BOOSTED ANALYSIS OF HIGGS PAIR PRODUCTION IN THE $b\bar{b}\tau^+\tau^-$ LEPHAD FINAL STATE

by

NICHOLAS LUONGO

A DISSERTATION

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

June 2023

DISSERTATION APPROVAL PAGE

Student: Nicholas Luongo

Title: Boosted Analysis of Higgs Pair Production in the $b\bar{b}\tau^+\tau^-$ Lephad Final State

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

Laura Jeanty	Chairperson
Eric Torrence	Advisor
Laura Jeanty	Core Member
Graham Kribs	Core Member
Thien Nguyen	Institutional Representative

and

Krista Chronister	Vice Provost of Graduate Studies

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

Degree awarded June 2023

DISSERTATION ABSTRACT

Nicholas Luongo

Doctor of Philosophy

Department of Physics

June 2023

Title: Boosted Analysis of Higgs Pair Production in the $b\bar{b}\tau^+\tau^-$ Lephad Final State

This dissertation presents the development of a boosted analysis in the search for the resonant production of a new heavy scalar X decaying to two Higgs bosons, which is predicted by some Beyond the Standard Model theories. The $b\bar{b}\tau^+\tau^-$ semi-hadronic decay channel of the Higgs bosons is considered. The analysis is developed using Monte Carlo simulated data and validated with 0.11 fb^{-1} of proton-proton collisions at $\sqrt{s} = 13$ TeV from the ATLAS detector at the Large Hadron Collider (LHC). Scalar X masses of 1, 1.6, and 2 TeV are considered and expected limits of 5.29×10^3, 23.42, and 18.60 fb, respectively, are placed on the $pp \to X \to HH$ cross section at 95% confidence level. These results are compared to existing resolved and boosted ATLAS $b\bar{b}\tau^+\tau^-$ analyses. A new method for di-τ identification and a kinematic neural network for event selection are also described. This dissertation contains previously published and unpublished material.

CURRICULUM VITAE

NAME OF AUTHOR: Nicholas Luongo

GRADUATE AND UNDERGRADUATE SCHOOLS ATTENDED:

University of Oregon, Eugene, Oregon
University of Michigan, Ann Arbor, Michigan

DEGREES AWARDED:

Doctor of Philosophy, Physics, 2023, University of Oregon
Bachelor of Science, Physics, 2014, University of Michigan
Bachelor of Science, Mathematics, 2014, University of Michigan

AREAS OF SPECIAL INTEREST:

Beyond the Standard Model
Machine Learning

PROFESSIONAL EXPERIENCE:

Graduate Research Assistant, University of Oregon, 2019-2023

Graduate Teaching Assistant, University of Oregon, 2017-2019

Business Intelligence Developer, Epic Systems, 2014-2017

GRANTS, AWARDS AND HONORS:

Group Technical Achievement Award, CAMPFIRE School, 2019

Outstanding Achievement in Physics Award, University of Michigan, 2014

PUBLICATIONS:

The ATLAS Collaboration (2022). Search for resonant and non-resonant Higgs boson pair production in the $b\bar{b}\tau^+\tau^-$ decay channel using 13 TeV pp collision data from the ATLAS detector. *arXiv:2209.10910*

The ATLAS Collaboration (2020). Deep Learning for Pion Identification and Energy Calibration with the ATLAS Detector. *ATLAS Pub Note*

ACKNOWLEDGEMENTS

Thank you to my advisor, Eric Torrence, for advice and guidance toward this accomplishment. I'd also like to thank the other faculty members and students of the UO ATLAS group for their advice and collaboration.

Thank you to all of my friends, especially my cohort-mates in the PhD program. I'm grateful to have been on this journey alongside all of you to share good times and commiseration.

Thank you to my family for their love, for always being there for me, and for showing interest in this very strange subject that I've chosen to pursue.

Thank you to Emilee Jackson for all of the love and support and joy that you have given me.

This dissertation is the product of unionized labor as part of the Graduate Teaching Fellows Federation, AFT Local 3544.

TABLE OF CONTENTS

Chapter Page

I. THEORETICAL OVERVIEW 1

1.1. The Standard Model 1
1.2. Electroweak Symmetry Breaking 6
1.3. The Higgs Boson 9
1.4. Beyond The Standard Model 12

II. EXPERIMENTAL SETUP 17

2.1. The Large Hadron Collider 17
2.2. The ATLAS Experiment 24
2.3. eFEX Tau Trigger 44

III. SIMULATION AND RECONSTRUCTION 57

3.1. Monte Carlo Samples 58
3.2. Object Definitions 63

IV. BOOSTED LEPHAD $b\bar{b}\tau^+\tau^-$ 80

4.1. Samples 83

Chapter Page

4.2. Objects and Event Preselection 85

4.3. System Definition 91

4.4. Signal Region 95

4.5. Comparison to Resolved Selection 103

4.6. Background Estimation 106

4.7. Validation Against Data Sample 112

4.8. Results 114

V. CONCLUSION 118

APPENDICES

A. ML4PIONS 119

A.1. Neural Networks 119

A.2. ML4Pions Overview 122

A.3. Contributions 124

B. RESOLVED $b\bar{b}\tau^+\tau^-$ SYSTEMATIC STUDIES 129

B.1. Parton Shower 129

B.2. Scale and PDF+α_s 130

C. RUN-3 TRIGGER STUDIES 133

Chapter Page

D. KINEMATIC NEURAL NETWORK DEVELOPMENT 136

D.1. Inputs 136

D.2. Architecture 137

REFERENCES CITED 141

LIST OF FIGURES

Figure Page

1.1. Particles of the Standard Model [1] 3

1.2. The Higgs potential 7

1.3. 4-lepton invariant mass in $H \rightarrow ZZ^* \rightarrow 4l$ with excess observed at 125 GeV 10

1.4. Predicted Higgs branching ratios with uncertainties [2] 11

1.5. Triangle (a) and box (b) ggF di-Higgs production diagrams 11

1.6. BSM ggF Di-Higgs production diagram involving heavy Higgs or graviton particle denoted X. 14

2.1. Aerial view of the Geneva region with the LHC and experiments shown 18

2.2. Diagram of the LHC complex [3] 19

2.3. Diagram of Linac4 showing the four different RF systems and the energies of hydrogen ions at each stage [4] 21

2.4. Schematic of an LHC dipole magnet [5] 22

2.5. Pileup and luminosity for runs 1, 2, and (partial) 3 [6]. 24

2.6. ATLAS cross-section diagram with subsystems labeled [7] 26

2.7. Inner detector with a track shown traversing the pixel, SCT, and TRT [8] 29

2.8. Central solenoid before installation [8] 32

2.9. Cross-section of the ATLAS calorimeter system showing the EM and hadronic calorimeter with barrel, endcap, and forward regions labeled. [7] 33

2.10. Dimensions and cell granularity of the EM1, EM2, and EM3 calorimeter layers in the barrel region. [9] 34

Figure Page

2.11. Stopping power of a μ^+ in Cu. Radiative energy loss becomes significant above 100 GeV. [2] 36

2.12. Cross-sectional view of the Muon Spectrometer showing the tracking, trigger, and magnet systems. [7] 37

2.13. The trajectories of muons traversing the three stations of MDTs in the barrel region of the MS. The red line indicates a 4 GeV muon and blue indicates a 20 GeV muon [8] 38

2.14. The subsystems and data flow of the ATLAS trigger system in Run 3. Also shown are the collision and data rates at each step in the process. [7] 41

2.15. Diagrams showing the decay products of a jet (a) and a τ_{had} (b). Compared to a jet, decay products of a τ_{had} are not common in the outer green region. 45

2.16. The coverage area of a single eFEX board. Algorithms are executed in the red and yellow regions. [10] 46

2.17. A single eFEX board [7] 47

2.18. The granularity of SuperCells in a single trigger tower for the four EM layers. 48

2.19. The inequalities checked in comparison to the central tower when determining if the central tower is a local maximum in E_{T} 48

2.20. Diagrams showing the seed SuperCells (red) and those additional SuperCells included in the E_{T} calculation (orange). (a) shows the coarse-grained layers (EM0, EM3, and HAD) and (b) shows the fine-grained layers (EM1 and EM2) 49

2.21. Diagrams showing the seed SuperCells (red) and those additional SuperCells included in the isolation calculation (orange) in the EM2 layer. (a) shows the core region and (b) shows the environment region. 49

2.22. ROC curves showing the signal and background efficiencies for several different reconstructed E_{T} algorithms as cuts on the E_{T} are performed. The algorithm as has been described in this paper performed best and is noted with an arrow. 51

2.23. ROC curves showing the signal and background efficiencies for several different r_{core} algorithms as cuts on the variable are performed. The line in pink shows the final definition of the algorithm. 52

2.24. Turn-on curves showing the efficiency of a truth-matched tau candidate being created as a function of true visible p_T . 12 GeV (a) and 20 GeV (b) Run-2 threshold curves are shown alongside Run-3 algorithm curves with E_T thresholds to achieve equal rates. The final Run-3 algorithm is labeled "Oregon Run3". Also shown is another algorithm considered for Run-3 labeled "TLV Run3". Ratio plots of the Run-3 algorithms over Run-2 are also shown. 54

2.25. Turn-on curves of Run-2 and Run-3 algorithms. The Run-3 algorithm is set to equal rate to Run-2 and then a cut on the isolation variable cut is applied. The percentage of candidates removed in the given E_T bin is shown. Also given are the rate reductions achieved by this additional isolation cut. 55

2.26. Turn-on curves of Run-2 and Run-3 algorithms. The Run-3 algorithm is set to equal rate to Run-2 and then a cut on the isolation variable cut is applied. The percentage of candidates removed in the given E_T bin is shown. Also shown are the rate reduction achieved by this additional isolation cut. 56

2.27. Primary tau trigger for Run-3. The rates for 21.3 (corresponding to equivalent Run-2 triggers) and master (Run-3) are shown. Triggers in blue are newly introduced in Run-3. [11] 56

3.1. Plots of PDFs for different valence and sea partons inside the proton. [2] 60

3.2. Diagram showing the main processes of a deep inelastic scattering. The processes are hard scattering (red), QCD radiation (orange), hadronization (light green), hadron decays (dark green), and QED radiation (yellow). [12] 62

3.3. Diagram showing the signatures left in ATLAS subsystems by various particles. [13] 64

3.4. Plots showing (top) the efficiencies of electron reconstruction as a function of true E_T on a single-electron sample at various steps in the reconstruction process and (bottom) a comparison of MC and data reconstruction efficiency on $Z \to ee$ events. [14] 66

3.5. Efficiency of truth muons and muons from light hadron decays as a function of η (left) and p_{T} (right) at Loose, Medium, and Tight WPs. [15] 68

3.6. The sequence of steps involved in topocluster creation and calibration. [16] 70

3.7. Probability that a topocluster is classified as EM-like as a function of its cluster depth λ_{clus}, signal cell density ρ_{cell}, and cluster energy. The line in red divides those clusters classified as EM-like or HAD-like. [16] 70

3.8. Diagram showing the clustering patterns of several different algorithms. The anti-k_t algorithm is shown on the bottom right. Regions shown are those in which an arbitrarily soft energy deposit would be added to jets of that color. [17] 72

3.9. Diagram showing the steps of the jet energy calibration when starting from EM-scale calorimeter jets. [18] 73

3.10. Diagram showing an example initial jet, clustering into subjets, and resulting trimmed jet. [19] 74

3.11. Distribution of DL1 output variables for b-jets, c-jets, and light-flavor jets. [20] 76

3.12. Efficiencies of reconstruction and reconstruction + identification for true 1-prong and 3-prong τ_{had} . [21] 78

3.13. Distribution of electron likelihood variable and efficiencies of signal τ_{had} and inverse electron background. [21] 79

4.1. Histograms of ΔR between truth lepton and τ_{had} for signal sample events. 81

4.2. Scatter plot showing truth Higgs p_{T} and ΔR of the its $\tau\tau$ decay products. The horizontal lines show $\Delta R = 0.2$ and 0.4. 82

4.3. A 1.6 TeV signal event viewed from along the beamline (a) and off the beamline (b). The red track is an electron, green deposits the EM calorimeter cells, and yellow deposits the hadronic calorimeter cells. . . . 84

4.4. Efficiencies of reconstruction and reconstruction + identification for e and τ_{had} in 1.6 TeV signal sample. 88

4.5. Efficiencies of requiring at least one reconstructed electron passing the ID working point for eHad signal samples. 89

4.6. Efficiencies of requiring the leading large-R jet p_{T} events be greater than some threshold. The horizontal line is 90% and the vertical line is 350 GeV. 90

4.7. Percentage of events in each sample that contain only electrons, only muons, and both. 92

4.8. Diagram showing the eHad di-τ object, which looks inside an $R = 1.0$ jet for reconstructed electrons and $R = 0.2$ subjets. 93

4.9. Distribution of max S_{KNN} per event for all events in the training samples. 98

4.10. Distribution of non-max S_{KNN} per event for all events in the training samples. 99

4.11. Histograms (top left) of the S_{KNN} variable and efficiency (top right), yield (bottom left), and sensitivity (bottom right) as a function of a $S_{\mathrm{KNN}} > x$ cut. The vertical line is at $S_{\mathrm{KNN}} = 0.96$. 100

4.12. Histograms (top left) of the m_{HH} variable and efficiency (top right), event number (bottom left), and sensitivity (bottom right) as a function of a $m_{\mathrm{HH}} > x$ cut. The vertical line is at 1000 GeV. 101

4.13. Histograms (top left) of the $H \rightarrow bb$ discriminant variable D_{Hbb} and efficiency (top right), event number (bottom left), and sensitivity (bottom right) as a function of a $D_{\mathrm{Hbb}} > x$ cut. The vertical line is at $D_{\mathrm{Hbb}} = 4$. 102

4.14. Efficiency of identifying true $H \rightarrow bb$ as a function of cut on D_{Hbb}. The horizontal dashed lines are the working point D_{Hbb} values of 1.76, 2.44, and 3.09. The vertical lines are their corresponding stated efficiencies of 50%, 60%, and 70%, respectively. 103

4.15. Histograms (top left) of the S_{KNN} variable and efficiency (top right), yield (bottom left), and sensitivity (bottom right) as a function of a $S_{\mathrm{KNN}} > x$ cut with the $m_{\mathrm{HH}} > 1$ TeV cut applied. The vertical line is at $S_{\mathrm{KNN}} = 0.96$. 104

4.16. Histograms of the distribution of event weights for JZW samples. 108

Figure Page

4.17. Ratio of events measured using Hbb-out and full selection methods in relaxed signal region as a function of S_{KNN}. 110

4.18. Scatter plots of $t\bar{t}$ and JZW events in the signal (D) and control (A,B,C) regions. 111

4.19. Ratio of events measured using ABCD and full selection methods in relaxed signal region as a function of S_{KNN}. 112

4.20. Distributions of m_{HH} for all background samples including JZW and data sample. 114

4.21. Ratio of events measured using ABCD and full selection methods in relaxed signal region as a function of S_{KNN} with data sample included. 115

4.22. Expected 95% confidence limits on $\sigma_{\mathrm{pp}\to\mathrm{X}\to\mathrm{HH}}$ for each X mass point with limited systematics included. 116

4.23. Observed and expected Run-2 limits on $\sigma_{\mathrm{pp}\to\mathrm{X}\to\mathrm{HH}}$ as a function of X mass from the resolved $b\bar{b}\tau^{+}\tau^{-}$ analysis [22]. 117

4.24. Observed and expected Run-2 limits on $\sigma_{\mathrm{pp}\to\mathrm{X}\to\mathrm{HH}}$ as a function of X mass from the boosted $b\bar{b}\tau^{+}\tau^{-}$ hadhad analysis [23]. 117

A.1. Diagram of a deep neural network with one hidden layer. 121

A.2. Diagram of a convolutional layer showing the application of a filter to one region of the input layer and the resulting value in the output layer. [24] 122

A.3. Difference in average energy deposition by π^{+} and π^{0} in the EMB2 layer of the calorimeter. [25] 123

A.4. DNN (a) and CNN (b) architectures for energy regression. [25] 124

A.5. Median energy response for combined CNN classification + DNN regression compared to LCW and EM energy scales. [25] 125

A.6. Topocluster image example before (a) and after (b) resizing to increase the granularity by a factor of two in both dimensions. 126

A.7. Ratio of τ_{had} topocluster energy to true τ_{had} visible p_{T} before and after application of energy calibration. 127

Figure Page

A.8. Normalized distributions of topocluster energies for τ_{had} and pion samples. 128

A.9. Ratio of τ_{had} (a) and pion (b) topocluster energy to truth energy value before and after application of energy calibration. 128

B.1. Normalized distribution of PNN score for samples using nominal and alternate PS generators at the 280 GeV mass point in the SLT channel. 130

B.2. $m_{\text{HH}}$of the 500 GeV signal sample in the SLT channel with scale variations and envelope. 131

B.3. Signed PS uncertainties as a function of the signal sample mass point. 132

C.1. Turn-on curves for HLT triggers requiring leading large-R jet p_{T} of 460 GeV and 420 GeV + 35 GeV mass cut for 1, 1.6, and 2 TeV signal samples. 134

C.2. Efficiencies of each trigger for all lephad signal samples. 134

C.3. Efficiencies of sets of combined triggers for lephad signal samples. + denotes the OR-ing together of the individual triggers. 134

C.4. Efficiencies of single and combined triggers for hadhad signal samples. + denotes the OR-ing together of the individual triggers. 135

D.1. Distribution of S_{KNN} variable with dropout applied between all layers. 138

D.2. Sensitivity as a function of S_{KNN} for network trained with only signal events (a) and with signal, $t\bar{t}$, and Z+jets events (b). Backgrounds used in sensitivity calculation are $t\bar{t}$ and Z+jets. 139

D.3. Sensitivity as a function of S_{KNN} for network trained with (a) and without (b) JZW events included in the training set. Backgrounds used in sensitivity calculation are $t\bar{t}$ and Z+jets. 140

LIST OF TABLES

Table Page

4.1. Processes and generators of samples used in this analysis. 85

4.2. Input event and output event numbers and preselection efficiencies for all samples. Efficiencies in the fiducial region are also provided for signal samples. In all cases, raw event numbers are used without luminosity or event weights applied. 91

4.3. Percentages of TM (network label 1) and non-TM (network label 0) events kept and discarded when performing channel sorting for signal, $t\bar{t}$, and Z+jets events. 97

4.4. Number of events in signal, $t\bar{t}$, and Z+jetssamples that pass the boosted and approximate resolved selections as well as the number passing both. 106

4.5. σ_{sample}, e_{filter}, N_{D}, and w_{sample} for all signal and background samples. 107

4.6. Selection efficiencies, number of events in the signal region, and associated errors for Z+jets and W+jets samples. 109

4.7. Control region definitions. - denotes that no value is specified 110

4.8. Events in control and signal regions plus statistical and systematic uncertainties for $t\bar{t}$ and JZW samples. 112

CHAPTER I

THEORETICAL OVERVIEW

1.1. The Standard Model

The formalization of our current understanding of all known forces and matter in the universe is collectively referred to as the Standard Model (SM). This model has been extremely successful in explaining the results of high energy particle physics experiments to date.

The Standard Model is a quantum field theory (QFT) that describes the existing particles and the forces that govern their interactions. The forces are the electromagnetic, weak nuclear, and strong nuclear forces. Gravity, the last of the known forces, is not explained by the Standard Model.

The particles of the SM are categorized by their properties and interactions. Firstly, they can be divided into bosons and fermions depending on whether they carry integer or half-integer spin, respectively.

Gauge bosons have spin 1 and act as carriers for the forces mentioned above. The photon is the gauge boson for the electromagnetic force, the W and Z for the weak nuclear force, and the gluon for the strong nuclear force. The Higgs boson is the only fundamental scalar (spin 0) particle in the SM.

The photon interacts with particles that carry an electromagnetic (EM) charge. However, it does not itself carry this charge, meaning that a particle cannot have its charge changed due to interaction with a photon.

The gluon interacts with particles that carry color charge, which is the charge of the strong nuclear force. The colors are red, green, and blue plus an

associated anti-color for each. Gluons carry color pairs that are a combination of a color and an anti-color, for example $r\bar{g}$. One characteristic of color charge is color confinement - that particles with a nonzero color charge cannot exist freely without forming a color singlet state immediately.

The $W^{\pm}$ and Z bosons interact with particles according to their values of weak isospin T_3 and weak hypercharge Y_W. Interactions with a $W^{\pm}$ boson are capable of changing a particle's flavor. They will also change the EM charge of a particle due to itself carrying a $\pm$ charge. The Z boson is not capable of changing the flavor or charge of particles interacting with it.

The fermions are further split into quarks and leptons depending on their interactions. Fermions are also simultaneously organized into three generations which have similar properties within themselves and are characterized by increasing particle mass when moving to higher generations.

Quarks interact through all three forces of the SM. Each of the three generations has two quarks, giving six flavors in total. They are the up and down, the charm and strange, and the top and bottom. The up, charm, and top quarks are referred to as up-type quarks and have electric charge of +2/3. The down, strange, and bottom quarks are referred to as down-type quarks and have electric charge of -1/3. Quarks carry weak isospin and color charge. Color confinement causes quarks to combine to form color singlet states, the most common of which are mesons, comprising two quarks, and baryons, comprising three quarks. These both fall under the category of hadrons, which is a general term for particles made up of quarks and held together by the strong nuclear force.

The proton and neutron are examples of hadrons. The proton is a bound state containing two up quarks and one down. In addition to these *valence* quarks

are a sea of other quarks and gluons being constantly exchanged between them. It is really these constituents that interact with each other during proton-proton collisions of interest. The fact that a proton's quarks and gluons often have a non-zero momentum in the rest frame of the proton itself must be considered when measuring proton-proton collisions.

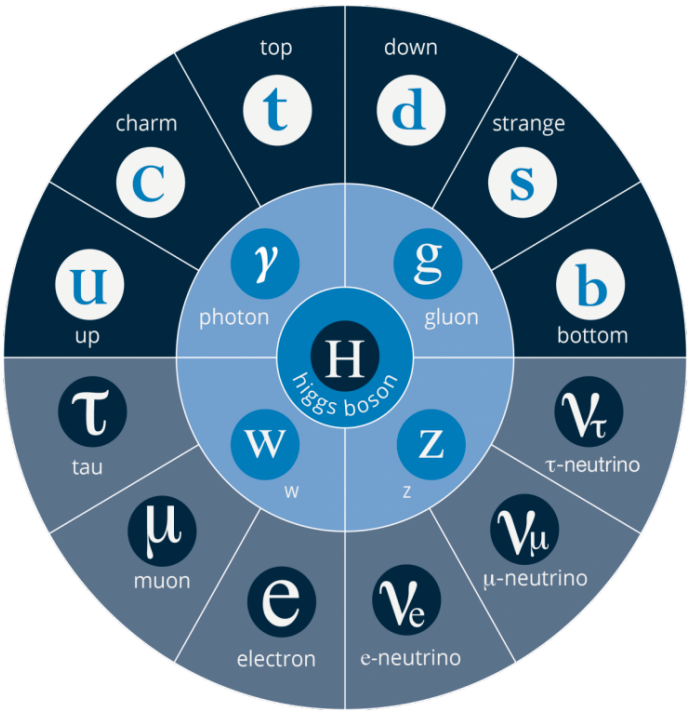


FIGURE 1.1. Particles of the Standard Model [1]

Leptons are fermions that do not interact through the strong nuclear force. They carry a conserved quantity called lepton number. Each lepton generation has one charged lepton and one lepton neutrino which is neutral. The names of the leptons are the electron, muon, and tau and associated neutrinos are the electron neutrino, muon neutrino, and tau neutrino, respectively.

Figure 1.1 shows all the particles of the SM. All charged particles shown also have a corresponding anti-particle, which has all the properties of the original except for charge, which is opposite that of the original.

1.1.1. The Lagrangian of the Standard Model

Formally, the Standard Model can be written down as a Lagrangian of its constituent quantum fields. Each particle described in the previous section has a corresponding field present in the Lagrangian.

There are also a number of free parameters of the model that can only be determined through experimentation. The exact definition of each parameter can depending on the formulation, but one places the number at 26 [26]. This includes the masses of all fundamental particles including the Higgs and neutrinos (13); Higgs vacuum expectation value to be described in Section 1.2; quark mixing angles of the CKM matrix (3); a CKM CP-violating phase; neutrino PMNS matrix parameters (4); QCD vacuum angle; and U(1), SU(2), and SU(3) gauge couplings (3).

This Lagrangian is invariant under a SU(3)×SU(2)$_{\mathrm{L}}$ ×U(1) gauge symmetry. The SU(3) symmetry corresponds to the strong nuclear force while the SU(2)$_{\mathrm{L}}$ ×U(1) portion corresponds to the weak nuclear and EM forces, referred to together as the *electroweak* sector. This will be described in more detail in Section 1.2.

The overall form of the Lagrangian [27] is:

$$\begin{aligned} & -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \overline{\psi}(i\gamma^{\mu}D_{\mu})\psi \\ & + (D_{\mu}\Phi)^{\dagger}D^{\mu}\Phi + \mu^{2}\Phi^{\dagger}\Phi - \frac{\lambda}{2}(\Phi^{\dagger}\Phi)^{2} \\ & + \overline{\Psi_{L}}y\psi_{R}\Phi + h.c. \end{aligned} \tag{1.1}$$

assuming a (+,-,-,-) metric. Different terms in the Lagrangian are responsible for different behaviors in the resulting theory.

The following term is responsible for the kinetic behavior of the gauge fields:

$$F_{\mu\nu}F^{\mu\nu} \tag{1.2}$$

The $F_{\mu\nu}$ term is called the field strength tensor for a given force. Through the $F_{\mu\nu}F^{\mu\nu}$ term in the Lagrangian it is responsible for the kinetic behavior of the gauge fields. Writing this term out for each force gives:

$$G^a_{\mu\nu}G^{a\mu\nu} + A^i_{\mu\nu}A^{i\mu\nu} + B_{\mu\nu}B^{\mu\nu} \tag{1.3}$$

where G is the gauge field for the strong nuclear force and $a = 1, .., 8$ for each gluon. A^i and B together make up the electroweak sector and $i = 1, 2, 3$ for each of the three A gauge fields.

The $\overline{\psi}(i\gamma^\mu D_\mu)\psi$ term contains the covariant derivative D_μ which has the form:

$$D_\mu = \partial_\mu + \frac{1}{2}ig'YB_\mu + \frac{1}{2}ig\sigma^i_L A^i_\mu + \frac{1}{2}ig_s T^a G^a_\mu \tag{1.4}$$

where g', g, and g_s are constants and Y, σ_L, and T are the generators of the U(1), $\mathrm{SU(2)_L}$, and SU(3) groups respectively. The first term provides the kinetic term for fermions and the rest are coupling terms between fermions and the gauge fields.

$\overline{\psi}(i\gamma^\mu D_\mu)\psi$ overall is the kinetic term for the fermion fields ψ. The other term relevant to fermions is $\overline{\Psi_L}y\psi_R\Phi$, which is referred to as a Yukawa coupling and will be discussed in the next section.

From experiment it is observed that only left-handed fermions participate in interactions with $\mathrm{SU(2)_L}$ gauge bosons. This motivates the grouping of left-handed particles into isospin doublets Ψ_L. Each lepton is paired with its

corresponding neutrino and each up-type quark with the down-type quark of the same generation.

Right-handed fermions remain in singlets ψ_R. This means that a mass term for fermions must contain a left-handed doublet that transorms under $\mathrm{SU(2)_L}$ with a right-handed singlet that does not. This makes it impossible for the overall mass term to be symmetric under the gauge transformation of $\mathrm{SU(2)_L}$.

Similarly to fermion mass terms, the mass terms of gauge bosons are also forbidden to maintain symmetry under the associated gauge transformation. This is respected for the EM and strong nuclear forces, whose gauge bosons, the photon and gluons, are indeed massless. However, the gauge bosons of the weak nuclear force, the $W^{\pm}$ and Z bosons, do have measured non-zero masses.

Explaining fermion masses and the masses of the $W^{\pm}$ and Z bosons are the primary motivations for the theory of electroweak symmetry breaking and the theorizing of the existence of the Higgs boson.

1.2. Electroweak Symmetry Breaking

The masses for fermions and weak nuclear force gauge bosons can be explained by first positing the existence of an unbroken electroweak sector with a gauge symmetry of $\mathrm{SU(2)\times U(1)}$. The corresponding gauge fields are A^i_μ where $i = 1, 2, 3$, and B_μ. Additionally, a complex doublet field called the Higgs field exists and couples to both the electroweak gauge fields and to all fermions except neutrinos.

The Lagrangian for the Higgs sector is the second line of Equation 1.1. The potential

$$V(\Phi) = -\mu^2 \Phi^\dagger \Phi + \frac{\lambda}{2}(\Phi^\dagger \Phi)^2 \tag{1.5}$$

has stationary values at $\Phi^\dagger\Phi = 0, \pm\frac{\mu^2}{\lambda}$. For $\mu^2 > 0$ and $\lambda > 0$ the value at 0 is unstable and the others are stable. Therefore, even if the field were to initially begin at 0 it would evolve to a non-zero state. The value $v = \sqrt{\frac{\mu^2}{\lambda}}$ is referred to as the *vacuum expectation value* (vev) of the Higgs field. Because all nonzero solutions are degenerate, we can without loss of generality choose it to be:

$$\Phi = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + H(x) \end{pmatrix} \tag{1.6}$$

with $v = 246$ GeV and $H(x)$ chosen in the real direction. Figure 1.2 shows the shape of the potential as a function of the non-zero coordinate.

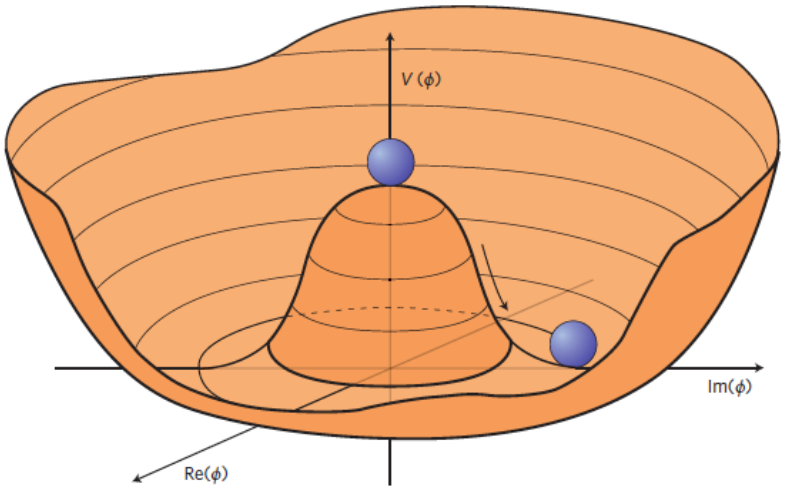


FIGURE 1.2. The Higgs potential

Plugging Equation 1.6 into Equation 1.1 yields many new terms. Included among them are:

$$\mathcal{L} \supset \frac{1}{2}\partial_\mu H \partial^\mu H - \lambda v^2 H^2 - \lambda v H^3 - \frac{1}{4}\lambda H^4 \tag{1.7}$$

The first two together describe a new scalar particle H with mass $m_H^2 = 2\lambda v^2$. The next two are tri- and quadri-linear terms of H coupling to itself. Note that the H^3 term can be rewritten using the new mass value as $\sqrt{\frac{\lambda}{2}} m_H H^3$. Once m_H is known, the only unknown in the H^3 and H^4 couplings is λ. This is not true for other terms in which λ only appears in a ratio with μ. Measurements of the H^3 and H^4 couplings therefore present a unique chance to measure the value of λ.

Other terms arise from the covariant derivative which couples the electroweak sector to the Higgs field. Introducing the Weinberg angle θ_W, which has the properties $\sin\theta_W = \frac{g'}{\sqrt{g^2+g'^2}}$ and $\cos\theta_W = \frac{g}{\sqrt{g^2+g'^2}}$ we can write these terms in the suggestive form:

$$\frac{v^2}{8}[g^2(A_\mu^1 \pm iA_\mu^2)^2 + (g^2+g'^2)(\cos\theta_W A_\mu^3 + \sin\theta_W B_\mu)^2] \tag{1.8}$$

These are mass terms for the $\frac{1}{\sqrt{2}}(A_\mu^1 \pm A_\mu^2)$ vector fields with a mass of $\frac{1}{2}vg$, as well as a mode that is a mix of the A_μ^3 and B_μ fields with a mass of $\frac{1}{2}v\sqrt{g^2+g'^2}$. The two vector bosons with the same mass are identified as the $W^\pm$ bosons and the remaining combination with a greater mass is the Z. The last linearly independent combination of these fields is $A_\mu = -\sin\theta_W A_\mu^3 + \cos\theta_W B_\mu$. As no such term exists in the set above, we can associate it with the massless photon.

The next set of terms are very similar to those in 1.8.

$$\begin{aligned}&\frac{vH}{4}[g^2(A_\mu^1 \pm iA_\mu^2)^2 + (g^2+g'^2)(\cos\theta_W A_\mu^3 + \sin\theta_W B_\mu)^2] + \\ &\frac{H^2}{8}[g^2(A_\mu^1 \pm iA_\mu^2)^2 + (g^2+g'^2)(\cos\theta_W A_\mu^3 + \sin\theta_W B_\mu)^2]\end{aligned} \tag{1.9}$$

The top line describes vertices with one Higgs and two vector bosons and the second line describes vertices with two Higgs and two vector bosons.

A similar process follows from the Yukawa couplings in Equation 1.1. Expanding this around the vev e.g. for the up quark yields:

$$\frac{1}{\sqrt{2}} v y_u \overline{u_L} u_R + \frac{1}{\sqrt{2}} H y_u \overline{u_L} u_R \tag{1.10}$$

where y_u is an entry in a matrix of Yukawa coupling. The first term is a mass term with $m = \frac{1}{\sqrt{2}} vy$ and the second is the coupling of the quark to the Higgs field. Equivalent terms exist for every fermion except the neutrinos. It can be seen that the same value y appears in both terms, meaning that more massive fermions couple more strongly to the Higgs field.

1.3. The Higgs Boson

The Higgs boson was initially proposed in 1964 by Higgs [28] and discovered in 2012 by the ATLAS and CMS experiments of the Large Hadron Collider at CERN [29][30]. This discovery led to the Nobel Prize in Physics 2013 being awarded jointly to Higgs and Englert, who independently proposed the theory of electroweak symmetry breaking [31].

The Higgs boson has a predicted lifetime of 1.62×10^{-22} seconds, so it decays before ever reaching the ATLAS and CMS detectors. Therefore its existence was instead inferred by searching for its decay products. The production cross section of the original particle has the following form when considering energies near its rest mass:

$$\sigma \propto \frac{1}{(E - m_0)^2 + (\Gamma/2)^2} \tag{1.11}$$

Here E is the energy of the collision and by conservation of energy is also equal to the invariant mass of all decay products, m_0 is the particle's rest mass, and Γ is the particle's width. From this it can be seen that as E approaches m_0, σ attains a maximum value. This peak is referred to as a *resonance* and it is a common way of discovering new particles, the Higgs included. Seeing as the Higgs excess is centered at 125 GeV, we conclude that to be its rest mass. The observed resonance for the Higgs that led to its discovery is shown in Figure 1.3.

The original analysis considered the $H \to ZZ^* \to 4l$, $H \to \gamma\gamma$, and $H \to WW^* \to e\nu\mu\nu$ channels due to having superior sensitivity over other channels. Figure 1.3 shows the excess observed in the invariant mass of 4-lepton systems in the $H \to ZZ^* \to 4l$ channel.

