogram).

Chapter 6

Machine Learning analysis

6.1 Overview

The LHC has the potential to tackle some of the most fundamental questions in modern physics, such as the nature of the mass, the dimensionality of space, the unification of fundamental forces, the particle nature of dark matter and the fine-tuning of the Standard Model. Addressing these questions requires a large-scale experimental program that can extract relevant information from the collected data. However, the data collected by HEP experiments are complex and high-dimensional, making traditional data analysis techniques insufficient. In HEP, these techniques involve making Boolean decisions followed by statistical analysis of the selected data, based on the distribution of a single observed quantity motivated by physics considerations, which cannot be easily extended to higher dimensions.

Machine learning is a powerful tool that is increasingly being used in high-energy physics to extract valuable information from large datasets generated by particle accelerators and detectors. Over the past few decades, particle

physicists have endeavoured to enhance the strength of their analyses by employing algorithms that make use of multiple variables at once. Various machine learning techniques, such as artificial neural networks, kernel density estimation, random forests and boosted decision trees have been utilized by physicists. For a while, boosted decision trees implemented in the TMVA software package were the standard machine learning approach in HEP. Although the boosted decision trees implemented in the TMVA software package were valuable for various data analysis tasks, they had limitations and were unable to perform as well as solutions designed by physicists, particularly when dealing with high-dimensional data.

The LHC experiments generate an enormous amount of data, both in terms of the number of collisions and the intricacy of each collision. The colliding proton beams are organized into bunches and intersect at a rate of approximately 40 MHz. Each collision can yield a significant number of particles, and the LHC detectors contain around 10^8 sensors to capture their information. These high data rates are crucial since the collisions that produce noteworthy results are infrequent.

HEP tasks can be easily transformed into machine-learning problems. In these problems, the aim is to find a function $g : a \rightarrow b$ that maps the observed data "a" to a lower-dimensional space of the target label "b", optimizing a specific metric called the loss function $L_{oss}[b, g(a)]$. In an ideal scenario, a learning algorithm would be able to find the optimal function that minimizes L_{oss} for all (a, b) pairs. However, due to the curse of dimensionality and an infinite

number of function choices, this is impractical. Instead, in supervised learning, the algorithm operates on a set of labelled training data that is sampled from the total collected data set. There are other machine learning approaches, such as unsupervised, semisupervised and weakly supervised learning, that alleviate or eliminate the requirement for labelled data in the training process. Consequently, the supervised algorithm minimizes the loss function directly concerning the model parameters.

6.2 Supervised Learning

Supervised machine learning is an approach in which an algorithm is trained on a set of labelled examples in order to learn a function that maps input data to corresponding output labels. This type of machine learning is widely used in HEP for tasks such as particle classification and event reconstruction. For example, in the search for new particles at the LHC, the data collected by the detectors consists of many collisions, each of which can produce a variety of different particles. By using supervised learning techniques, physicists can train algorithms to recognize the different types of particles based on their observable characteristics, such as their energy and momentum.

Another application of supervised learning in HEP is in the reconstruction of collision events. After a collision, the particles produced by the collision interact with the detector, leaving behind a pattern of signals that can be used to reconstruct the trajectory of the particles. By training algorithms on simulated data that includes the true trajectories of the particles, physicists can

use supervised learning to improve the accuracy of event reconstruction. This can be especially useful for rare or elusive events that are difficult to detect using traditional techniques.

Supervised machine learning techniques can also be applied to anomaly detection in HEP. Anomaly detection is the process of identifying rare events that deviate significantly from the expected background distribution. In HEP, anomalies can be indicative of new and interesting physics phenomena, such as the discovery of a new particle or an unexpected decay channel. Anomaly detection is particularly challenging in HEP due to the high dimensionality of the data and the rarity of the interesting events.

One of the most powerful tools for supervised learning in high-energy physics is deep neural networks (DNNs). DNNs are composed of multiple layers of interconnected nodes or neurons and are capable of learning complex representations of data. They are particularly well-suited for high-energy physics applications, where the input data can be very high-dimensional and complex.

6.3 Applications of machine learning in particle physics and COVID-19

Machine learning has become an indispensable tool in particle physics, enabling scientists to tackle complex data analysis tasks and uncover new insights. One

prominent application of machine learning in particle physics is in particle identification. Identifying and distinguishing between different types of particles is crucial for understanding the underlying physics processes. Machine learning algorithms, such as deep neural networks, have been developed and trained on large datasets to accurately classify particles based on their energy deposits, trajectories, and other relevant detector information. These algorithms can effectively discriminate between particles with similar signatures, improving the precision and efficiency of particle identification in experiments like the Large Hadron Collider (LHC). By leveraging the power of machine learning, physicists can enhance their ability to identify rare particles and study their properties with greater accuracy.

Another important application of machine learning in particle physics is in event reconstruction. Particle physics experiments generate vast amounts of data, and reconstructing the interactions of particles within detectors is a challenging task. Machine learning algorithms have been employed to reconstruct the trajectories, energies, and properties of particles from the detector signals. By training models on simulated data and comparing the predictions with the actual measurements, machine learning algorithms can accurately reconstruct particle tracks and vertices, allowing for a more detailed understanding of the physics processes involved. These techniques have proven particularly valuable in experiments like the LHC, where the high collision rates and complex detector geometries require efficient and accurate event reconstruction methods. Machine learning-based event reconstruction has greatly improved the precision and speed of data analysis, enabling physicists to extract valuable

information from large datasets and explore new physics phenomena.

Machine learning has also found significant applications in addressing the challenges posed by the COVID-19 pandemic, including in the field of particle physics. In addition to its traditional applications in particle physics experiments, machine learning techniques have been adapted and applied to analyze and understand the spread of the virus, predict infection rates, and develop strategies for containment and mitigation. By leveraging the power of machine learning algorithms, researchers have been able to analyze large-scale epidemiological data, identify patterns, and make predictions that inform public health decisions and interventions. These techniques have been instrumental in modelling the dynamics of the pandemic, optimizing resource allocation, and guiding vaccination strategies, highlighting the versatility and impact of machine learning in tackling complex and pressing societal challenges.

Machine learning has demonstrated its potential in various areas beyond epidemiological analysis. One notable application is in the development of diagnostic tools. Machine learning algorithms have been trained on vast datasets of medical images, such as chest X-rays and computed tomography (CT) scans, to aid in the detection and classification of COVID-19-related lung abnormalities. These algorithms can assist healthcare professionals by providing faster and more accurate assessments, enabling timely diagnoses and appropriate patient management.

6.4 Deep Neural Networks

DNNs have been used in a wide range of high-energy physics applications, from particle identification to jet reconstruction. For example, in the multi-purpose experiments at CERN, a deep neural network was used to identify jets produced in the collision of protons. The network was trained on simulated data and was able to achieve a very high level of accuracy in identifying jets.

One of the key advantages of deep neural networks is their ability to automatically extract features from raw data. This is particularly useful in high-energy physics, where the data can be very complex and difficult to interpret. By learning these features automatically, deep neural networks can often achieve better performance than traditional machine learning algorithms [101].

However, training deep neural networks requires a large amount of data and computational resources. This can be a significant challenge in high-energy physics, where the data can be very large and complex. To overcome this challenge, researchers often use advanced techniques such as distributed training, in which the network is trained on multiple computers at the same time.
Another challenge in using deep neural networks in high-energy physics is the need for careful validation and testing. The performance of the network must be evaluated on both simulated and real data to ensure that it is robust and reliable. Moreover, the network must be carefully calibrated to ensure that its predictions are accurate [102].

Despite these challenges, deep neural networks have already had a significant

impact on high-energy physics, and are likely to become even more important in the future. Ongoing research is focused on developing new techniques for training and testing deep neural networks, as well as applying them to new areas of high-energy physics research [103].

One exciting area of research is the use of deep neural networks for anomaly detection. Anomalies are events that are unexpected or rare and can be a sign of new physics. Deep neural networks are well-suited for this task, as they are able to learn complex patterns in the data and can often detect subtle deviations from the expected behaviour.

Another area of research is the use of deep neural networks for the real-time analysis of data. In high-energy physics experiments, data is generated at an incredibly fast rate, and it is often necessary to make decisions in real time based on the data. Deep neural networks can be used to analyze the data in real time and make decisions on the fly, allowing researchers to respond quickly to unexpected events or discoveries [103].

6.4.1 Exploration of Deep Neural Network Architecture

In this study, we employed deep neural network models to tackle the task at hand, with each model tailored to a specific mass range of the heavy pseudo-scalar (A). Here, we focus on describing the architecture of one of the models, denoted as *Model_4*, which is designed for $m_A = 400\,\text{GeV}$. The model was implemented using the Sequential API in Keras.

Model_4 consists of multiple layers and each layer is defined using the Dense layer, which is a fully connected layer in the neural network. The activation function employed throughout the model is the rectified linear unit (ReLU), known for its ability to capture nonlinear relationships effectively. The first layer of the model for $2SS\ell$ (3ℓ) comprises 39 (88) nodes and it expects an input with 11 (12) dimensions. The values in the brackets are for the 3ℓ channel.

The architecture of *Model_4* can be summarized as follows: the initial layer, with 39 (88) nodes, utilizes the ReLU activation function and takes in an input of 11 (12) dimensions. This is followed by several consecutive layers, each containing 39 (88, 49 & 36) nodes and also utilizing the ReLU activation function. These intermediate layers help the model learn hierarchical representations of the input data, extracting relevant features at different levels of abstraction. The model contains a total of 13 (17) hidden layers, each with 39 (six with 88, six with 49 and five with 36) nodes and ReLU activation.

The final layer of *Model_4* consists of a single node and uses the sigmoid activation function. This indicates that the model's output is binary, with values ranging between 0 and 1, making it suitable for binary classification tasks. The sigmoid activation function ensures that the output is interpreted as a probability-like score.

To train the model, we compiled it using the binary cross-entropy loss function, which is commonly used for binary classification problems. The Adam optimizer, known for its efficiency and adaptability, was chosen with a specific

learning rate of 0.001 (0.0003253). Additionally, we included accuracy as a metric to evaluate the model's performance.

For the training process, we utilized the fit function, providing the training data denoted as X2_train (X3_train) and corresponding labels Y2_train (Y3_train). The model was trained for 35 (50) epochs with a batch size of 32 (46). We enabled verbose mode to track the training progress, and the data was shuffled during training. To assess the model's performance and monitor potential overfitting, 30% of the training data was allocated for validation.

It's important to note that the above description provides insights into the architecture and training process of *Model_4*. However, it should be mentioned that there are two additional DNN models developed for $m_A = 500\,\text{GeV}$ and $m_A = 600\,\text{GeV}$, which are not covered in this particular explanation. It should be noted that the other two models are not far different from the $m_A = 400\,\text{GeV}$ model. They have similarities and just a few differences in them, but they have the same logic. The tool used to measure the performance of the model are explained in the following section.