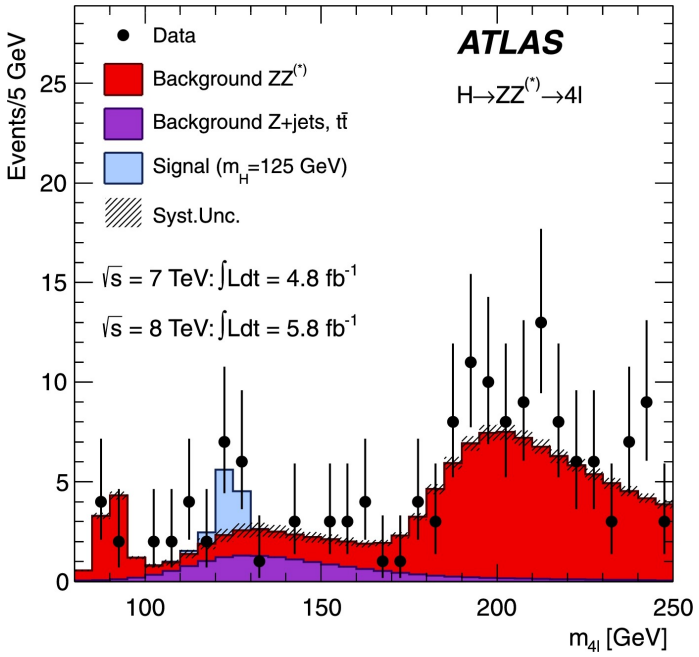


FIGURE 1.3. 4-lepton invariant mass in $H \to ZZ^* \to 4l$ with excess observed at 125 GeV

Other values of interest are the branching ratios of the Higgs, which are the fractions of its decays to given final states. Predicted SM values are shown in Figure 1.4. At $m_H = 125$ GeV, our decay modes of interest are $H \to b\bar{b}$ with a branching ratio of $58.2\%^{+1.2\%}_{-1.3\%}$ and $H \to \tau^+\tau^-$ with a branching ratio of $6.27\%^{+1.6\%}_{-1.6\%}$ [2].

Decay channel	Branching ratio	Rel. uncertainty
$H \to \gamma\gamma$	2.27×10^{-3}	2.1%
$H \to ZZ$	2.62×10^{-2}	$\pm$1.5%
$H \to W^+W^-$	2.14×10^{-1}	$\pm$1.5%
$H \to \tau^+\tau^-$	6.27×10^{-2}	$\pm$1.6%
$H \to b\bar{b}$	5.82×10^{-1}	$^{+1.2\%}_{-1.3\%}$
$H \to c\bar{c}$	2.89×10^{-2}	$^{+5.5\%}_{-2.0\%}$
$H \to Z\gamma$	1.53×10^{-3}	$\pm$5.8%
$H \to \mu^+\mu^-$	2.18×10^{-4}	$\pm$1.7%

FIGURE 1.4. Predicted Higgs branching ratios with uncertainties [2]

Di-Higgs production refers to processes in which two Higgs bosons are produced before they each decay independently. The most common SM production mode is gluon-gluon fusion, whose Feynman diagrams are shown in Figure 1.5.

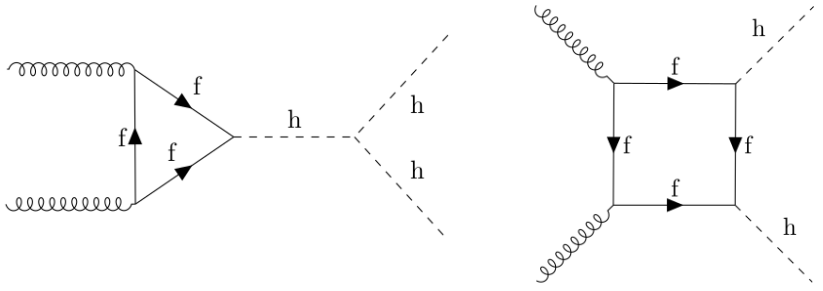


FIGURE 1.5. Triangle (a) and box (b) ggF di-Higgs production diagrams

As mentioned in Section 1.2, the sensitivity to the trilinear coupling and therefore λ of the process in Figure 1.5 (a) motivates much interest in the precise measurement of the SM di-Higgs production cross section.

The top quark is the most common particle to take the place of f in the diagram as its large mass corresponds to the largest coupling to the Higgs. In calculating the total cross section of the di-Higgs production, the two diagrams above interfere with each other, resulting in a very small overall cross section.

1.4. Beyond The Standard Model

Despite the incredible successes of the Standard Model, there are several known phenomena that it is not able to explain. As mentioned before, the SM cannot currently incorporate strongly-coupled gravity and does not provide a reason for dark energy at the scale that it is observed.

Dark matter (DM) is another substance about which very little is known. It is a theoretical form of matter which interacts gravitationally but not electromagnetically with the matter of the SM. One source of evidence of dark matter comes from galaxy rotation curves. If light-emitting matter constitutes the majority of the mass in a galaxy, one would expect the velocity of stars to decrease as their radius increase outside the central bulge. However, observed curves are flat which suggests an additional source of non-luminous mass distributed in a halo about the galaxy [32]. Other sources of evidence are a smaller than predicted baryonic density fluctuations [2], anisotropy power spectrum of the CMB [2], and the mass distributions in the merging of the Bullet cluster [33].

The baryon asymmetry problem is the name given to the observation that almost the entirety of the observed universe is in the form of matter as

opposed to anti-matter. In theory, equal amounts should have been produced at the beginning of the universe and indeed, they are produced equally in particle colliders. Processes that can contribute to producing the asymmetry are those that violate baryon number and those that violate C and CP symmetry [34]. Some BSM theories contain additional sources for these beyond those found in the SM [35].

Extensions of the SM have therefore been proposed to help explain these shortcomings through new undiscovered particles and forces. Because the Higgs is the newest and therefore least understood and measured of the SM particles, these new theories often propose extensions of the so-called Higgs sector or tie the new particles to the Higgs boson in some way. A new, very massive particle may also couple strongly to the Higgs boson, making it an effective tool to probe these theories.

One example is a family of theories called the 2-Higgs Doublet Model (2HDM) [36], which appears in minimal supersymmetric theories among others. In these models, there is another complex doublet Higgs field in addition to the previously described SM Higgs doublet.

If both doublets acquire vevs during electroweak symmetry breaking, then two massive scalars appear instead of the one in the SM. The two scalars will have unequal masses, with the less massive denoted h and the more massive denoted X, and may couple together through Xhh terms that arise in the Lagrangian after EWSB. If the SM Higgs is taken to be h and the heavier X can be produced in the proton-proton collisions of the LHC, it may decay into pairs of SM Higgs as shown in Figure 1.6 so long as $m_X > 2m_h$.

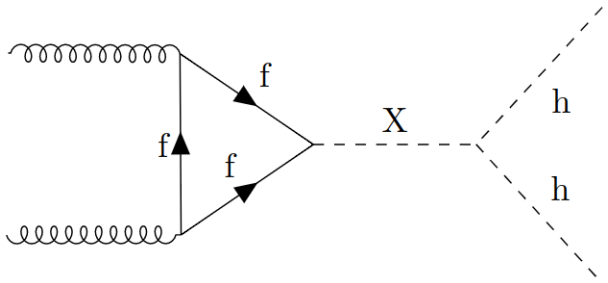


FIGURE 1.6. BSM ggF Di-Higgs production diagram involving heavy Higgs or graviton particle denoted X.

Another model that allows for an equivalent diagram is the bulk Randall-Sundrum model, in which the heavy Higgs is replaced by the spin-2 Kaluza-Klein excitation of the graviton [37].

The existence of the new scalar X decaying to two Higgs bosons will enhance the overall di-Higgs production cross section beyond that of the SM. In reality, current ATLAS analyses are not sensitive enough to measure the small SM di-Higgs production cross section, meaning dominant sources of observed events come from other SM processes. The major SM processes considered here are $t\bar{t}$, Z+jets, W+jets, and di-jet production. An observed excess over these is evidence of the existence of the X scalar. Such an excess is most likely to be observed as a resonance in events whose reconstructed di-Higgs mass is close to the mass of the X.

Because the ATLAS detector was designed with the detection of Higgs decays as a top priority, it is also well suited for di-Higgs final states. From the Higgs decay modes in Figure 1.4, it can be seen that the Higgs bosons will decay to the $b\bar{b}\tau^+\tau^-$ final state 7.3% of the time. Of these events, the two τ then decay in the $\tau_{lep}\tau_{had}$ channel 46% of the time, giving an overall $b\bar{b}\tau_{lep}\tau_{had}$ final state 3.3%

of the time, which places it among the more common of the possible di-Higgs final states.

The existence of high-p_{T} leptons and b-jets in this final state allow for better separation between $X \rightarrow HH \rightarrow b\bar{b}\tau^{+}\tau^{-}$ and SM processes compared to some other di-Higgs channels. The lepton also provides a good object on which to trigger. A reasonable reconstruction of the important di-Higgs mass discriminating variable is possible despite the presence of undetectable neutrinos. These characteristics together make the $b\bar{b}\tau^{+}\tau^{-}$ channel competitive in sensitivity with other di-Higgs decay channels that may be more distinct from SM processes or have larger cross sections.

Those SM processes listed above have the potential to produce some combination of b-jets, jets of other flavors which may be mistaken for b-jets, leptons, and τ_{had} in the final state. Requiring these objects with the intention of selecting $b\bar{b}\tau^{+}\tau^{-}$ events will also select events from these SM processes to some extent. In addition, these processes all have relatively large cross sections which increases the amount that they will contaminate a selection intended for $b\bar{b}\tau^{+}\tau^{-}$ events.

ATLAS analyses have been performed to measure di-Higgs production in many different final states based on the Higgs decay channels. The following analyses searched for resonant and/or non-resonant production and were performed with Run-2 data at $\sqrt{s} = 13$ TeV. A resolved $b\bar{b}\tau^{+}\tau^{-}$ analysis in the $\tau_{\mathrm{lep}}\tau_{\mathrm{had}}$ and $\tau_{\mathrm{had}}\tau_{\mathrm{had}}$ channels was performed [22] as was a boosted analysis in the $\tau_{\mathrm{had}}\tau_{\mathrm{had}}$ channel [23]. Results of these analyses in the form of observed and expected limits will be compared against the expected limits produced by this

analysis. Other ATLAS analyses have been performed in the $b\bar{b}l\nu l\nu$ [38], $b\bar{b}b\bar{b}$ [39][40], and $b\bar{b}\gamma\gamma$ [41] final states.

CHAPTER II

EXPERIMENTAL SETUP

The results presented pertain to the ATLAS experiment located at the Large Hadron Collider (LHC). Section 2.1 will describe the LHC and section 2.2 will outline the ATLAS detector and subsystems. The LHC and ATLAS are currently in Run-3, which began in July 2022. Unless stated otherwise, given values and machine descriptions are those for Run-3.

2.1. The Large Hadron Collider

The Large Hadron Collider (LHC) is a particle accelerator located underground on the Franco-Swiss border near the city of Geneva, Switzerland. Inside the LHC, two beams of protons circulate in opposite directions within a 27 km in circumference ring before being directed into head-on collisions with one another.

The proton beams are caused to intersect with each other at four points spaced evenly around the ring of the LHC. It is at each of these points that one of four major experiments observes the resulting proton-proton collisions. The experiments are ATLAS [8], CMS [42], ALICE [43], and LHCb [44]. ATLAS and CMS are general-purpose detectors, ALICE studies heavy-ion collisions, and LHCb focuses on processes containing b-quarks.

The LHC is currently in Run-3 of its operation schedule, which began in July 2022. Each run is a separate data-taking phase between which upgrades and maintenance are done on the LHC and its experiments. In Run-3, the center-of-mass (COM) energy and luminosity will achieve their highest values yet.

FIGURE 2.1. Aerial view of the Geneva region with the LHC and experiments shown

2.1.1. Proton Lifecycle

The protons that will be collided begin as hydrogen atoms that are caused to capture an additional electron to make negatively charged hydrogen ions H^-.

$$H + e^- = H^- \tag{2.1}$$

These ions then enter the Linac4 linear accelerator where they are accelerated to 160 MeV [4]. Upon entering the next stage of the accelerator, the ions are passed through a carbon foil that strips off the electrons leaving bare protons. After leaving the Linac4 system the protons enter a series of synchrotron systems - the Proton Synchrotron Booster (PSB), Proton Synchrotron (PS), Super Proton Synchrotron (SPS), and finally the LHC itself.

Protons are accelerated in all of these systems by way of radio frequency (RF) cavities of various types. In the synchrotrons, the beams are bent around the circular paths by dipole magnets and focused more tightly by quadrupole and

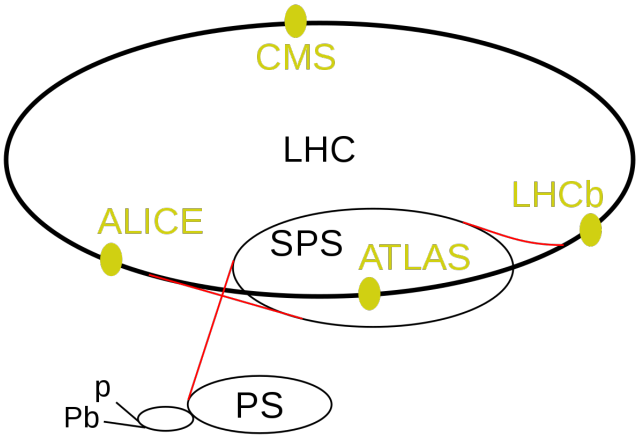


FIGURE 2.2. Diagram of the LHC complex [3]

higher order multipole magnets. These systems will be discussed in more detail in Section 2.1.2.

The beams are accelerated in the PSB, PS, SPS, and LHC to 2 GeV, 26 GeV, 450 GeV, and 6.8 TeV respectively [45][46]. Energy must be imparted to increase their speed but also to restore energy lost by synchrotron radiation, a consequence of any relativistic charged particle undergoing an acceleration perpendicular to its direction of motion. The magnitude of energy lost in this way scales as $\frac{1}{m^4}$. This is a significant advantage of colliding protons compared to, say, the less massive electron.

It is in the PS that the protons are formed into bunches, concentrated groups evenly spaced apart, each containing approximately 10^{11} protons each. Before entering the LHC from the SPS, they are split into two beams that travel in opposite direction around the LHC ring. Once the beams are at their final energy they are directed to cross at the four interactions points where the results will be observed by the main experiments. Collisions have a total center-of-mass energy of 13.6 TeV, double that of each individual beam.

2.1.2. RF and Magnet Systems

Protons being charged particles, they can be accelerated using electrical and magnetic fields. The force felt by a charged particle in a magnetic and electric field is governed by the Lorentz force:

$$\boldsymbol{F} = q(\boldsymbol{E} + \boldsymbol{v} \times \boldsymbol{B}) \tag{2.2}$$

where q is the particle's charge, $\boldsymbol{v}$ is its velocity, $\boldsymbol{E}$ is the electric field, and $\boldsymbol{B}$ is the magnetic field. The electric field term will be most relevant for increasing the energy of a particle and the magnetic field term for bending a particle in the circular path of the accelerators.

Radio frequency (RF) cavities are used in both the linear accelerator and synchrotrons to create electrical fields that accelerate the protons. These cavities are shaped in such a way that when electromagnetic waves are directed into them a resonance effect causes large oscillating electric fields to develop. With careful timing, a charged particle traveling through these cavities will experience a pull while entering the cavity and a push while exiting. The electric field can also shape the bunches longitudinally by encountering them on the crest of the waveform, giving a slight extra acceleration to the protons behind the center and less to those ahead of it. In this way both ends are directed closer to the center, compressing the bunch.

Different designs are used depending on the energies of the particles involved, with the four different subsystems of the Linac4 shown in Figure 2.3. In the LHC, there are 8 RF cavities per beam which deliver 2 MV at a frequency of 400 MHz. As the cavities are superconducting, they must be maintained at a temperature

of 4.5K. From injection at 450 GeV to maximum energy, a given proton will pass through the cavities roughly 10^7 times over the course of around 20 minutes [47].

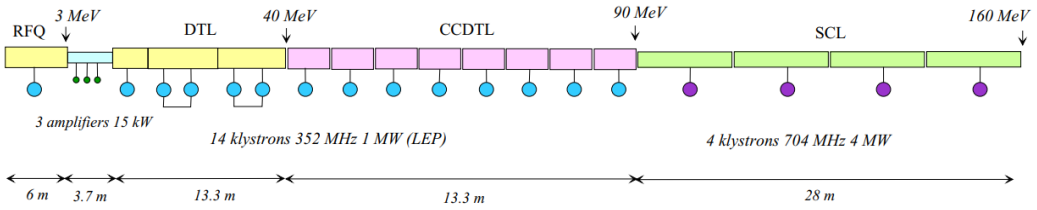


FIGURE 2.3. Diagram of Linac4 showing the four different RF systems and the energies of hydrogen ions at each stage [4]

The two main types of magnet featured in the accelerator systems are dipole and quadrupole magnets. Higher order multipole magnets are also present that provide corrections to the produced field. Magnets are superconducting and have an operating temperature of 1.9K which is achieved with liquid helium.

Dipoles in the LHC provide a uniform magnetic field to bend the proton beams at 1232 points around the circular accelerators. This means that although their broad trajectory appears ring-like, it is more accurately described as many straight segments with bends at the dipoles. Each of these dipoles produces a magnetic field of 8.3T. The PSB, PS, and SPS similarly bend the beam using 32, 100, and 744 dipole magnets with field strengths of 0.87, 1.24, and 2.02 T, respectively [48].

The quadrupoles are used to focus and compress the proton beams as much as possible, allowing for a greater chance of collision. A single quadrupole can only focus along a single axis, but overall focusing can be achieved by arranging a series with different orientations. Special quadrupole groups called *inner triplets* are positioned just before the interaction points to reduce the beam size by around a factor of ten at the point of collision.

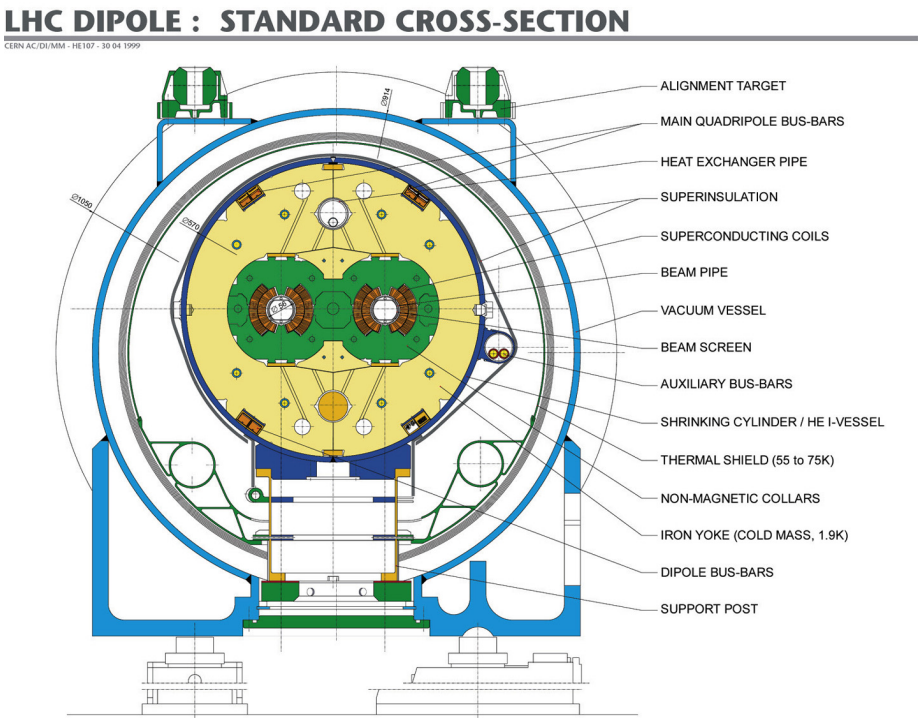


FIGURE 2.4. Schematic of an LHC dipole magnet [5]

2.1.3. Proton-Proton Collisions

At the time of collision, each beam is made up of 2808 bunches equally spaced around the LHC ring. With the bunches traveling at nearly the speed of light, they collide at each interaction point with a frequency of 40 MHz, giving a time between collisions of 25 ns. The number of interactions that are produced from these collisions is given by:

$$N = \sigma L \tag{2.3}$$

where σ is the cross section of the proton-proton interaction process and L is called the *luminosity*. This equation is valid for a total number of events over a given time in which case L is called the integrated luminosity or a rate of events in

which case L is called the instantaneous luminosity. The cross section has no time dependence and is the same in both cases.

The instantaneous luminosity of the LHC collisions can be represented as:

$$L \propto \frac{N^2 f_{rev}}{4\epsilon\beta^*} \tag{2.4}$$

where N is the number of protons in a bunch, f_{rev} is the frequency of revolution around the LHC ring, ϵ is the emittance, and β^* is the beta function at the collision point. The emittance is a measure of how well the beam is prepared in terms of the protons having uniform momentum and being spaced closely together. Beta is a measure of the ability of the magnet system to focus the beam into as small an area as possible. Both ϵ and β^* have units of length. Omitted are a gamma factor and geometric factor from the beams colliding at a slight angle. With nominal values of $N = 1.15 \times 10^{11}$, $f_{rev} = 40$MHz, $\epsilon = 3.75\mu$m, $\beta^* = 0.55$m, and allowing for the fact that some bunch positions are left empty to allow for the safe dumping of the beam, one can make an approximation of the nominal instantaneous luminosity of $10^{34}\text{cm}^{-2}\text{s}^{-1}$ [47]. This was quickly surpassed after the beginning of LHC operation and expectations are to double it for Run-3 [49]. The total integrated luminosity of Run-3 is expected to be 350 fb^{-1}, more than both previous runs combined.

The cross section σ of proton-proton collisions can be calculated using the tools of QFT. In general, we are specifically interested in the cross section of inelastic collisions i.e. those for which new particles are created. The inelastic cross section at 13 TeV is 78.1 ± 2.9 mb (1 barn = $10^{-28}m^2$) [50].

The number of actual interactions between two protons at each bunch crossing is expected to be ∼60 in Run-3. This is much higher than the average

of around 33 for Run-2 which can be seen in Figure 2.5. This means that if one of these interactions produces a final state of interest, the detector signature will also include the outgoing particles from all the others as well. This effect is referred to as *run123pileuplumi*, and it can be in-time, from the same bunch crossing as the interaction of interest, or out-of-time, from the other crossings in a window before or after whose effects may overlap in the detector.

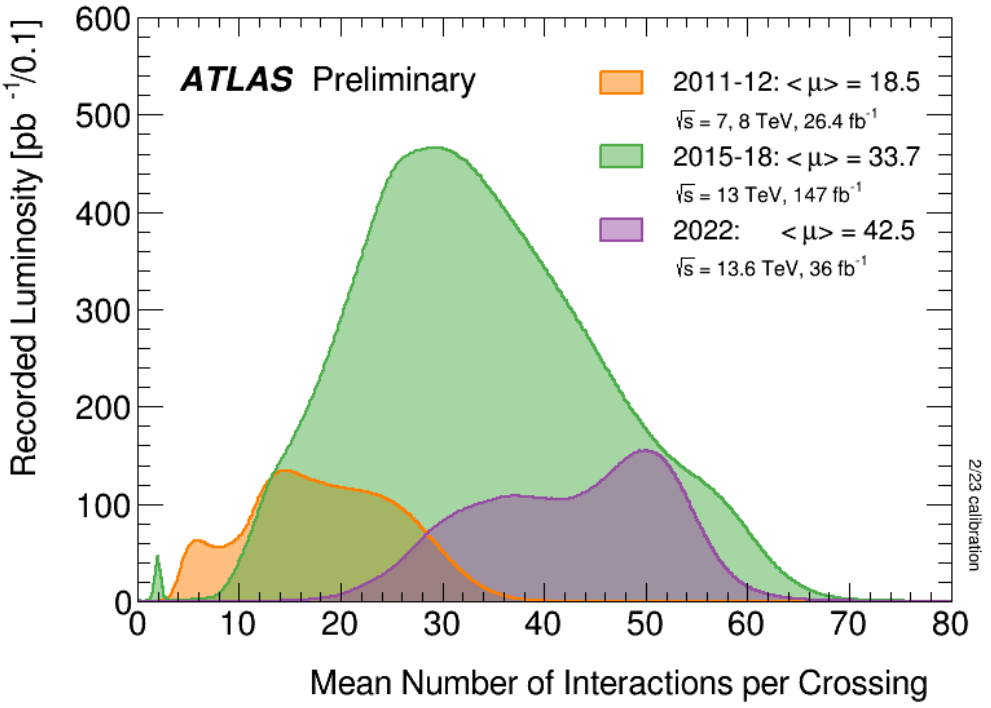


FIGURE 2.5. Pileup and luminosity for runs 1, 2, and (partial) 3 [6].

2.2. The ATLAS Experiment

ATLAS [8] is a cylindrical general purpose detector with a forward-backward symmetry that is centered at one of the LHC interaction points. It is made up of several subsystems, each designed to measure certain properties or types of particles. This sensitivity to various signatures was partly motivated by the

expectation that were the Higgs boson to be discovered, it would have decay modes to many different final states. The main detector components of ATLAS are the Inner Detector (ID), calorimeter, and Muon Spectrometer (MS). Each of these is divided into central cylindrical barrel components which face perpendicular to the beam line and disks in outer endcap regions that face parallel to the beam line. Another major subsystem is the trigger, which takes input data from the calorimeter and MS to determine which events will be written to disk and which will be discarded.

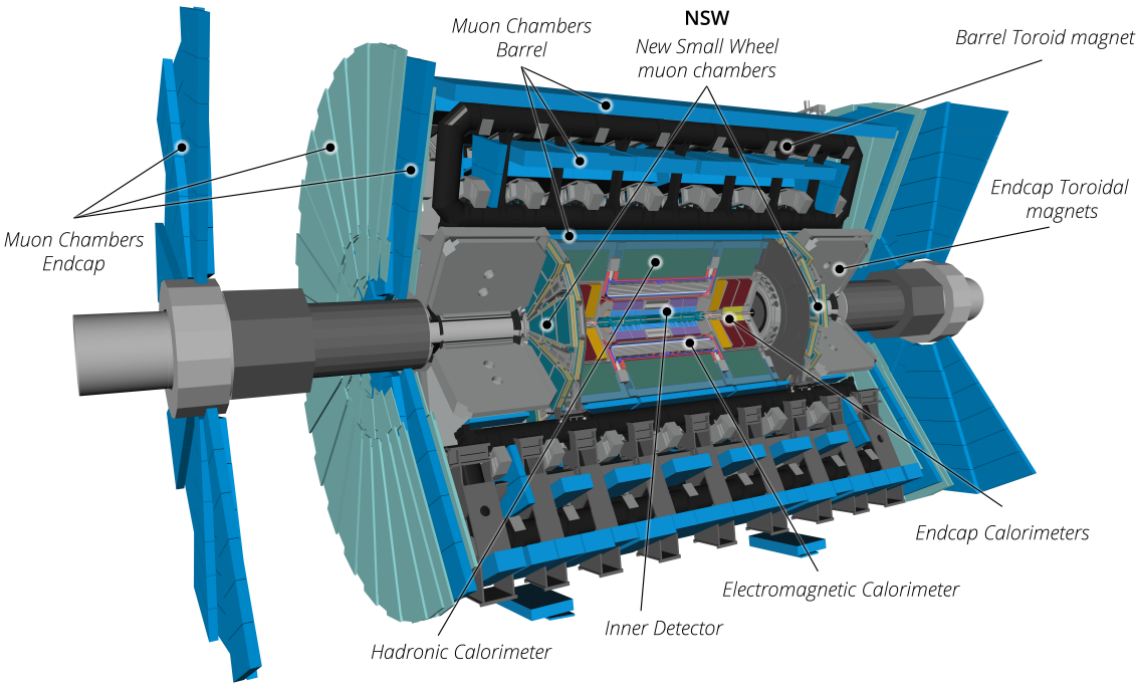


FIGURE 2.6. ATLAS cross-section diagram with subsystems labeled [7]

2.2.1. Coordinate System

Consider an (x, y, z) coordinate system whose origin is at the ATLAS interaction point. The z coordinate is parallel to the beam line and has a positive direction pointing counterclockwise when looking down at the LHC from above. The x coordinate points toward the center of the ring and y points up toward the surface of the Earth.

While this is a valid coordinate system, it does not take advantage of the cylindrical symmetry of our detector. Instead consider a system in which ϕ measures the azimuthal coordinate around the beam and a new variable y called the *rapidity* is defined as:

$$y \equiv \frac{1}{2} \ln \left(\frac{E + p_z c}{E - p_z c} \right) \tag{2.5}$$

This variable is useful because a Lorentz boost along the beam axis only adds a constant to the rapidity, so the difference in rapidity of any two objects remains invariant. A downside is that the calculation relies on the total energy and longitudinal momentum of the object, which may not always be measured with sufficient accuracy.

In place of rapidity, another variable called the *pseudo-rapidity*, denoted η, is used which closely approximates the rapidity and relies only on θ, the angle relative to the beam axis. To derive η, start by rewriting $E = \sqrt{p^2c^2 + m^2c^4}$ and $p_z = p \cos\theta$ [51] :

$$y = \frac{1}{2} \ln \left(\frac{pc \left(1 + \frac{m^2c^4}{p^2c^2}\right)^{\frac{1}{2}} + pc\cos(\theta)}{pc \left(1 + \frac{m^2c^4}{p^2c^2}\right)^{\frac{1}{2}} - pc\cos(\theta)} \right) \tag{2.6}$$

Considering the highly relativistic limit in which $pc \gg mc^2$ allows for us to apply the binomial theorem. After simplification and the use of a half-angle identity we arrive at:

$$y \approx \eta \equiv -\ln \tan \frac{\theta}{2} \tag{2.7}$$

A very common measure of the distance between two particles in the ATLAS detector is ΔR, which is defined:

$$\Delta R = \sqrt{\Delta\phi^2 + \Delta\eta^2} \tag{2.8}$$

This is invariant under longitudinal boosts because both $\Delta\eta$ and ϕ are themselves invariant.

Other variables of interest for particles measured by ATLAS are those defined purely in the transverse direction such as transverse momentum p_{T} and missing transverse energy $E_{\mathrm{T}}^{\mathrm{miss}}$. In addition to also being invariant under longitudinal boost, a high p_{T} is the sign of a collision with high momentum transfer as the initial p_{T} of the protons is zero.

2.2.2. Inner Detector

The portion of ATLAS that is closest to the beam pipe and therefore the first portion that particles pass through is referred to as the Inner Detector (ID)[52]. The subsystems of the ID are the Pixel Detector [53], Semiconductor Tracker (SCT)[54], and Transition Radiation Tracker (TRT)[55]. The entire ID is immersed inside of a 2T magnetic field that is provided by a solenoid magnet.

The purpose of the ID is to measure the charge, position, and momentum of electrically charged particles. When a charged particle passes through the ID it imparts ionization energy to create free charges. These charges are then collected and provide a *hit* - a single point that a charged particle is known to have passed through. By combining hits together, the trajectory of the charged particle can be reconstructed. The path of a single particle is called a *track*. Due to the magnetic field, these tracks will have a curvature. The direction of curvature reveals the charge of the particle and the degree of curvature as measured by the *sagitta* - or deviation from a straight line - is a function of its momentum. Tracks can be extrapolated back to the beam line to determine the point of origin or *vertex* associated with the particle. If the track does not intersect with the beam line, then it may have been produced at a secondary vertex and be the product of the decay of a parent particle.

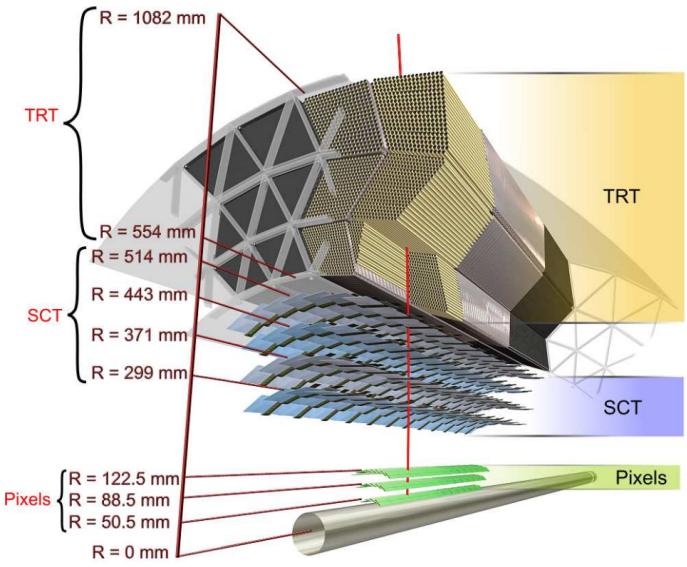


FIGURE 2.7. Inner detector with a track shown traversing the pixel, SCT, and TRT [8]

The Pixel Detector is the innermost subsystem with a distance of 33.25-122.5 mm from the beam line. It takes its name from the 92 million silicon pixels which are arranged in four layers. Hits in the Pixel Detector are measured with a precision of almost 10μm. The closest layer to the beam line is the Insertable B-Layer [56] which was installed in 2014 around a new smaller beam pipe. Its close proximity to the collision point allows it to better reconstruct secondary vertices. This is most useful in identifying τ-leptons and jets that are initiated by a b-quark, called b-jets. Outside of the IBL are three more barrel layers and on each side are four endcap disks. An individual pixel in the non-IBL (IBL) layer is approximately 50 μm wide and 300 μm (250 μm) long.

The SCT lies outside of the Pixel Detector in the radial distance of 30cm-52cm from the beam line. It detects charged particles using silicon microstrips as opposed to the smaller pixels. There are four layers in the barrel region that cover $|\eta| < 1$ and nine endcap modules on each side out to $\eta = 2.5$, which are arranged with the goal of producing four hits for each charged particle passing through them. There are 61 m^2 of silicon in the SCT which is connected to 6.3 million total readout channels [54]. This allows for the position in the bending plane of a single hit to be measured with a precision of 16 μm.

The TRT is the outermost subsystem of the ID and the largest in radial distance, with $56 < r < 107$cm [8]. It also has a barrel region ($|\eta| < 1$) with endcaps ($1 < |\eta| < 2$) on either side. It is not made up of silicon, but 300,000 drift tubes, also called straws. The cylindrical straw walls are made of a polyimide material and kept at a negative voltage. In the barrel (endcap) each straw is 144cm (37cm) long. They are filled with a gas mixture of 70% xenon, 27% CO_2,

and 3% O_2. In the middle is a grounded gold and tungsten wire. Surrounding the straws is the radiator medium made of polypropylene.

A charged particle traversing the TRT will emit transition radiation when encountering a boundary between media with different dielectric constants. These photons will then ionize the Xenon in the drift tubes, producing free electrons which will be drawn to the inner wire where they will be collected and the resulting signal measured. Within $|\eta| < 2$, a track can be expected to cross 36 straws. The larger size of the TRT and greater number of hits allow for it to better measure the bend of a particle's trajectory in the magnetic field and therefore determine its momentum. The energy of the transition radiation photons produced by an electron will be greater (8-10 keV) than those produced by a pion (2 keV) [55]. By considering the number of hits above a threshold energy along a track, the TRT can discriminate between the two species.

A 2T magnetic field that bends the charged particles as they pass through the ID is provided by the central solenoid [57] aligned along the beam axis. Over 9km of superconducting niobium-titanium wires are embedded in pure aluminum and carry a current of 7.73 kA. The wires must be cooled to 4.5 K, which is achieved through the use of liquid helium. When fully operational, the solenoid has a stored energy of 40 MJ. The solenoid is 2.3 m in diameter and its total length is 5.3 m. A major consideration affecting the design of the solenoid system was to minimize the interaction of particles with the solenoid before entering the calorimeter. To eliminate unnecessary walls between systems, the solenoid sits in the same vacuum vessel as the EM calorimeter.

FIGURE 2.8. Central solenoid before installation [8]

2.2.3. Calorimeter

Beyond the ID and solenoid magnet is the ATLAS calorimeter system, which is tasked with measuring the energy and position of both charged and neutral particles passing through it, absorbing them in the process. It can also determine an event's missing transverse energy ($E_{\mathrm{T}}^{\mathrm{miss}}$), the imbalance of measured energy in the transverse plane. The only particles it does not capture are neutrinos, which pass directly through ATLAS, and muons, which do not deposit an appreciable amount of energy in the calorimeters. It is broken up into two distinct calorimeters: the electromagnetic (EM) calorimeter immediately outside of the solenoid and the hadronic calorimeter outside of that. Each of these has barrel, endcap, and forward portions which allow detection all the way out to $|\eta| = 4.9$. Both calorimeters are made up of cells which each provide an energy measurement, allowing for the resolution of showers created by traversing particles.

The EM calorimeter [9] is primarily intended to measure electrons and photons. It is a sampling calorimeter made of alternating layers of absorbing lead-stainless-steel and active liquid argon (LAr). The lead-stainless-steel is chosen for its high Z and therefore its low radiation length - the average distance a particle will travel before undergoing a radiative process. This causes electrons and photons to produce EM showers of more and progressively lower energy electrons and photons. When these products pass through the LAr layers they free electrons from the argon which are collected and measured until the shower is fully absorbed. To increase the number of times that particles cross the boundaries between materials, they are arranged in an accordion structure as shown in Figure 2.10. In total, the EM calorimeter has about 70,000 readout channels for passing information to computing systems.

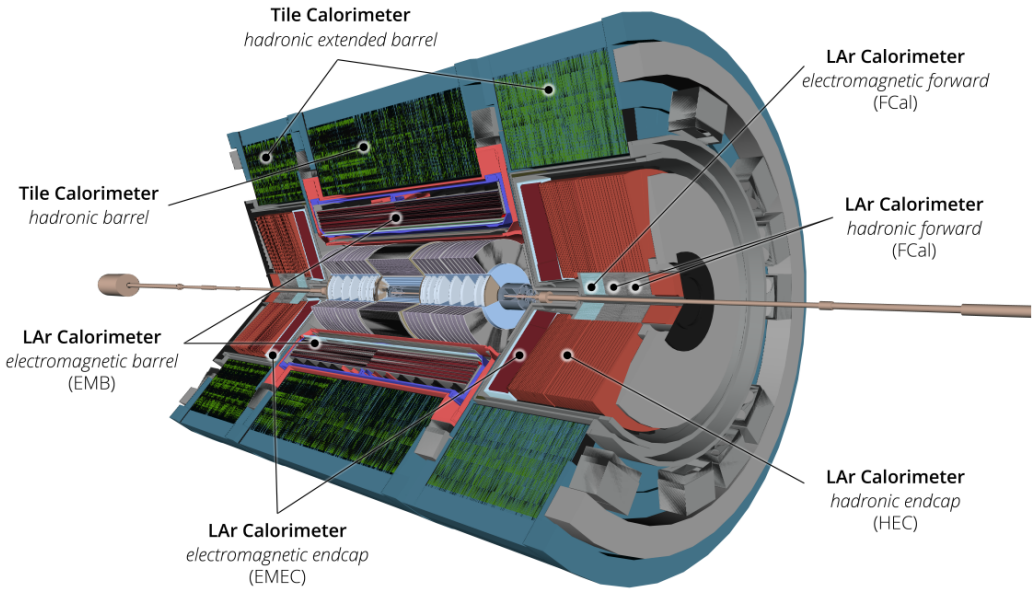


FIGURE 2.9. Cross-section of the ATLAS calorimeter system showing the EM and hadronic calorimeter with barrel, endcap, and forward regions labeled. [7]

There are four layers in the barrel ($|\eta| < 1.475$) of the EM calorimeter. The closest to the beam line is called the presampler and sits in front of the

solenoid magnet, allowing for measurement of the energy lost to the material that sits in front of the rest of the calorimeter. Three more layers make up the remainder of the barrel, each with varying depths and cell granularity in η and ϕ shown in Figure 2.10. The very fine granularity in η of EM1 is intended to allow for distinguishing between π^0 and photons. EM2 also has a relatively fine granularity due to being the thickest layer and collecting the greatest portion of an EM shower. EM3 is more coarse and thinner as it only collects the remnants of a shower that continues past EM2. The total depth is such that the EM calorimeter can fully absorb EM showers from electrons and photons and limit their punch-through to the hadronic calorimeter and muon system.

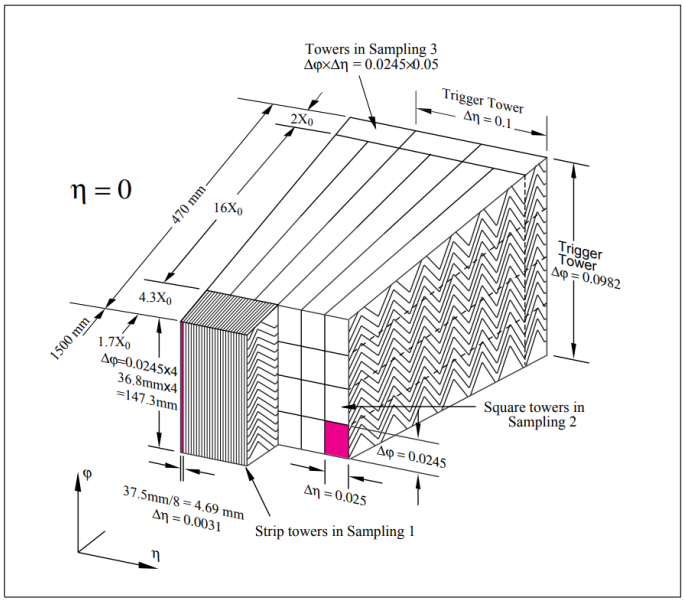


FIGURE 2.10. Dimensions and cell granularity of the EM1, EM2, and EM3 calorimeter layers in the barrel region. [9]

The EM end-caps (EMEC) are cylindrical *wheels* that lie at higher eta on either side of the barrel. They also contain the same four layers as the barrel, but the distribution and granularity change as a function of η. In total, the EMEC exists in the range $1.375 < |\eta| < 3.2$. The depth varies, but is $> 24X_0$ for $|\eta| >$

1.475. Inside the same cryostat as the EMEC (and other hadronic calorimeters to be described shortly) is the forward calorimeter (FCal). This covers the range $3.2 < |\eta| < 4.9$ and is roughly 10 radiations lengths in depth. It must be radiation hard due to the intense flux of incident particles at this high η. The first layer of the FCal is designed for EM signatures and uses copper as the absorber material in place of lead-stainless-steel.

The hadronic calorimeter sits behind the EM calorimeter and is intended to capture energy from particles undergoing hadronic interactions with the nuclei of detector materials. Pions, kaons, protons, and neutrons will generally punch through the EM calorimeter to be fully absorbed by the hadronic calorimeter. An important measure of depth for hadron showers is the interaction length λ. A hadron traversing a material has a $1 - \frac{1}{e}$ chance of undergoing an interaction with a nucleus through the strong nuclear force.

The barrel portion of the hadronic calorimeter is the only part of the ATLAS calorimeter system that is not made up of the previously described LAr materials. It is called the Tile calorimeter (TileCal) [58], taking its name from the plastic scintillating tiles that serve as the active material. The absorber material is steel, chosen for its low interaction length. The central barrel region extends out to $\eta = 1.0$ and two extended barrel components reach out to $\eta = 1.7$ and sit outside of the end-cap regions. The barrel has three layers which have a combined depth of 7.4λ at $\eta = 0$. Cells in the TileCal have a granularity of 0.1×0.1 in $\Delta\eta \times \Delta\phi$ except for the outer regions of the third layer where it is 0.2×0.1.

The hadronic end-cap (HEC) and hadronic portions of the FCal also use liquid-argon as the active material but instead of lead-stainless-steel for the absorber use copper and tungsten, respectively. Both are arranged with two

separate layers. The HEC cells have a granularity of 0.1×0.1 for $|\eta| < 2.5$ and 0.2×0.2 for $|\eta| < 3.1$. In the FCal, the geometry changes to a metal matrix in which tubes of the same metal are regularly spaced and the LAr resides in small gaps between the tubes and their surroundings.

2.2.4. Muon Spectrometer

The outermost subsystem of ATLAS, the Muon Spectrometer [59] (MS), is designed to observe the muons which have traveled through the rest of ATLAS largely unimpeded to this point. Being charged particles, muons leave tracks in the ID much as electrons do. However, they do not produce EM showers in the calorimeters in contrast to electrons. Energy loss from ionization is a function of the $\beta\gamma$ of the incident particle as shown in Figure 2.11. Muon $\beta\gamma$ values cause energy loss from ionization to fall near the minimal possible value, prompting the label of *minimum ionizing particle* (MIP).

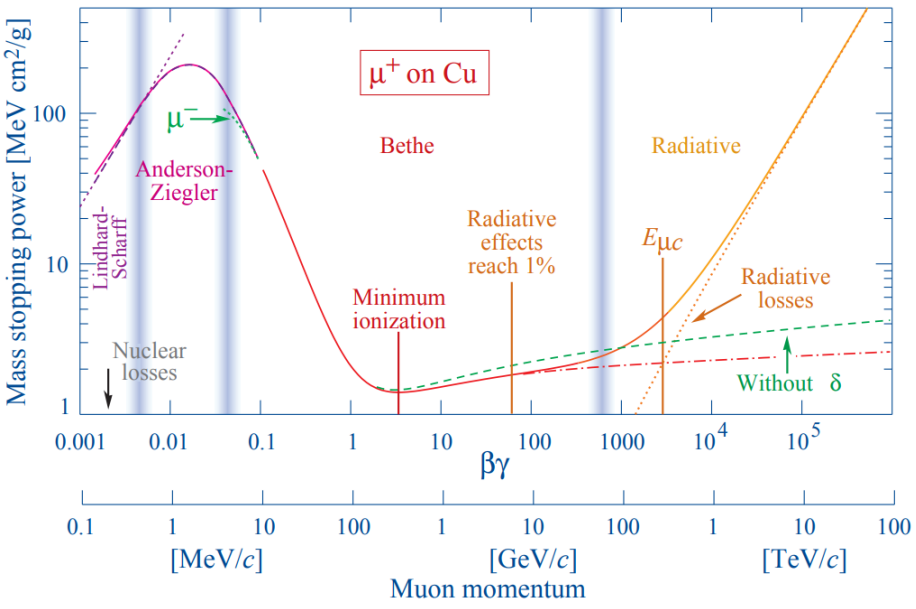


FIGURE 2.11. Stopping power of a μ^+ in Cu. Radiative energy loss becomes significant above 100 GeV. [2]

The MS measures the trajectory of muons passing through it using a number of different systems. It sits inside a magnetic field allowing for momentum determination similarly to the ID. The barrel toroid sits outside of the calorimeters and end-cap toroids. It is made up of eight coils evenly spaced radially around the beam line. The coils have inner and outer diameters of 9.4 m and 20.1 m, respectively [8]. It is 25.3 m in length and weighs 830 tons in total. The end-cap toroids sit inside the ends of the barrel toroid and provide the magnetic field for higher η particles. They are smaller than the barrel but have a very similar structure with eight coils each as well. Both barrel and end-caps use a wire of Al/Nb/Ti/Cu that must be cooled to 4.6K and can achieve a magnetic field of up to 3.5T.

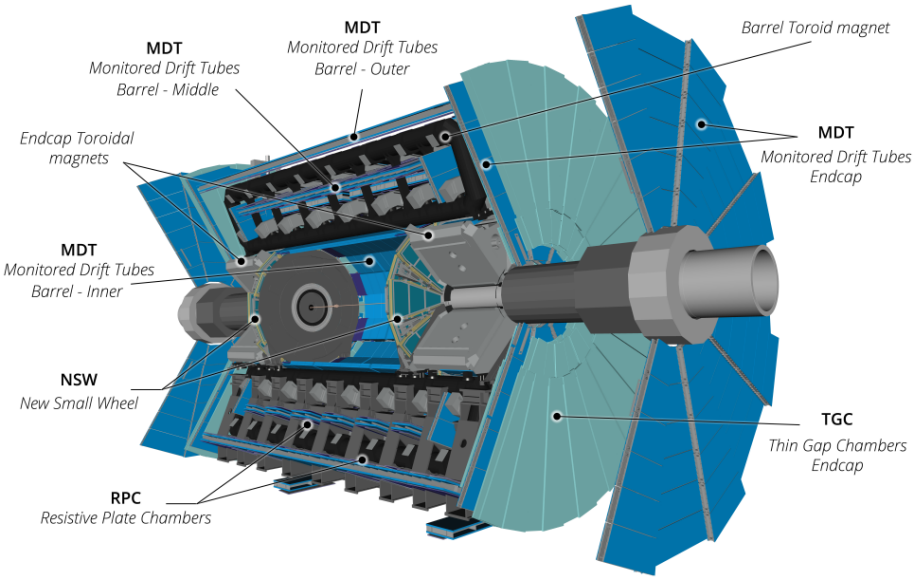


FIGURE 2.12. Cross-sectional view of the Muon Spectrometer showing the tracking, trigger, and magnet systems. [7]

In the barrel, the path of the muon in η is reconstructed mainly using three layers of Monitored Drift Tubes (MDT). These are tubes with a diameter of 30 mm filled with Ar/CO_2 with a tungsten-rhenium wire in the middle. Muons

traveling through the tube will release electrons from the gas which will drift towards the wire due to an applied electric field. By grouping tubes that collect a signal together, the path of the muon can be determined. These are grouped into three stations, the innermost of which contains four MDTs and the outer two of which have three each.

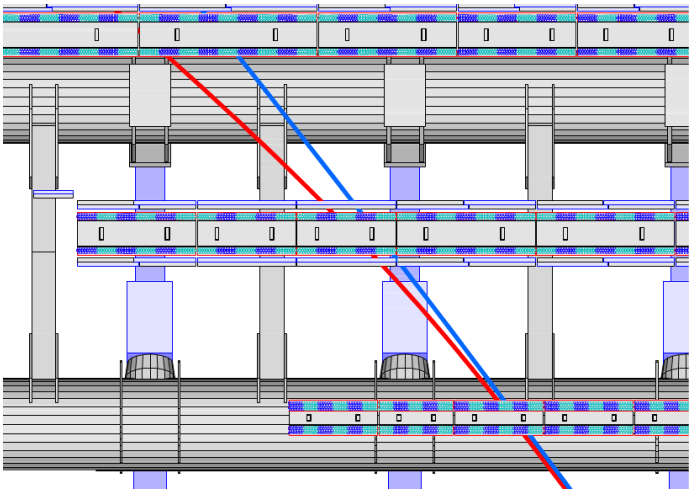


FIGURE 2.13. The trajectories of muons traversing the three stations of MDTs in the barrel region of the MS. The red line indicates a 4 GeV muon and blue indicates a 20 GeV muon [8]

.

The barrel region also contain gaseous detectors called Resistive Place Chambers (RPCs) whose purpose is to measure muon energy and relay information to the ATLAS trigger system (described in 2.2.5. These are only present in the middle and outermost stations. The RPCs in the middle station are used for a $p_{\mathrm{T}} > 6$ GeV trigger and RPCs in both stations together form a high-p_{T} > 20 GeV trigger. The multiple RPC hits allow for quickly determining if a muon originated from the interaction point, though with limited resolution compared to the MDTs. All triggering chambers of the MS measure in both the η and ϕ coordinates.

The end-cap regions ($1 < |\eta| < 2.7$) also use MDTs for their momentum measurement for $|\eta| < 2$. The end-caps are also made up of three stations in the form of cylindrical wheels whose faces are perpendicular to the beam direction. The end-cap toroids are positioned only in between the first and second wheels, so the muon travels in a straight line through each station as opposed to following a curved trajectory throughout all three as in the barrel. The momentum determination therefore requires a point-angle instead of a sagitta calculation.

For $|\eta| > 2$, Cathode Strip Chambers (CSCs) are used instead of MDTs. CSCs are made up of smaller cells, which is advantageous for keeping a low average occupancy with the higher rate of particles at such high η. The CSCs are placed on the inner tracking wheel such that eight hits are expected for each traversing particle. Triggering is also provided in the end-caps for $|\eta| < 2.4$ by Thin Gap Chambers (TGCs), also chosen for their ability to handle rate and their radiation hardness [59]. These are placed along the inner and middle wheels. The inner wheel TGCs provide a measurement of the ϕ coordinate for tracking while the middle wheel TGCs provide triggering for high and low p_{T} muons as in the barrel.

During the Phase-I upgrade, the inner stations of the end-caps were replaced with the New Small Wheels (NSWs) [60]. This was motivated by the decreased tracking and triggering performance of the existing small wheel under the increased luminosity conditions of Run-3 and beyond. Two new chamber technologies are used in the NSWs: small-strip TGC (sTGC) detectors for triggering and micro-mesh gaseous structure (Micromegas) detectors for tracking. Without the increased performance of the NSWs over the original small wheels, the trigger rates for muon would have become unacceptably high and either

required a raising of the p_{T} threshold or the introduction of a prescale. Either of these options would have reduced drastically the sensitivity of ATLAS to important processes of interest such as $H \to \mu\mu$.

2.2.5. Trigger System

As mentioned before, the rate of bunch crossings by the LHC is 40 MHz. Ideally, ATLAS would be able to collect the output data of every single collision for later analysis. However, this would require the ability to write 120 TB of data to disk every second [7]. This being unrealistic, a decision must be made whether or not to write any given event output to disk or not. This process is handled by the *trigger* system. It does so by searching first for distinctive signatures that correspond to certain particles or event characteristics and later for events that contain combinations of these. The trigger lies inside of the broader Trigger and Data Acquisition System (TDAQ) that is also responsible for the ferrying of data to different systems throughout the process and eventually writing it to disk.

Significant effort went into upgrading the ATLAS trigger for Run-3 [10], and its status as of the beginning of Run-3 will be described here. The upgrades involved the creation of a number of new electronics boards specializing in certain signatures as well as incorporating data from other components newly installed for Run-3 such as the NSW. There are two steps to the trigger chain, the first being the Level-1 Trigger and the second being the High Level Trigger (HLT).

The Level-1 Trigger [61] is a hardware trigger that uses inputs from the calorimeter system and the muon system to make an initial determination of whether the output event should be processed further. The target latency within which a decision must be made is 2 μs with 500 ns contingency. A large portion

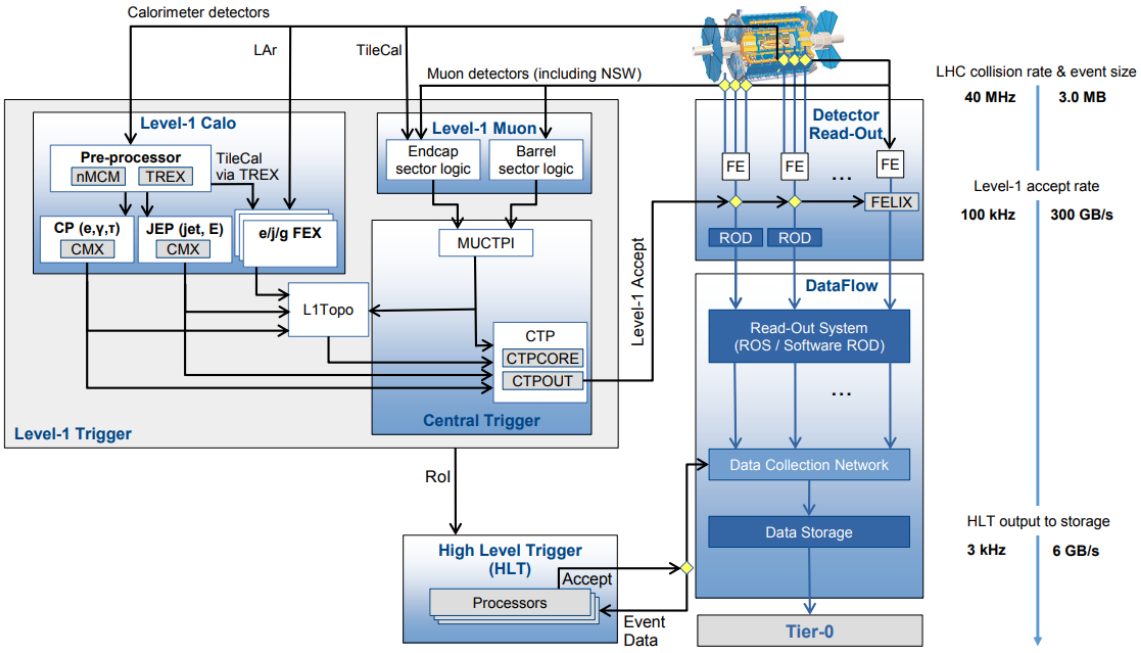


FIGURE 2.14. The subsystems and data flow of the ATLAS trigger system in Run 3. Also shown are the collision and data rates at each step in the process. [7]

of this time is spent sending signals to the trigger hardware and back, leaving only a fraction for analysis and decision making [62]. This short amount of time is the reason for the implementation of the Level-1 Trigger in specialized hardware. The two systems are at first considered in separate Level-1 Calorimeter (L1Calo) and Level-1 Muon (L1Muon) systems whose outputs are eventually combined and considered together.

The L1Calo system is mainly designed to identify events with high-E_{T} electrons, photons, jets, and hadronically decaying τ-leptons as well as those with large $E_{\mathrm{T}}^{\mathrm{miss}}$ or $\mathrm{H_T}$, which is the scalar sum of the transverse energy of all objects in the event. Newly added for Run-3 are three feature extractor boards: the eFEX, gFEX, and jFEX. The eFEX identifies e, γ, and τ; the jFEX identifies jets, $E_{\mathrm{T}}^{\mathrm{miss}}$, and τ; the gFEX identifies large-R jets, $E_{\mathrm{T}}^{\mathrm{miss}}$, and $\mathrm{H_T}$.

Energies are provided from the LAr calorimeter portions already digitized, whereas the Tile calorimeter energies are digitized by the Tile REceiver eXchange (TREX) board before reaching the FEXes. The eFEX has access to cells at the SuperCell level, whose granularity of $\Delta\eta \times \Delta\phi = 0.025 \times 0.1$ in the first and second barrel layers and 0.025×0.1 in other layers. The jFEX granularity is 0.1×0.1 and the gFEX is 0.2×0.2. In most instances, the granularity has been improved from previous runs. All boards use Field Programmable Gate Arrays (FPGA) to sum up cell energies over a defined area for the energy of the object of interest. A secondary *isolation* value, which measures how distinct this energy deposition is from its surroundings, is also calculated in some cases. For a detailed example of this, see Section 2.3.

The L1Muon system relies on the RPCs in the barrel, TGCs in legacy wheels, and sTGC and Micromegas in the NSW. The trigger begins with hits in the RPC2 layer in the barrel and the M3 TGC layer in the endcaps. From there it compares the path of a muon candidate to that of a muon originating from the IP and having infinite momentum i.e. a straight path. It then searches for hits in other layers in a cone of allowance around the theoretical track hits. The greater the deviation from this theoretical track, the lower the momentum of the actual muon. Adjusting the allowed deviation then also adjusts the p_{T} threshold of the trigger.

For both L1Calo and L1Muon, when a trigger is satisfied a Trigger Object (TOB) is created with information on the E_{T} , η, and ϕ of the object. These are then sent to the L1 Topological Trigger (L1Topo), which considers how the different objects relate to each other and may issue its own accept if certain conditions are met e.g. an event has an electron and a muon close together.

Outputs of L1Calo, L1Muon, and L1Topo are finally sent to the Central Trigger Processor (CTP). The CTP contains the L1 trigger items - a single specification that, if satisfied, allows for the event to be passed to later trigger stages. Each trigger item has an associated rate at which it accepts events. Together, the rate of all L1 trigger items cannot exceed the rate at which the L1 system is capable of passing data to the HLT. For Run-3 the maximum allowed Level-1 accept rate is 100 kHz. For those triggers that would have an unacceptably high rate, a *prescale* factor may be applied in which only a fraction of the events that meet its requirements actually cause a firing of the trigger. Another tool for controlling rate is the addition of p_{T} thresholds to individual objects being triggered on.

The HLT is a software trigger that has access to additional information from the events passed by the Level-1 trigger. The η and ϕ coordinates as well as type of all TOBs are delivered, and the HLT is then provided full granularity tracking, calorimeter, and muon data from a Region-of-Interest (RoI) around each TOB. In this way the HLT can approximate the selections that may be done offline after the data has been stored to disk. The HLT must reduce the incoming rate of 100 kHz to an average output rate of 3 kHz writing to disk. The HLT also contains a menu of triggers items, each of which is seeded by one or more Level-1 triggers along with additional requirements on quantities reconstructed in the HLT.

The data acquisition portion of the TDAQ system is responsible for delivering relevant event data to the trigger and eventually to disk storage. Two systems functions in parallel for Run-3: a new software-based system for new detector components in Run-3 and a legacy system for everything else.

The legacy system begins by transferring event data from front-end electronics to specialty hardware boards called Readout Drivers (RODs) each tailored for the detector system feeding it data. After preliminary processing and formatting, a Level-1 accept will cause the RODs to pass along their data to a single Read-Out System (ROS). The ROS buffers data as needed while the Data Collection Network passes RoI information to the HLT as requested. Once an HLT accept is received, the buffered data flows to Data Storage where it is prepared for writing to disk.

The NSW, LAr, and L1Calo systems will use a more unified framework that is less dependent on custom hardware for the first step of the data acquisition chain. The Front-End LInk eXchange (FELIX) boards are capable of accepting front-end read-out from all detector systems and routing it to an analogue of the RODs implemented in software (SW RODs). The SW RODs will still perform detector-specific data processing and buffering. It will also send output to the ROS where it will be merged with output from legacy systems RODs.

2.3. eFEX Tau Trigger

As mentioned previously, the eFEX board of the Level-1 trigger has a dedicated algorithm for reconstructing tau leptons. Presented below is the work of developing and testing the performance of said algorithm for use during Run-3. A bitwise simulation was also coded in the ATLAS software framework Athena to mimic the flow of data from one module to another during experimental conditions.

Only hadronically-decaying tau leptons (τ_{had}) are of interest here. A tau may also decay leptonically to either an electron or muon as well as a tau neutrino.

In these instances the leptons are indistinguishable from those produced directly from the interaction point and are therefore handled by electron and muon triggers. In total, a tau will decay hadronically approximately 65% of the time. The final state in this case is a tau neutrino, one or three charged pions, and ≥ 0 neutral pions. Of hadronic decays, the most common mode is to a tau neutrino, neutral pion, and charged pion 38% of the time [2].

(a) (b)

FIGURE 2.15. Diagrams showing the decay products of a jet (a) and a τ_{had} (b). Compared to a jet, decay products of a τ_{had} are not common in the outer green region.

A τ_{had} showering in the ATLAS calorimeter is expected to deposit energy in both the EM and hadronic layers. The neutral pions will decay to $\gamma\gamma$ and be absorbed in the EM layers. Charged pions will deposit energy in the EM as well as the hadronic layers. The shape of the shower is variable due to the many available decay modes and therefore number of particles depositing energy. The shower profile may therefore look similar to a jet, which is also made up of a variable number of charged and neutral hadrons. One difference is that generally the hadron decay products of a τ_{had} will be clustered together more closely when they hit the detector, creating a smaller and more dense area of energy deposition with a lack of energy in a surrounding ring. One extra consideration is that the neutrino produced in the tau decay will pass through the detector unnoticed, carrying away some portion of the energy of the original tau. This means that the

measured E_{T} of the tau will be lower than its truth energy. For this reason, the *visible* distinction is included ($\tau_{had-vis}$) when one is referring to the non-neutrino decay products.

2.3.1. eFEX

The eFEX is chosen to reconstruct τ_{had} because it receives the full SuperCell granularity from the calorimeters. Because of this, the regions it considers can be fine-tuned to fully capture the signature of the tau while avoiding that of nearby pileup. This will be important for the calculation of the E_{T} of the τ_{had} as well as an isolation variable which helps distinguish τ_{had} from background processes.

An eFEX board is shown in Figure 2.17. A single board contains four Xilinx Processor FPGAs that have the τ_{had} as well as e/γ algorithms implemented in their firmware. Each FPGA is responsible for a 0.6×1.0 region of the detector with extra overlap between adjacent regions to ensure full coverage. In total there are 24 eFEX boards which cover the full ϕ range and $|\eta| < 2.5$.

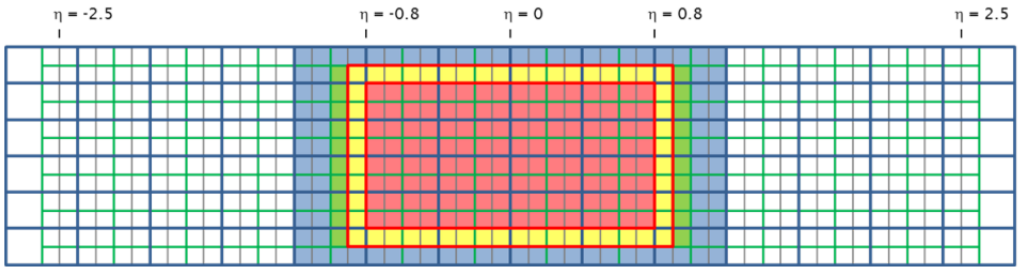


FIGURE 2.16. The coverage area of a single eFEX board. Algorithms are executed in the red and yellow regions. [10]

2.3.2. τ_{had} Algorithm

The τ_{had} algorithm works by using a sliding 0.3×0.3 window which corresponds to 3×3 towers. A *tower* is a collection of all SuperCells in each layer

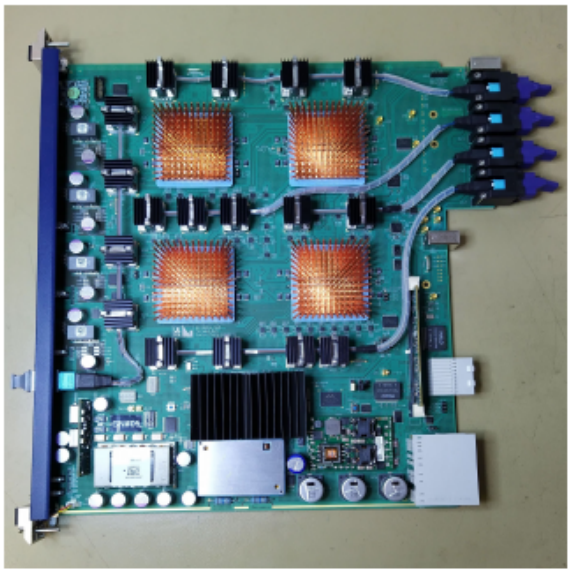


FIGURE 2.17. A single eFEX board [7]

that lie within the same 0.1×0.1 region. There is a single SuperCell per tower for each coarse-grained layer (EM0, EM3, and HAD) and four SuperCells per layer for each fine-grained layer (EM1 and EM2). In the fine-grained layers, each SuperCell is 0.025×0.1 as opposed to 0.1×0.1 in $\Delta\eta \times \Delta\phi$.

For each 3×3 region, the central tower is considered a seed if it has E_{T} > 1 GeV and it has a greater or equal E_{T} than all eight adjacent towers. The inequality checked is strict or inclusive such that if multiple adjacent towers have equal energies, at most one seed will result. See Figure 2.19 for the comparison.

After a seed tower is identified, a seed SuperCell is chosen for each layer. For the coarse-grained layers, this is simply the only SuperCell in the tower. The seed SuperCell for fine-grained layers is chosen by looking to EM2 and finding the SuperCell with the greatest E_{T} , again accounting for adjacent SuperCells with equal E_{T} . The seed SuperCell for EM1 is that in the same position as the seed SuperCell for EM2.

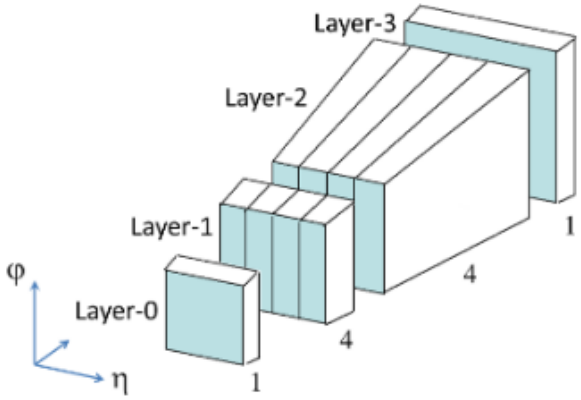


FIGURE 2.18. The granularity of SuperCells in a single trigger tower for the four EM layers.

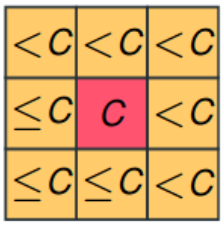
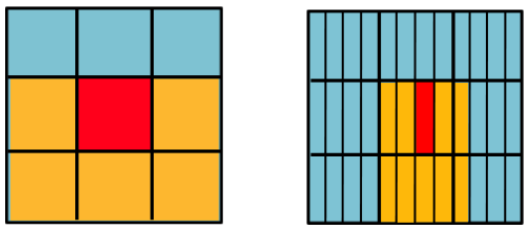


FIGURE 2.19. The inequalities checked in comparison to the central tower when determining if the central tower is a local maximum in E_{T} .

The E_{T} and isolation calculations depend on summing the energy of cells in defined regions with respect to the seed SuperCell in each layer. These regions are asymmetric in ϕ, so the first step is to determine the orientation in this direction. This is decided by considering the SuperCells directly adjacent in ϕ to the seed SuperCell in the EM2 layer. Of these two, it is in the direction of greater energy that the cells will be included in the energy sums. The reconstructed E_{T} is the sum of 3×2 regions in each of the coarse-grained layers and 5×2 regions in the fine-grained layers. The isolation variable called r_{core} is the ratio of two regions

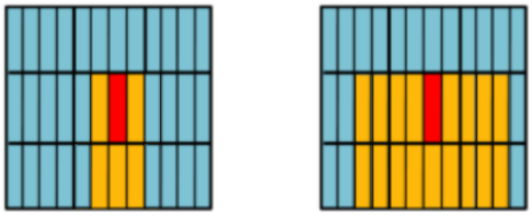
defined solely in the EM2 layer. The numerator of the variable is a sum of a *core* 3×2 region and the denominator is the sum of an *environment* 9×2 region.

$$r_{\text{core}} = \frac{E_{\text{T},3\times2}}{E_{\text{T},9\times2}} \tag{2.9}$$



(a) (b)

FIGURE 2.20. Diagrams showing the seed SuperCells (red) and those additional SuperCells included in the E_{T} calculation (orange). (a) shows the coarse-grained layers (EM0, EM3, and HAD) and (b) shows the fine-grained layers (EM1 and EM2)



(a) (b)

FIGURE 2.21. Diagrams showing the seed SuperCells (red) and those additional SuperCells included in the isolation calculation (orange) in the EM2 layer. (a) shows the core region and (b) shows the environment region.

2.3.3. Development

The size of regions to be summed over for both the reconstructed E_{T} and isolation variable were chosen for their ability to best separate true taus from QCD jets. In both cases, several different sizes of region were tested and the

performance of the relevant variable compared. Performance on signal τ_{had} is determined by using a Monte Carlo (MC) simulated $Z \to \tau\tau$ sample. Background is provided by a $JZ0W$ sample of low-p_{T} QCD processes.

The reconstructed E_{T} is intended to best capture as much of the energy deposited into the calorimeter by the τ_{had} decay products as possible. It must do this while also allowing for rejection of QCD jets by applying a cut on the reconstructed E_{T} . Therefore it is advantageous for the E_{T} algorithm to maximize the energy of true τ_{had} while minimizing that of jets. The broader energy deposit of jets means that this can be accomplished by choosing a region size that is large enough to capture the majority of a τ_{had} energy while also being small enough to exclude energy from jets.

The isolation variable also seeks to take advantage of this difference in shower size between τ_{had} and jets. A τ_{had} is expected to have a value of the variable closer to unity as most of its energy will fall within the core region. A jet will have only a portion of its energy in the core region and more captured in the larger environment region, leading to an r_{core} value less than unity.

For each sample, all τ_{had} candidates are reconstructed for each event and considered as a single group. The reconstructed E_{T} and isolation variable are calculated but no thresholds are applied for filtering yet. For the signal sample, τ_{had} candidates are considered only if they are within $\Delta\text{R} < 0.2$ of a truth τ_{had} and that truth τ_{had} must have $p_{\text{T}} > 20$ GeV.

Distributions are made of the reconstructed E_{T} values for each sample. Cuts of increasing value X are then applied to both distributions of the form E_{T} $>$ X GeV. For each cut an efficiency can be calculated by dividing the number of τ_{had} candidates that pass the cut by the total number of candidates for that

sample. This results in an efficiency per cut value for both the signal sample and background sample. By plotting these two sets of efficiencies against each other a receiver operating characteristic (ROC) curve is formed. The ROC curves of potential E_{T} region specifications are compared and the algorithm was chosen that best preserves signal efficiency while rejecting background. Figure 2.22 shows the performance of several E_{T} region definitions.