6.4.2 Response distribution of DNN

In this study, we conduct a Monte Carlo simulation of proton-proton collisions at the LHC. The signal and Standard Model background events are generated using Madgraph5, with the NNPDF3.0 parton distribution functions. The UFO model files required for the Madgraph analysis were obtained

from FeynRules after proper implementation of the model. After the parton-level analysis, the parton showering and hadronization are performed using Pythia8. We use Delphes(v3) for the detector-level simulation following the showering/hadronization process.

The DNN used is a binary classification algorithm that categorizes events into two categories: 0 for the Standard Model production of four top quarks and 1 for Beyond the Standard Model production. The signature used to explain the results reported by ATLAS with b-jets follows the production mechanism $pp \to t\bar{t}A \to t\bar{t}t\bar{t}$, which results in an excess of multi-lepton final states associated with b-tagged jets in the two channels of interest.

The DNN used in this study was optimized with the best hyper-parameters, and the output distributions obtained from the model for testing data are shown in figure 6.1 and 6.2. The output distributions of the DNN demonstrate that the Standard Model production of four top quarks and Beyond the Standard Model production of four top quarks are close to each other, making it challenging to separate the two. This is further highlighted by the Receiver Operating Characteristic (ROC) curves shown in figure 6.3, which have a low Area Under the Curve (AUC) of barely above 50%. The ROC curves were obtained using test data samples that were not used in the training of the DNNs.

This implies that the DNN is not able to discriminate effectively between

the two types of events, SM and BSM. The ROC curve is a graphical representation of the performance of a binary classifier system, and AUC is a measure of how well the classifier can distinguish between the two classes. An AUC of 50% means that the classifier is not better than a random guess. Therefore, the results suggest that the DNN is not able to differentiate between the SM and BSM four top quark productions.

It is possible that the small difference in the input features of the SM and BSM events is not enough for the DNN to make a distinction. Additionally, the small sample size of events in the simulation could be a factor in the DNN's poor performance. Furthermore, the DNN may not be complex enough to capture the subtle differences between the two types of events.

It is also possible that the current study's DNN structure, architecture or even the dataset used may not be suitable for this task, thus new methods or techniques may be needed to be explored, such as using other types of neural networks or incorporating more advanced data pre-processing techniques to extract more relevant features from the data.

In conclusion, the DNN is not able to distinguish effectively between the SM and BSM four top quark production events, which is reflected in the low AUC of the ROC curve. This highlights the need to explore other techniques or methods to improve the discrimination capabilities of the DNN.

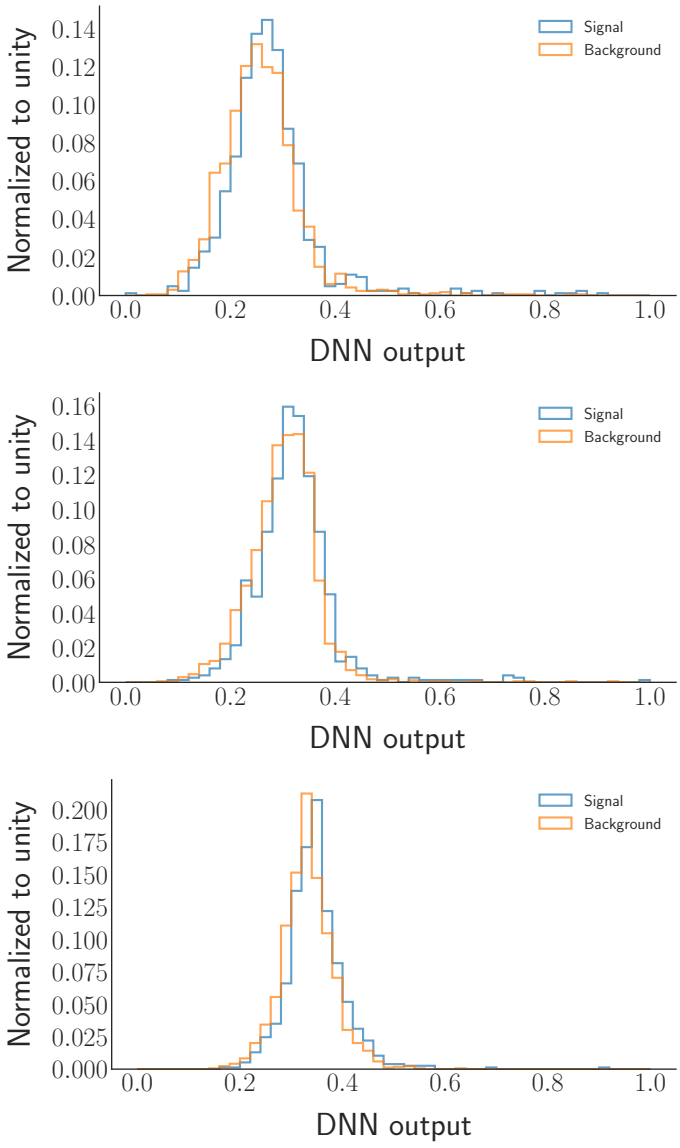


FIGURE 6.1: The DNN output distributions for the three mass ranges for $2SS\ell$ channel. The top is for m_A=400 GeV, the middle is for m_A=500 GeV and the bottom is for m_A=600 GeV.

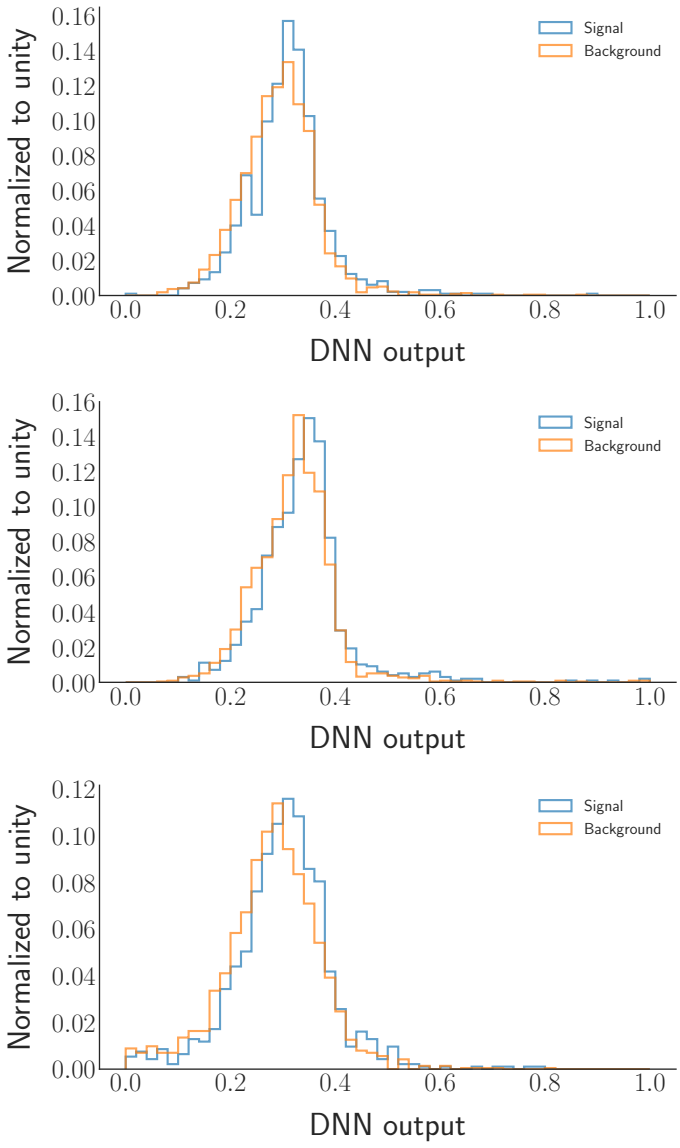


FIGURE 6.2: The DNN output distributions for the three mass ranges for 3ℓ channel. The top is for m_A=400 GeV, the middle is for m_A=500 GeV and the bottom is for m_A=600 GeV.

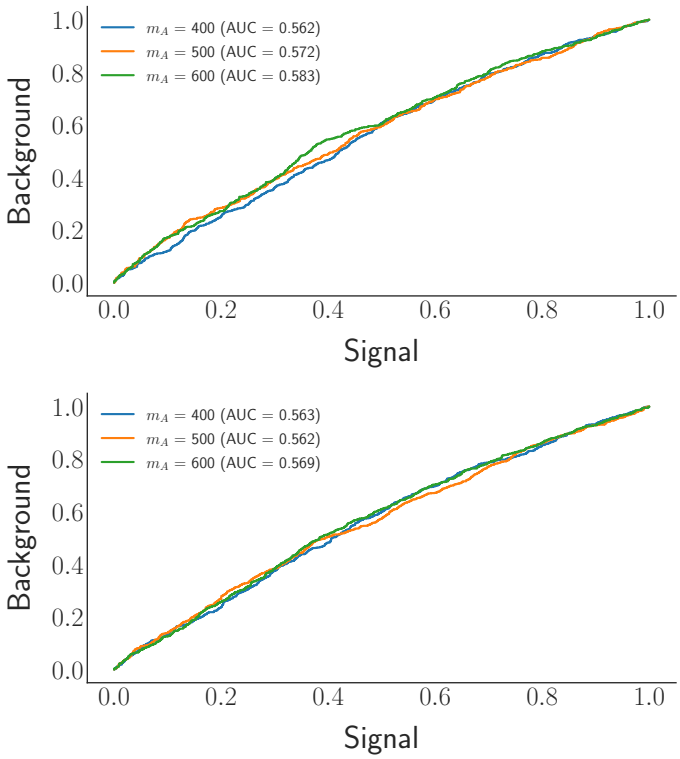


FIGURE 6.3: The ROC curves obtained from the DNN models for signal and background. The graphs correspond to $2SS\ell$ (top) and 3ℓ (bottom) channels.

Chapter 7

Conclusions

This thesis presents a study of the production of four top quarks at the Large Hadron Collider during Run 2 of proton proton collisions at center of mass energy of $\sqrt{s} = 13\,\text{TeV}$ using two categories of multi-lepton channels: two same-sign leptons and three isolated leptons. The signature used to produce the four top quarks was $pp \to t\bar{t}A$ with $A \to t\bar{t}$ leading to four top quarks. The heavypseudo-scalar can also decay in this manner $A \to ZH$. The dominant signature between the two was found to be $A \to t\bar{t}$. Since four top quark is a rare process, the best way to analyse this process was through utilising its finale states, thus going through the route of multi-lepton channels. Due its large background, it was feasible to choose to DNNs to try to classify the SM four top quark production from the one produced via the associate production of a heavypseudo scalar.

A total of 12 kinematic variables were proposed as input variables fed to the DNN algorithm to train on them. From these input variables, it was observed that there was not much discrimination between the four top-quark productions in the SM and BSM used in this study. To further investigate this, a

multivariate algorithm was trained with these twelve features. However, still, no discrimination between the SM and BSM four top-quark signals was observed when extracting the DNN response distributions after training. This was also illustrated by the AUC of the ROC curves being marginally higher than 50%, thus confirming that the SM and BSM four top quark production are indistinguishable from one another using the proposed twelve kinematic variables. The only discrimination between the SM and BSM signals would be the difference in their cross-sections.