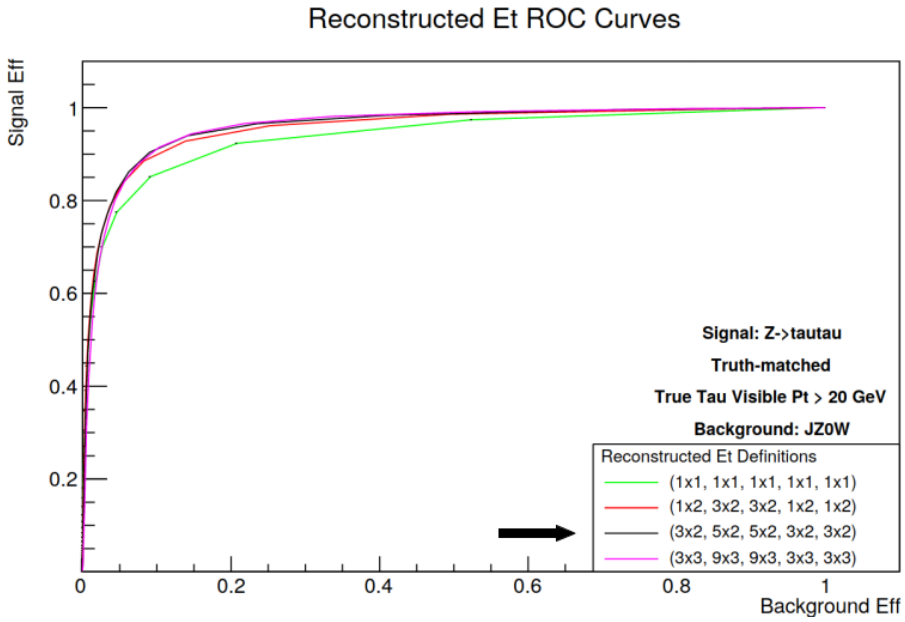


FIGURE 2.22. ROC curves showing the signal and background efficiencies for several different reconstructed E_{T} algorithms as cuts on the E_{T} are performed. The algorithm as has been described in this paper performed best and is noted with an arrow.

The exact specifications for the core and environment regions of the r_{core} isolation variable were tested in a similar manner. The variable was calculated for all τ_{had} candidates from the same signal and background samples. Cuts of the form $r_{\mathrm{core}} > \mathrm{X}$ are applied for increasing X and the efficiencies for signal and background are also plotted as ROC curves. It is also advantageous to consider curves in bins of reconstructed E_{T} , as it is most important for analyses that high-

p_{T} taus be kept. Figure 2.23 shows ROC curves of various core and environment definitions specifically in the bin of 30 GeV $< p_{\mathrm{T}} <$ 40 GeV.

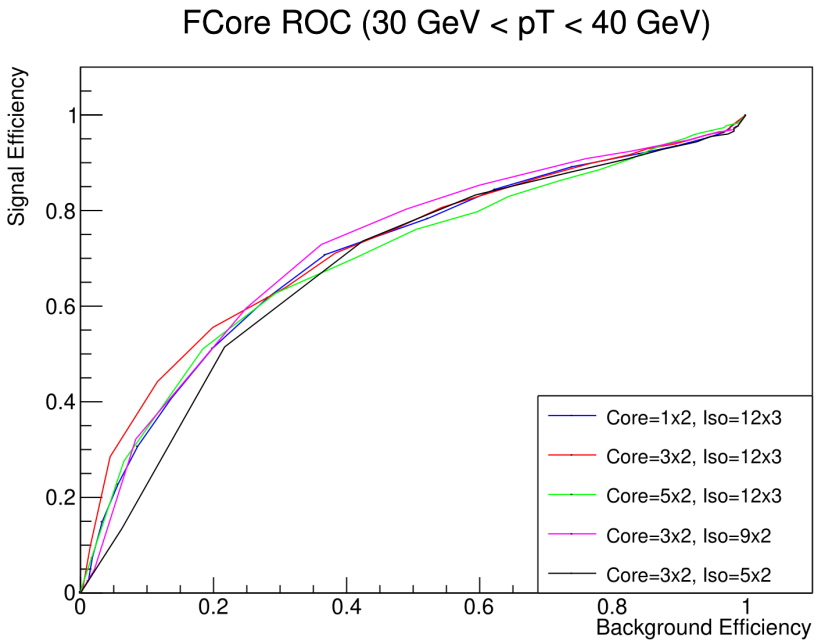


FIGURE 2.23. ROC curves showing the signal and background efficiencies for several different r_{core} algorithms as cuts on the variable are performed. The line in pink shows the final definition of the algorithm.

2.3.4. Performance

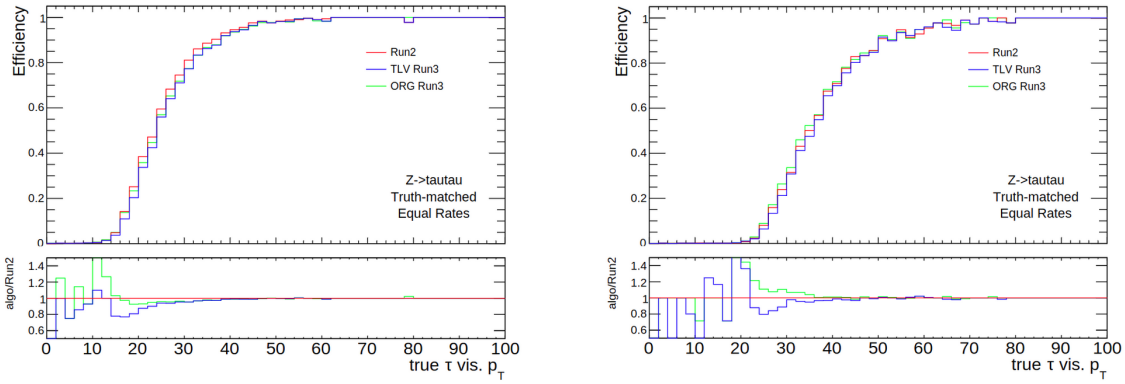
The performance of the τ_{had} algorithm can be determined by creating a *turn-on* curve, which shows the efficiency of true τ_{had} accepted by the trigger as a function of true visible p_{T} . As τ_{had} with a higher p_{T} are generally easier to reconstruct, the efficiency curve will increase as a function of true p_{T} and ideally achieve full efficiency at high p_{T} values.

Additionally, it is important that the rate of acceptance of the τ_{had} trigger is within the limits allowed by the overall Level-1 trigger menu. The overall rate is driven by the low-E_{T} QCD background, so a minimum E_{T} threshold on all

tau candidates is necessary to accomplish this. This of course also sets a lower bound on the sensitivity of any downstream analysis to τ_{had} as well, so great effort is made to push this threshold as low as possible.

The Run-2 tau algorithm and rate provide useful benchmarks against which the Run-3 algorithm can be tested. For the purposes of comparisons between the runs, an approximation of the Run-2 algorithm was implemented over Run-3 conditions. In Run-2, the E_{T} benchmarks of 12 and 20 GeV were used in the trigger menu. The E_{T} threshold of the Run-3 algorithm can then be set to achieve the same of efficiency on $JZ0W$ events as the 12/20 GeV thresholds with the Run-2 algorithm. In this way, a comparison of the algorithms' performance can be made at equal rates. The increase in beam energy from Run-2 to Run-3 results in more pileup in each collision, which can make it more difficult to distinguish τ_{had} in the calorimeter data. The move from tower granularity to SuperCell granularity means that the Run-3 algorithm can be more finely tuned to capture τ_{had} energy while discriminating against jets. Both of these effects together lead to the performance of the Run-3 algorithm being very similar to that of the Run-2 algorithm. See Figure 2.24 for turn-on curves comparing Run-2 and Run-3 algorithms at equal rates. A 12.5 GeV E_{T} threshold for the Run-3 algorithm achieves equal rate to the 12 GeV threshold for Run-2. A 20.5 GeV threshold for Run-3 achieves equal rate to the 20 GeV threshold for Run-2.

After the E_{T} threshold has been set, a cut can also be placed on the r_{core} isolation value to further reduce rate and improve the purity of the sample of τ_{had} candidates. This will necessarily also decrease the efficiency for any given true p_{T} point on a turn-on curve. Figure 2.25 shows turn-on curves with such an isolation



(a) (b)

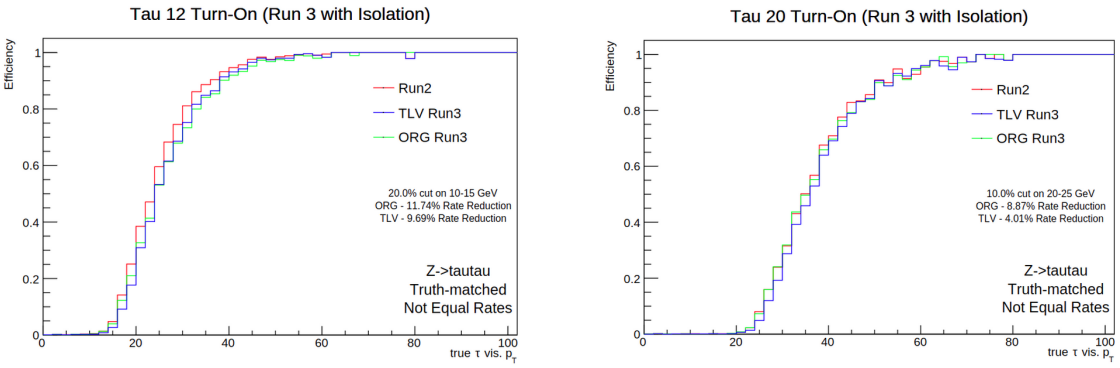
FIGURE 2.24. Turn-on curves showing the efficiency of a truth-matched tau candidate being created as a function of true visible p_{T} . 12 GeV (a) and 20 GeV (b) Run-2 threshold curves are shown alongside Run-3 algorithm curves with E_{T} thresholds to achieve equal rates. The final Run-3 algorithm is labeled "Oregon Run3". Also shown is another algorithm considered for Run-3 labeled "TLV Run3". Ratio plots of the Run-3 algorithms over Run-2 are also shown.

cut applied. As we are most interested in preserving high-p_{T} taus, the cut is only applied to τ_{had} in low reconstructed E_{T} bins.

2.3.5. Bitwise Simulation

After a τ_{had} is identified in eFEX, a TOB is created and loaded with information about the E_{T} , isolation, and position of the tau candidate. The TOB itself is made up of 32 bits and the bits are dedicated to different values as shown in Figure 2.26. A bitwise simulation of the data received from calorimeters, TOB creation, and output of TOBs to L1Topo was coded in Athena for testing purposes. The simulation is performed mainly at the FPGA level where the production algorithms are coded in firmware.

The 12-bit energy field requires that the values that the eFEX receives in steps of 25 MeV must be rescaled to the 100 MeV energy scale of the TOB.

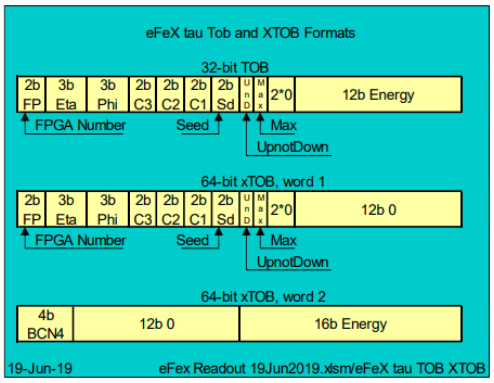


(a) (b)

FIGURE 2.25. Turn-on curves of Run-2 and Run-3 algorithms. The Run-3 algorithm is set to equal rate to Run-2 and then a cut on the isolation variable cut is applied. The percentage of candidates removed in the given E_{T} bin is shown. Also given are the rate reductions achieved by this additional isolation cut.

With 12 bits, the maximum energy value that may be sent is 409.5 GeV and any greater energy is simply set to that maximum value. The minimum E_{T} threshold necessary to create a TOB is a configurable value that can be read from the trigger menu. The UnD (UpnotDown) bit relays the ϕ orientation of the E_{T} and isolation clusters as discussed in Section 2.3.2. The Eta and Phi fields are indices for which tower in the FPGA region is the seed tower. That information along with the FPGA number and hardwired information of which board the candidate originates from allows the absolute η and ϕ coordinates to be calculated downstream. The C1 and C2 fields are set aside to carry the isolation information - C1 for the r_{core} variable and C2 for a second isolation variable, r_{had}, which reflects the fraction of candidate energy deposited in the hadronic calorimeter layer. Instead of sending literal values, the eFEX can receive three values from the trigger menu per isolation variable corresponding to Low, Medium, and High thresholds and fill 1, 2, or 3 into the Cx bits, respectively, if that threshold is met.

Primary tau triggers for Run-3 are shown in Figure 2.27.



(a)

TOB/xTOB Fields	Content. Refer to reference [4] for algorithm details
FP	FPGA Number 0-3 within the eFeX module. FPGAs are ordered in increasing eta. Each FPGA covers an eta region of 0.4
Eta	Eta position of the TOB trigger tower relative to the FPGA algorithm origin, in units of 0.1. Value 0 to 5.
Phi	Phi position of the TOB trigger tower relative to the algorithm origin, in units of $\pi/32$ (≈ 0.1)
Had (eγ only)	Count of number of Hadronic thresholds satisfied. Thresholds are assumed to be in increasing order
f3 (eγ only)	Count of number of f3 algorithm thresholds satisfied.
Re (eγ only)	Count of number of Reta thresholds satisfied.
Sd (eγ only)	Seed eta position within TOB trigger tower, in units of 0.025
C1 (τ only)	Count of number of τ condition 1 thresholds satisfied.
C2 (τ only)	Count of number of τ condition 2 thresholds satisfied.
C3 (τ only)	Count of number of τ condition 3 thresholds satisfied.
UnD	UpnotDown - single bit set if seed includes the supercell above (higher phi) rather than below the central supercell of the seed.
Max	Single bit set if seed supercell is a local maximum (i.e. there is no more energetic supercell in the tower.
Energy (TOB)	12-bit cluster energy in units of 100 MeV.
Energy(xTOB)	16-bit cluster energy in units of 25 MeV.
BCN4	4 lowest bits of BCN of data from which xTOBs are calculated

(b)

FIGURE 2.26. Turn-on curves of Run-2 and Run-3 algorithms. The Run-3 algorithm is set to equal rate to Run-2 and then a cut on the isolation variable cut is applied. The percentage of candidates removed in the given E_{T} bin is shown. Also shown are the rate reduction achieved by this additional isolation cut.

Category	Trigger	Rate 2e34	
		21.3	master
TAU	L1_TAU100	1.4 kHz	1.4 kHz
1tau	HLT_tau160_mediumRNN_tracktwoMVABDT	43 Hz	40 Hz
LLP tau	HLT_tau180_mediumRNN_tracktwoLLP	-	47 Hz
2TAU	L1_TAU60_2TAU40	3.1 kHz	3.1 kHz
2tau	HLT_tau80_mediumRNN_tracktwoMVABDT_tau60_mediumRNN_tracktwoMVABDT_03dRAB	9 Hz	9 Hz
LLP tau	HLT_tau80_mediumRNN_tracktwoLLP_tau60_mediumRNN_tracktwoLLP_03dRAB	-	9 Hz
2TAU L1Topo (Legacy)	L1_TAU60_DR-TAU20ITAU12I	1.3 kHz	1.3 kHz
2tau	HLT_tau80_mediumRNN_tracktwoMVABDT_tau35_mediumRNN_tracktwoMVABDT_03dRAB30	6 Hz	2 Hz
2TAU+1J	L1_TAU25IM_2TAU20IM_2J25_3J20	2.5 kHz	2.6 kHz
2tau	HLT_tau40_mediumRNN_tracktwoMVABDT_tau35_mediumRNN_tracktwoMVABDT_03dRAB	17 Hz	15 Hz
2TAU+1J L1Topo (Legacy)	L1_DR-TAU20ITAU12I-J25	4.3 kHz	5.8 kHz
2tau	HLT_tau35_mediumRNN_tracktwoMVABDT_tau25_mediumRNN_tracktwoMVABDT_03dRAB30	62 Hz	51 Hz
2TAU+2J	L1_TAU20IM_2TAU12IM_4J12p0ETA25	2.1 kHz	2.5 kHz
2tau	HLT_tau35_mediumRNN_tracktwoMVABDT_tau25_mediumRNN_tracktwoMVABDT_03dRAB	26 Hz	23 Hz

FIGURE 2.27. Primary tau trigger for Run-3. The rates for 21.3 (corresponding to equivalent Run-2 triggers) and master (Run-3) are shown. Triggers in blue are newly introduced in Run-3. [11]

CHAPTER III

SIMULATION AND RECONSTRUCTION

The primary method of searching for new physics in ATLAS is by attempting to find discrepancies between observed collisions and simulations of the known SM processes. This generally proceeds by identifying objects with properties of the theoretical BSM process that are not commonly found in SM processes. *Objects* here refers to patterns in the data that can be correlated with a species of particle or physics concept. The creation of objects from the raw output of detector systems is called *reconstruction*. One might think that searching in data for events with these properties is all one must do to confirm or deny the existence of the new physics. However, there are often SM processes that produce similar or identical signatures in the detector, which are referred to as *backgrounds*.

The output of any given proton-proton collision is inherently probabilistic due the quantum nature of the underlying processes. Furthermore, the ATLAS detector does not perfectly measure and identify the characteristics of the truth outgoing particles. Therefore, it is impossible to know with certainty whether any given observed event is the result of a SM or BSM process. What one can do is make statistical statements about the confidence of the observed results contradicting or supporting the SM as we know it. This requires a very precise understanding of the relevant BSM and SM processes which motivates the use of Monte Carlo (MC) simulations.

3.1. Monte Carlo Samples

To attempt to faithfully simulate the nature and properties of outgoing particles from a collision, a Monte Carlo process is employed which incorporates the sampling of many random variables while determining the final products of a simulated event. In this way, simulated samples are able to reproduce the distributions of objects and object variables without exactly recreating any individual event as it appears in data.

MC events are divided into samples based on their underlying physics process. If events were to be simulated together, then the vast majority of simulated events would be low-p_{T} QCD processes that are not needed for every analysis. This would be a waste of both time and computational resources. By separating processes, the samples chosen can be reweighted according to the cross section of each process to ensure a realistic mix of events. MC samples are useful both because they largely reproduce the reconstructed properties of data events of the same underlying process and because they give access to the true particle objects produced. This allows testing of methods that seek to identify e.g. which electron in the output event corresponds to the truth electron produced in the collision. The process of MC sample generation involves two steps: the generation of the truth particles produced in the event and the simulation of the particles traveling through ATLAS and producing a response in the detector.

3.1.1. Generation

Truth generation of an event involves the determination of all final-state particles that result from the proton-proton collision. This must take into account the initial state of the colliding protons plus their constituents, the process to be

simulated, decays of intermediate particles, and the pileup environment distinct from the process of interest. This process is performed by a number of specialized software packages called *generators*. The generators of interest in the work presented here are MadGraph5 for signal processes, Powheg for $t\bar{t}$, Pythia for JZW, and Sherpa for Z+jets and W+jets.

The primary process being simulated results from a hard scatter interaction involving incoming protons. *Hard scatter* here means that there is a large momentum transfer involved, producing final states with large final momentum and potentially creating intermediate systems of large mass. This is the process represented by a Feynman diagram such as that shown in Figure 1.6 and whose probability of occurrence can be calculated with the use of a matrix element (ME). The ME includes the summation of all diagrams with the desired incoming and outgoing particles. Tree level diagrams (if possible) contribute the most to the overall ME while diagrams with increasing loops contribute less and less while being more numerous. It is practical then to only calculate the ME to a certain order, denoted leading order (LO), next-to-leading order (NLO), next-to-next-to-leading order (NNLO), and so on. During calculation of the ME, divergences may occur in the UV and IR regimes. These are managed by the introduction of renormalization and factorization scales, μ_R and μ_F, respectively.

Although the particles collided by the LHC are protons, the particles that participate in a hard scatter interaction are the quarks and gluons, collectively called partons, that make up the proton. Each proton is made up of three permanent *valence* quarks and a multitude of other ephemeral *sea* quarks and gluons. Each parton only carries a fraction of the 6.8 TeV proton energy. The exact amount is determined by Parton Distribution Functions (PDFs), which give

the probability of finding a parton species with a fraction x $\in [0, 1]$ of the overall proton's momentum.

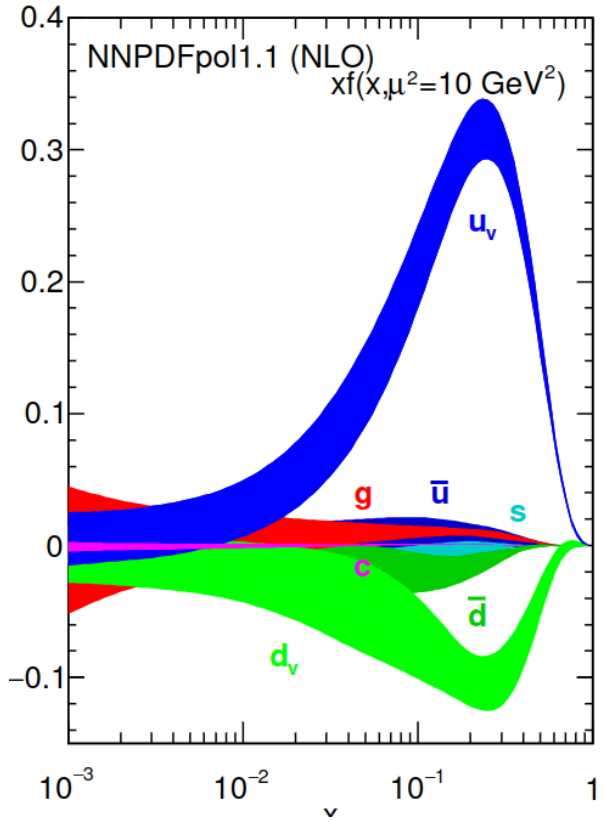


FIGURE 3.1. Plots of PDFs for different valence and sea partons inside the proton. [2]

Once the states of the initial partons that will undergo the hard scattering process are known, the calculation of the outgoing particles can be determined based on the previously mentioned ME. There may also be interactions between remaining partons of the same protons that must be accounted for. In total, this gives all outgoing particles from the two protons that are responsible for the hard scattering interaction, called the *underlying event*. These are a combination of stable particles that will survive to either interact with the detector (e, μ, π) or

pass through it (ν), unstable particles that will decay before reaching the detector (τ, bosons), and particles that carry color charge (quarks, gluons).

Unstable particles will decay and any unstable decay products will also decay until only stable particles remain. In parallel with this process is the parton showering (PS) of the particles carrying color charge, which proceeds in several steps. Hadronization is a consequence of the color confinement of the strong nuclear force, which is the observation that only particles with neutral color charge can propagate freely. A free quark or gluon will therefore quickly form into a colorless hadron. The energy for this is drawn from the strong nuclear force, whose strength increases with the distance between any two color-charged particles. The exception to this process is the top quark, whose lifetime is so short at 5×10^{-25} s [2] that it decays before hadronizing. All color-charged particles are also capable of radiating gluons, which will go on to hadronize themselves. The result is that a single prompt quark or gluon can create a complex shower of hadrons called a jet when it encounters the ATLAS detector.

Gluons may be radiated by partons before or after the hard scattering process occurs. The results of this happening before (after) hard scattering is referred to as Initial State Radiation (Final State Radiation) or ISR (FSR). In a similar fashion, electrically charged particles may also radiate photons at any point in this process.

Separate from the generation of the hard scatter process is that of the interactions producing pileup elsewhere in the event. This comes from all other proton-proton collisions during the same event and from events nearby in time whose signatures may effect the measurement of the detector. These pileup interactions of minimum-bias QCD processes are generated on their own and

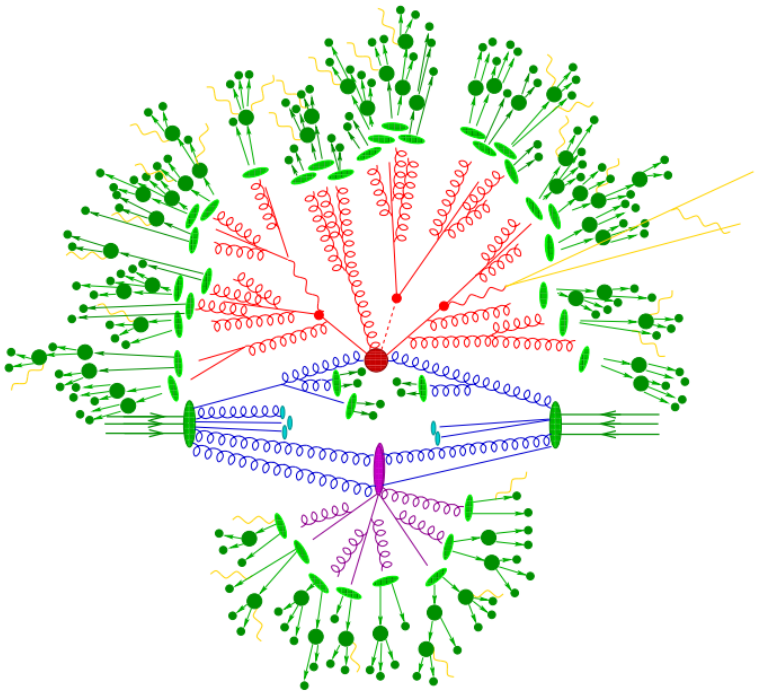


FIGURE 3.2. Diagram showing the main processes of a deep inelastic scattering. The processes are hard scattering (red), QCD radiation (orange), hadronization (light green), hadron decays (dark green), and QED radiation (yellow). [12]

then mixed in with the hard scatter products to create the full record of particles created in the event.

3.1.2. Detector Simulation and Reconstruction

The second stage of Monte Carlo simulation is the interaction of the generated particles with the ATLAS detector. This is done by recreating the detector to great detail inside the simulation program. Particles are then run through this simulated detector and interactions between the various materials in each detector subsystem are modeled. This process is also probabilistic and relies on the averaging over many potential interactions between the traversing particle and those that make up the detector material.

Once the MC simulation has determined the interactions of the truth particles with the detector, it can determine how those interactions are translated into the electronics signals output by the detector subsystems. At this point, the same system is used to reconstruct the MC data as is used to reconstruct observed event data.

3.2. Object Definitions

In order to better facilitate physics analysis, the outputs of cell energies and track hits from calorimeter and tracking systems must be analyzed and converted into physics objects. These objects may be particles such as electrons or other notable characteristics of the event such as $E_{\mathrm{T}}^{\mathrm{miss}}$. Each object takes inputs from the detector subsystems that are most sensitive to its signatures. Outlined in this section are the object definitions most relevant to the analysis presented.

3.2.1. Primary Vertex

Objects can be associated to the same collision by assigning them to a vertex i.e. a point of origin [63]. Vertices are created by considering ID tracks that have $p_{\mathrm{T}} > 400$ MeV, $|\eta| < 2.5$, and fulfill various requirements for number of hits in different ID subsystems [64]. Tracks are then grouped together based on whether they originate from nearby points. From these, a primary vertex is chosen that has a maximum value of $\Sigma p_{\mathrm{T}}^{\mathrm{track}}$.

3.2.2. Electrons

Electrons leave a relatively simple signature in the ATLAS detector. One characteristic is a narrow shower of energy primarily contained in the EM

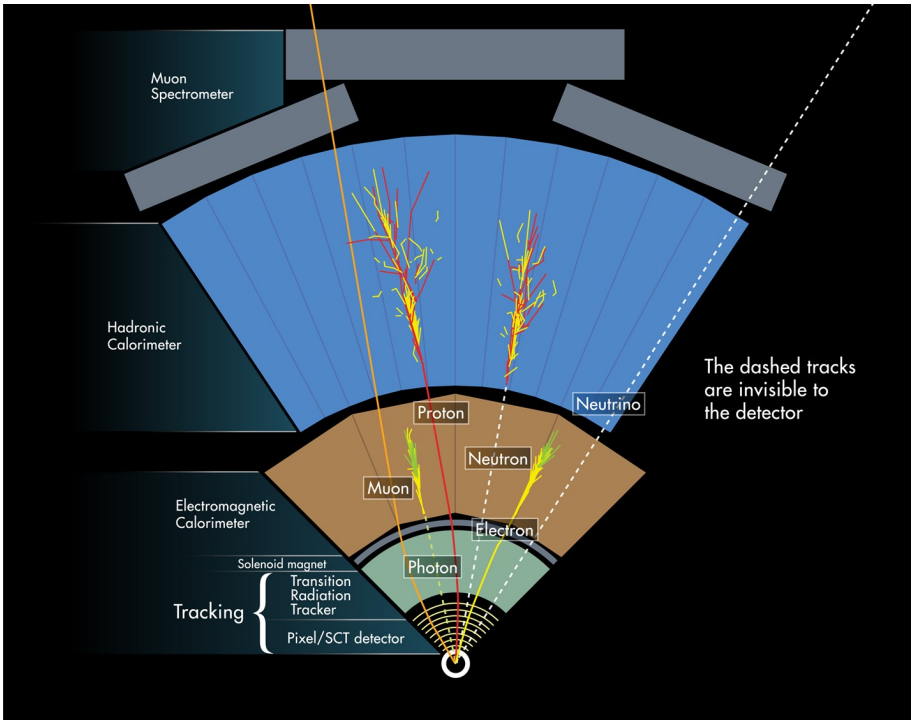


FIGURE 3.3. Diagram showing the signatures left in ATLAS subsystems by various particles. [13]

calorimeter. This is due to the bremsstrahlung radiation of photons. These photons then go on to pair produce into electron-positron pairs. This process repeats with the average energy of individual particles decreasing until they can be fully absorbed. Seeds in the calorimeter are identified by a sliding window algorithm with a cluster size of 3×5 in $\eta \times \phi$ and a cluster E_{T} threshold of 2.5 GeV [14].

The EM shower of an electron looks very similar to that created by an incident photon. However, a photon will not leave tracks in the ID as an electron does. An electron candidate then requires the existence of at least one track in the ID that is close to the seed cluster when extrapolated to the calorimeter. These tracks have a Gaussian-sum filter applied to them which account for energy losses due to bremsstrahlung which also change the curvature of the track. If multiple

tracks are found, the best is chosen based on position, track quality, and vertex matching.

Reconstructed clusters are formed from seed clusters by adding surrounding cell energies until the region is 3×7 in the barrel and 5×5 in the endcap. The energy in this cluster is then calibrated based on MC and data samples of $Z \rightarrow ee$ [65]. At this stage of the process the electrons are referred to as *reconstructed* electrons.

Further requirements can be placed on these objects to remove other particles that have mistakenly been reconstructed as electrons. This will necessarily remove some amount of true electrons in the process. This is called *identification* (electron ID or muon ID to distinguish from the Inner Detector), and is applicable to other reconstructed objects as well. Electron ID is a likelihood-based method using both track and cluster information. Four cuts are defined on the resulting discriminant variable, called VeryLoose, Loose, Medium, and Tight in order of increasing strictness.

Isolation is another important way to distinguish true electrons from background, as prompt electrons are expected to exist in a region without surrounding objects. Calorimeter and track-based isolation is determined by looking in a region of dR around the track or cluster and measuring the activity. Fewer tracks or clusters in this region signal a greater isolation and therefore a greater chance of the object being a true electron. This strategy fails, however, in cases where the electron is the decay product of a boosted system and so will have other child particles nearby. Various isolation working points (WPs) exist that use the tracking information, cluster information, or a combination of both.

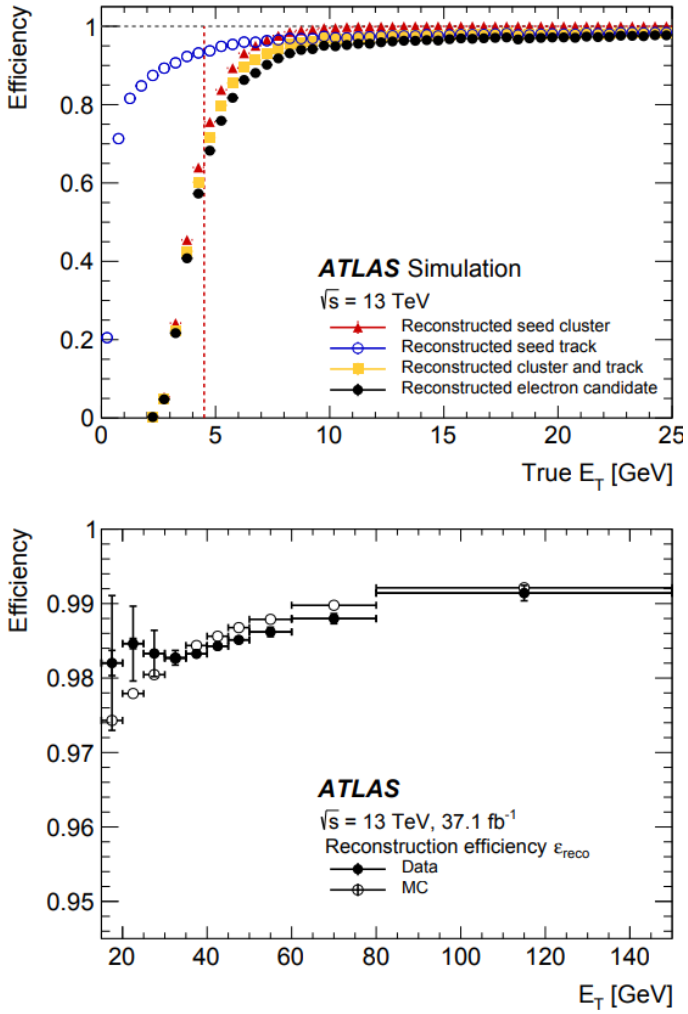


FIGURE 3.4. Plots showing (top) the efficiencies of electron reconstruction as a function of true E_{T} on a single-electron sample at various steps in the reconstruction process and (bottom) a comparison of MC and data reconstruction efficiency on $Z \to ee$ events. [14]

3.2.3. Muons

The most important subsystems for reconstructing muons are the specialized Muon Spectrometer and Inner Detector. Because muons are MIPs, their energy deposits in the calorimeter are small but can nevertheless also be used in reconstruction. There are several different combinations of signatures across systems that can result in a reconstructed muon.

Tracks are initially reconstructed independently in the ID and MS and then compared [15]. Pairs of tracks in the two different subsystems that are consistent with a single muon are matched together. Considerations are made for how energy loss throughout the path will change the curvature similarly to electrons in the ID. Muons reconstructed in this way are called combined muons (CM).

Inside-out muons (IO) are created when there is a standalone ID track and no full MS track. It can instead be matched to three independent hits with looser criteria. This is especially useful for $|\eta| < 0.1$ which has limited MS coverage. MS-extrapolated muons (ME) work in the opposite direction, from a standalone MS track which does not match an ID track. This is done for η values for which the MS has coverage but the ID does not ($2.5 < |\eta| < 2.7$). Segment-tagged (ST) muons involve an ID track extrapolated to match one MS segment. Finally, calorimeter-tagger (CT) muons search for energy deposits in the calorimeter that match the extrapolation of an ID track.

Muon ID WPs are defined similarly to electrons. In addition to Loose, Medium, and Tight WPs, there are High-p_{T} and Low-p_{T} definitions as well. In general, the ID requirements reject the decays of light hadrons which might be mistaken for muons. These decays fail to meet the ID requirements because of in-flight decays that produce lower-quality tracks. Isolation cuts, which consider the

energy of a cone around the muon compared to the muon p_{T} , are more effective at rejecting non-prompt muons from heavy hadron decays. These decays are likely to result in other products in the vicinity and therefore reduce isolation.

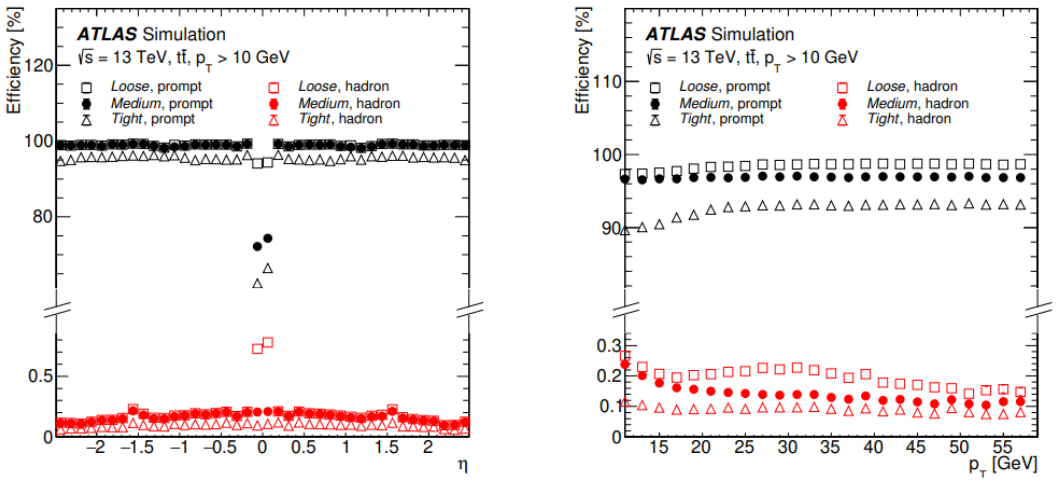


FIGURE 3.5. Efficiency of truth muons and muons from light hadron decays as a function of η (left) and p_{T} (right) at Loose, Medium, and Tight WPs. [15]

3.2.4. Jets

A jet is the name given to the spray of particles that are created by a gluon or quark following the hadronization described in Section 3.1.1. Although QCD is the driving force behind this process, further decays of some hadrons ($\pi_0 \rightarrow \gamma\gamma$) can lead to EM-showering particles. Therefore, a jet will appear as a wide "blob" of energy in the calorimeter with portions in both the EM and hadronic calorimeters. Charged hadrons will also create tracks in the ID.

3.2.4.1. Topoclusters and the Anti-k_t Algorithm

In order to identify which energy deposits in the calorimeter are most likely to belong to the same particle, the concept of topological clusters (topoclusters)

is employed [16]. The assumption is that a particle traversing the detector will deposit energy in a 3-D contiguous set of cells. To reconstruct these clusters, energetic seed cells are found and adjacent cells whose energies are above some threshold are successively added to the developing cluster.

The general topoclustering algorithm considers the measured cell energy divided by the average noise value of that cell. This noise value is a combination of electronic noise and the energy from pileup.

$$\varsigma_{cell} = \frac{E_{cell}}{\sigma_{noise,cell}} \tag{3.1}$$

Seed cells are those for whom ς_{cell} exceeds some value S. The growth step then considers all adjacent cells to those currently in the cluster and compares ς_{cell} to another value N. All cells with energies over N are added to the cluster and the growth step repeats. If clusters overlap during this step then the clusters are merged together. This continues until no more cells are added in a growth step. Finally, the last clustering step adds adjacent cells with ς_{cell} greater than P to the cluster. To prevent multiple close by particles from being included in the same cluster, splitting is performed if multiple local maxima are found within the same cluster. ATLAS uses as default a system of $S = 4$, $N = 2$, and $P = 0$.

The ATLAS calorimeter measures energies at the EM scale, meaning for EM objects like electrons, photons, and muons it will measure energies that are close to that of the true particle. However, hadronic energies will be underestimated which necessitates calibration. The ATLAS calorimeter is considered a non-compensating calorimeter for this reason. There is also generally energy lost due to the clustering process and to inactive material such as supports that can be accounted for.

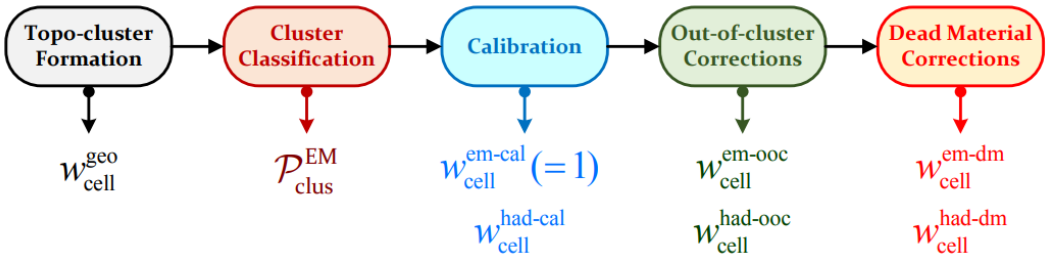


FIGURE 3.6. The sequence of steps involved in topocluster creation and calibration. [16]

Calibration is done at the topocluster scale. First, the cluster is classified as either being the result of EM or hadronic energy (or some mix) and assigned a value of $0 < \mathcal{P}_{\text{clus}}^{\text{EM}} < 1$. This value is determined by considering cluster depth in the calorimeter and average cell signal density. A cell-by-cell reweighting is then performed taking into account $\mathcal{P}_{\text{clus}}^{\text{EM}}$ and a hadronic calibration factor binned in calorimeter layer, cell η, signal cell density, and cluster energy at the EM scale. Out-of-cluster and inactive material corrections are made last and based on MC single-particle simulation. See Appendix A for work exploring the potential of deep learning models to replace the classification and calibration of topoclusters.

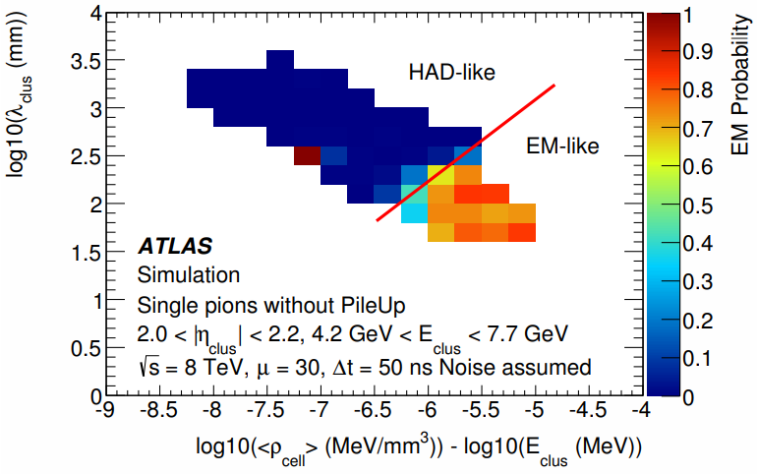


FIGURE 3.7. Probability that a topocluster is classified as EM-like as a function of its cluster depth λ_{clus}, signal cell density ρ_{cell}, and cluster energy. The line in red divides those clusters classified as EM-like or HAD-like. [16]

There are many potential algorithms for determining which clusters should be grouped together into a jet. The primary method in ATLAS is called the *anti-k_t* method [17]. The general effect of the algorithm is to create circular regions of inclusion to a jet, with different shapes created if two jets are close to one another.

The algorithm for clustering into jets proceeds by considering all combinations of two objects denoted i and j. Distances values are then calculated:

$$d_{ij} = \min(k_{ti}^{-1}, k_{tj}^{-1}) \frac{\Delta_{ij}^2}{R^2} \tag{3.2}$$

$$d_{iB} = k_{ti}^{-1} \tag{3.3}$$

where $k_{ti} = p_{\mathrm{T}}$ for i, Δ_{ij} is the dR between i and j, and R is a parameter that controls the maximum size of jets created. If the mininimum of all distance values is a d_{ij}, then the i and j objects are combined and replaced with a single object. If a d_{iB} value is the minimum, then the i object is called a jet and removed from consideration. The process continues until all objects are included in jets. The most commonly used jets in ATLAS are those with $R = 0.4$ and those with $R = 1.0$, also referred to as *large-R* jets.

Jets can also be created by clustering tracks measured in the ID. The same anti-k_{T} process applies and the resulting jets are referred to as *track-jets*. Track jets can be linked to jets in the calorimeter through *ghost-association*, which assigns an arbitrarily small energy deposit to the extrapolated position of the track-jet to the calorimeter. If this track-jet deposit would be clustered into a calorimeter jet, then the two are considered ghost-associated.

One important property of any effective jet clustering algorithm is called infrared and collinear safety (IRC). Collinear safety means that the algorithm

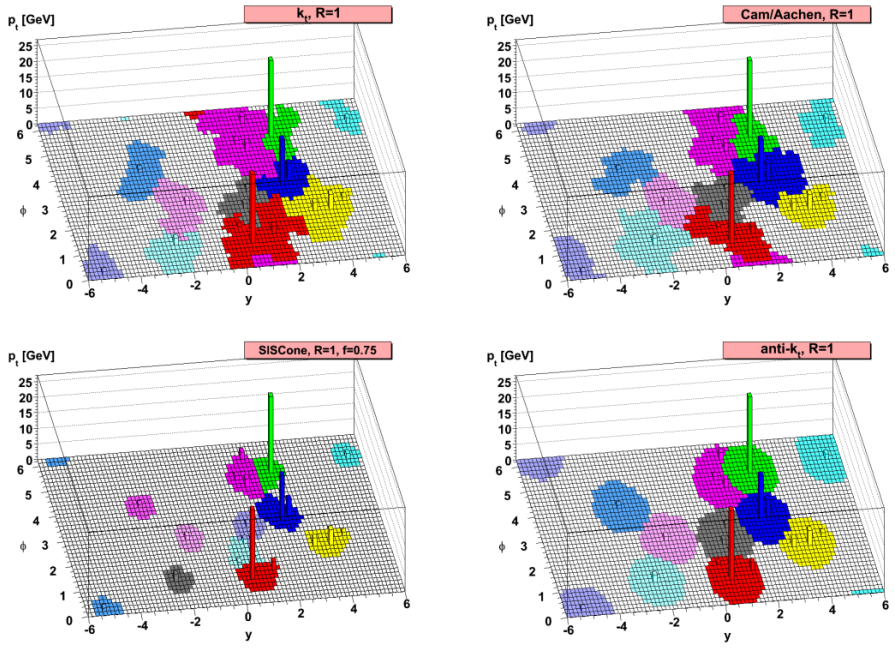


FIGURE 3.8. Diagram showing the clustering patterns of several different algorithms. The anti-k_t algorithm is shown on the bottom right. Regions shown are those in which an arbitrarily soft energy deposit would be added to jets of that color. [17]

outcome (final jets) should not be affected by the splitting of a jet into child jets that are collinear but together maintain the total energy and direction of the original. Infrared safety means that the addition of arbitrarily soft objects inside a jet should not change its boundary or other constituents. The anti-k_t algorithm satisfies both of these properties.

3.2.4.2. Jet Calibration

Once the constituents of a jet are determined, then a number of calibrations are made to correct its origin, energy, and direction [18]. This can be for either EM- or LCW-scale objects. The first of these is a vertex correction, which changes the vertex of the jet to the primary vertex from the default of the center of the detector. *in situ* energy corrections from MC samples is then performed in two

steps. First, a pile-up correction is applied which removes energy based on the average energy density in that event and area of the jet. Next is a correction of the observed energy change based on the number of primary vertices and μ to account for in-time and out-of-time pileup, respectively.

Absolute corrections are then made to the energy and η of the jets. MC samples are again used, this time comparing the reconstructed jet observables to their truth jet counterparts. Other characteristics of the jet that have been found to influence the detector response such as initiating particle and shower profile are also corrected. Finally, a last round of calibration is applied only to data events to correct for differences in response between MC and data.

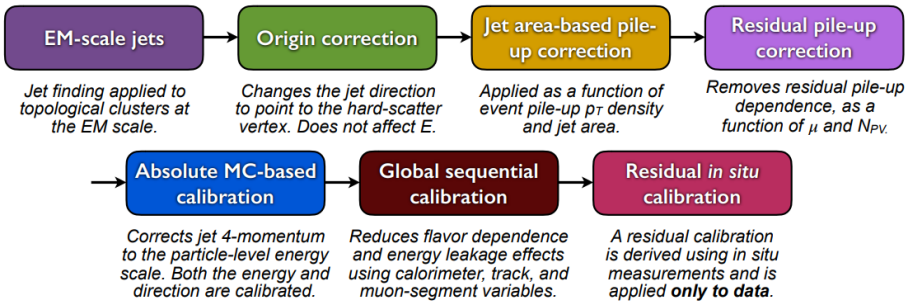


FIGURE 3.9. Diagram showing the steps of the jet energy calibration when starting from EM-scale calorimeter jets. [18]

3.2.4.3. Jet Trimming

A method that is used to further reduce the effect of pileup in jets is called *trimming* [66]. This method relies on the fact that energy deposits from pileup are usually much softer than those from the underlying jet. Because a larger jet size increases the chance of pileup contamination, this is most often done for $R = 1.0$ jets.

To remove the pileup, jet clustering is run again over the jet constituents with the R paramter set to some smaller value R_{sub}, creating what are called *subjets*. The energy of these subjets is then compared with the energy of the original jet. Any subjets with energy less than some defined cutoff value f_{cut} are then removed.

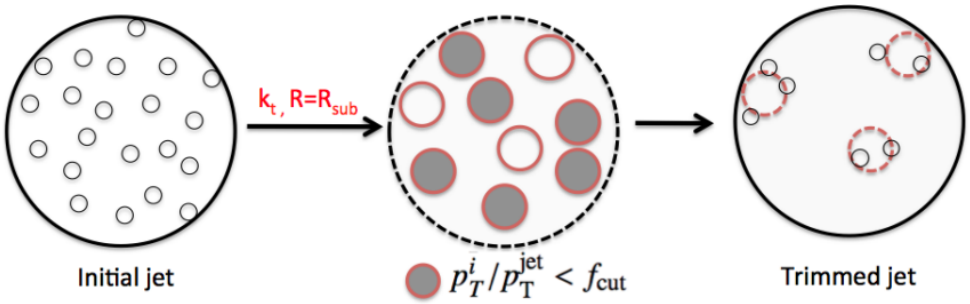


FIGURE 3.10. Diagram showing an example initial jet, clustering into subjets, and resulting trimmed jet. [19]

3.2.4.4. *b*-tagging

Of special interest in many analyses in ATLAS are those jets that are initiated by a b-quark, called b-jets. The process of classifying a jet as a b-jet as opposed to one initiated by another flavor of quark or a gluon is called b-tagging [20]. There are multiple algorithms that are used in ATLAS for b-tagging, but they all rely on the high mass, high decay multiplicity, and long lifetime of b-hadrons. The most difficult to distinguish from are charm quark c-jets, as they also have a large mass compared to light-flavor jets.

When a sufficienctly long-lived particle travels some distance before decaying into charged particles, this shows up as a *displaced vertex* in the ID. Displaced here refers to the tracks intersecting at some point off of the beam line. The

average mean lifetime of a B meson is on the order of 10^{-12} s, leading to a distance travelled of a few mm in the detector before decaying.

Tracks that result from a B meson decay can found in two ways. One is to consider pairs of tracks in a jet and perform a χ^2 fit to determine how well the tracks form a secondary vertex. Individual tracks can also be found by choosing those with large impact parameter d_0, the minimum distance in the transverse plane from the primary vertex at any point along the backwards extrapolated track.

The outputs of the two methods above, along with jet variables such as the energy, mass, and track multiplicity form the inputs to high-level multivariate algorithms. Of most interest here is the DL1 algorithm [67], which uses a deep neural network to classify the jet, outputting the probabilities that the jet is either a b-jet, a c-jet, or a light-flavor jet. Various working points are defined that can target a particular efficiency of accepting true b-jets and deliver a stated rejection of non-b-jets as evaluated on MC samples.

3.2.4.5. Variable-R Jets

A simple alteration that can be made to any jet clustering algorithm is to modify the size parameter R to be dependent on the p_{T} of the jet. A popular new definition for R_{eff} is:

$$R_{\mathrm{eff}} = \frac{\rho}{p_{\mathrm{T}}} \tag{3.4}$$

where ρ is a tuneable dimensionless constant. Jets clustered using this method are called *variable-radius* (VR) jets [68]. As low- (high-)p_{T} jets can grow arbitrarily large (small) under this system, there is often a minimum and

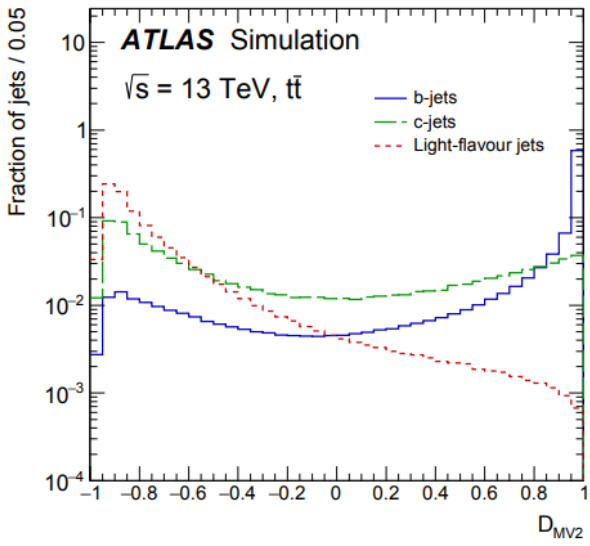


FIGURE 3.11. Distribution of DL1 output variables for b-jets, c-jets, and light-flavor jets. [20]

maximum R_{eff} allowed to provide bounds on the size of jets created. Any jet then that would have an R_{eff} outside of these bounds is instead set to the bound value.

VR jets are useful in the context of boosted objects with close-together decay products. In the context of $H \to b\bar{b}$, as the boost of H increases, the p_{T} of each b-jet increases and their ΔR decreases. With a fixed radius, both jets would eventually fall within that radius and be combined into a single jet. VR jets will decrease R_{eff} of each jet as its p_{T} increases, improving the chance that they will continue to be reconstructed separately.

3.2.5. τ_{had}

Hadronically-decaying taus share many of the same calorimeter and tracking signatures are jets. This is because the decay products (charged and neutral pions) are the same as those found in jets. In fact, the reconstruction process for τ_{had} uses many of the same techniques as that for jets, while also exploiting some of

the τ_{had} properties that can distinguish them from jets like long lifetime, low decay multiplicity, and narrow energy deposits.

The τ_{had} reconstruction algorithm is seeded with $R = 0.4$ anti-k_t jets [69]. A vertex is chosen as the track associated to the jet with greatest momentum fraction. Other tracks must then pass cuts based on their impact parameters relative to the vertex and number of hits in the tracker. Distinction is made between a *core* region ($\Delta R < 0.2$) and an *isolation* region ($0.2 < \Delta R < 0.4$) around the vertex.

Topoclusters within $\Delta R < 0.2$ of the seed jet barycenter are summed to calculate the direction and energy of the τ_{had} . The η and ϕ of the τ_{had} are defined in relation to the chosen vertex. An energy calibration is applied, starting with the sum of LC-scale topocluster energies. The amount of energy from pileup, E_{pileup}, is found to depend linearly on the number of primary vertices in the event, N_{PV}. Once this is subtracted, a detector response calibration $\mathcal{R}$ binned in energy, η, and number of prongs n_p is applied to restore the τ_{had} energy to that of the true $\tau_{\text{had-vis}}$ energy in MC samples.

$$E_{\text{calib}} = \frac{E_{\text{LC}} - E_{\text{pileup}}}{\mathcal{R}(E_{\text{LC}} - E_{\text{pileup}}, |\eta|, n_p)} \tag{3.5}$$

Identification of τ_{had} for rejection of jets is provided by Boosted Decision Trees (BDT) that have been trained to distinguish between $Z/\gamma* \rightarrow \tau\tau$ signal events and di-jet background [21]. Separate BDTs are trained for true 1-prong and 3-prong τ_{had} . Input variables to the BDT convey information about the EM and hadronic distribution of energy, mass of all track 4-momenta, and track distribution in the core and isolation regions, among others. Cuts on the BDT output scores define Loose, Medium, and Tight WPs for 1-prong (3-prong) τ_{had}

with corresponding signal efficiencies of 0.6 (0.5), 0.55 (0.4), and 0.45 (0.3), respectively. See Figure 3.12 for reconstruction and ID performance.

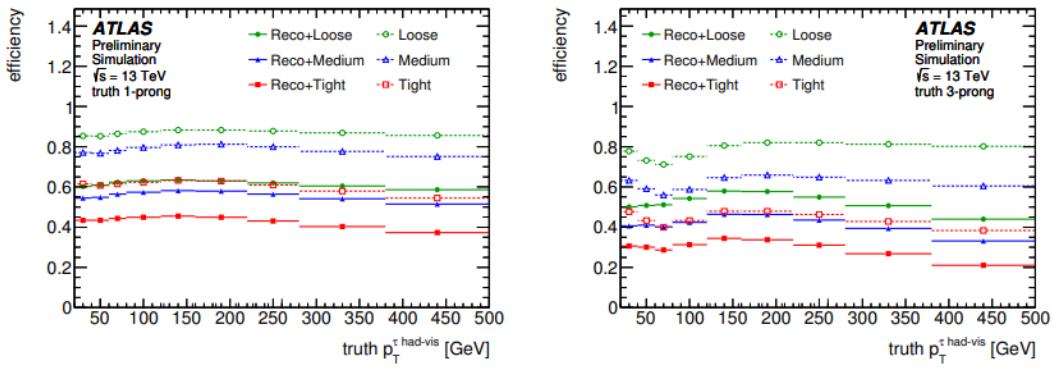


FIGURE 3.12. Efficiencies of reconstruction and reconstruction + identification for true 1-prong and 3-prong τ_{had} . [21]

Even with the BDT ID above, there is a possibility of electrons being mistaken for 1-prong τ_{had} . Shower shape information and number of hits in different regions of the tracker are used in creation of a likelihood discriminant between the two. In the event that a τ_{had} and an electron are reconstructed within $\Delta R < 0.4$ of one another, the τ_{had} is rejected if the electron has a likelihood value above some threshold that parametrized in η^{trk} and p_{T} to retain a constant 95% Loose τ_{had} efficiency.

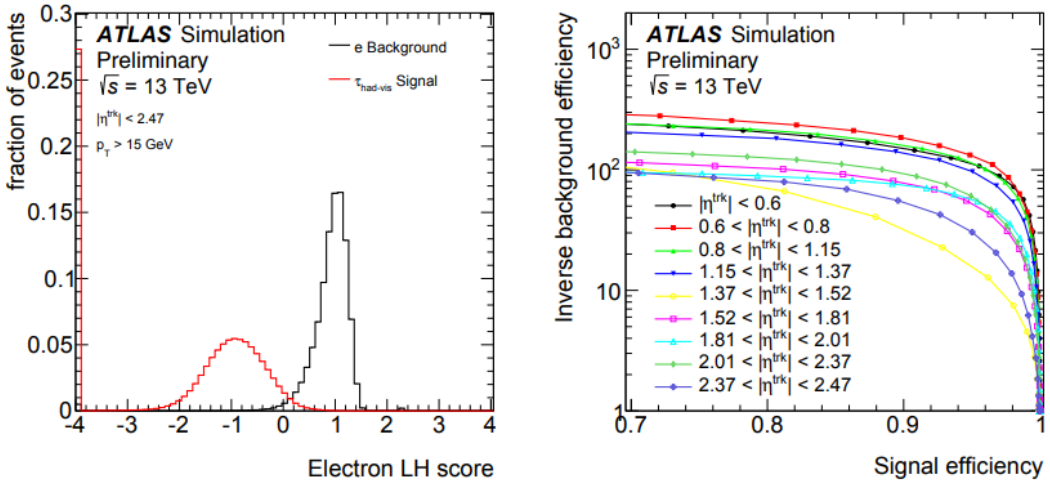


FIGURE 3.13. Distribution of electron likelihood variable and efficiencies of signal τ_{had} and inverse electron background. [21]

CHAPTER IV

BOOSTED LEPHAD $b\bar{b}\tau^+\tau^-$

The primary work presented here is the development of an analysis searching for a new heavy scalar decaying to two Higgs bosons. One Higgs boson then decays to a $b\bar{b}$ quark pair and the other decays to a $\tau^+\tau^-$ pair. The decays of the τ leptons can further be specified such that one decays hadronically and the other decays leptonically to an electron or muon, referred to as the *lephad* channel as opposed to the *hadhad* channel where both τ leptons decay hadronically. As shown in Figure 1.4, the $b\bar{b}\tau^+\tau^-$ branching ratio is reasonable at 7.3% and performs well at intermediate X mass values compared to the most common $b\bar{b}b\bar{b}$ channel due to more manageable background [70].

The heavy scalar process is simulated at high mass values of 1, 1.6, and 2 TeV to study various levels of boost in the system. In the fully boosted regime, the $b\bar{b}$ detector signatures fall within the size of a single large-R jet and the performance of resolving each individually using standard jets degrades. The same is also true of resolving the individual τ leptons in the $\tau^+\tau^-$ system.

The primary backgrounds considered are $t\bar{t}$, Z+jets, W+jets, and QCD dijet processes. These are chosen for their inclusion of some combination of τ_{had} , leptons, and jets in final states. In several cases, samples of specific decay channels are chosen to ensure this. Sample details are given in Section 4.1.

The final state that is selected for in this search is a single lepton (e or μ), a single τ_{had} that is close by the lepton, and a large-R jet that is far from the lepton and τ_{had} in ΔR to contain the decay products of the $H \to b\bar{b}$ system. The separation of variables is chosen because the two Higgs bosons are produced back-

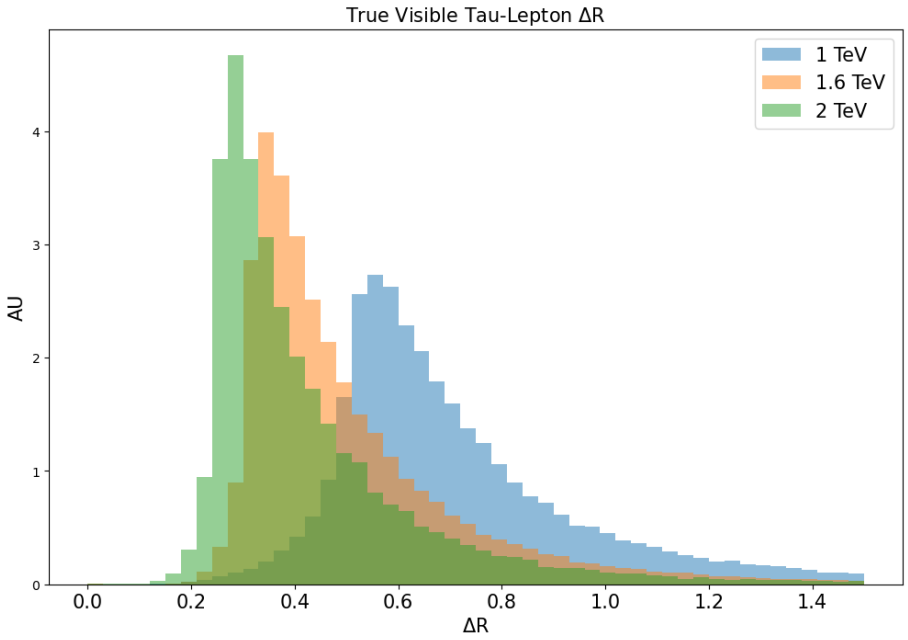


FIGURE 4.1. Histograms of ΔR between truth lepton and τ_{had} for signal sample events.

to-back or nearly so and therefore their decay products will appear in opposite sides of the detector. Details are given in Section 4.2.

The choice of final objects to represent an event proceeds in several steps. First, a requirement is made that events contain at least one lepton of a specific ID WP and a large-R jet with high p_{T} . For the leading lepton of each species in the event, an object nearby to that lepton is chosen as the τ_{had} candidate and a large-R jet at high separation is chosen for the $H \to b\bar{b}$ candidate.

A deep neural network is trained to separate signal events at all three mass WPs from $t\bar{t}$and Z+jetsbackground. The trained network is then applied to all samples and the output score is used in the final selection. Also used in the selection are an $H \to b\bar{b}$ tagger discriminant and the mass of the reconstructed HH system.

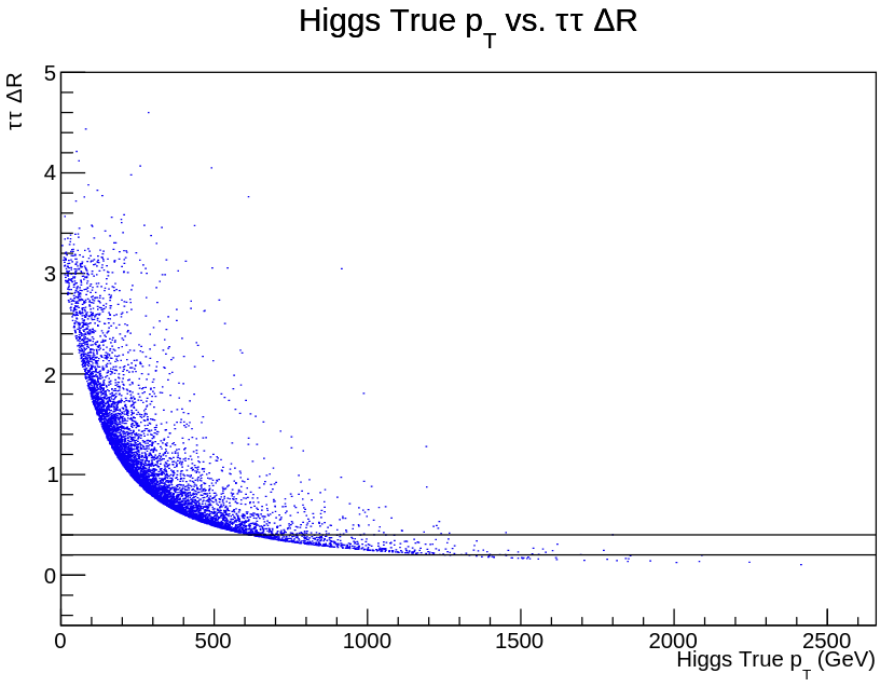


FIGURE 4.2. Scatter plot showing truth Higgs p_{T} and ΔR of the its $\tau\tau$ decay products. The horizontal lines show $\Delta R = 0.2$ and 0.4.

A signal region is defined based on the three variables above in Section 4.4. For those samples that are not adequately simulated by MC, ABCD regions are defined to give an estimate of their contributions. Based on the estimated number of events from all considered background processes, a limit can be placed on the overall $X \to HH$ cross section for each mass WP. These results are presented and compared to existing limits of other analyses in Section 4.8.

A simulated 1.6 TeV eHad signal event is shown using the VP1 software [71] in Figure 4.3. The reconstructed ID tracks and calorimeter hits are shown with cuts on p_{T} and E_{T} applied for easier viewing. The red track corresponds to an electron, the green cells are EM calorimeter hits, and the yellow cells are hadronic calorimeter hits. If the electron and nearby deposit are taken to be the $H \to \tau^+\tau^-$ system and the large opposite calorimeter deposits to be the combined $H \to b\bar{b}$

system, then the approximate back-to-back production and boost of the two Higgs bosons can be seen.

4.1. Samples

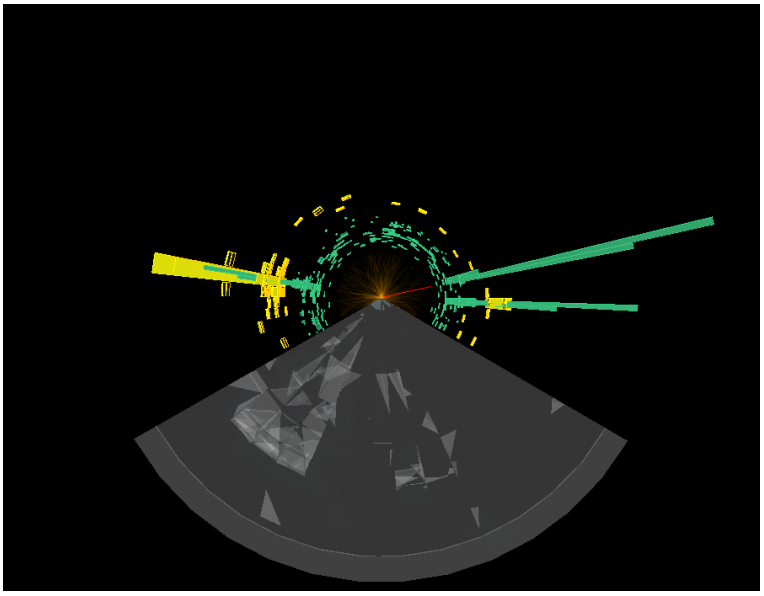
The MC and data samples used in this analysis are presented in Table 4.1. All MC samples are simulated using $\sqrt{s} = 13$ TeV proton-proton collision under Run-2 conditions in 2016.

All $X \to HH \to bb\tau\tau$ signal samples are generated with MadGraph5 2.6.1 [72] [73]. PS and hadronization is performed by Pythia 8.235 [74] for the 1 and 2 TeV samples and Herwig 7.7.1.3 [75] for the 1.6 TeV sample. For all mass points the NNPDF23 set of PDFs are used with the A14 tune [76] for 1 and 2 TeV and the MMHT tune [77] for 1.6 TeV. Decays involving c- and b-hadrons are handled using EvtGen 1.6.0 [78]. The Higgs mass is set to $m_H = 125$ GeV and the narrow width approximation is used with the width of the scalar X set to 10 MeV.

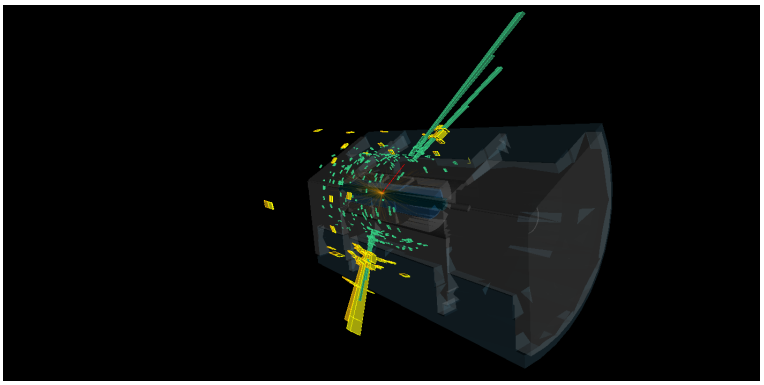
The $t\bar{t}$ background sample is generated with Powheg with the decay channels required to produce at least one lepton. Pythia 8.230 is used for PS and hadronization and EvtGen 1.6.0 for b- and c- hadron decays. The NNPDF23LO set of PDFs are used with the A14 tune.

The Z+jets background sample requires the Z boson to decay to two τ leptons. The maximum H_{T} of the primary vertex is required to be $280 < H_{\mathrm{T}}^{\mathrm{max}} < 500$ GeV. It also has a b-filter to ensure that at least one b-jet is present in the final state. It is generated with Sherpa 2.2.1 [79] which also handles the PS and hadronization modeling. The PDF set is NNPDF30NNLO [80].

Three W+jets background samples are considered and separated by the decay channel of the W boson ($\tau\nu$, $e\nu$, and $\mu\nu$). The same b-filter specification



(a)



(b)

FIGURE 4.3. A 1.6 TeV signal event viewed from along the beamline (a) and off the beamline (b). The red track is an electron, green deposits the EM calorimeter cells, and yellow deposits the hadronic calorimeter cells.

and cut on the $H_{\mathrm{T}}^{\mathrm{max}}$ are used as for the Z+jets sample. SHERPA 2.2.1 is also used for generation, PS, and hadronization and NNPDF30NNLO is also the PDF set.

JZW dijet background samples are organized by leading jet p_{T} in the event. They are named for increasing jet p_{T} ($JZ0W$ lowest, $JZ1W$ greater, etc). Considered in this analysis are the $JZ3W$, $JZ4W$, $JZ5W$, and $JZ6W$ slices. Events in each slice are assigned a weight in order to recover accurate distributions. Events for all samples are generated with PYTHIA 8.186 [81] which is also used for PS and hadronization. b- and c-hadron decays are handled by EVTGEN 1.2.0. The NNPDF23LO PDF set is used with the A14 tune.