This study highlights the potential of AI tools and techniques in providing new insights into complex processes that can be leveraged in other fields. The COVID-19 pandemic, for example, has necessitated international collaboration and the utilization of big data and AI-based techniques to inform public health strategies. By utilizing these techniques, it is possible to gain a more accurate, timely and locally nuanced understanding of the pandemic, including vaccination strategies, social vulnerabilities and new variants of concern. This information can then be used to inform critical public health interventions and help contain, control and mitigate the pandemic. This highlights the importance of AI tools not just in top quark studies, but also in solving real-world problems like the COVID-19 pandemic.

Appendices

Appendix A

Applications of ML techniques to model COVID-19

A.1 Introduction

The COVID-19 pandemic has resulted in over 6.8 million deaths and has had significant medical and economic consequences due to lockdown measures. The strategy to end the pandemic relied heavily on the few drugs available and vaccines, which have been developed rapidly through global efforts. However, the abbreviated development process has increased uncertainties about the long-term efficacy and safety of vaccines and has resulted in increased scrutiny in post-marketing surveillance for biologically plausible side effects.

With the WHO (World Health Organisation) struggling to supply vaccines in developing countries, it is important to consider ways to distribute the vaccines in a manner that maximizes the chances of immunity. With insufficient vaccines in developing countries, it is crucial to consider the delicate ways in

which to roll out the vaccines to achieve herd immunity. This is further complicated by factors such as community members not being willing to consume the vaccine.

The community of South Africa expected front-line workers such as healthcare workers and police officers to be given priority in receiving the vaccine due to them being exposed to the virus during its peak. However, the incubation of the virus is also influenced by comorbidities, such as asthma, diabetes, and hypertension. Machine learning techniques may provide a solution to take into account multiple comorbidities to advise on the distribution of vaccines. The comorbidities are of concern due to their potential influence on the deaths related to COVID-19.

A.2 Covid-19 Data

The spread of the Covid-19 virus around the country is influenced by various factors, including non-compliance with non-pharmaceutical interventions such as social distancing, wearing masks, sanitizing and avoiding crowded places. Despite these measures being implemented to curb the spread of the virus, non-compliance has resulted in crowded hospitals and increased reliance on hospital data as a source of information about the progression of the virus. Hospitalization data provides a more comprehensive understanding of the Covid-19 situation, including details about patient age, gender, and pre-existing health conditions.

The hospital data we will examine is from hospitals in Gauteng province, South Africa and is collected by the Gauteng Department of Health. It provides various information on patients who were hospitalized during the initial stage of the Covid-19 pandemic. The data includes the ward in which patients were admitted, which reflects their level of criticality, such as the general ward for less severe patients, the Intensive Care Unit (ICU) for critically ill patients, and the High Care (HC) ward for the most critical patients. Additionally, the data contains details such as patient ethnicity, age, and gender, which can be used to identify any potential biases or correlations between Covid-19 and existing medical conditions. The dataset also includes information on pre-existing medical conditions, or co-morbidities, which helps to understand their impact on the severity of Covid-19 during the incubation period.

The GDOH data was analyzed and it was found that patients in HC and ICU were more prone to fatalities from Covid-19. These patients were classified as the Severe class, while those in the general ward were labelled as the General ward class. The data analysis also revealed important features to consider in the model-building process, such as the patient's ethnic group, age, gender, and pre-existing comorbidities. This would help determine if the virus showed bias towards any of these features.

Patients from ages 0 to 100 years are recorded in the COVID-19 data, but the age distribution peaks at age 60 for both severe cases and general ward cases as seen in figure A.1. The gender distribution (-1 represents the males, +1 represents the females) is balanced with slightly more males in severe cases and

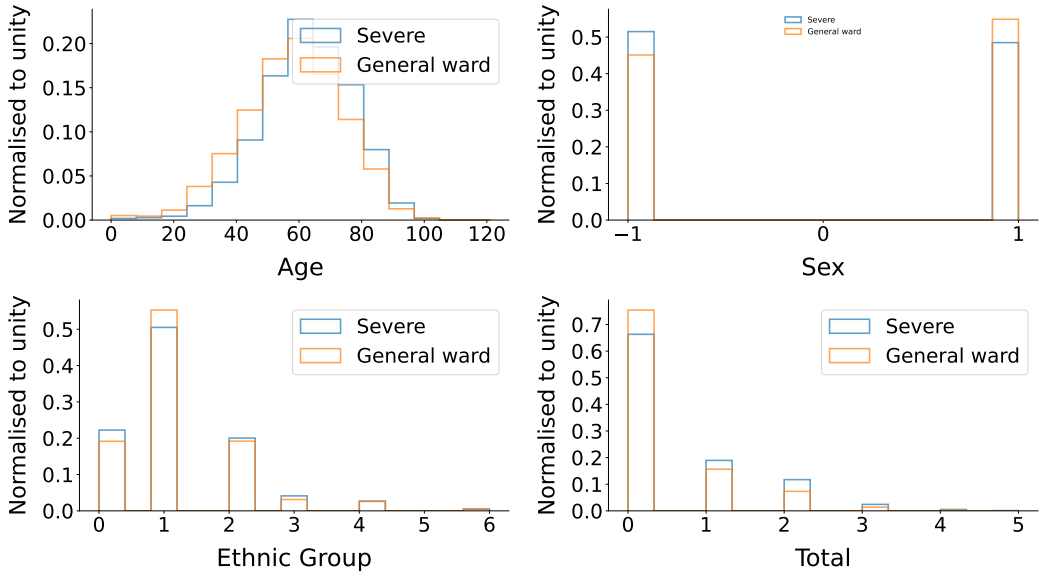


FIGURE A.1: The distribution plots of Age, Gender (sex), Ethnic group and the total sum of pre-existing comorbidities upon arrival at the hospital ward.

more females in general ward cases, with the difference between the two not being significant. The data revealed 5 ethnic groups: Black Africans, Whites, Indians, Coloureds, and Asians. Unknown or undisclosed ethnic groups were classified as unknown. The dominant ethnic group in South Africa is Black Africans and the distribution shows no bias towards a specific ethnic group. The distribution of ethnic groups is as follows: 0 represents unknown ethnic groups, 1 represents Black Africans, 2 represents Whites, 3 represents Indians, 4 represents Coloureds and 5 represents Asians. The distribution of the total number of pre-existing co-morbidities gives information regarding the health status of the patients upon arrival at the hospital. As seen in the distributions, a patient can have a maximum of approximately five comorbidities out of the total eleven considered here. With most of the patients having zero comorbidities.

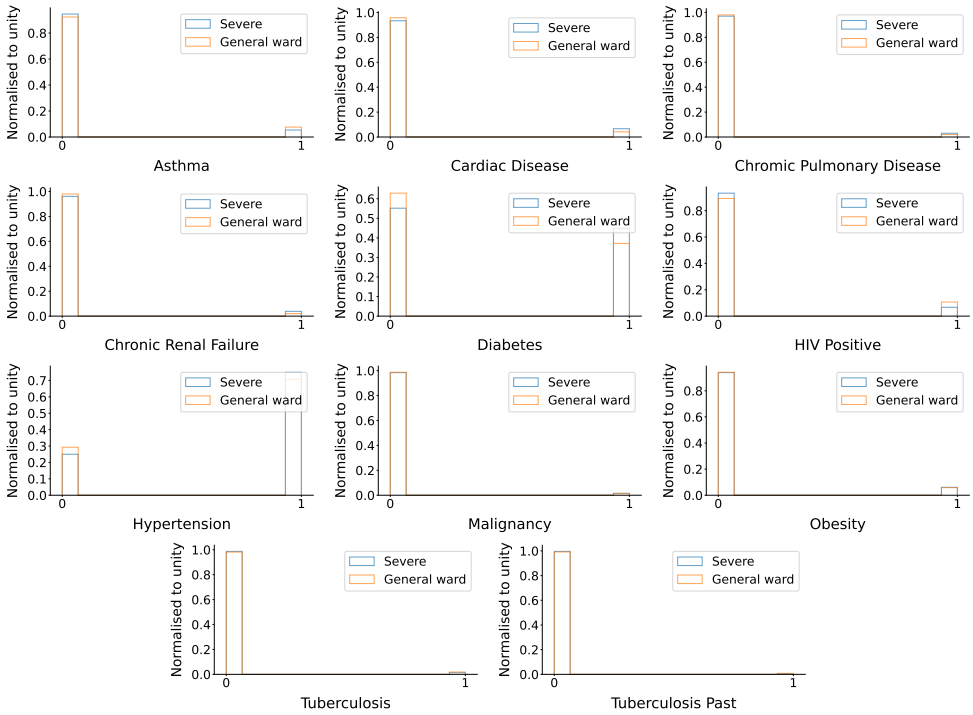


FIGURE A.2: The distributions of comorbidities that are partof input variables are used to train the DNN. The blue histogram represents severe category, while the orange histogram represents general ward category. The peak at zero denotes patients who do not have a comorbidity, while the peak at one denotes patients with comorbidity.

In the GDOH data, 11 co-morbidities were identified: Asthma, Diabetes, Cardiac Disease, Chronic Pulmonary Disease, Chronic Renal Failure, HIV Positive, Hypertension, Malignancy, Tuberculosis, Tuberculosis Past, and Obesity. The distribution of these co-morbidities is displayed in figure A.2, with 1 representing "yes" to having the disease and 0 representing "no" to not having the disease for all co-morbidities.

roc

From figure A.2, we see that hypertension and diabetes have a significant impact on Covid-19's severity, while asthma, cardiac disease, obesity, and HIV have a relatively lesser impact. Most of these comorbidities play a crucial role in the severity of Covid-19. If you look at the sum of all comorbidities in figure A.1 (bottom right distribution), you can see that most general ward patients have fewer comorbidities than severe class patients. This indicates that patients with multiple comorbidities are more likely to be classified as having severe cases of Covid-19. The only exception is when a patient has zero comorbidities, then the number of general ward cases is greater than the severe class. But for one or more comorbidities, the severe class is always higher than the general ward.

A.3 Leveraging ML techniques used in physics to model COVID-19 vaccine strategies

The knowledge and skills gained from analyzing top quarks in high-energy physics can be applied to the study of the COVID-19 pandemic and the roll-out of vaccines. By utilizing ML techniques, we can gain insights and make predictions about the spread of the virus and the efficacy of vaccine distribution efforts. The four top quark analysis was an 11 and 12-dimensional problem, where we applied DNNs to classify the SM four top quark production from the BSM four top quark production. Similarly here, we have a binary classification problem, which requires classifying the patients according to the level of severity based on contracting the virus. Here we have a slightly increased dimensionality, which consists of 14 input variables. We have eleven

co-morbidities together with ethnicity, gender and age. As we know that only a limited supply of vaccines were made available for developing countries, it is extremely important to distribute it to people who are in need of it in order to minimize the mortality surrounding this virus.