Process	Generator	PS and Hadronization	PDF	Tune
$X \to HH \to bb\tau\tau$ (1 TeV)	MADGRAPH5 2.6.1	PYTHIA 8.235	NNPDF23	A14
$X \to HH \to bb\tau\tau$ (1.6 TeV)	MADGRAPH5 2.6.1	HERWIG 7 7.7.1.3	NNPDF23	MMHT
$X \to HH \to bb\tau\tau$ (2 TeV)	MADGRAPH5 2.6.1	PYTHIA 8.235	NNPDF23	A14
$t\bar{t}$	POWHEG	PYTHIA 8.230	NNPDF23LO	A14
Z+jets	SHERPA 2.2.1	SHERPA 2.2.1	NNPDF30NNLO	-
W+jets ($\tau\nu$)	SHERPA 2.2.1	SHERPA 2.2.1	NNPDF30NNLO	-
W+jets ($e\nu$)	SHERPA 2.2.1	SHERPA 2.2.1	NNPDF30NNLO	-
W+jets ($\mu\nu$)	SHERPA 2.2.1	SHERPA 2.2.1	NNPDF30NNLO	-
$JZ3W$	PYTHIA 8.186	PYTHIA 8.186	NNPDF23LO	A14
$JZ4W$	PYTHIA 8.186	PYTHIA 8.186	NNPDF23LO	A14
$JZ5W$	PYTHIA 8.186	PYTHIA 8.186	NNPDF23LO	A14
$JZ6W$	PYTHIA 8.186	PYTHIA 8.186	NNPDF23LO	A14

TABLE 4.1. Processes and generators of samples used in this analysis.

All generated particles are input to the GEANT4 [82] [83] simulation of the ATLAS detector and this output is processed through the same reconstruction chain as data.

4.2. Objects and Event Preselection

From considering the final state particles of the signal process, a "naive" object selection would select for events containing one lepton (e or μ), one τ_{had} , and two b-tagged jets. Indeed, this is the strategy for the resolved selection to

be described in Section 4.5. However, in the boosted regime the decay products of each Higgs boson overlap in the detector which presents problems for resolving each object individually. Presented in this section are the objects required in event preselection and the justification when those objects differ from the naive objects. In general, preselection requirements and cut values are chosen to preserve around 90% of signal events while removing as much background as possible.

At times, it will be of use to refer only to the events in which the leptonically decaying τ decays as $\tau \to e\nu$ or $\tau \to \mu\nu$. In these cases the events with a e (μ) in the final state will be referred to as *eHad* (*muHad*).

4.2.1. Truth-matching

A process that is useful for studying how an event's reconstructed objects compare to the truth particle record is called *truth-matching*. This method involves searching for a reconstructed object with some defined properties that lies within a defined region of a truth particle or object. Often, the proximity definition is based on the ΔR between the truth and reconstructed objects. If the proximity requirement is satisfied, it is assumed that the reconstructed object is the result of the signature of the truth particle being faithfully collected and reconstructed by the ATLAS detector. For the remainder of this paper, $\Delta R < 0.2$ will be used in truth-matching unless otherwise stated.

4.2.2. Fiducial Region

When assessing the efficiency of a selection being applied to signal samples, a *fiducial region* is used. This region of phase space involves truth objects that have a reasonable chance of being measured and reconstructed by the detector. In

practice, this means minimum requirements on the p_{T} and ensuring that the η of the truth particles is within the region in which the detector has sensitivity. The fiducial region cuts are:

- $\tau_{\text{had-vis}}$ $p_{\mathrm{T}} > 20$ GeV
- $\tau_{\text{had-vis}}$ $|\eta| < 2.5$
- e/μ $p_{\mathrm{T}} > 20$ GeV
- e/μ $|\eta| < 2.5$
- Di-τ_{vis} $p_{\mathrm{T}} > 300$ GeV

Di-τ_{vis} here refers to the combined 4-vector of the $\tau_{\text{had-vis}}$ and e/μ. Because the fiducial region is defined based on signal truth objects, it is not applied to background samples.

4.2.3. τ_{had}

As described in Section 3.2, τ_{had} reconstruction relies on energy deposits in the calorimeter as well as tracking from the Inner Detector. τ_{had} ID is partially based on distribution of clusters and tracks in core and isolation regions, meaning that it is sensitive to nearby objects in the detector. The boosted topology of the $H \rightarrow \tau\tau$ will cause the τ_{had} and lepton to be close together. The impact of this can be seen in the Figure 4.4. It shows the efficiency of truth-matching τ_{had} and e of various ID working points as a function of truth τ_{had} -e ΔR . Only the 1.6 TeV signal sample is shown and only those events that have been confirmed in the truth record to be the $\tau \rightarrow e\nu$ decay channel are included.

Because the Loose τ_{had} ID is so detrimental to selection efficiency, there is no requirement on τ_{had} reconstruction or ID applied in the preselection. τ_{had} without ID are used later in the analysis, but they are assumed to be common enough that

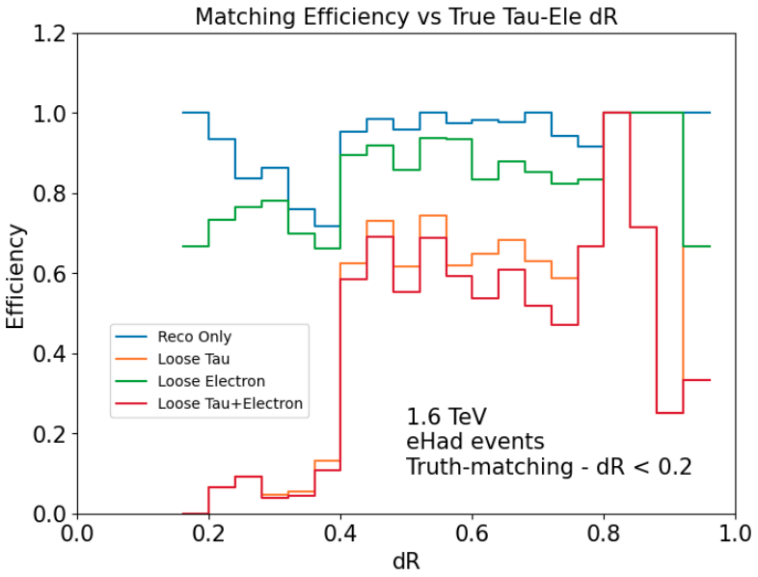


FIGURE 4.4. Efficiencies of reconstruction and reconstruction + identification for e and τ_{had} in 1.6 TeV signal sample.

an explicit requirement is not needed. This is validated later when there are no instances in which a τ_{had} is searched for but not found.

4.2.4. Electrons

Electrons are also susceptible to not being reconstructed or properly identified due to their signatures overlapping with the nearby τ_{had} . Figure 4.5 shows the efficiency of requiring a reconstructed electron plus the given electron ID working point. The Medium WP, which is used in the $b\bar{b}\tau^+\tau^-$ resolved analysis, removes too much of the signal sample to be used in the preselection. The Loose WP is chosen instead, which preserves 98%, 97%, and 96% of the 1, 1.6, and 2 TeV eHad signal events in the fiducial region, respectively. The slightly higher efficiency at lower signal mass points supports the assumption that a greater boost in the $H \to \tau\tau$ system leads to greater difficulty in identifying electrons.

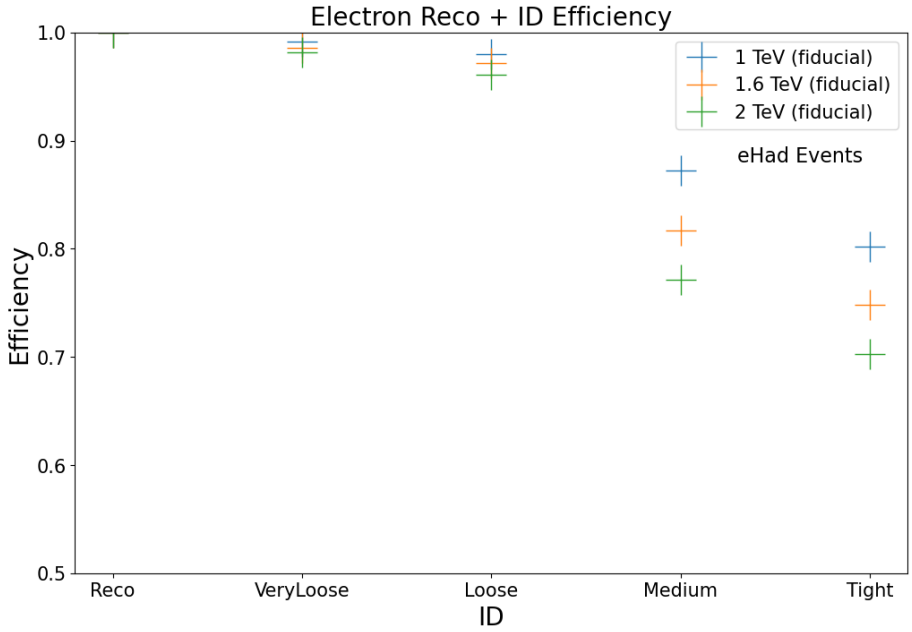


FIGURE 4.5. Efficiencies of requiring at least one reconstructed electron passing the ID working point for eHad signal samples.

4.2.5. Muons

Muon reconstruction and identification relies more heavily on MS and ID tracking information than calorimeter energy. Deposits from a nearby τ_{had} in the calorimeter therefore do not have a strong effect on muon ID as it does for electrons. A Tight ID is used for muons, which preserves 99% of muHad events in the fiducial region for all signal samples.

4.2.6. Large-R Jets

Events will be required to contain at least one large-R jet to contain the boosted $H \to bb$ system. Cutting on the p_{T} of the leading large-R jet is an effective tool for removing a large portion of background events while preserving most of the signal events. Figure 4.6 shows the efficiency of a leading large-R jet $p_{\text{T}} > X$ cut at various values for signal, $t\bar{t}$, and Z+jets samples.

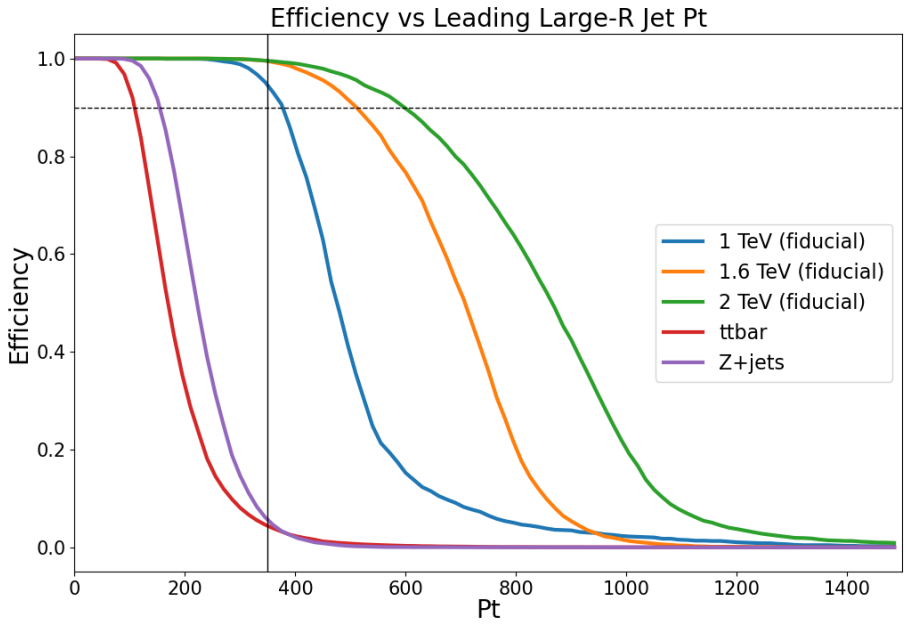


FIGURE 4.6. Efficiencies of requiring the leading large-R jet p_{T} events be greater than some threshold. The horizontal line is 90% and the vertical line is 350 GeV.

A cut requiring the leading large-R jet $p_{\mathrm{T}} >$ 350 GeV has efficiencies of 95%, 99%, and 99% in the fiducial region for 1, 1.6, and 2 TeV signal samples, respectively. For $t\bar{t}$ and Z+jets background samples, efficiencies are 4% and 5%, respectively.

4.2.7. Preselection

Using the objects above, a preselection can be defined. A single preselection is used for all signal samples for simplicity. The requirements of the preselection are:

- $\geq$ 1 Loose electron

  OR

  $\geq$ 1 Tight muon
- $\geq$ 1 Large-R jet

– Leading large-R jet $p_{\mathrm{T}} > 350$ GeV

As no ATLAS derivation existed that applied this preselection, a custom derivation was created for this analysis.

Sample	Input Events	Output Events	Presel. Efficiency (Fiducial)
$X \to HH \to bb\tau\tau$ (1 TeV)	100,000	72,278	72.3% (88.1%)
$X \to HH \to bb\tau\tau$ (1.6 TeV)	130,000	114,567	88.1% (95.6%)
$X \to HH \to bb\tau\tau$ (2 TeV)	60,000	53,826	89.7% (94.8%)
$t\bar{t}$	1,460,000	55,697	3.8%
Z+jets	490,000	18148	3.7%
W+jets ($\tau\nu$)	200,000	2,601	1.3%
W+jets ($e\nu$)	200,000	5,839	2.9%
W+jets ($\mu\nu$)	200,000	5,495	2.7%
$JZ3W$	100,000	3,734	3.7%
$JZ4W$	100,000	20,134	20.1%
$JZ5W$	100,000	28,771	28.7%
$JZ6W$	100,000	37,073	37.1%

TABLE 4.2. Input event and output event numbers and preselection efficiencies for all samples. Efficiencies in the fiducial region are also provided for signal samples. In all cases, raw event numbers are used without luminosity or event weights applied.

4.3. System Definition

After an event has passed preselection, the variables for final selection must be defined. To do this, a representative lepton, τ_{had} , and large-R jet variable are chosen per event per lepton decay channel. Because the preselection allows for events that contain both species (e/μ) of final state lepton, in these cases a set of objects, referred to here as an e- or μ- *system* or *lepton system* more generically, will be chosen for each. This is required for roughly $20 - 25\%$ of events across samples as shown in Figure 4.7. For an event in which only one species is present, only the relevant lepton system will be created.

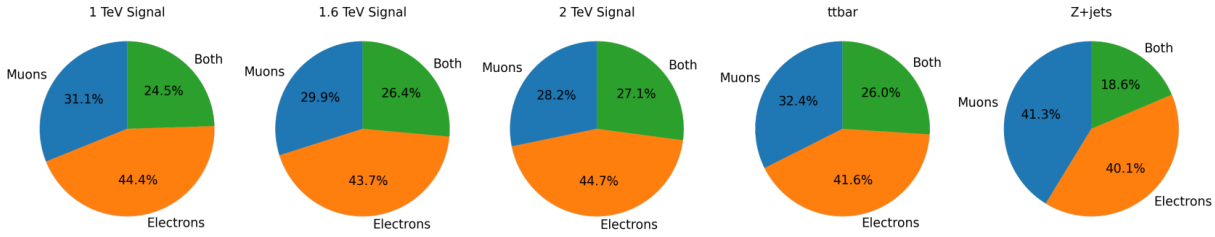


FIGURE 4.7. Percentage of events in each sample that contain only electrons, only muons, and both.

The algorithms for constructing an e-system and a μ-system are similar. In both cases, the leading lepton of the particular species is chosen. Around this lepton are the τ_{had} and large-R jet then chosen. The algorithms for defining each system were determined by evaluating their efficiency in choosing truth-matched reconstructed objects in signal samples.

4.3.1. e-system Leptons

The selection of the electron and τ_{had} for the e-system is more complex than that of the μ-system because of the reconstruction and identification problems posed by an electron and τ_{had} in close proximity to one another. Care is also taken to avoid the case in which electrons and τ_{had} are reconstructed on top of one another i.e. the same detector signature is reconstructed as both objects.

The e-system algorithm at one point looks for an $R = 0.2$ subjet if no suitable reconstructed τ_{had} candidate is found. Subjets are drawn from the ATLAS eHad di-τ object [84], which attempts to reconstruct and identify $\tau\tau$ systems by first searching for large-R jets with $p_{\text{T}} > 300$ GeV. Inside each of these jets, electrons and $R = 0.2$ subjets with $p_{\text{T}} > 15$ GeV are reconstructed. A di-τ object is created from every electron-subjet pair that has $\Delta R > 0.1$ where the subjet

is among the two leading subjets. The p_{T} of the di-τ is the p_{T} of the combined electron-subjet system.

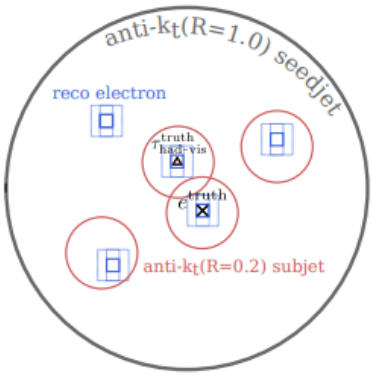


FIGURE 4.8. Diagram showing the eHad di-τ object, which looks inside an $R = 1.0$ jet for reconstructed electrons and $R = 0.2$ subjets.

The e-system is chosen as follows.

1. The leading electron is the candidate electron
2. The closest reconstructed τ_{had} to the candidate electron in ΔR is the candidate τ_{had}
3. If the candidate electron-τ_{had} $\Delta R < 0.1$, then the closest τ_{had} with $0.1 < \Delta R < 1$ to the candidate electron and $p_{\mathrm{T}} > 20$ GeV is the candidate τ_{had}
4. If no other τ_{had} is found but another electron exists within $\Delta R < 1.0$, the closest is the candidate electron
5. If candidate electron-τ_{had} $\Delta R < 0.2$, then consider the di-τ object with greatest p_{T} containing the candidate electron. The subjet of this di-τ is the candidate τ_{had} .

The candidate objects after these steps are the electron and τ_{had} objects of the e-system. τ_{had} will be used to refer to the τ_{had} /subjet object even if a subjet was chosen as the candidate.

4.3.2. μ-system Leptons

The algorithm for selection of μ-system leptons is much simpler. It also looks for pairs of muons and τ_{had} , preferably within an $R = 1.0$ region. Because τ_{had} reconstruction is more effective near a muon there is no referencing of ATLAS di-τ objects.

The algorithm is:

1. The leading muon is the candidate muon
2. The leading τ_{had} in $0.2 < \Delta R < 1.0$ from the candidate muon is the candidate τ_{had}
3. If no τ_{had} is found in the previous ΔR range, then the closest τ_{had} in ΔR in the event is the candidate muon

The candidate objects after these steps are the muon and τ_{had} objects of the μ-system.

4.3.3. Large-R Jets

For both e- and μ-systems, the process of choosing the large-R jet of the system to represent the $H \rightarrow b\bar{b}$ signature is the same. The leading and sub-leading large-R jets in p_{T} are found in the event. Of these, the jet that is further in ΔR from the chosen lepton of the system is the chosen large-R jet. This is intended to enforce the condition that the two Higgs bosons are produced back-to-back or nearly so in the detector.

4.3.4. System Truth-matching Performance

The percentage of relevant events for which the objects selected for each system are truth-matched is similar for both channels. In this instance $\Delta R < 0.1$ is used for truth-matching of leptons and τ_{had} and $\Delta R < 0.2$ is used for matching a large-R jet to the truth Higgs which decays to $b\bar{b}$. The e-system objects are truth-matched for 76.8% (88.1%) of eHad events in the 1 TeV (2 TeV) samples. The μ-system objects are truth-matched for 75.6% (86.3%) of muHad events in the 1 TeV (2 TeV) samples.

4.4. Signal Region

The signal region, which will be used to set expected limits on the signal process, is defined using three variables: the output variable of a kinematic neural network S_{KNN}, the mass of the di-Higgs system m_{HH}, and the discriminant of an $H \to bb$ tagger applied to the large-R jet of the system. This first requires a single e- or μ-system is chosen to represent each event, which will also be performed by the same neural network.

Cuts on these three variables are chosen in order to maximize overall sensitivity in the signal region for the 1.6 TeV X mass point in order to prioritize the boosted regime. The sensitivity defined in Equation 4.1 for S signal events and B background events. Maximizing sensitivity corresponds to placing the strictest limits in the background-only hypothesis used in this analysis.

$$s = \frac{S}{\sqrt{B}} \tag{4.1}$$

4.4.1. Kinematic Neural Network

The kinematic neural network (KNN) is used to classify events as signal- or background-like based almost exclusively based on the kinematic properties of an input lepton, τ_{had} , and large-R jet. It also performs channel sorting when an e- and μ-system both exist for a given event.

The KNN is a feed-forward dense neural network that accepts 19 input features listed below and outputs a single 0-1 classifier value. The inputs are the 4-vector components of each object in a chosen system plus a simple 0/1 flag denoting whether the input system is an e- (0) or μ- (1) system. The lepton and τ_{had} 4-vectors are summed and included separately as this is expected to be a useful discriminating quantity. The network is arranged with six hidden layers of [38, 57, 57, 57, 57, 38] neurons each. A sigmoid function is applied after each layer and the last layer has a 20% dropout applied to prevent overfitting.

The network was implemented and trained using PyTorch [85]. Training is performed using the SGD optimizer [86] with a binary crossentropy loss function.

Inputs to the neural network:

- e/μ p_{T} , η , ϕ , E
- τ_{had} p_{T} , η , ϕ , E
- Di-τ p_{T} , η , ϕ , E , m
- Large-R jet p_{T} , η , ϕ , E , m
- e/μ system flag

All signal, $t\bar{t}$, and Z+jets MC samples are used in the training of the neural network. In total around 300,000 events are used. Each event may contribute either 1 or 2 lepton systems as inputs based on the leptons present in the event.

Individual systems are assigned a training label of 1 if all reconstructed objects are truth-matched to their signal truth objects, otherwise they are given a label of 0. All systems derived from background events are labeled 0. This allows the network to discriminate against improperly chosen reconstructed objects of signal events as well as against background events. 2-fold cross validation is used in training, with separate networks being trained on even (odd) events and applied to odd (even) events for performance benchmarking.

For more information on the development and training of the neural network, see Appendix D.

4.4.1.1. Channel Sorting

If an event has both an e- and μ-system, one must be chosen to represent the entire event. This process is referred to here as *channel sorting*. We desire that the most signal-like system be chosen, which is accomplished by selecting the system with greatest S_{KNN} for each event.

TM Systems Kept	99.1%
TM Systems Discarded	0.9%
Non-TM Systems Kept	66.4%
Non-TM Systems Discarded	33.6%

TABLE 4.3. Percentages of TM (network label 1) and non-TM (network label 0) events kept and discarded when performing channel sorting for signal, $t\bar{t}$, and Z+jets events.

A simpler method of channel sorting would be to simply choose the system corresponding to the lepton with greater p_{T} . For signal events, both the leading lepton and S_{KNN} channel sorting techniques can be applied and compared to the truth record to determine the efficiency of sorting an event into the correct channel. The S_{KNN} technique sorts into the correct channel for 92% of events

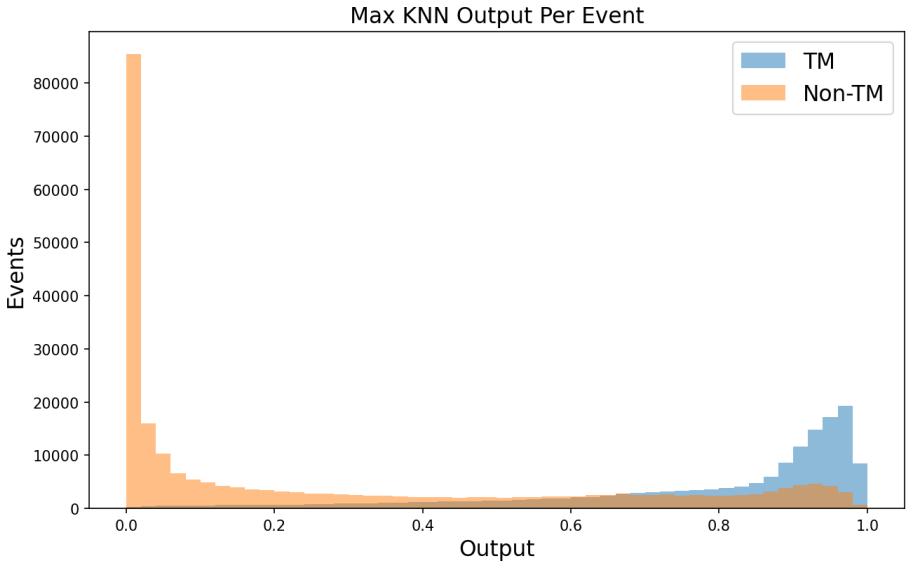


FIGURE 4.9. Distribution of max S_{KNN} per event for all events in the training samples.

whereas the leading lepton technique sorts to the correct channel for 80% of events.

For the remainder of the description of this analysis, all variables referenced are properties of the system chosen by the channel sorting process.

4.4.1.2. Event Classification

After channel sorting, there is now a single system and S_{KNN} value for each event. A cut can now be placed on this variable to reject background. Figure 4.11 shows the distribution of S_{KNN} values for each sample as well as how the efficiency, yield, and sensitivity change as a cut on S_{KNN} is varied. The yield value for each sample includes a uniform event weighting that scales each sample to the same overall integrated luminosity as described in Section 4.6.

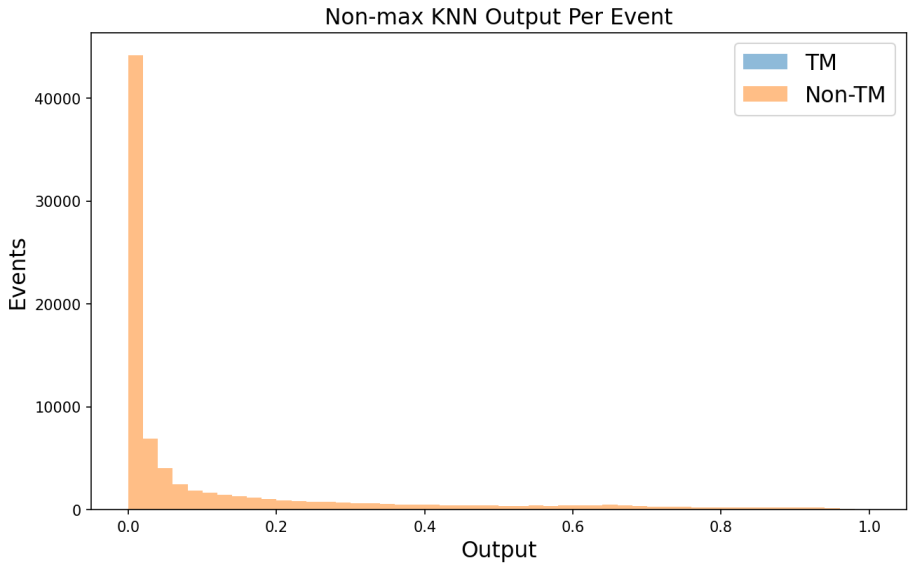


FIGURE 4.10. Distribution of non-max S_{KNN} per event for all events in the training samples.

The sensitivity of the 1.6 and 2 TeV samples are similarly maximized at a S_{KNN} value of 0.96. As can be seen from yield and efficiency curves, the Z+jetsevents are particularly well rejected by a cut on this variable.

4.4.2. m_{HH}

Because the $X \rightarrow HH$ signal will produce a resonance at the X mass, the mass of the di-Higgs system m_{HH} is an effective variable for rejecting background. m_{HH} here is defined as the mass of the 4-vector that results from summing those of the lepton, τ_{had} , and large-R jet.

In order to maximize the sensitivity of the 1.6 TeV sample, the cut on m_{HH} for the signal region definition is set at 1 TeV. Histograms and efficiency, yield, and sensitivity as a function of m_{HH} are shown in Figure 4.12.

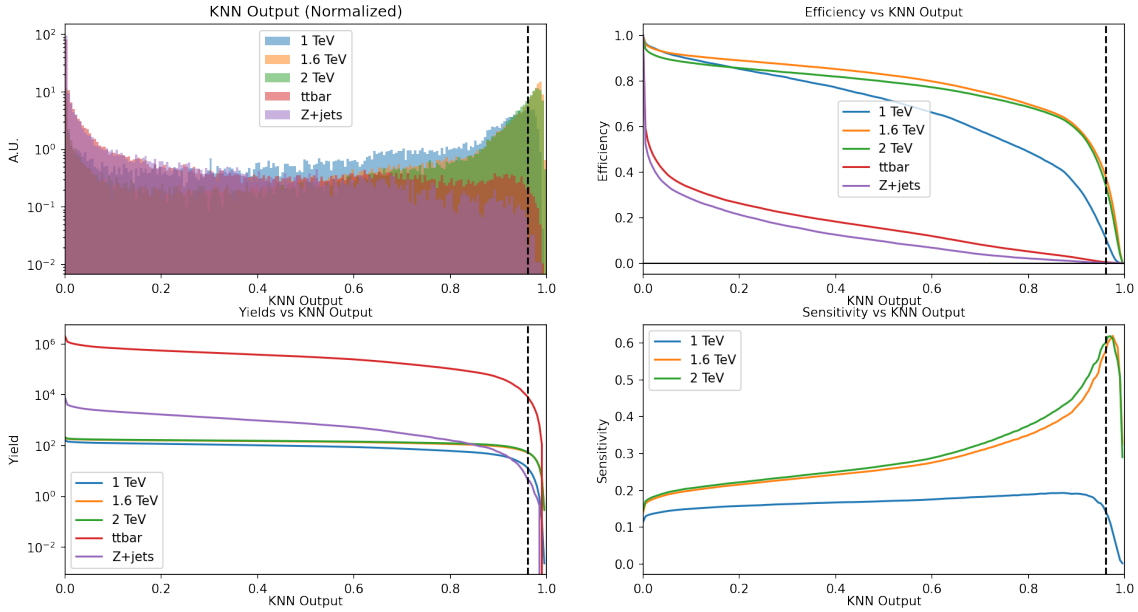


FIGURE 4.11. Histograms (top left) of the S_{KNN} variable and efficiency (top right), yield (bottom left), and sensitivity (bottom right) as a function of a $S_{\text{KNN}} > x$ cut. The vertical line is at $S_{\text{KNN}} = 0.96$.

4.4.3. $H \to bb$ Tagger

The third variable used for signal region definition is the discriminant value output by an $H \to bb$ tagger [87]. The tagger operates on a large-R jet to determine if it captures a true $H \to bb$ decay. These large-R jets are first trimmed by removing $R = 0.2$ subjets containing 5% or less of the total energy of the large-R jet. Variable-radius track-jets are then reconstructed and those that are ghost-associated with the large-R jet are kept. The DL1r b-tagging algorithm is applied to the track-jets. The number and b-tagging working point of these track-jets are used as inputs to a boosted decision tree along with the large-R jet kinematics and substructure.

The base outputs of the tagger are three values p_{H}, p_{QCD}, and p_{top} for relative classification as the result of Higgs, QCD, or top processes. These three

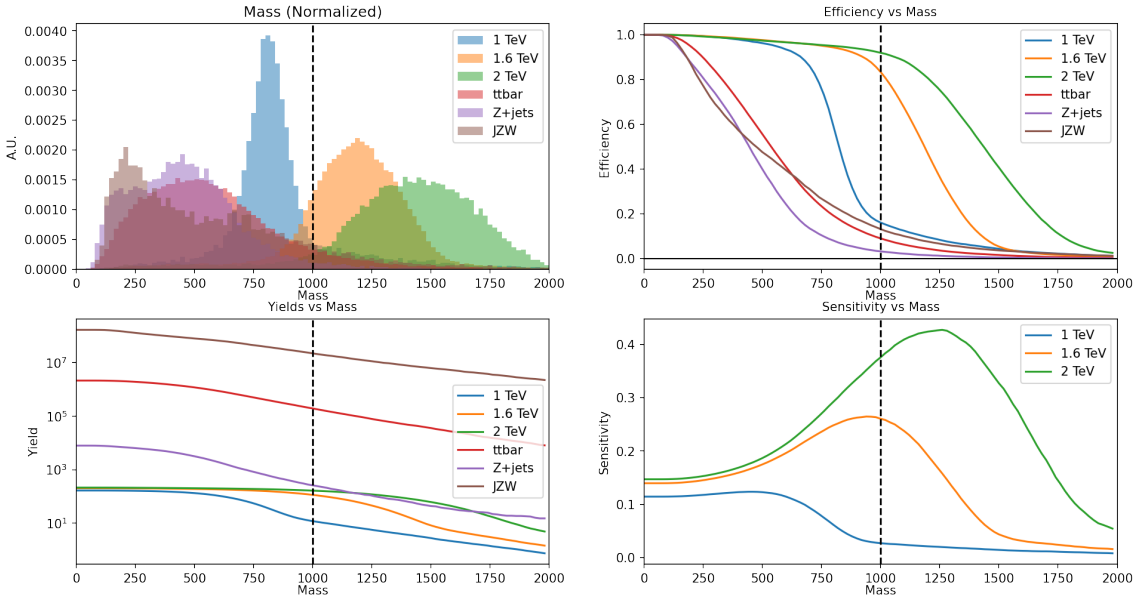


FIGURE 4.12. Histograms (top left) of the m_{HH} variable and efficiency (top right), event number (bottom left), and sensitivity (bottom right) as a function of a $m_{\mathrm{HH}} > x$ cut. The vertical line is at 1000 GeV.

variable along with a top fraction f_{top} can be combined for the single discriminant D_{Hbb}.

$$D_{\mathrm{Hbb}} = \ln\left(\frac{p_{\mathrm{H}}}{(1-f_{\mathrm{top}})p_{\mathrm{QCD}} + f_{\mathrm{top}}p_{\mathrm{top}}}\right) \tag{4.2}$$

Stated efficiency working points are provided for the $H \to bb$ tagger in the form of percentages of true $H \to bb$ processes preserved at various cuts on the D_{Hbb} discriminant. This was validated against large-R jets truth-matched to $H \to bb$ in signal events. The results are in agreement with the stated working points as shown in Figure 4.14.

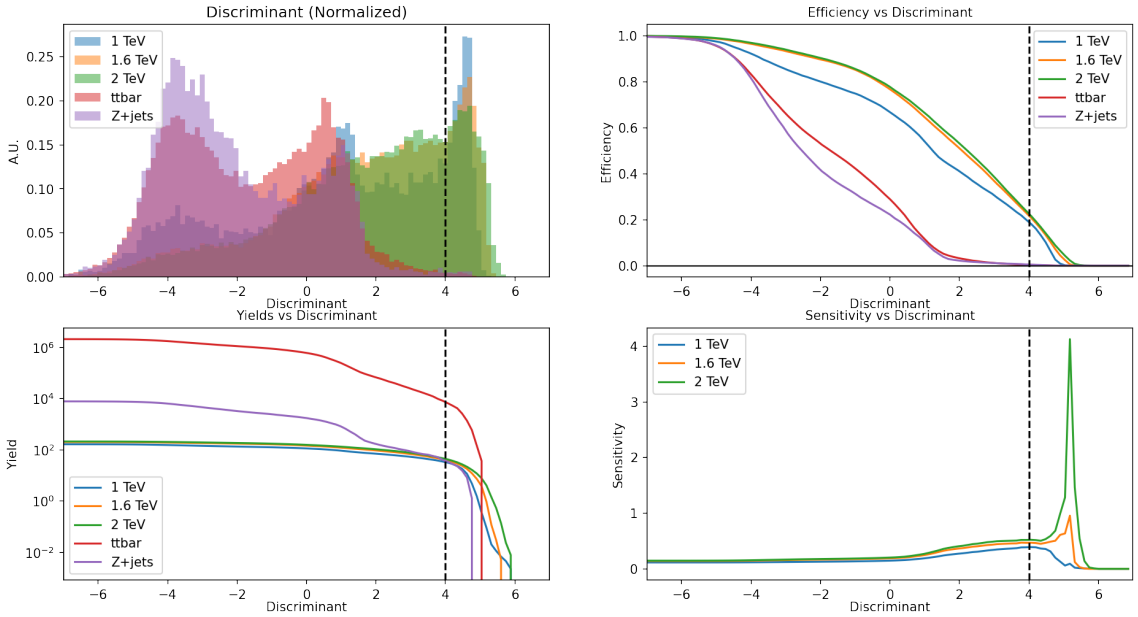


FIGURE 4.13. Histograms (top left) of the $H \rightarrow bb$ discriminant variable D_{Hbb} and efficiency (top right), event number (bottom left), and sensitivity (bottom right) as a function of a $D_{\text{Hbb}} > x$ cut. The vertical line is at $D_{\text{Hbb}} = 4$.

4.4.4. Full Selection

To this point, cut values on the three signal region variables have been determined in isolation. It must be verified that each of these values is still optimal in the presence of the others. We will also see that we are left with very low statistics when all three of these cuts are applied. The full signal selection cuts are:

- $m_{\text{HH}} > 1$ TeV
- $S_{\text{KNN}} > 0.96$
- $D_{\text{Hbb}} > 4$

First we check that the S_{KNN} optimal cut is unchanged after the m_{HH} cut has already been applied. In Figure 4.15 it can be seen that the $S_{\text{KNN}} = 0.96$ still

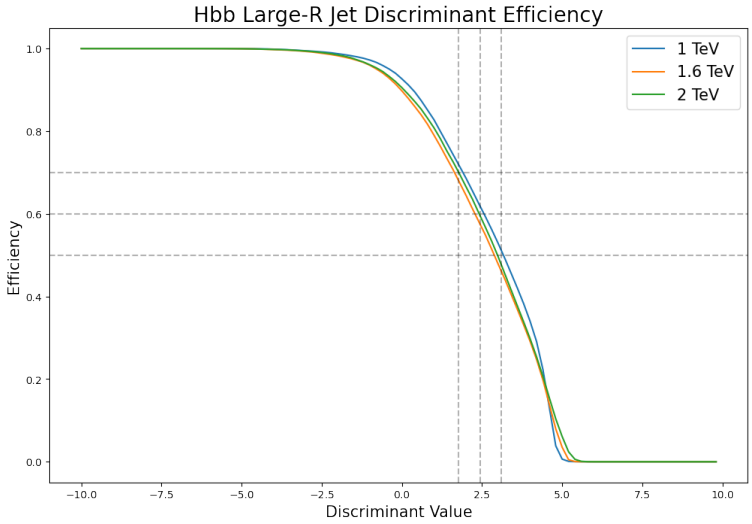


FIGURE 4.14. Efficiency of identifying true $H \rightarrow bb$ as a function of cut on D_{Hbb}. The horizontal dashed lines are the working point D_{Hbb} values of 1.76, 2.44, and 3.09. The vertical lines are their corresponding stated efficiencies of 50%, 60%, and 70%, respectively.

achieves the maximum possible sensitivity for both 1.6 and 2 TeV signal samples. Adding the D_{Hbb} cut on top of the previous two however, reduces the available statistics to the point that no events are present in the signal region for some background samples. Other methods described in Section 4.6 are used to estimate the number of events in the signal region.

4.5. Comparison to Resolved Selection

A resolved analysis searching for a heavy scalar in the $HH \rightarrow bb\tau\tau$ decay mode exists and is optimized for lower mass points [22]. It is important to determine whether or not the boosted selection described here has overlap with resolved selection if it were to be simply extended to higher X mass points. This can be tested by applying an approximation of the resolved selection and finding what portion of events are also selected by the boosted selection.

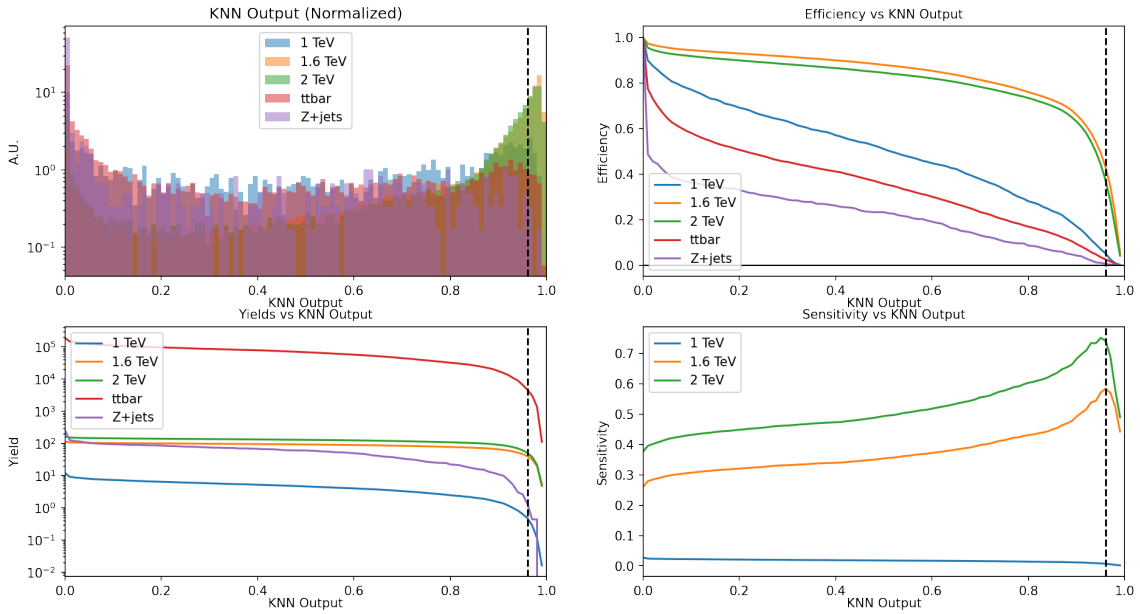


FIGURE 4.15. Histograms (top left) of the S_{KNN} variable and efficiency (top right), yield (bottom left), and sensitivity (bottom right) as a function of a $S_{\text{KNN}} > x$ cut with the $m_{\text{HH}} > 1$ TeV cut applied. The vertical line is at $S_{\text{KNN}} = 0.96$.

The resolved $\tau_{\text{lep}}\tau_{\text{had}}$ analysis selects one lepton (e/μ), one τ_{had} , and two b-tagged jets with p_{T} and $|\eta|$ restrictions along with event-level requirements. There are two different trigger channels which may cause an event to be accepted and the exact values of these cuts are dependent on which triggers the event satisfies, either the single-lepton triggers (SLT) or lepton-plus-tau-triggers (LTT). To simplify this comparison, a single set of cuts is derived from the combination of these two trigger channels. As the SLT channel contains the majority of total events, whenever there is a difference between the trigger channels the SLT values are chosen.

The set of cuts used here to approximate the resolved selection are shown below. This excludes trigger and vertex requirements. The resolved analysis further uses a BDT to classify signal and background events. The BDT is not

reproduced here, which results in the the below being an overestimate of the number of events that would pass the full resolved selection.

- 1 Loose lepton (e/μ)
- 1 Tight e
  - $p_{\mathrm{T}} > 25$ GeV
  - $|\eta| < 2.47$, $1.37 < |\eta| < 1.52$

  OR

  1 Medium μ
  - $p_{\mathrm{T}} > 21$ GeV
  - $|\eta| < 2.7$
- 1 Loose τ_{had}
  - $p_{\mathrm{T}} > 20$ GeV
  - $|\eta| < 2.47$, $1.37 < |\eta| < 1.52$
- Opposite-sign charge between lepton and τ_{had}
- $m_{\tau\tau}^{\mathrm{MMC}} > 60$ GeV
- 2 b-tagged jets
  - Leading (subleading) $p_{\mathrm{T}} > 45(20)$ GeV
  - $m_{\mathrm{bb}} < 150$ GeV

Table 4.4 shows the number of events passing the resolved, boosted, and both selections along with the percentage of events passing the boosted selection that also pass the resolved selection for each sample. The maximum of the overlap percentages is 2.50% for the 1.6 TeV sample. From this it can concluded that there is not a significant overlap in the resolved and boosted selections. This shows that the boosted selection is identifying signal events beyond those of the resolved selection and that accepting events from both selections in a future analysis would not add significant complexity from understanding their overlap.

Sample	N_{boosted}	N_{resolved}	N_{overlap}
$X \to HH \to bb\tau\tau$ (1 TeV)	23	4001	0
$X \to HH \to bb\tau\tau$ (1.6 TeV)	6747	2053	173
$X \to HH \to bb\tau\tau$ (2 TeV)	3919	413	29
$t\bar{t}$	1	147	0
Z+jets	0	35	0

TABLE 4.4. Number of events in signal, $t\bar{t}$, and Z+jetssamples that pass the boosted and approximate resolved selections as well as the number passing both.

4.6. Background Estimation

An estimate of the total number of background events in the signal region is needed in order to set expected limits on the $X \to HH$ process. Different methods are used to estimate the number of events depending on the particular background process. Those processes for which the MC modeling is considered accurate are estimated by a direct selection with consideration of the very low statistics available in the signal region. Those without accurate MC modeling are estimated using an ABCD method on control regions adjacent to the signal region.

A Run-2 integrated luminosity of 139 fb^{-1} is assumed for all event numbers. Sample weights are applied to all events in a sample uniformly to achieve the effective luminosity. This weighting takes into account the cross section of the process when it was generated, the luminosity being scaled to, the generator filter efficiency if e.g. only certain decay modes are included in a sample, and the number of events run through the initial derivation step of this analysis.

$$w_{\text{sample}} = \frac{\sigma_{\text{sample}} \times L \times e_{\text{filter}}}{N_{\text{D}}} \tag{4.3}$$

with w_{sample} weight applied to the sample, σ_{sample} the cross section of the process as it was simulated, L the luminosity (in this case 139 fb^{-1}), e_{filter} the generator filter efficiency, and N_{D} the number of derivation events.

For JZW samples, individual MC event weights are also applied to obtain accurate distributions.

$$w_{\text{event}} = w_{\text{sample}} \times w_{\text{event}}^{\text{MC}} \tag{4.4}$$

with w_{event} the weight applied to the event and $w_{\text{event}}^{\text{MC}}$ the MC event weight taken from the sample.

The w_{sample} and input values for each sample are listed in Table 4.5. The event weights for JZW samples are shown in figure 4.16.

Sample	σ_{sample} (fb)	e_{filter}	N_{D}	w_{sample}
$X \to HH \to bb\tau\tau$ (1 TeV)	8.29	0.20006	100000	0.0023
$X \to HH \to bb\tau\tau$ (1.6 TeV)	8.29	0.20006	130000	0.0018
$X \to HH \to bb\tau\tau$ (2 TeV)	8.29	0.20697	60000	0.0040
$t\bar{t}$	730000	0.5439	1480000	37.29
Z+jets	8680	0.1762	490000	0.43
W+jets ($\tau\nu$)	71944	0.1360	200000	6.80
W+jets ($e\nu$)	72077	0.1387	200000	6.95
W+jets ($\mu\nu$)	72063	0.1314	200000	6.58
$JZ3W$	2.65×10^{10}	0.00032	100000	11791.62
$JZ4W$	2.55×10^{8}	0.00053	100000	188.07
$JZ5W$	4.55×10^{6}	0.00092	100000	5.85
$JZ6W$	257530	0.00094	100000	0.34

TABLE 4.5. σ_{sample}, e_{filter}, N_{D}, and w_{sample} for all signal and background samples.

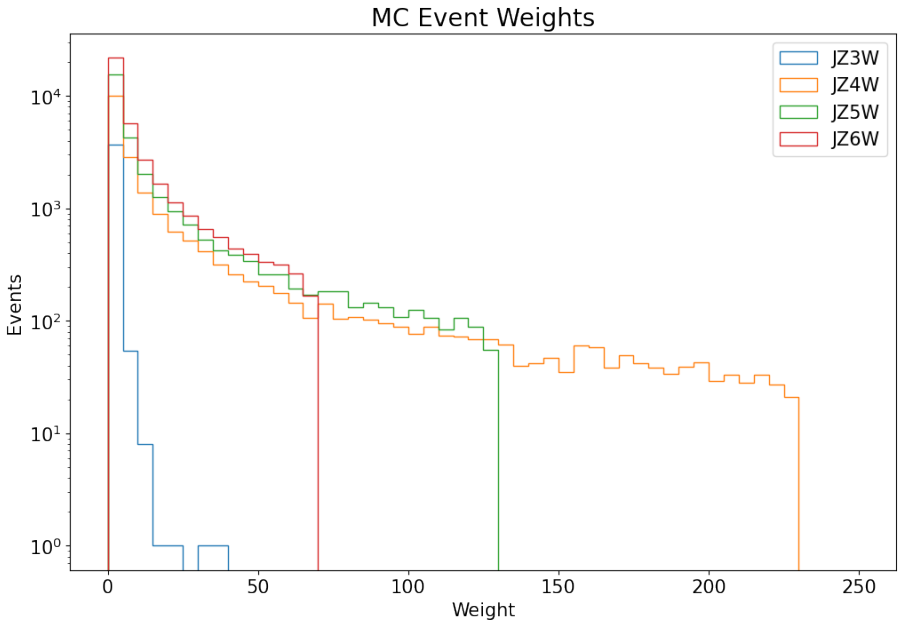


FIGURE 4.16. Histograms of the distribution of event weights for JZW samples.

4.6.1. Z+jets, W+jets

The Z+jets and W+jets samples are assumed to provide reliable simulation and reconstruction of events. Therefore, by applying the signal region selection with proper normalization one can determine their contribution.

Instead of directly applying all three cuts, the efficiency of the combined $m_{\mathrm{HH}} + S_{\mathrm{KNN}}$ cut is multiplied by the efficiency of the isolated D_{Hbb}. This will be referred to as the *Hbb-out* method. The $D_{\mathrm{Hbb}} = 4$ cut value is confirmed to still be the optimal value when the efficiencies of the combined $m_{\mathrm{HH}} + S_{\mathrm{KNN}}$ cuts are applied to the samples first.

This product can then be multiplied by the number of events in the sample being evaluated to obtain an estimate of the number of events in the signal region.

$$N_{\mathrm{signal}} = N_{\mathrm{sample}} \times e_{m_{\mathrm{HH}}+S_{\mathrm{KNN}}} \times e_{D_{\mathrm{Hbb}}} = \frac{N_{m_{\mathrm{HH}}+S_{\mathrm{KNN}}} \times N_{D_{\mathrm{Hbb}}}}{N_{\mathrm{sample}}} \tag{4.5}$$

A systematic uncertainty is attached to this value and is propagated from the initial uncertainties as:

$$\sigma_{\text{syst}} = N_{\text{signal}} \times \sqrt{\left(\frac{\sigma_{m_{\text{HH}}+S_{\text{KNN}}}}{N_{m_{\text{HH}}+S_{\text{KNN}}}}\right)^2 + \left(\frac{\sigma_{D_{\text{Hbb}}}}{N_{D_{\text{Hbb}}}}\right)^2 + \left(\frac{\sigma_{\text{total}}}{N_{\text{total}}}\right)^2} \tag{4.6}$$

where $\sigma_{\text{x}} = \sqrt{N_{\text{x}}}$ is Poisson standard deviation.

The two efficiency measurements, the resulting number of events in the signal region, and statistical and systematic errors for each sample are shown in Table 4.6.

Sample	$e_{m_{\text{HH}}+S_{\text{KNN}}}$	$e_{D_{\text{Hbb}}}$	N_{signal}	σ_{stat}	σ_{syst}
Z+jets	1.7×10^{-4}	4.7×10^{-3}	6.1×10^{-3}	7.8×10^{-2}	3.6×10^{-3}
W+jets ($\tau\nu$)	0	4.2×10^{-3}	0	0	0
W+jets ($e\nu$)	1.7×10^{-4}	5.0×10^{-3}	3.4×10^{-2}	1.9×10^{-1}	3.5×10^{-2}
W+jets ($\mu\nu$)	5.5×10^{-4}	6.2×10^{-3}	1.2×10^{-1}	3.5×10^{-1}	7.4×10^{-2}

TABLE 4.6. Selection efficiencies, number of events in the signal region, and associated errors for Z+jets and W+jets samples.

A check of the errors can be made by relaxing some of the cuts and seeing how the ratio of signal region events between the Hbb-out method and full selection behave with increased statistics. In this test the D_{Hbb} cut was relaxed to 0 and the S_{KNN} cut was ranged from 0 to 0.9 in increments of 0.1. The results are shown in Figure 4.17.

From the fluctuations of the ratio, one could apply a 50% error to the number of events in the signal region. This being less than the errors shown in Table 4.6, the errors in the table are kept.

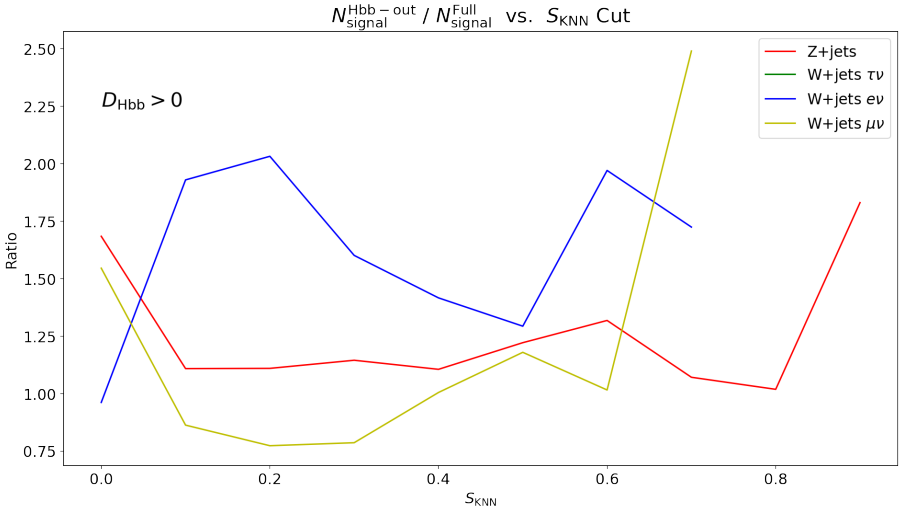


FIGURE 4.17. Ratio of events measured using Hbb-out and full selection methods in relaxed signal region as a function of S_{KNN}.

4.6.2. $t\bar{t}$, JZW

For the backgrounds for which MC samples may not be accurate models of the true processes, an ABCD method is used to approximate their signal region contamination. This involves creating control regions in two uncorrelated variables and comparing the ratios of events between different regions.

The control regions are defined in the S_{KNN} and D_{Hbb} variables once the m_{HH} cut has already been applied. The $S_{\mathrm{KNN}} > 0.96$ cut is altered to $0.2 < S_{\mathrm{KNN}} < 0.96$ and the $D_{\mathrm{Hbb}} > 4$ cut is altered to $-2 < D_{\mathrm{Hbb}} < 4$. Region definitions are summarized in Table 4.7 and the distribution of $t\bar{t}$ and JZW events in the signal and control regions is shown in Figure 4.18.

Region	Min S_{KNN}	Max S_{KNN}	Min D_{Hbb}	Max D_{Hbb}
D (signal)	0.96	-	4	-
C (control)	0.2	0.96	4	-
B (control)	0.96	-	-2	4
A (control)	0.2	0.96	-2	4

TABLE 4.7. Control region definitions. - denotes that no value is specified

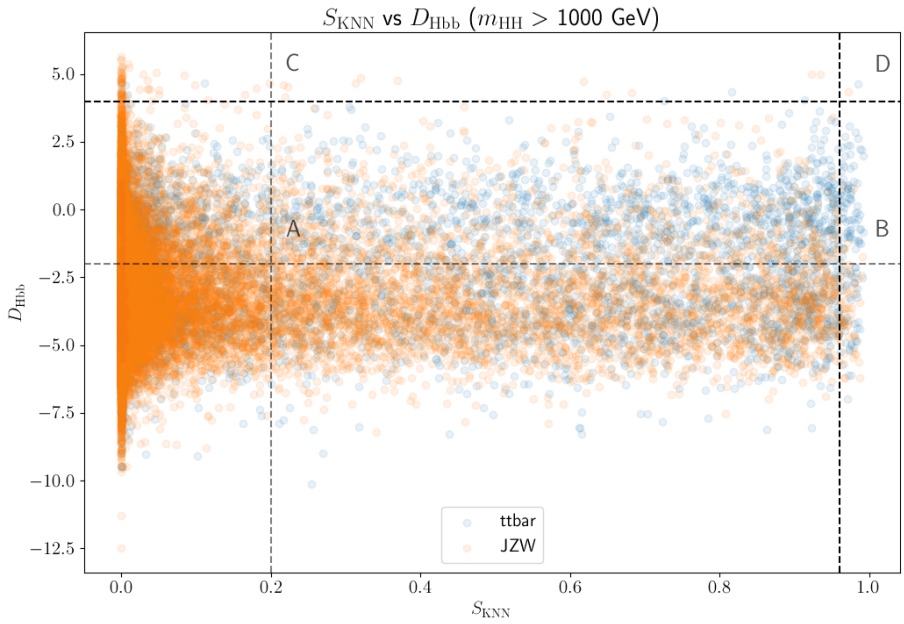


FIGURE 4.18. Scatter plots of $t\bar{t}$ and JZW events in the signal (D) and control (A,B,C) regions.

To estimate the number of events in the signal region D, the ratio of events in the A and B regions is found and then applied to the number of events in the C region:

$$N_{\mathrm{D}} = N_{\mathrm{C}} \times \frac{N_{\mathrm{B}}}{N_{\mathrm{A}}} \tag{4.7}$$

with associated uncertainty:

$$\sigma_{\mathrm{D}}^{\mathrm{syst}} = N_{\mathrm{D}} \times \sqrt{\left(\frac{\sigma_{\mathrm{C}}}{N_{\mathrm{C}}}\right)^2 + \left(\frac{\sigma_{\mathrm{B}}}{N_{\mathrm{B}}}\right)^2 + \left(\frac{\sigma_{\mathrm{A}}}{N_{\mathrm{A}}}\right)^2} \tag{4.8}$$

The number of events shown in each control region as well as the derived signal region events plus statistical and systematic uncertainties are shown in Table 4.8.

Sample	N_{A}	N_{B}	N_{C}	N_{signal}	σ_{stat}	σ_{syst}
$t\bar{t}$	4.43×10^4	2.87×10^3	1.12×10^2	7.24	2.69	0.70
JZW	7.63×10^5	1.15×10^2	2.72×10^3	4.12	2.03	0.14

TABLE 4.8. Events in control and signal regions plus statistical and systematic uncertainties for $t\bar{t}$ and JZW samples.

A similar check can be performed here as with the Z+jets and W+jets backgrounds. In this case the D_{Hbb} cut is set to 2 and the S_{KNN} cut ranges from 0.5 to 0.96. An error of roughly 30% read from Figure 4.19 is similar or less than that in Table 4.8.

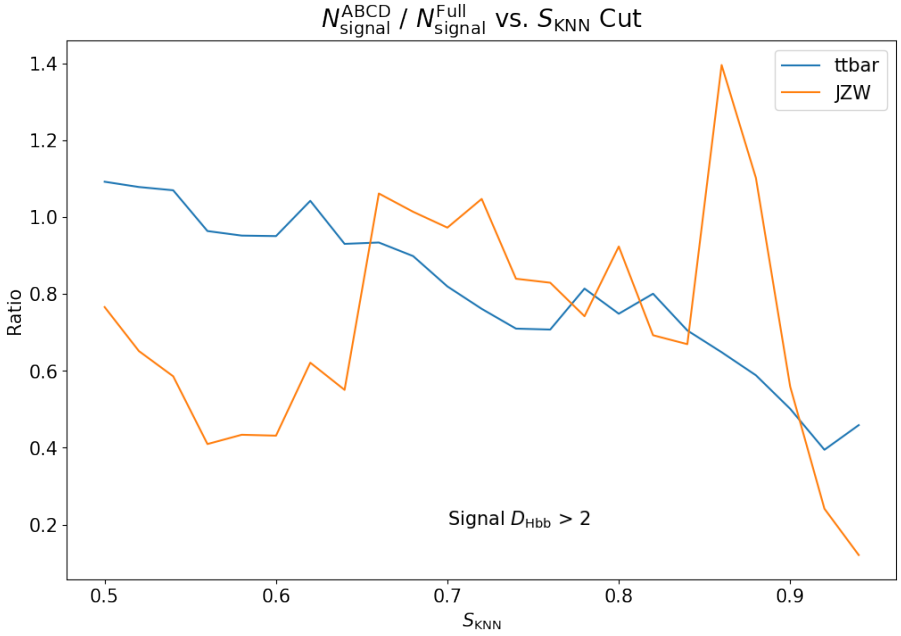


FIGURE 4.19. Ratio of events measured using ABCD and full selection methods in relaxed signal region as a function of S_{KNN}.

4.7. Validation Against Data Sample

In order to validate the performance of this analysis against real data collected by the ATLAS detector, a sample from Run-2 was chosen. In general, the signal region selection is not applied to data events until the development of

the analysis is fully completed. This practice, called *blinding*, aims to prevent the analysis from being tailored to the data collected. Were blinding not performed, the chance of a normal statistical fluctuation in the data being mistaken for a real signal (a false positive) increases dramatically. Blinding increases the chance that the analysis is sensitive to the properties of the true process rather than these statistical fluctuations.

Although the signal region is not formally blinded in this case, a similar principle is adhered to by the fact that this data sample represents a very small fraction of the total luminosity of Run-2. Because of this, statistical fluctuations in this sample are assumed to be largely uncorrelated with those of the full Run-2 dataset. Furthermore, because of the assumed small cross section of the process being tested, it is unlikely that any significant signal will be found with such low statistics.

The data sample in question was collected during a single run of data-taking during 2018 and corresponds to .11 fb^{-1} of integrated luminosity. This is only $7.9 \times 10^{-2}\%$ of the total integrated luminosity of Run-2. The run was taken from a GoodRunsList, meaning that all detector subsystems were verified to be functioning properly during the collection of this data.

Before validating against the data sample, the main background samples being considered were $t\bar{t}$, Z+jets, and W+jets. All samples were normalized to the .11 fb^{-1} of the data sample and the distributions of the signal region variables were compared. It was observed that the distributions from the background samples did not align with that of the data sample. This was expected as the cuts approached the values in the signal region definition, as it would mean that non-dominant backgrounds are being removed.

Given the orders of magnitude difference between the numbers of events in the selected background samples and data sample, another overlooked background process was required. The JZW samples were added for consideration as a result. The JZW distribution matches the data distribution much better at low and high values of the signal region variables. The example of m_{HH} is shown in Figure 4.20.

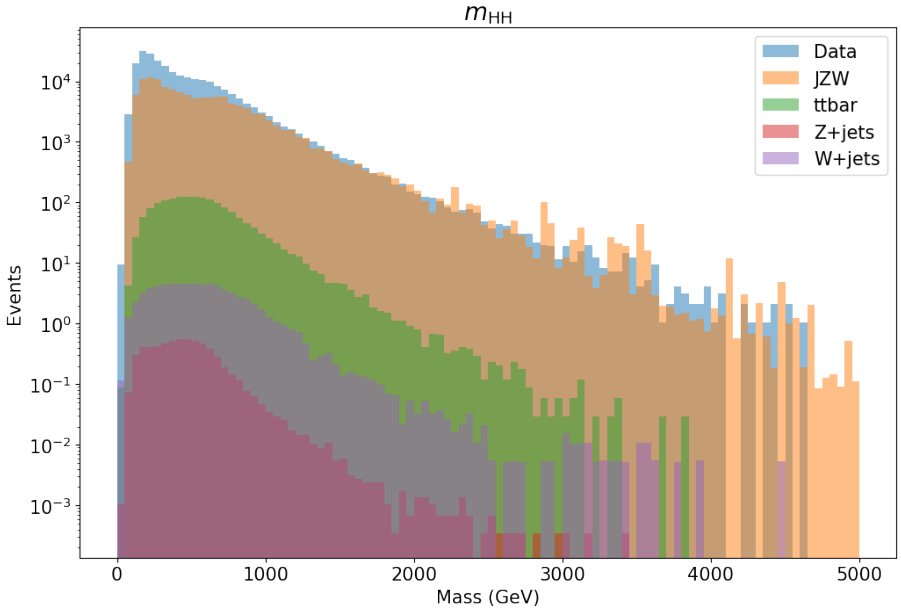


FIGURE 4.20. Distributions of m_{HH} for all background samples including JZW and data sample.

Another check performed with the data sample was the validation on the error of signal region events for the ABCD method in Section 4.6. Adding the data sample to this check, we can see this that the error is at the same scale as that seen in the MC background samples.

4.8. Results

Considering estimates for all background processes, the number of expected events in the signal region for Run-2 is:

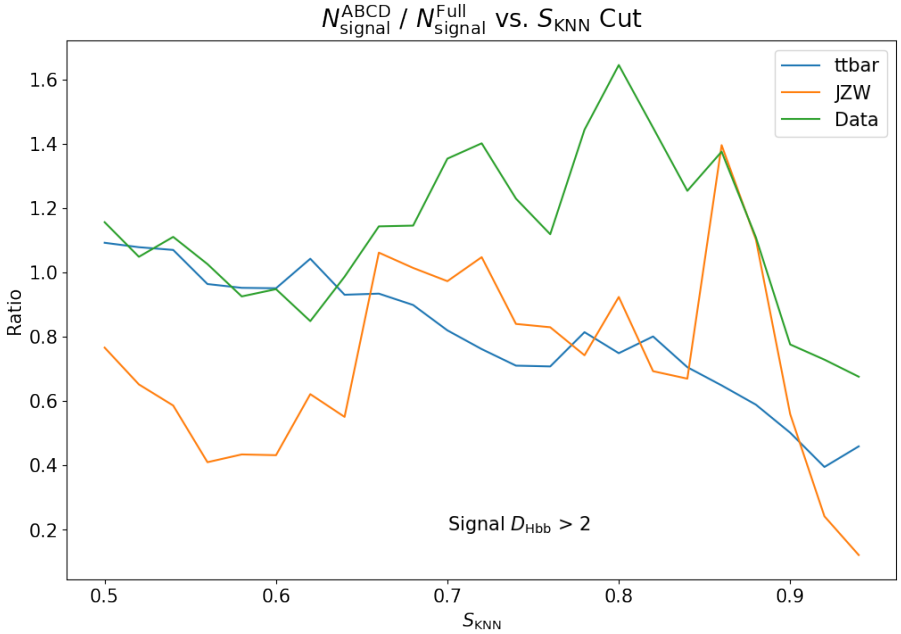


FIGURE 4.21. Ratio of events measured using ABCD and full selection methods in relaxed signal region as a function of S_{KNN} with data sample included.

$$N_{\text{signal}} = 11.47 \pm 3.39\ (\text{stat}) \pm 0.72\ (\text{syst}) \tag{4.9}$$

With the background-only hypothesis, this value can be used to set expected limits on the $X \to HH$ cross section at 95% confidence by calculating how many events would need to be produced by the signal process to exceed that confidence threshold. Needed to recover a cross section from signal region events are the efficiency of selection ϵ, $HH \to b\bar{b}\tau^+\tau^-$ branching ratio, and luminosity L of Run-2.

$$\sigma_{\text{exp}} = \frac{N_{\text{signal}}}{\epsilon \times L \times \text{BR}(HH \to b\bar{b}\tau^+\tau^-[lh])} \tag{4.10}$$

where $\epsilon = 2.30 \times 10^{-2}$%, 5.20%, and 6.54% for 1, 1.6, and 2 TeV mass points, respectively. The luminosity used is 139 fb^{-1} and the branching ratio is 3.27%.

The number of expected signal region events is varied according to its systematic uncertainty, which allows for errors to be placed on the expected cross section. An expected limit is set independently for each X mass point. Limits for all X mass points are shown in Figure 4.22 with linear interpolation performed for mass points between those tested.

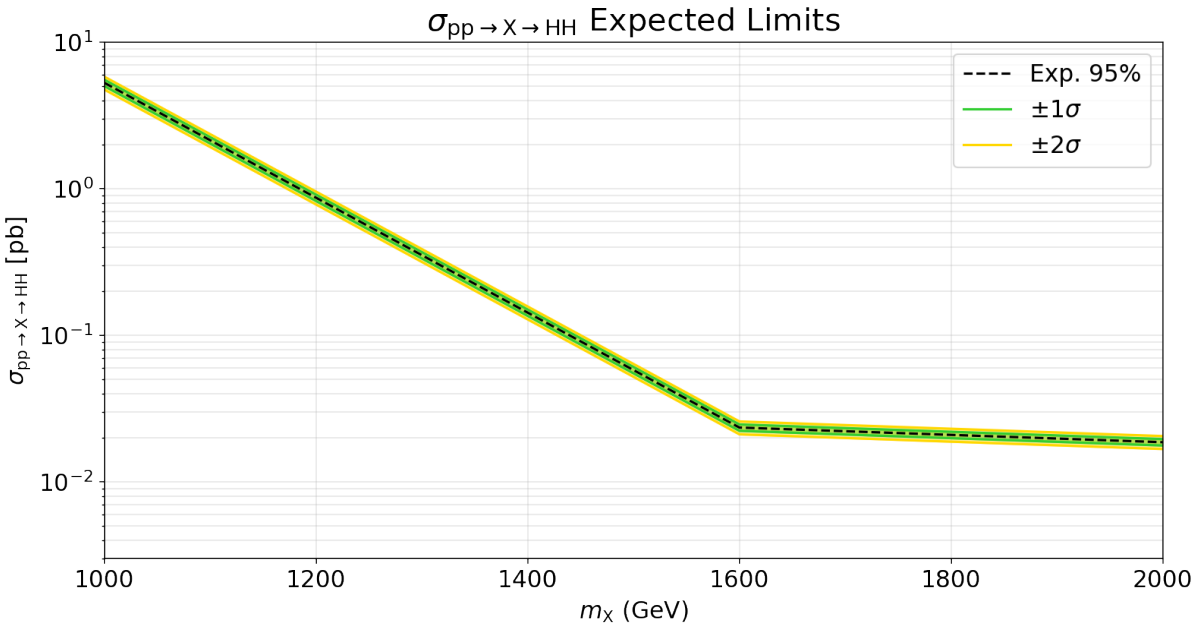


FIGURE 4.22. Expected 95% confidence limits on $\sigma_{pp\to X\to HH}$ for each X mass point with limited systematics included.

Expected limits for the 1, 1.6, and 2 TeV mass points are 5.29×10^3 fb, 23.42 fb, and 18.60 fb, respectively. The 1.6 TeV value for which this analysis was optimized compares to the roughly 60 fb observed limit placed at 1.6 TeV by the resolved lephad analysis as shown in Figure 4.23 [22]. The boosted hadhad analysis set a 49.5 fb observed limit at 1.6 TeV as shown in Figure 4.24 [23].

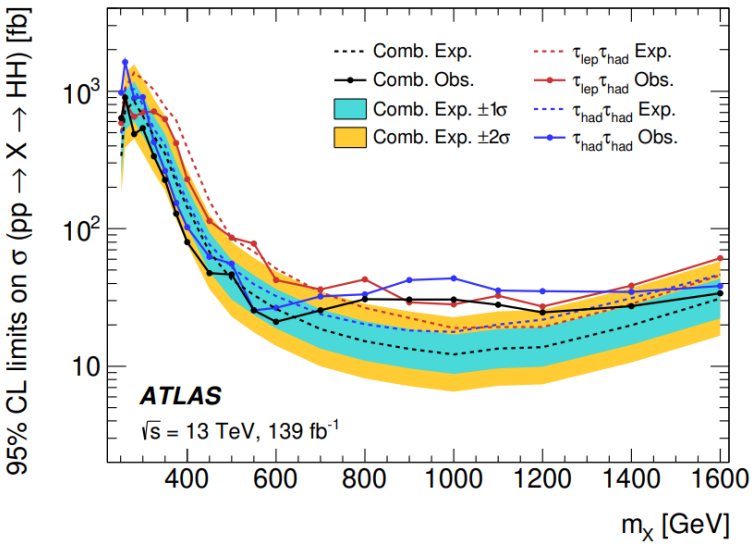


FIGURE 4.23. Observed and expected Run-2 limits on $\sigma_{\mathrm{pp}\to\mathrm{X}\to\mathrm{HH}}$ as a function of X mass from the resolved $b\bar{b}\tau^{+}\tau^{-}$ analysis [22].

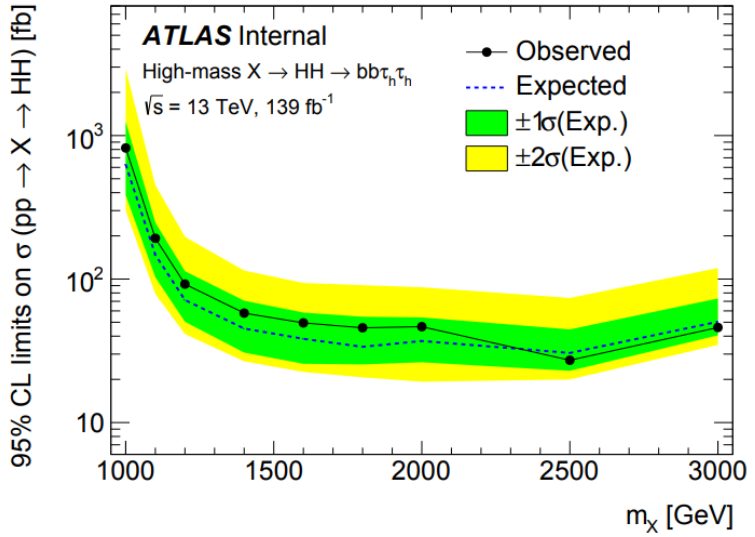


FIGURE 4.24. Observed and expected Run-2 limits on $\sigma_{\mathrm{pp}\to\mathrm{X}\to\mathrm{HH}}$ as a function of X mass from the boosted $b\bar{b}\tau^{+}\tau^{-}$ hadhad analysis [23].