We use DNN algorithm to identify the destination of Covid-19 patients. Our model has 10 layers, including 2 input (50 nodes, 14 dimensions, Relu activation) and 1 output layer (1 node, Sigmoid activation) for binary classification (Severe class or General ward class). The 8 hidden layers each have 50 nodes and Relu activation. The model's goal is to predict the patient outcome based on input variables used to train this model, with the output layer linking to the target variables using a learning rate of 0.003 and the Adams optimizer. This resulted in excellent model performance, which will be shown in the next section.

A.3.1 DNN classification results

The training and testing data were split into 70% and 30% respectively. The model achieved an accuracy of 65% when training with the data and the response distributions of both training and testing are shown in figure A.3. The task of our model is to perform binary classification between severe class and general ward class. From the response distributions in figure A.3 the model is able to perform the classification task as we see strong separation between the two classes.

The model has been trained using hospitalization data and now it's time to

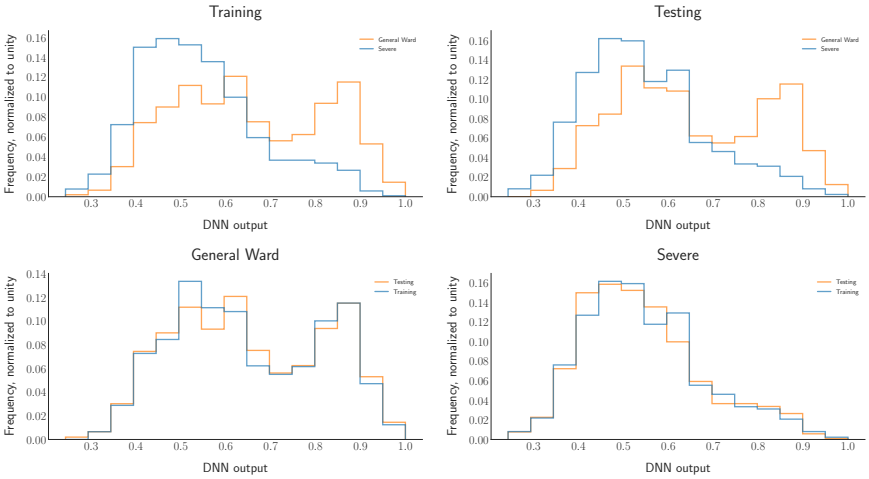


FIGURE A.3: The response distributions of the DNN after training (top left) and testing (top right). The bottom distributions show the comparison between traning and testing data for general ward (bottom left) and severe (bottom right) class.

test it on sample data representing the health status of different age groups in South Africa. The results, shown in figure A.4, indicate that the DNN accurately predicts on this sample. The majority of the population is composed of youth and young adults with few comorbidities, which do not significantly impact COVID-19 severity, thus the green histogram peaks towards the general ward (close to 1). Therefore it is a small fraction of the population that would fall in the category of severe class. This population is the target group that needs to be identified and given first preference when distributing the limited vaccines.

The severity level of the DNN response varies across its distribution, making it possible to segment the distribution into equal parts for a deeper understanding of the model's results. The distribution is divided into 5 segments of 20%, forming 5 stages for vaccine roll-out. The classes are ranked from 1 (most

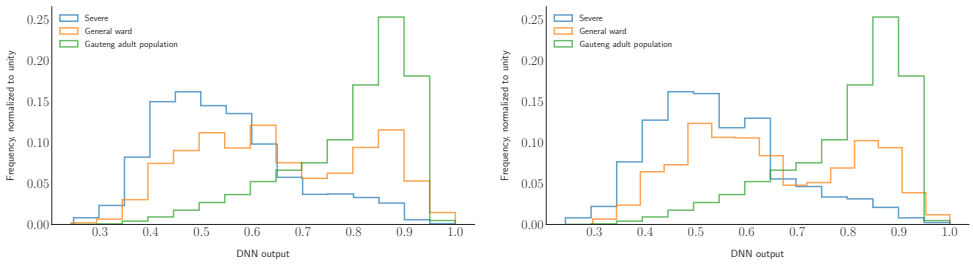


FIGURE A.4: The DNN response distributions for prediction on the data sample that represents the population at large (Gauteng population). The distribution on the left shows training data and the one on the right shows testing data.

severe) to 5 (least severe). As we confirmed that diabetes and hypertension are major comorbidities impacting COVID-19, we will focus on these and age to better determine the target age group.

The first row of figure A.5 displays the age distribution. As we move from the left graph, we see that the peak of the age starts at 70 years for the first 20%. This peak represents the age group most likely to experience critical outcomes from COVID-19, suggesting that individuals over 60 years old should be prioritized for vaccination to reduce fatalities. The fifth graph in the first row shows the peak at 30 years, indicating that younger adults are less likely to experience severe outcomes from COVID-19. Based on these age distributions, it is concluded that COVID-19 primarily affects the elderly rather than young adults.

The class distributions of age in figure A.6 reveal that females are at a higher risk, as seen in Class 1 where over 75% of individuals are females. As we move to Class 5, more males are found to be in better health than females. This

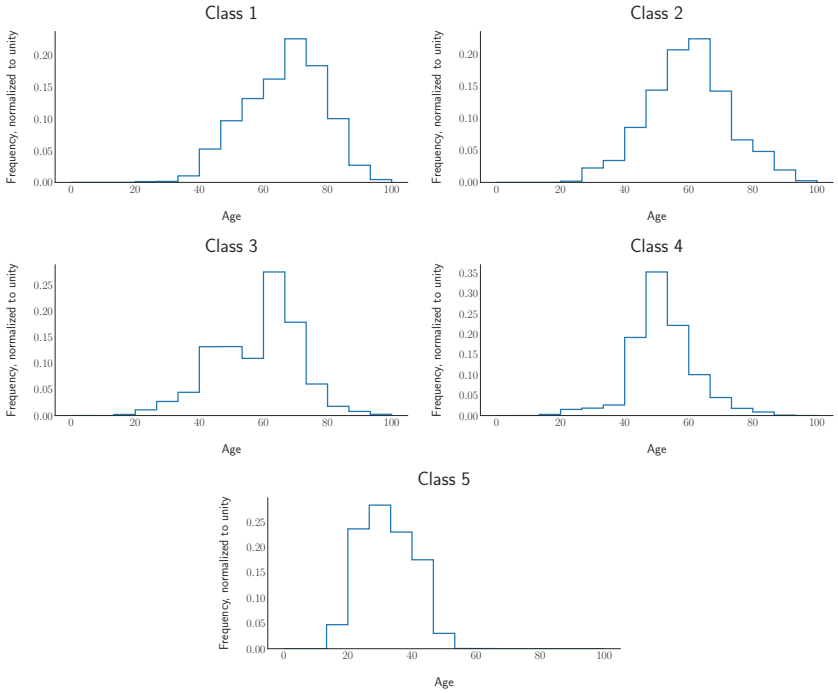


FIGURE A.5: The five classes distributions for Age. These distributions illustrate the level of severity in 20% intervals.

information is crucial in understanding the impact of COVID-19 on different genders and can be used to target the right population with the right resources. Further research and analysis can also be conducted to uncover the reasons behind this trend and implement more effective measures to address it.

The distributions in figure A.7 and A.8 reveal the two most prevalent comorbidities being diabetes and hypertension. In the most severe segment (class 1), 70% of the population have both conditions, while in the least severe segment (class 5), no individuals have either. Our model's input variables highlight that factors such as age, gender, diabetes and hypertension play a role in the severity of COVID-19. The model helps identify which age groups are at a higher risk of fatalities from the virus.

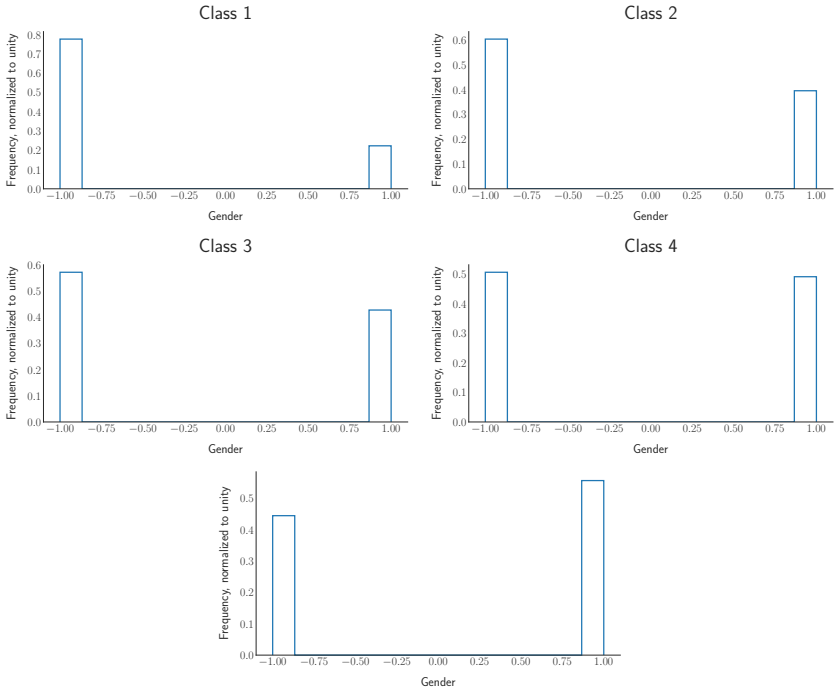


FIGURE A.6: The five classes distributions for gender. These distributions illustrate the level of severity in 20% intervals.

The impact of vaccination on reducing Covid-19 severity is depicted in figure A.9 through the ROC curve. Our analysis reveals that vaccinating 20% of the Gauteng population would result in a reduction of illness severity by over 80%. This approach could lead to quicker immunity attainment compared to waiting for 60 million plus vaccine doses for the entire population, which may not be feasible due to production and distribution constraints.

By utilizing this classification model, we can determine who requires vaccinations, allowing us to maximize the use of vaccine supply and minimize expenses. This optimization can redirect funds to other crucial areas to support the overall economy of South Africa. With the help of AI, we can ensure an

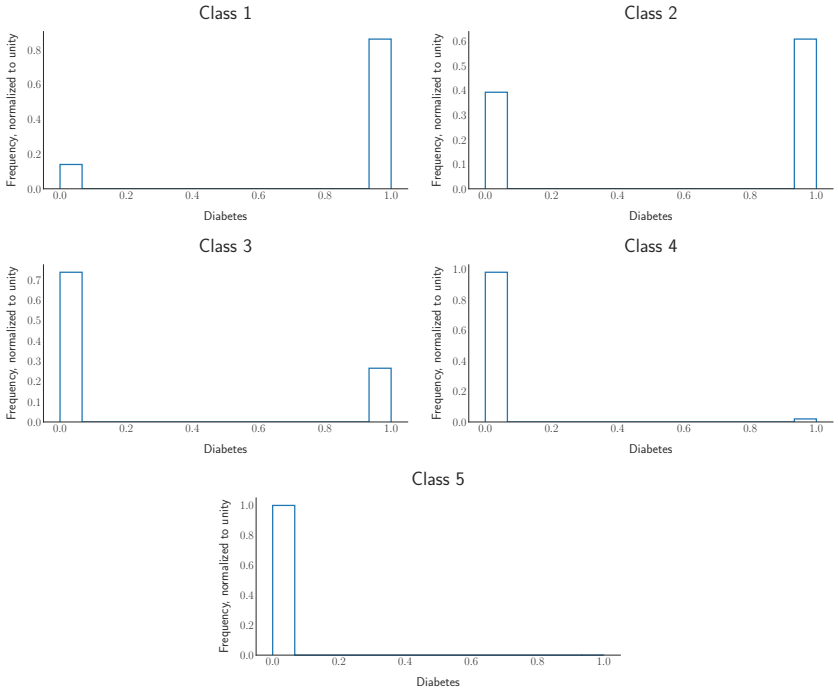


FIGURE A.7: The five classes distributions for Diabetes. These distributions illustrate the level of severity in 20% intervals.

effective and efficient rollout of vaccines to the population. Additionally, by classifying the population based on their risk factors, the government can prioritize the most vulnerable and at-risk individuals, and make sure they receive the vaccine first. This will help to reduce the spread of the virus and control the severity of the illness, particularly among those who are most vulnerable to serious illness or death. Furthermore, by using AI, we can also ensure that vaccine distribution is equitable and reaches all parts of the country, including rural and remote areas. This can help to prevent the spread of the virus and ensure that everyone has access to the life-saving vaccine, regardless of their location.

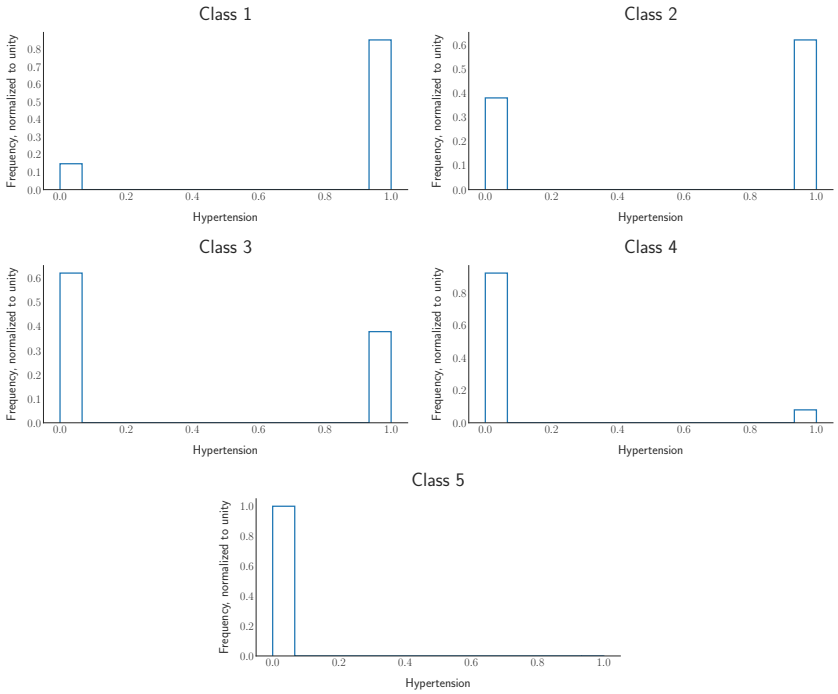


FIGURE A.8: The five classes distributions for hypertension. These distributions illustrate the level of severity in 20% intervals.

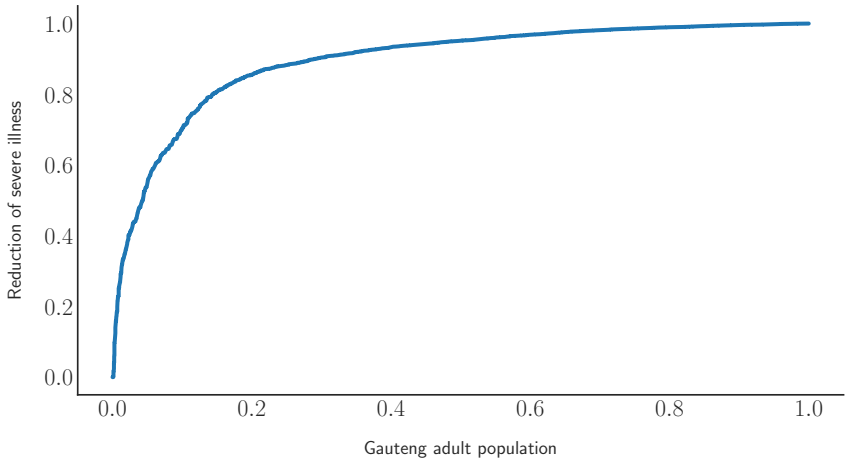


FIGURE A.9: A ROC curve showing the possible outcome of illness reduction when using this model to roll-out vaccines.

A.4 Transferring COVID-19 vaccines strategies to Botswana

A.4.1 COVID-19 in Botswana

Botswana confirmed its first three cases of COVID-19 in March 2020 and since then, the number of positive cases has significantly increased, along with the number of deaths. Despite facing challenges such as a shortage of healthcare workers, intensive care facilities, and ventilators, the government had implemented measures such as public education, social distancing, community lockdowns, travel restrictions and border control to address the pandemic. The government had also declared a state of public emergency and rolled out a vaccination program that has faced administrative challenges. Several organizations, including the Ministry of Health and Wellness, Presidential Taskforce and WHO, have provided support in developing key documents, providing technical guidance and facilitating engagement with other agencies and partners to find effective ways of controlling the pandemic.

Botswana has some of the world's highest rates of comorbidities, increasing the severity of COVID-19 complications. Compared to other middle-income populations, the country has lower inpatient care facilities, posing additional pressure on its health system. Despite these challenges, Botswana's investment in the health sector has enabled access to needed health services without significant severity for many citizens, assisting the country's comprehensive response to the pandemic. However, there is a need for a clear vaccine rollout strategy to ensure the effective vaccination of the population.

A.4.2 Age structures between South Africa (Gauteng) and Botswana

Demographic models remain widely used in population dynamics and ecology because they involve simple mathematical tools and are suitable for individual enumeration or census methods. Age structure is essential in predicting endemic and epidemic infections like measles, influenza and COVID-19, which have age-dependent contact rates, and ignoring age-structured contacts can lead to misunderstanding epidemiological data and costly policy errors. Therefore, age structure is a crucial risk factor that should be considered in understanding epidemic settings and is an integral part of disease forecasting.

Based on data from Statistics South Africa and Botswana, Figures A.10, A.11 and A.12 show the estimated percentage of the total population residing in Botswana and Gauteng provinces for the year 2019. Gauteng has the largest population, while Botswana has the highest proportion of people aged 0-4 years, and Gauteng has the highest proportion of people aged 30-34 years. The population percentage ratio of Botswana and Gauteng, calculated by dividing Botswana's percentage over Gauteng's percentage in both males and females, shows that Botswana's proportion is higher than Gauteng's if the ratio is greater than 1 and vice versa. In contrast, a ratio of 1 means the proportions are equal.

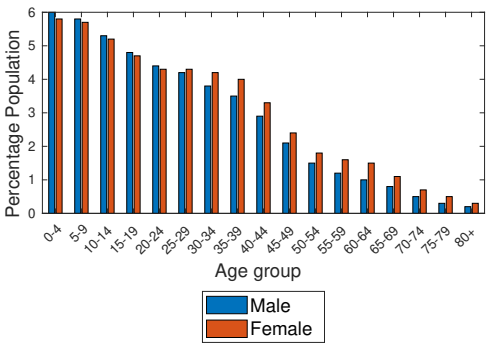


FIGURE A.10: Age structures of the populations of Botswana

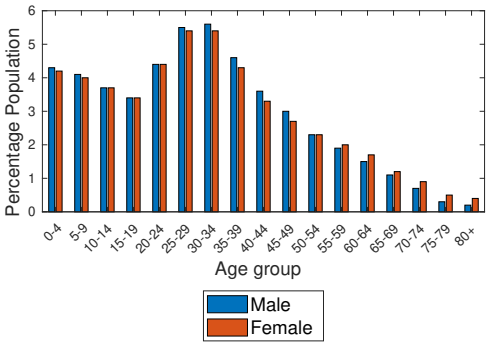


FIGURE A.11: Age structures of the population's of Gauteng (South Africa)

A.4.3 Comorbidities in Botswana

Studies suggest that COVID-19 patients with underlying health conditions or comorbidities, especially elderly individuals, have a slower progression towards recovery and an increased risk of death. Unfortunately, there is limited information regarding COVID-19-related comorbidities and prevalence rates in Botswana, which could help understand patterns and influence vaccination strategies. Data provided by the Global Burden of Disease study indicates high HIV/AIDS prevalence rates among the 15-49 and 50-69 aged population

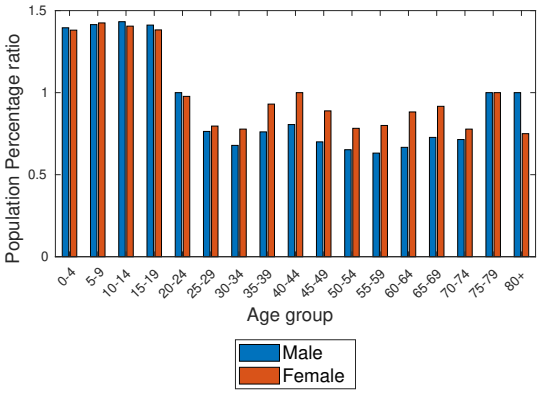


FIGURE A.12: Age structures of the population's ratio of Botswana and Gauteng - South Africa

groups in Botswana, while among the older population (70+ years), Botswana has higher incidents of lower respiratory infections, Ischaemic Heart disease and Stroke. A comparative analysis of Botswana and Gauteng data shows almost identical numbers of comorbidities (by type) amongst the older population (50–69 and 70+ age groups).

Figures A.13 and A.14 provide additional data about hypertension/high blood pressure, a key comorbidity. The figures indicate higher percentages of South Africans with hypertension compared to Botswana. However, the difference recorded between the years 2000 and 2015 was not significant compared to the 1975-1995 period. Figure A.14 suggests that hypertension was predominantly high in females compared to their male counterparts between 1975 and 2000. From 2005, a greater proportion of males had high blood pressure compared to females.

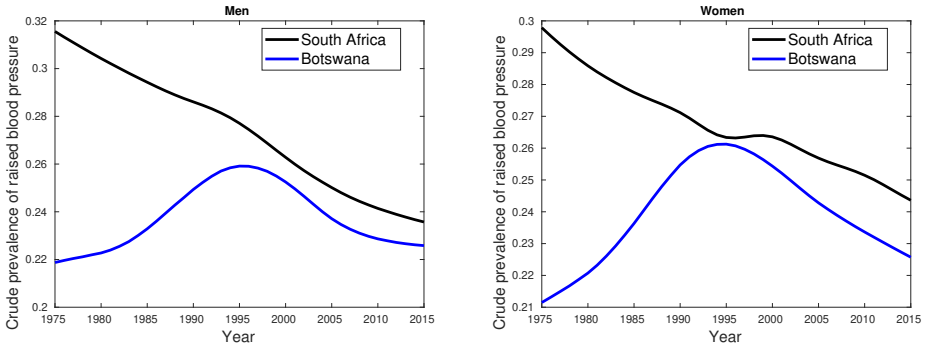


FIGURE A.13: Crude prevalence of raised blood pressure for males and female trends in Botswana and South Africa from 1975 to 2015.

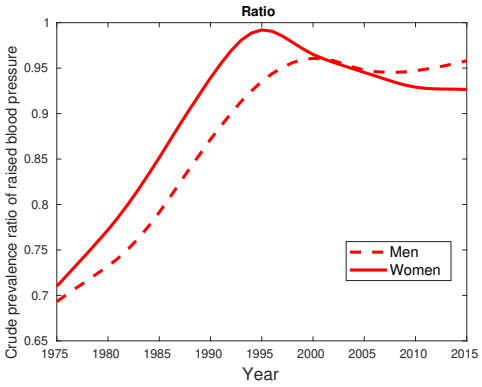


FIGURE A.14: Crude prevalence of raised blood pressure for male and female ratios between Botswana and South Africa from 1975 to 2015.

A.4.4 Comparative study between Botswana and South Africa

To identify individuals who can achieve herd immunity quickly through vaccines, a smart algorithm that uses deep learning is recommended due to the 14-dimensional nature of the problem. Input variables such as age, gender,

ethnicity and comorbidities are used to train the algorithm, which has distinctive hyper-parameters, hidden layers and activation functions to yield high-efficiency results. Adopting the same model used in Sec. A.3 is likely to reveal similar behavioural characteristics of COVID-19, including the two significant comorbidities. However, the primary objective is to compare mortality rates in South Africa and Botswana.

The study aims to identify target groups for COVID-19 vaccination using input variables that include 11 comorbidities, age, gender and ethnicity. The distributions of input variables for South Africa and Botswana are compared as seen in figure A.15. The results shown in figure A.15 indicate that the major comorbidities that increase the severity of COVID-19 are similar in both countries, which suggests that the health status of South Africa is comparable to that of Botswana. Since these two countries show similarities in their health status, it is reasonable to conclude that the first target group to receive vaccination must have both hypertension and diabetes, which are the major comorbidities in both countries. That said, the roll-out strategy for vaccination should be similar to that outlined in sec. A.3, with the more severe group comprising mainly of elderly people. To confirm the severity of the virus, the mortality rates of the virus in both countries need to be compared.

A.4.5 Mortality rates between South Africa and Botswana

We used the mortality data for Botswana and South Africa in a DNN to classify the severity of COVID-19. The DNN's output distributions in figure A.16 show that mortality is similar to the severe class, indicating that the severity

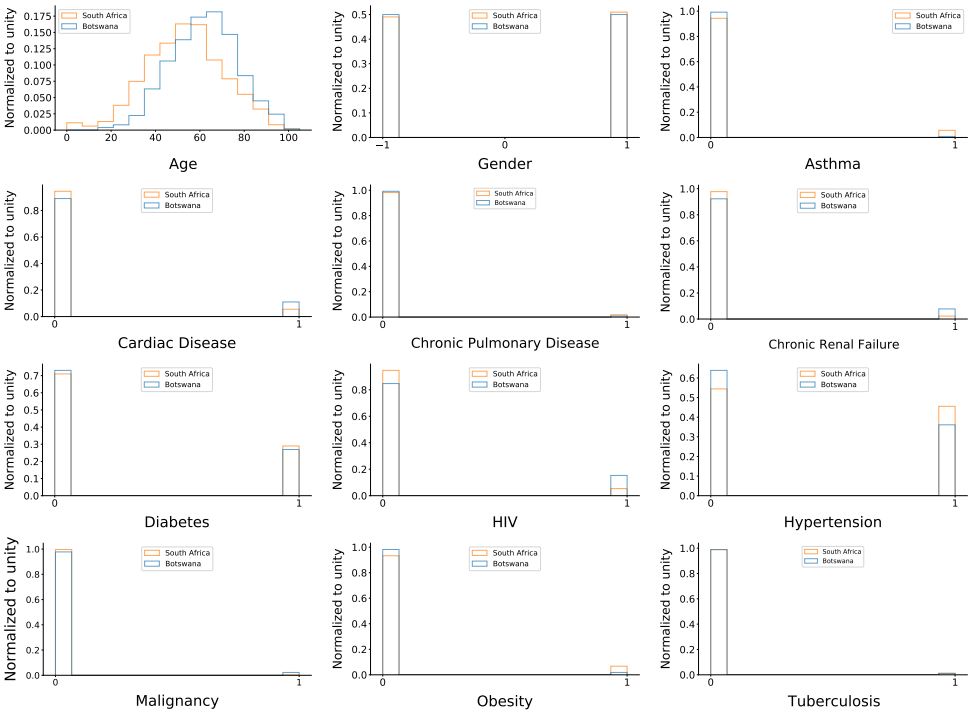


FIGURE A.15: The distributions of the input variables are used to train the DNN. The blue histogram represents Botswana data, while the orange histogram represents South African data. For histograms related to comorbidities, the peak at zero denotes patients who do not have a comorbidity, while the peak at one denotes patients with comorbidity.

of COVID-19 is comparable between these two countries. Based on the two critical comorbidities, we split the mortality population into five classes as seen in figure A.17, with class 1 being the most severe and class 5 the least severe, similar to how it was done in Sec. A.3. We were able to confirm that the elderly need to be the first priority group when distributing the vaccines, with over 70% of class 1 being above 60 years of age and having both hypertension and diabetes. As we move from class 1 to class 5, the age peak is shifted from the elderly to young adults, and the proportion of people with hypertension and diabetes decreases, with only <1% of class 5 having these comorbidities.

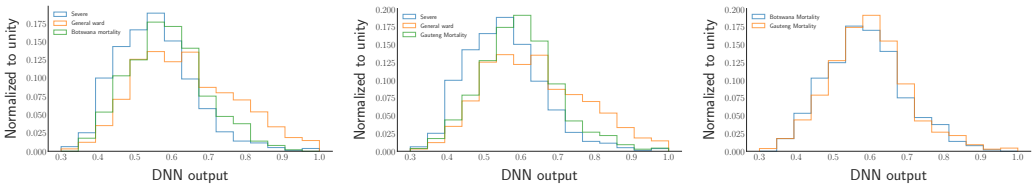


FIGURE A.16: The DNN output distributions for Mortality behind COVID-19. The distribution on the left represents the mortality of Botswana, and the one on the right represents the mortality of Gauteng (South Africa).

Therefore, prioritizing the elderly with comorbidities for vaccination will reduce the severity and mortality behind the virus and reaching heard immunity rapidly. The DNN output distributions provide insights into target groups that should receive the limited supply of vaccines, making it possible to minimise its spread while sustaining lives. Therefore it can be confirmed that Botswana can follow the suimilar strategy for distribution the vaccines to the population while prioritize those who are in need of it.

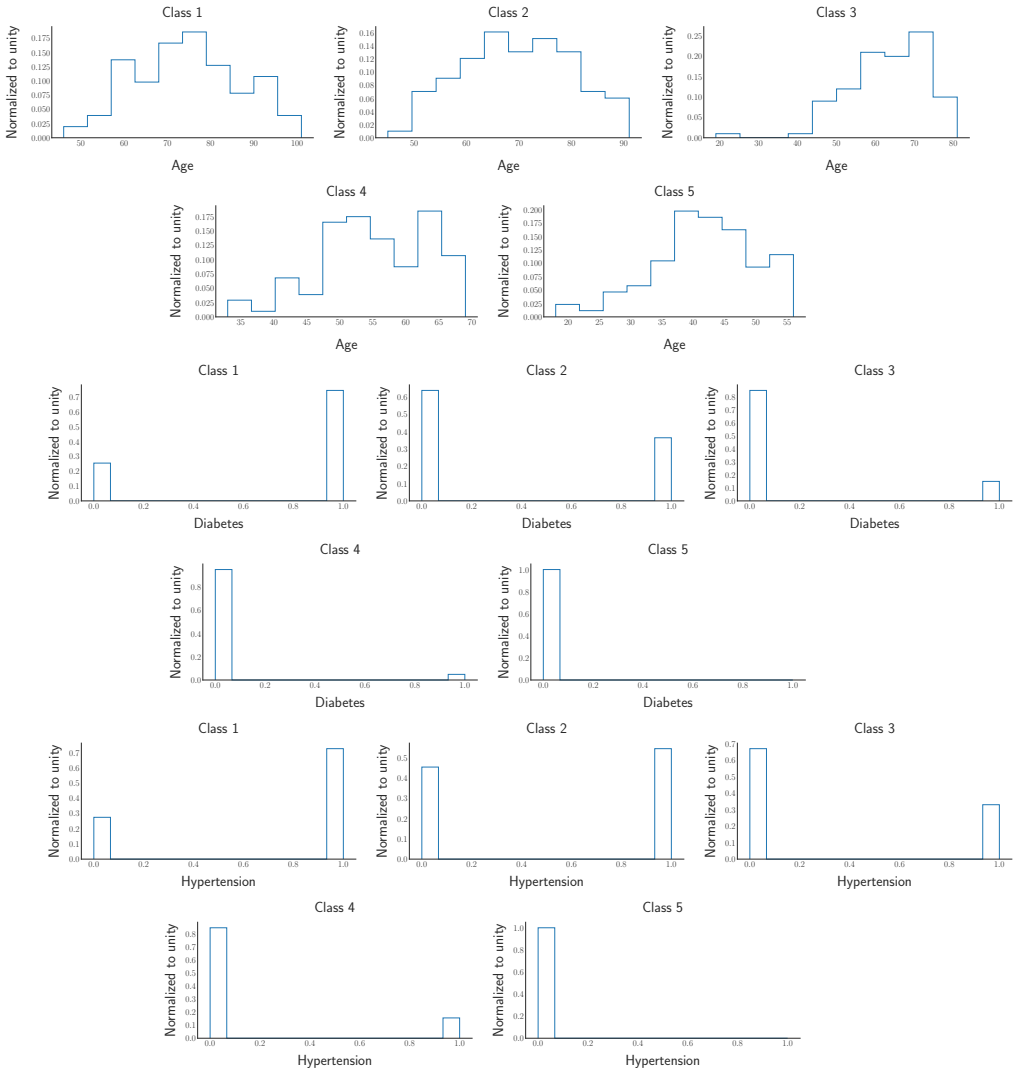


FIGURE A.17: Five classes were obtained from DNN output distributions. The distributions include Age, Diabetes and Hypertension. The first (last) row is the most (least) severe target group which is Class 1 (Class 5).

Bibliography

[1] Georges Aad et al. “Evidence for tt¯ tt¯ production in the multilepton final state in proton–proton collisions at s= 13 TeV with the ATLAS detector”. In: *The European Physical Journal C* 80.11 (2020), p. 1085.

[2] Georges Aad et al. “Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC”. In: *Physics Letters B* 716.1 (2012), pp. 1–29.

[3] Serguei Chatrchyan et al. “Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC”. In: *Physics Letters B* 716.1 (2012), pp. 30–61.

[4] Peter Ware Higgs. “Broken symmetries, massless particles and gauge fields”. In: *Phys. Lett.* 12 (1964), pp. 132–133.

[5] François Englert and Robert Brout. “Broken symmetry and the mass of gauge vector mesons”. In: *Physical review letters* 13.9 (1964), p. 321.

[6] Georges Aad et al. “Evidence for the spin-0 nature of the Higgs boson using ATLAS data”. In: *Phys. Lett. B* 726 (2013), pp. 120–144. DOI: 10.1016/j.physletb.2013.08.026. arXiv: 1307.1432 [hep-ex].

[7] Serguei Chatrchyan et al. “Study of the Mass and Spin-Parity of the Higgs Boson Candidate Via Its Decays to Z Boson Pairs”. In: *Phys. Rev.*

Lett. 110.8 (2013), p. 081803. DOI: 10.1103/PhysRevLett.110.081803. arXiv: 1212.6639 [hep-ex].

[8] Stefan von Buddenbrock et al. "The Madala hypothesis with Run 1 and 2 data at the LHC". In: *Journal of Physics: Conference Series*. Vol. 889. 1. IOP Publishing. 2017, p. 012020.

[9] Stefan von Buddenbrock. "Exploring LHC Run 1 and 2 data using the Madala hypothesis". In: *Journal of Physics: Conference Series*. Vol. 878. 1. IOP Publishing. 2017, p. 012030.

[10] Stefan Von Buddenbrock et al. "The compatibility of LHC Run 1 data with a heavy scalar of mass around 270\, GeV". In: *arXiv preprint arXiv:1506.00612* (2015).

[11] Stefan von Buddenbrock et al. "Phenomenological signatures of additional scalar bosons at the LHC". In: *Eur. Phys. J. C* 76.10 (2016), p. 580. DOI: 10.1140/epjc/s10052-016-4435-8. arXiv: 1606.01674 [hep-ph].

[12] Yaquan Fang et al. "Impact of additional bosons on the exploration of the Higgs boson at the LHC". In: *International Journal of Modern Physics A* 32.34 (2017), p. 1746010.

[13] Gustavo Castelo Branco et al. "Theory and phenomenology of two-Higgs-doublet models". In: *Physics reports* 516.1-2 (2012), pp. 1–102.

[14] Margarete Mühlleitner et al. "Phenomenological comparison of models with extended Higgs sectors". In: *Journal of High Energy Physics* 2017.8 (2017), pp. 1–55.

[15] Mukesh Kumar et al. "The impact of additional scalar bosons at the LHC". In: *Journal of Physics: Conference Series*. Vol. 802. 1. IOP Publishing. 2017, p. 012007.

[16] Stefan von Buddenbrock et al. "Multi-lepton signatures of additional scalar bosons beyond the Standard Model at the LHC". In: *J. Phys. G* 45.11 (2018), p. 115003. DOI: 10.1088/1361-6471/aae3d6. arXiv: 1711.07874 [hep-ph].

[17] Stefan von Buddenbrock et al. "The emergence of multi-lepton anomalies at the LHC and their compatibility with new physics at the EW scale". In: *Journal of High Energy Physics* 2019.10 (2019), pp. 1–41.

[18] G. Basti. "The Philosophy of Nature of the Natural Realism. The Operator Algebra from Physics to Logic". In: *Philosophies* 7 (Oct. 2022), p. 121. DOI: 10.3390/philosophies7060121.

[19] Georges Aad et al. "Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC". In: *Phys. Lett. B* 716 (2012), pp. 1–29. DOI: 10.1016/j.physletb.2012.08.020. arXiv: 1207.7214 [hep-ex].

[20] Serguei Chatrchyan et al. "Observation of a New Boson at a Mass of 125 GeV with the CMS Experiment at the LHC". In: *Phys. Lett. B* 716 (2012), pp. 30–61. DOI: 10.1016/j.physletb.2012.08.021. arXiv: 1207.7235 [hep-ex].

[21] John Ellis, Mary K Gaillard, and Dimitri V Nanopoulos. "A historical profile of the Higgs boson". In: *The standard theory of particle physics* (2016), pp. 255–274.

[22] Nguyen Ky and Nguyn Vân. "Latest results on the Higgs boson discovery and investigation at the ATLAS-LHC". In: *Journal of Physics: Conference Series* 627 (June 2015). DOI: 10.1088/1742-6596/627/1/012011.

[23] Peter W. Higgs. "Broken symmetries, massless particles and gauge fields". In: *Phys. Lett.* 12 (1964), pp. 132–133. DOI: 10.1016/0031-9163(64)91136-9.

[24] F. Englert and R. Brout. "Broken Symmetry and the Mass of Gauge Vector Mesons". In: *Phys. Rev. Lett.* 13 (1964). Ed. by J. C. Taylor, pp. 321–323. DOI: 10.1103/PhysRevLett.13.321.

[25] Peter W. Higgs. "Broken Symmetries and the Masses of Gauge Bosons". In: *Phys. Rev. Lett.* 13 (1964). Ed. by J. C. Taylor, pp. 508–509. DOI: 10.1103/PhysRevLett.13.508.

[26] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble. "Global Conservation Laws and Massless Particles". In: *Phys. Rev. Lett.* 13 (1964). Ed. by J. C. Taylor, pp. 585–587. DOI: 10.1103/PhysRevLett.13.585.

[27] LHC Group et al. "Standard Model Higgs-Boson Branching Ratios with Uncertainties". In: *The European Physical Journal C* 71 (Sept. 2011), pp. 1–29. DOI: 10.1140/epjc/s10052-011-1753-8.

[28] Georges Aad et al. "Evidence for $t\bar{t}t\bar{t}$ production in the multilepton final state in proton–proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector". In: *Eur. Phys. J. C* 80.11 (2020), p. 1085. DOI: 10.1140/epjc/s10052-020-08509-3. arXiv: 2007.14858 [hep-ex].

[29] Stefan von Buddenbrock et al. “Constraints on a 2HDM with a singlet scalar and implications in the search for heavy bosons at the LHC”. In: *J. Phys. G* 46.11 (2019), p. 115001. DOI: 10.1088/1361-6471/ab3cf6. arXiv: 1809.06344 [hep-ph].

[30] Lyndon Evans and Philip Bryant. “LHC machine”. In: *Journal of instrumentation* 3.08 (2008), S08001.

[31] Mariusz Sapi and Agnieszka Priebe. “Symulacje war tości progowych Monitorów S Wiązki za pomoc ą Geant4”. In: (Feb. 2023).

[32] The Collaboration et al. “Studies of the performance of the ATLAS detector using cosmic-ray muons”. In: *European Physical Journal C* 71 (Nov. 2010), pp. 1–36. DOI: 10.1140/epjc/s10052-011-1593-6.

[33] ATLAS Inner Detector. “Technical Design Report”. In: *CERN= LHCC* (1997), pp. 98–13.

[34] Georges Aad et al. “The ATLAS Inner Detector commissioning and calibration”. In: *The European Physical Journal C* 70.3 (2010), pp. 787–821.

[35] ATLAS Collaboration et al. *ATLAS inner detector: Technical Design Report, 2. Technical Design Report ATLAS. CERN, Geneva, 1997.*

[36] ATLAS collaboration et al. “Technical design report for the ATLAS inner tracker pixel detector”. In: *CERN, Geneva, Tech. Rep. CERN-LHCC-2017-021. ATLAS-TDR-030* (2017).

[37] A Abdesselam et al. “The ATLAS semiconductor tracker end-cap module”. In: *Nuclear Instruments and Methods in Physics Research Section*

A: Accelerators, Spectrometers, Detectors and Associated Equipment 575.3 (2007), pp. 353–389.

[38] A Abdesselam et al. "The barrel modules of the ATLAS semiconductor tracker". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 568.2 (2006), pp. 642–671.

[39] Ashfaq Ahmad et al. "The silicon microstrip sensors of the ATLAS semiconductor tracker". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 578.1 (2007), pp. 98–118.

[40] E Abat et al. "The ATLAS TRT barrel detector". In: *Journal of Instrumentation* 3.02 (2008), P02014.

[41] E Abat et al. "The ATLAS TRT end-cap detectors". In: *Journal of Instrumentation* 3.10 (2008), P10003.

[42] E Abat et al. "The ATLAS Transition Radiation Tracker (TRT) proportional drift tube: design and performance". In: *Journal of Instrumentation* 3.02 (2008), P02013.

[43] D Attree et al. "The Evaporative Cooling System for the ATLAS Inner Detector". In: *Journal of Instrumentation* 3 (July 2008). DOI: 10.1088/1748-0221/3/07/P07003.

[44] Marija Marjanović. "ATLAS Tile calorimeter calibration and monitoring systems". In: *IEEE Transactions on Nuclear Science* 66.7 (2019), pp. 1228–1235.

[45] Atlas Collaboration et al. *ATLAS tile calorimeter: Technical design report*. CERN, 1996.

[46] A Airapetian et al. "ATLAS calorimeter performance". In: (1996).

[47] Coll ATLAS. *Tile Calorimeter Technical Design Report*. Tech. rep. Report CERNLHCC/96-42, 1996.

[48] Atlas Collaboration et al. *ATLAS muon spectrometer: Technical design report*. 1997.

[49] Joao Pequenao. *Computer generated image of the ATLAS Muons subsystem*. Tech. rep. 2008.

[50] L Adamczyk et al. *Technical design report for the ATLAS forward proton detector*. Tech. rep. 2015.

[51] Georges Aad et al. "Luminosity determination in pp collisions at \sqrt {s} = 7 TeV using the ATLAS detector at the LHC". In: *The European Physical Journal C* 71.4 (2011), pp. 1–37.

[52] Stefan Ask. "Status of the forward physics projects in ATLAS". In: *arXiv preprint arXiv:0706.0644* (2007).

[53] Steffen Kaiser et al. "Search for the Higgs Boson in the Process $pp \to Hqq$, $H \to WW$ with the ATLAS Detector". In: (Feb. 2023).

[54] Morad Aaboud et al. "Performance of the ATLAS trigger system in 2015". In: *The European Physical Journal C* 77.5 (2017), pp. 1–53.

[55] Rajagopalan S. "The Performance of the ATLAS Trigger System in the LHC proton-proton Collisions". In: *Physics Procedia* 37 (Dec. 2012). DOI: 10.1016/j.phpro.2012.02.503.

[56] ATLAS collaboration et al. “Luminosity determination in collisions at = 13TeV using the ATLAS detector at the LHC”. In: ATLAS-CONF-2019-021, 2019,: https://cds. cern. ch/record/2677054. 2019.

[57] G Avoni et al. “The new LUCID-2 detector for luminosity measurement and monitoring in ATLAS”. In: *Journal of Instrumentation* 13.07 (2018), P07017.

[58] Georges Aad et al. “Performance of electron and photon triggers in ATLAS during LHC Run 2”. In: *The European Physical Journal C* 80.1 (2020), pp. 1–41.

[59] ATLAS collaboration et al. “Performance of the ATLAS muon triggers in Run 2”. In: *arXiv preprint arXiv:2004.13447* (2020).

[60] ATLAS collaboration et al. *Vertex reconstruction performance of the ATLAS detector at s= 13 TeV*. Tech. rep. ATL-PHYS-PUB-2015-026, 2015.

[61] Georges Aad et al. “Electron and photon performance measurements with the ATLAS detector using the 2015–2017 LHC proton-proton collision data”. In: (2019).

[62] Georges Aad et al. “Muon reconstruction performance of the ATLAS detector in proton-proton collision data at $\sqrt{s} = 13\,\text{TeV}$”. In: *The European Physical Journal C* 76.5 (2016), pp. 1–30.

[63] Sea Agostinelli et al. “GEANT4—a simulation toolkit”. In: *Nuclear instruments and methods in physics research section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 506.3 (2003), pp. 250–303.

[64] Elzbieta Richter-Was, Luc Poggioli, and D Froidevaux. *ATLFAST 2.0 a fast simulation package for ATLAS*. Tech. rep. CERN-ATL-PHYS-98-131, 1998.

[65] Johan Alwall et al. "The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations". In: *Journal of High Energy Physics* 2014.7 (2014), pp. 1–157.

[66] Richard D Ball et al. "Parton distributions for the LHC Run II". In: *Journal of High Energy Physics* 2015.4 (2015), pp. 1–148.

[67] Torbjörn Sjöstrand et al. "An introduction to PYTHIA 8.2". In: *Computer physics communications* 191 (2015), pp. 159–177.

[68] Atlas Collaboration et al. *ATLAS Pythia 8 tunes to 7 TeV data*. Tech. rep. LHC/ATLAS Experiment, 2014.

[69] Enrico Bothmann et al. "Event generation with Sherpa 2.2". In: *SciPost Physics* 7.3 (2019), p. 034.

[70] Ansgar Denner, Stefan Dittmaier, and Lars Hofer. "Collier: a fortran-based complex one-loop library in extended regularizations". In: *Computer Physics Communications* 212 (2017), pp. 220–238.

[71] Steffen Schumann and Frank Krauss. "A parton shower algorithm based on Catani-Seymour dipole factorisation". In: *Journal of High Energy Physics* 2008.03 (2008), p. 038.

[72] Stefan Höche et al. "QCD matrix elements+ parton showers. The NLO case". In: *Journal of High Energy Physics* 2013.4 (2013), pp. 1–22.

[73] Daniel de Florian et al. “Handbook of LHC Higgs cross sections: 4. Deciphering the nature of the Higgs sector”. In: *arXiv. org* (2016).

[74] Simone Alioli et al. “A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX”. In: *Journal of High Energy Physics* 2010.6 (2010), pp. 1–58.

[75] Stefano Frixione et al. “Single-top hadroproduction in association with a W boson”. In: *Journal of High Energy Physics* 2008.07 (2008), p. 029.

[76] Michał Czakon and Alexander Mitov. “Top++: a program for the calculation of the top-pair cross-section at hadron colliders”. In: *Computer Physics Communications* 185.11 (2014), pp. 2930–2938.

[77] Georges Aad et al. “Measurement of W±and Z-boson production cross sections in pp collisions at s= 13 TeV with the ATLAS detector”. In: *Physics Letters B* 759 (2016), pp. 601–621.

[78] Atlas Collaboration et al. *The simulation principle and performance of the ATLAS fast calorimeter simulation FastCaloSim*. Tech. rep. ATL-PHYS-PUB-2010-013, 2010.

[79] ATLAS Collaboration et al. “Muon reconstruction performance of the ATLAS detector in proton–proton collision data at s= 13 TeV (2016)”. In: *arXiv preprint arXiv:1603.05598* ().

[80] M Cacciari, GP Salam, and G Soyez. “The anti-kt jet clustering algorithm. JHEP04, 063 (2008). doi: 10.1088”. In: *arXiv preprint arXiv:0802.1189* (2008).

[81] Atlas Collaboration et al. “Topological cell clustering in the ATLAS calorimeters and its performance in LHC Run 1”. In: (2017).

[82] ATLAS Collaboration et al. "Jet energy scale measurements and their systematic uncertainties in proton–proton collisions at sqrt (s)= 13 TeV with the ATLAS detector (2017)". In: *arXiv preprint arXiv:1703.09665* ().

[83] ATLAS collaboration et al. "Performance of pile-up mitigation techniques for jets in pp collisions at s= 8 TeV using the ATLAS detector (2015)". In: *arXiv preprint arXiv:1510.03823* ().

[84] ATLAS collaboration et al. "Selection of jets produced in 13 TeV proton–proton collisions with the ATLAS detector". In: ATLAS-CONF-2015-029. 2015.

[85] ATLAS collaboration et al. "Atlas b-jet identification performance and efficiency measurement with $t\bar{t}$ events in pp collisions at $\sqrt{s}$ =13 TeV". In: *arXiv preprint arXiv:1907.05120* ().

[86] Morad Aaboud et al. "Performance of missing transverse momentum reconstruction with the ATLAS detector using proton-proton collisions at $\sqrt{s} = 13\,\text{TeV}$". In: *The European Physical Journal C* 78.11 (2018), pp. 1–46.

[87] Qing-Hong Cao, Shao-Long Chen, and Yandong Liu. "Probing Higgs Width and Top Quark Yukawa Coupling from $t\bar{t}$H and $t\bar{t}t\bar{t}$ Productions". In: *arXiv preprint arXiv:1602.01934* (2016).

[88] Qing-Hong Cao et al. "Limiting top-Higgs interaction and Higgs-boson width from multi-top productions". In: *arXiv preprint arXiv:1901.04567* (2019).

[89] Kevin Greif and Kevin Lannon. "Physics Inspired Deep Neural Networks for Top Quark Reconstruction". In: *EPJ Web of Conferences* 245 (Jan. 2020), p. 06029. DOI: 10.1051/epjconf/202024506029.

[90] Stefan Von Buddenbrock et al. "Constraints on a 2HDM with a singlet scalar and implications in the search for heavy bosons at the LHC". In: *Journal of Physics G: Nuclear and Particle Physics* 46.11 (2019), p. 115001.

[91] "Search for heavy Higgs bosons decaying to a top quark pair in proton-proton collisions at $\sqrt{s} = 13\,\text{TeV}$". In: (2019).

[92] Morad Aaboud et al. "Search for Heavy Higgs Bosons A/H Decaying to a Top Quark Pair in pp Collisions at $\sqrt{s} = 8$ TeV with the ATLAS Detector". In: *Phys. Rev. Lett.* 119.19 (2017), p. 191803. DOI: 10.1103/PhysRevLett.119.191803. arXiv: 1707.06025 [hep-ex].

[93] J. H. Kühn, A. Scharf, and P. Uwer. "Weak Interactions in Top-Quark Pair Production at Hadron Colliders: An Update". In: *Phys. Rev. D* 91.1 (2015), p. 014020. DOI: 10.1103/PhysRevD.91.014020. arXiv: 1305.5773 [hep-ph].

[94] Georges Aad et al. "Search for heavy Higgs bosons decaying into two tau leptons with the ATLAS detector using pp collisions at $\sqrt{s} = 13$ TeV". In: *Phys. Rev. Lett.* 125.5 (2020), p. 051801. DOI: 10.1103/PhysRevLett.125.051801. arXiv: 2002.12223 [hep-ex].

[95] Albert M Sirunyan et al. "Search for additional neutral MSSM Higgs bosons in the $\tau\tau$ final state in proton-proton collisions at $\sqrt{s} = 13$ TeV".

In: *JHEP* 09 (2018), p. 007. DOI: 10.1007/JHEP09(2018)007. arXiv: 1803.06553 [hep-ex].

[96] Georges Aad et al. "Search for a heavy Higgs boson decaying into a Z boson and another heavy Higgs boson in the $\ell\ell bb$ and $\ell\ell WW$ final states in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector". In: *Eur. Phys. J. C* 81.5 (2021), p. 396. DOI: 10.1140/epjc/s10052-021-09117-5. arXiv: 2011.05639 [hep-ex].

[97] Albert M Sirunyan et al. "Search for new neutral Higgs bosons through the H$\rightarrow$ ZA $\rightarrow \ell^+\ell^-$b$\bar{\text{b}}$ process in pp collisions at $\sqrt{s} = 13$ TeV". In: *JHEP* 03 (2020), p. 055. DOI: 10.1007/JHEP03(2020)055. arXiv: 1911.03781 [hep-ex].

[98] "Search for heavy neutral Higgs bosons produced in association with b-quarks and decaying to b-quarks at $\sqrt{s} = 13$ TeV with the ATLAS detector". In: (Mar. 2019).

[99] Georges Aad et al. "Search for heavy neutral Higgs bosons produced in association with b-quarks and decaying into b-quarks at $\sqrt{s} = 13$ TeV with the ATLAS detector". In: *Phys. Rev. D* 102.3 (2020), p. 032004. DOI: 10.1103/PhysRevD.102.032004. arXiv: 1907.02749 [hep-ex].

[100] Morad Aaboud et al. "Search for heavy resonances decaying into a W or Z boson and a Higgs boson in final states with leptons and b-jets in 36 fb^{-1} of $\sqrt{s} = 13$ TeV pp collisions with the ATLAS detector". In: *JHEP* 03 (2018). [Erratum: JHEP 11, 051 (2018)], p. 174. DOI: 10.1007/JHEP03(2018)174. arXiv: 1712.06518 [hep-ex].

[101] Jeffrey Dean et al. "Large scale distributed deep networks". In: *Advances in neural information processing systems* 25 (2012).

[102] Kyle Cranmer, Juan Pavez, and Gilles Louppe. "Approximating likelihood ratios with calibrated discriminative classifiers". In: *arXiv preprint arXiv:1506.02169* (2015).

[103] Pierre Baldi, Peter Sadowski, and Daniel Whiteson. "Searching for exotic particles in high-energy physics with deep learning". In: *Nature communications* 5.1 (2014), p. 4308.