CHAPTER V

CONCLUSION

This dissertation presented the development of an analysis to search for a new heavy scalar X decaying to two Higgs bosons, focusing on the $b\bar{b}\tau^+\tau^-$ semi-hadronic decay channel.

The analysis searches for the resonant production of di-Higgs events created by the decay of the heavy scalar. It was developed using Monte Carlo simulation samples and validated using data with integrated luminosity of 0.11 fb^{-1} collected by the ATLAS detector at the Large Hadron Collider. A new method of identifying semi-hadronic di-τ objects and a kinematic neural network are developed and used in the final selection. Expected limits are set on the $pp \to X \to HH$ cross section for the 1, 1.6, and 2 TeV mass points considered.

Also covered in this dissertation are the development of the Run-3 eFEX τ_{had} trigger, contributions to the ML4Pions project, systematic studies of the signal process for the resolved $b\bar{b}\tau^+\tau^-$ analysis, and Run-3 trigger studies in the lephad and hadhad channels.

APPENDIX A

ML4PIONS

The topocluster calibration as described in Section 3.2 requires many steps to first identify hadronic or EM topoclusters and then adjust their energy accordingly. This section outlines the ML4Pions project [25], which attempts to do this with deep learning, and my contributions to the work. It first gives a general overview of neural networks.

A.1. Neural Networks

This section will include general information about neural networks, including an explanation of dense and convolutional layers which are relevant to both the $b\bar{b}\tau^+\tau^-$ analysis KNN and the ML4Pions work.

A neural network is a machine learning algorithm that can be used to approximate a wide range of classification and regression maps between input and output data of arbitrary dimension. The algorithm learns the relationship between input and output variables through the process of *training*. These neural networks are use a *supervised* learning algorithm, meaning that they train on data that has been labeled with the desired output for a given input.

Neural networks are made up of *nodes* or *neurons* with inputs and outputs that are connected together. Nodes are often arranged in layers of varying configurations, the two types of which to be described here are *dense* layers and *convolutional* layers. Layers in between the input and output layers are called *hidden* layers. The value of a given node is the sum of the values of each input

node multiplied by a weight term plus a bias term. An activation function is also commonly applied to introduce non-linearity into the system.

To calculate the value of the ith node $a_i^{(n+1)}$ in the $n+1$ layer:

$$a_i^{(n+1)} = \sigma\left(\sum_{j=1}^{N} w_{ij}^{(n)} a_j^{(n)} + b_i^{(n+1)}\right) \tag{A.1}$$

where $w_{ij}^{(n)}$ is the weight connecting the jth node in layer n to the ith node in layer $n+1$, $a_j^{(n)}$ is the jth node in layer n, $b_i^{(n+1)}$ is the bias term for the ith node in layer $n+1$, and N is the number of nodes in layer n. σ may be, for example, the sigmoid function:

$$\sigma(x) = \frac{1}{1+e^{-x}} \tag{A.2}$$

Or the ReLU function:

$$\sigma(x) = \begin{cases} 0 & \text{if } x \leq 0 \\ x & \text{if } x > 0 \end{cases} \tag{A.3}$$

Weights and biases are assigned values which are used to propagate information from the inputs of the network to the outputs. The networks considered in this paper are feedforward, meaning this is a direct path without any cycles. Values are randomly initialized and updated during training. The output of a network can be a real value for regression or passed through a function with restricted range like the sigmoid for classification.

Training is the process of aligning the output values of a set of inputs with the provided labels. This is done by altering the values of weights and biases. The measure of how well the network does this is represented by the loss. In the case of

regression, the loss is often the sum of the squared difference between the output value and label for all samples. For classification, the loss may be the binary cross-entropy of predicted outputs and labels.

A learning rule is used to update weights for a given input/output set. There are many ways to do this, but the most commonly used is some form of gradient descent. This calculates the gradient of the loss function with respect to each weight and updates it in the direction that decreases the loss. This is possible because every transformation from the inputs to outputs is differentiable, making the derivative of the loss with respect to each weight calculable.

A dense or fully connected layer of neurons is one in which every input neuron is connected to every output neuron. This does not preserve any structure between the input features.

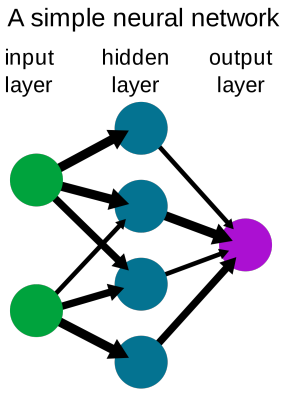


FIGURE A.1. Diagram of a deep neural network with one hidden layer.

A convolutional layer is designed to process data with a 2-dimensional grid structure like images. It convolves smaller filters over the grid that identify characteristics such as edges in a translation invariant way. With multiple layers in sequence, more complex and abstract structures can be identified and used in the learning task.

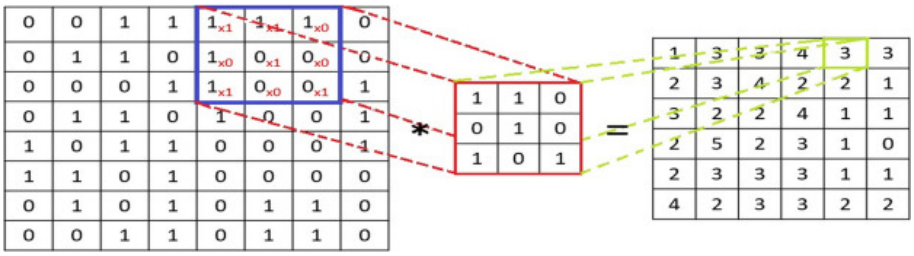


FIGURE A.2. Diagram of a convolutional layer showing the application of a filter to one region of the input layer and the resulting value in the output layer. [24]

A.2. ML4Pions Overview

Pions are ubiquitous in the events captured by the ATLAS detector and the most common particle found in jets. They can either be hadronically showering charged pions π^+ or electromagnetically showering neutral pions π^0. The ability to correctly classify a cluster as originating from a charged or neutral pion and correcting the cluster energy accordingly is an important step to general hadronic calibration for ATLAS.

Topoclusters are generated from ParticleGun simulations of single pions interacting with the ATLAS detector over a wide range of energies. The topoclusters are represented as images of the energy deposited in cells of six layers of the calorimeter. A model trained on these images can potentially exploit differences in shower shape to help it classify the topocluster and determine its true energy. An example of such differences is shown in Figure A.3.

Deep neural networks are the model chosen to perform this task. Several types of model architecture are tested, being constructed with either dense layers, convolutional layers, or both. Models made of dense layers are referred to as DNN and those containing convolutional layers are referred to as CNN. A third architecture, called DenseNet, combines properties of both.

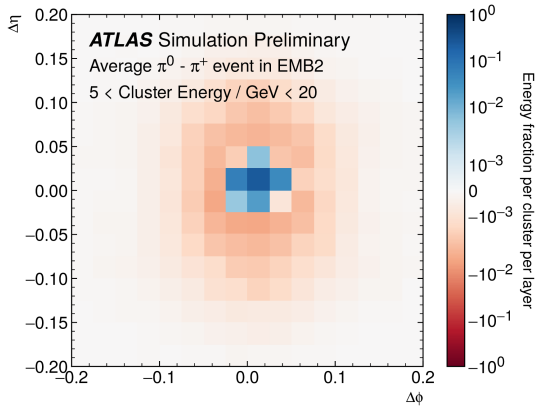
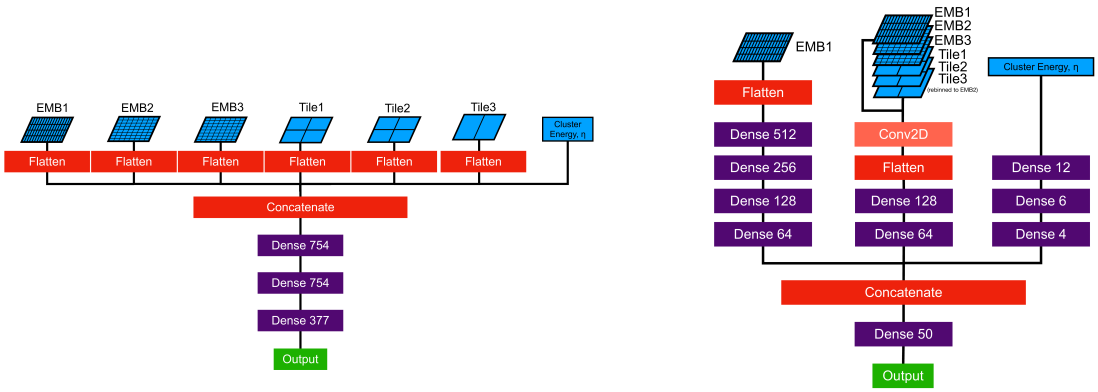


FIGURE A.3. Difference in average energy deposition by π^+ and π^0 in the EMB2 layer of the calorimeter. [25]

For DNN architectures, the topocluster images are flattened before being fed into dense layers. Convolutional layers are intended to take advantage of the image structure of the topoclusters. This allows the different layers of the calorimeter to be treated similarly to the RGB channels of traditional images.

Separate networks are trained for the classification task, regression for hadronic topoclusters, and regression for EM topoclusters. The training label for classification is determined by whether the cluster was created by a π^+ or π^0 and the label for regression is the true energy deposited in the cluster. Two regression networks (a third is not pictured) are shown in Figure A.4.

The trained networks are shown to perform better at both classification and regression than the current LCW method over a wide range of pion energies. The most effective method was found to be the output of the CNN classification model coupled with the DNN energy regression. The median energy response of this method compared with that of the LCW and EM energy scales is shown in Figure A.5.



(a) (b)

FIGURE A.4. DNN (a) and CNN (b) architectures for energy regression. [25]

A.3. Contributions

My work on the ML4Pions project involved the preparation of the image data and exposure of other variables for use in training. I also performed preliminary studies of how networks trained on pions could also be applied to the energy regression of τ_{had} .

A.3.1. Data Preparation

As shown in Figure A.4, the CNN architecture requires that the calorimeter layers be combined into channels of the same image. This requires that the layer images all be the same granularity. However, the cells of the calorimeter are different granularities depending on the layer. In practice, so as not to lose information all layers have their granularity increased to that of the finest layer that they are grouped with. My work resizes layer images while conserving the total energy of the layer. The granularity can only be increased in such a way that

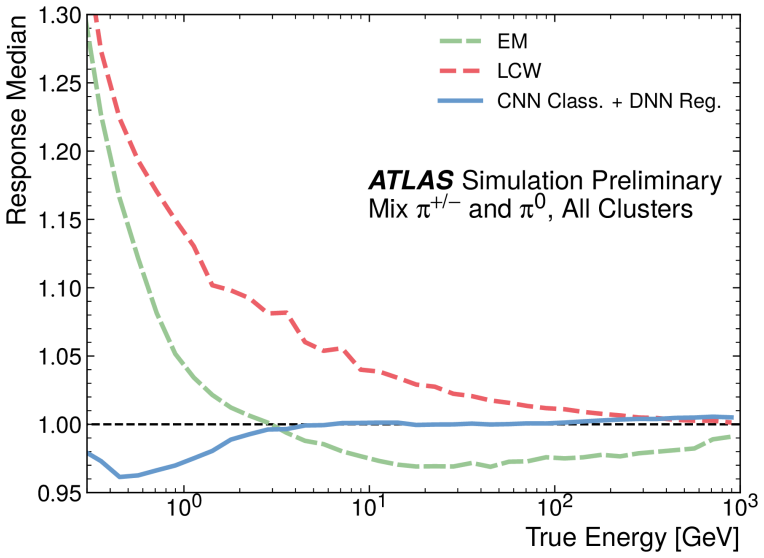


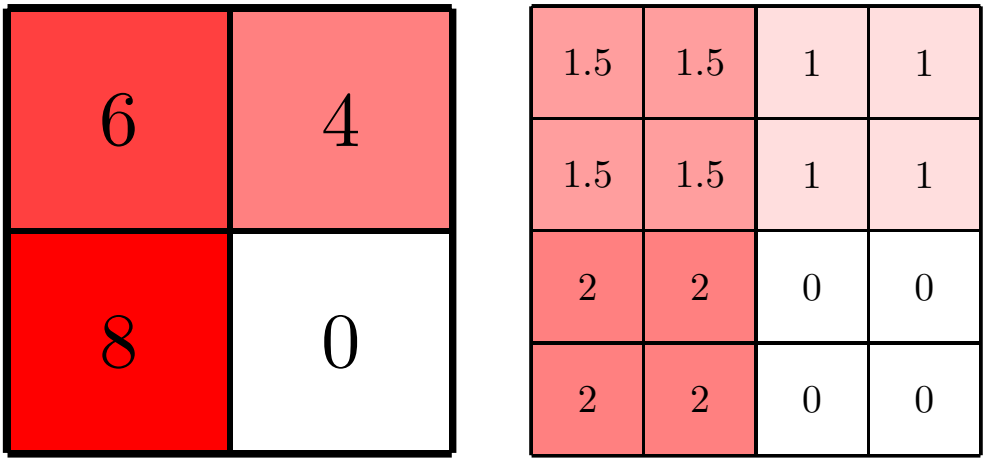
FIGURE A.5. Median energy response for combined CNN classification + DNN regression compared to LCW and EM energy scales. [25]

each cell is replaced with a group of equally-sized cells sharing the same boundary. An example is shown in Figure A.6.

A.3.2. τ_{had} Studies

Because τ_{had} decay only to pions, in theory the deep learning models that correctly calibrate the energy of individual pions may also do so for those produced in a tau decay. The τ_{had} may decay to one or three $\pi^{\pm}$ and zero or more π^0. This transferability was tested by applying the trained models to topoclusters that result from τ_{had} in $\gamma \rightarrow \tau\tau$ events, as no ParticleGun samples of single τ_{had} are available.

The τ_{had} sample contains pile-up which may overlay some of the energy deposits of the pions. It also does not carry the same truth topocluster energy that is used as a regression training label for the pions. To attempt to correct for this, the true visible p_{T} of the τ_{had} is used instead. For this to most closely align



(a) (b)

FIGURE A.6. Topocluster image example before (a) and after (b) resizing to increase the granularity by a factor of two in both dimensions.

with the truth topocluster energy, the τ_{had} are required to decay in the $\tau \rightarrow \pi^{\pm}\nu$ mode and for only a single topocluster to have been created from that $\pi^{\pm}$.

The initial results of applying the pion-trained networks to τ_{had} is shown in Figure A.7. It can be seen that many calibrated topocluster energies are too high (greater than 1). The network is therefore not calibrating the τ_{had} topoclusters properly in the way it does the pion topoclusters.

One reason this might be is that the distribution of topocluster energies between the τ_{had} and pion samples are different. Training on pion samples could then cause the network to focus heavily on an uncommon phase space for τ_{had} , degrading performance elsewhere.

In a simple attempt to remove the largest discrepancy between the distributions, a $E > 1$ GeV cut is placed on the topocluster energy of pion topoclusters before retraining the network on the new set of events. The results of this newly trained network on both pion and τ_{had} topoclusters is shown in Figure A.9.

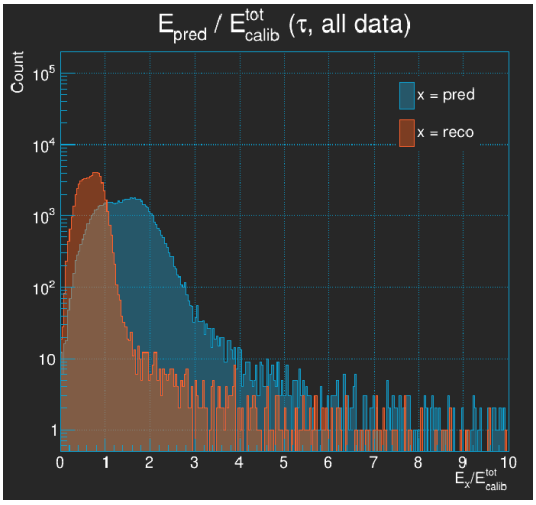


FIGURE A.7. Ratio of τ_{had} topocluster energy to true τ_{had} visible p_{T} before and after application of energy calibration.

The calibration is much more accurate for τ_{had} topoclusters and in line with that of pion topoclusters, with both distributions centered closer to 1 after the calibration is applied. This shows that the pion-trained networks can be applied to τ_{had} topoclusters, with the potential to expand this work to τ_{had} decays producing multiple pions and multiple topoclusters.

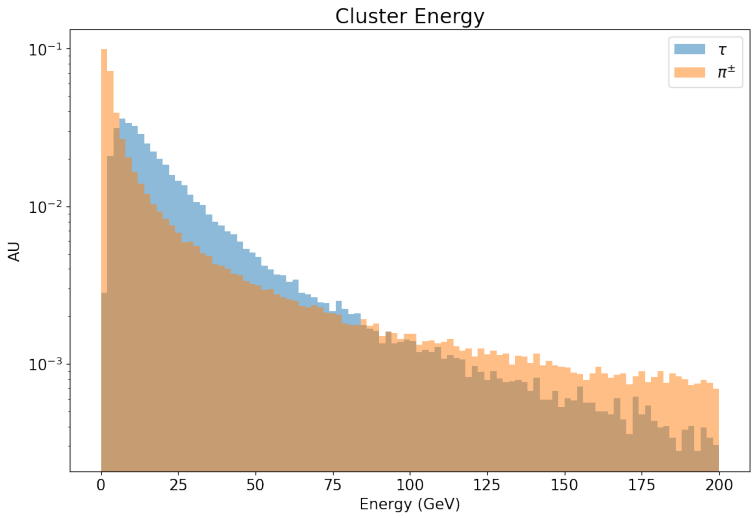
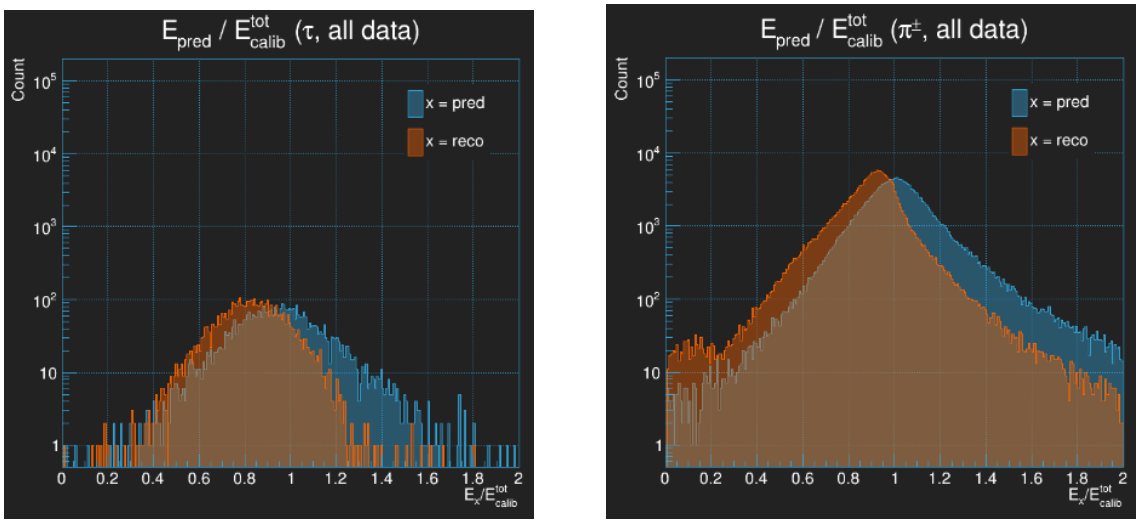


FIGURE A.8. Normalized distributions of topocluster energies for τ_{had} and pion samples.



(a) (b)

FIGURE A.9. Ratio of τ_{had} (a) and pion (b) topocluster energy to truth energy value before and after application of energy calibration.

APPENDIX B

RESOLVED $b\bar{b}\tau^+\tau^-$ SYSTEMATIC STUDIES

I performed systematic uncertainty studies of the signal process of the resolved $b\bar{b}\tau^+\tau^-$ ATLAS analysis. The systematics studies are the parton shower, scale, and PDF uncertainties. These final uncertainties are meant to reflect theory uncertainties on inputs to the sample generation and how those affect the final results of the analysis.

B.1. Parton Shower

Herwig is used for parton showering in the nominal signal samples for this analysis. To study the PS uncertainty, alternate samples are produced at the 280 GeV, 500 GeV, and 1 TeV mass points which use Pythia for parton showering. A normalization uncertainty is then calculated by comparing the acceptance between the two generators and a fit line in the parametric neural network (PNN) output variable is calculated between the two to account for shape effects. This was done for both the SLT and LTT trigger channels. A plot of the 280 GeV PNN score is shown in Figure B.1.

The uncertainty between the nominal and alternative PS generators was greatest in the 280 GeV sample for the SLT channel and the 1 TeV mass point for the LTT channel. Because only events with high PNN score will be selected, the uncertainty was calculated using only the bins containing different percentages of signal events. The final uncertainty was chosen using the last bins with 85% of signal. For mass points not checked, the uncertainty was linearly interpolated between the three chosen mass points and extrapolated to a 251 GeV mass point.

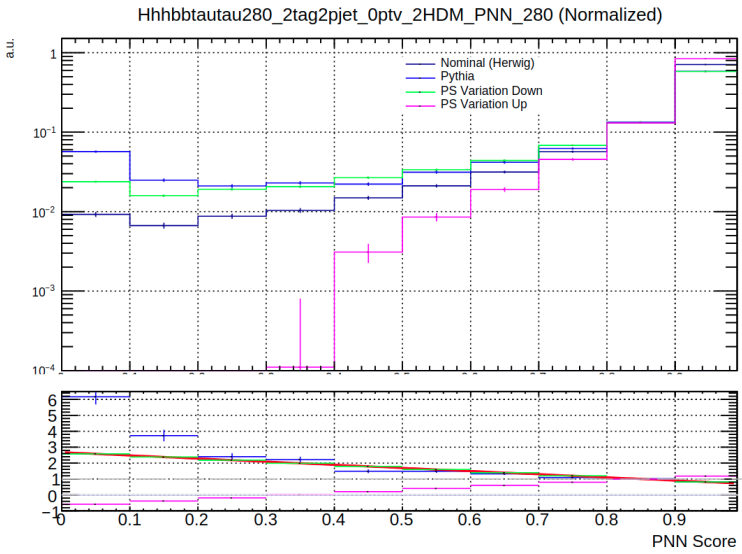


FIGURE B.1. Normalized distribution of PNN score for samples using nominal and alternate PS generators at the 280 GeV mass point in the SLT channel.

B.2. Scale and PDF+α_s

For the scale and PDF uncertainties, full samples with these variations were not available as with the PS samples. Finding the difference in output after running through the analysis framework was therefore not an option. Instead, truth samples were generated at the 500 GeV and 1 TeV mass points which contained alternate event weights for each variation. An approximation of the selection was then applied to the truth objects and the acceptance uncertainty when compared to the nominal weighting was calculated. Shape effects were also searched for in distributions of PNN input variables.

The scale variations were determined by first varying the renormalisation and factorisation scales by factors of two in the following combinations:

$$\{\mu_r, \mu_f\} \times \{0.5, 0.5\}, \{1, 0.5\}, \{0.5, 1\}, \{1, 1\}, \{2, 1\}, \{1, 2\}, \{2, 2\} \tag{B.1}$$

For each of these combinations, the acceptance is calculated as the sum of all accepted event weights divided by the sum of all event weights. An acceptance uncertainty is then found by comparing to the nominal scale values. To check for shape effects, an envelope is created of histograms in the PNN input variables. An example histogram of $m_{\mathrm{HH}}$is shown in Figure B.2.

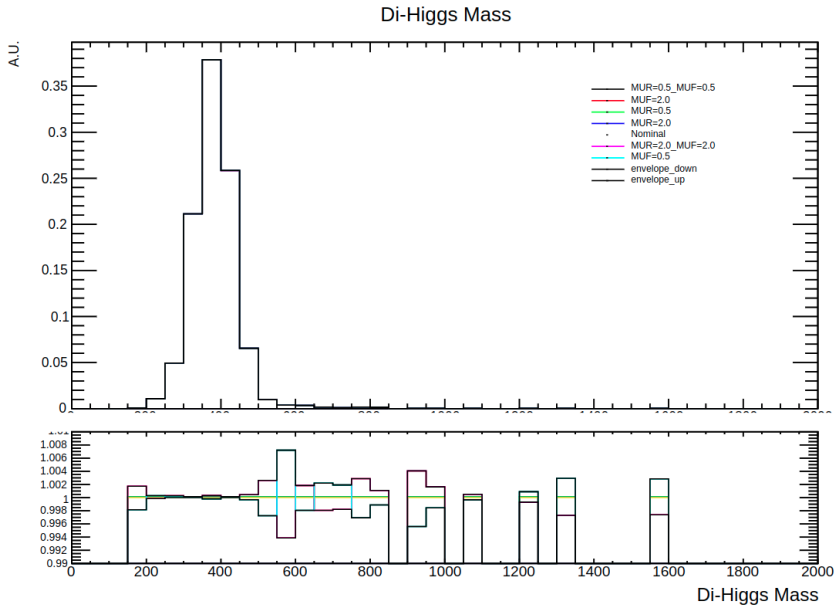


FIGURE B.2. $m_{\mathrm{HH}}$of the 500 GeV signal sample in the SLT channel with scale variations and envelope.

The greatest acceptance uncertainty over mass points and trigger channels was 0.2%, which was applied uniformly to all samples and channels. No shape effect was found in the PNN input variables.

PDF and α_{s} uncertainties are calculated independently and them summed in quadrature to give a single quoted PDF+α_{s} uncertainty. 100 PDF variations are included in the truth samples whose discrepancies with the nominal set are summed to obtain the PDF uncertainty. Variations of α_{s} are performed from the nominal and alternate PDFs and this allows the calculation of the α_{s} uncertainty.

$$\delta^{\mathrm{pdf}}\sigma = \sqrt{\sum_{k=1}^{N_{\mathrm{mem}}} (\sigma^{(k)} - \sigma^{(0)})^2} \tag{B.2}$$

$$\delta^{\alpha_{\mathrm{s}}}\sigma = \frac{\sigma(\alpha_{\mathrm{s}} = 0.1195) - \sigma(\alpha_{\mathrm{s}} = 0.1165)}{2} \tag{B.3}$$

$$\delta^{\mathrm{PDF}+\alpha_{\mathrm{s}}}\delta = \sqrt{(\delta^{\mathrm{pdf}}\sigma)^2 + (\delta^{\alpha_{\mathrm{s}}}\sigma)^2} \tag{B.4}$$

The greatest variation in PDF+α_{s} was seen at the 500 GeV mass point in the LTT channel, shown in Figure B.3. The total uncertainty was 2.28%, dominated by the PDF uncertainty. This was applied to all mass points and trigger channels.

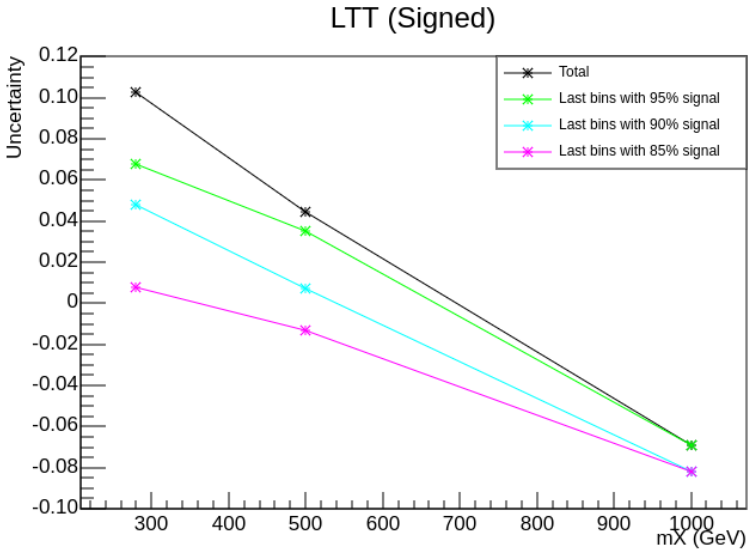


FIGURE B.3. Signed PS uncertainties as a function of the signal sample mass point.

APPENDIX C

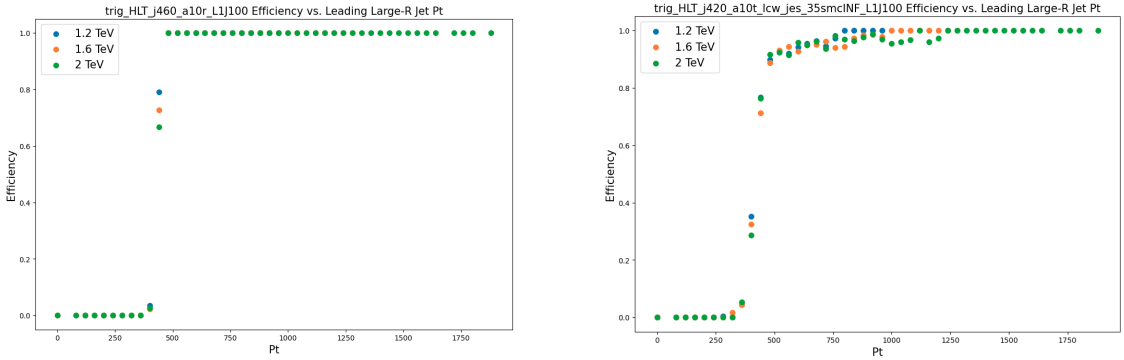
RUN-3 TRIGGER STUDIES

I performed studies on the efficiency of various HLT triggers on boosted lephad and hadhad $b\bar{b}\tau^+\tau^-$ events for Run-3. New triggers are introduced that involve large-R jets and may be useful in the boosted regime. The same 1, 1.6, and 2 TeV mass points are used as for the $b\bar{b}\tau^+\tau^-$ analysis in Chapter IV.

The efficiency was evaluated by creating turn-on curves of trigger efficiency as a function of the leading large-R jet p_{T} of the event and by overall efficiency of events passing the trigger. The efficiency of large-R jets triggers was also tested in combination with lepton triggers.

An example of two triggers tested is shown in Figure C.1. The first requires a large-R jet with $p_{\mathrm{T}} >$ 460 GeV and the second requires a large-R jet with p_{T} $>$ 420 GeV and jet mass $>$ 35 GeV. In both cases the jets are seeded by a 100 GeV jet at Level-1. Lowering the p_{T} requirement may allow for some efficiency at lower values while adding the mass requirement could keep the rate in check. However, the mass cut also removes some events at higher p_{T} which are desirable to preserve.

The efficiency of various single and combined triggers for the lephad (hadhad) channel are shown in Figures C.2, C.3 (C.4). For both channels, the lowering of p_{T} and addition of mass cuts reduced the overall efficiency. Triggers with multiple large-R jets were also tested and performed relatively poorly. In both decay channels, the best results were achieved by OR-ing a large-R jet trigger with those of leptons, $e/\mu/\tau$ in the lephad case with 89% efficiency and τ in the hadhad case with 80% efficiency.



(a) (b)

FIGURE C.1. Turn-on curves for HLT triggers requiring leading large-R jet p_{T} of 460 GeV and 420 GeV + 35 GeV mass cut for 1, 1.6, and 2 TeV signal samples.

Trigger (Lephad)	1.2 TeV Efficiency	1.6 TeV Efficiency	2 TeV Efficiency
L1_J100	0.9652	0.9858	0.992
L1_SC111-CJ15	0.9666	0.9866	0.9908
L1_TAU100	0.9076	0.9582	0.977
L1_TAU60_2TAU40	0.911	0.953	0.9604
HLT_j460_a10r_L1SC111	0.6618	0.8488	0.909
HLT_j420_a10t_lcw_jes_35smcINF_L1SC111	0.6444	0.8166	0.8846
HLT_2j330_a10t_lcw_jes_35smcINF_L1SC111	0.2198	0.3462	0.4208
HLT_e26_lhtight_nod0_ivarloose	0.2912	0.2812	0.2506
HLT_mu26_ivarmedium	0.2904	0.3048	0.3112
HLT_tau160_mediumRNN_tracktwoMVA_L1TAU100	0.5464	0.794	0.8912

FIGURE C.2. Efficiencies of each trigger for all lephad signal samples.

Trigger (Lephad)	1.2 TeV Efficiency	1.6 TeV Efficiency	2 TeV Efficiency
HLT_j460_a10r + HLT_e26	0.745	0.8844	0.929
HLT_j460_a10r + HLT_mu26	0.7614	0.8948	0.9368
HLT_j460_a10r + HLT_tau160	0.7552	0.8902	0.9366
HLT_j460_a10r + HLT_e26 + HLT_mu26	0.8428	0.9294	0.9566
HLT_j460_a10r + HLT_e26 + HLT_mu26 + HLT_tau160	0.8898	0.9486	0.97
HLT_j460_a10r + HLT_j420_10t_lcw_jes_35smcINF	0.7022	0.8688	0.9214

FIGURE C.3. Efficiencies of sets of combined triggers for lephad signal samples. + denotes the OR-ing together of the individual triggers.

Trigger (Hadhad)	1.2 TeV Efficiency	1.6 TeV Efficiency	2 TeV Efficiency
L1_J100	0.9676	0.9868	0.9926
L1_SC111-CJ15	0.9676	0.9862	0.992
L1_TAU100	0.9236	0.9654	0.981
L1_TAU60_2TAU40	0.9538	0.9816	0.9844
HLT_j460_a10r_L1SC111	0.6824	0.8586	0.9124
HLT_j420_a10t_lcw_jes_35smcINF_L1SC111	0.673	0.8458	0.9066
HLT_2j330_a10t_lcw_jes_35smcINF_L1SC111	0.4662	0.6526	0.7536
HLT_j360_a10t_lcw_jes_60smcINF_j360_a10t_lcw_jes_L1SC111	0.448	0.6812	0.8004
HLT_j370_a10t_lcw_jes_35smcINF_j370_a10t_lcw_jes_L1SC111	0.424	0.6656	0.7948
HLT_j460_a10r + HLT_j420_a10t_lcw_jes_36smcINF	0.7184	0.8736	0.9236
HLT_tau160_mediumRNN_tracktwoMVA_L1TAU100	0.6944	0.8612	0.9228
HLT_tau80_medium1_tracktwo_L1TAU60_tau60_medium1_tracktwo_L1TAU40	0.3566	0.6408	0.7896
HLT_j460_a10r + HLT_tau160	0.8082	0.918	0.949

FIGURE C.4. Efficiencies of single and combined triggers for hadhad signal samples. + denotes the OR-ing together of the individual triggers.

APPENDIX D

KINEMATIC NEURAL NETWORK DEVELOPMENT

This appendix includes information about the process of arriving at the final inputs, architecture, and performance of the KNN used in Chapter IV.

D.1. Inputs

The inputs to the KNN are intentionally restricted to kinematic properties of the lepton, τ_{had} , and large-R jet containing the $H \rightarrow b\bar{b}$ products. This calls for the E, p_{T} , η, and ϕ of each to be included. The mass of the large-R jet is also included to help the network identify when the large-R jet is the result of a decaying Higgs. Because the masses of leptons are filled with the same default particle mass instead of calculated from observables, they were not included. The combined di-τ system variables are simply the 4-vector sum of the lepton and τ_{had} inputs already included, plus the di-τ mass as that is dependent on the topology and should relate to the Higgs mass. With arbitrary data and training time, the network would perhaps reconstruct and use these variables without any need to specify them explicitly. Their inclusion is intended then to reduce the amount of training necessary by giving the network access to the variables at the beginning of training.

The only network input that is not a kinematic flag is the e/μ flag. This is included because the data is the combination of two similar but distinct distributions, those of eHad and muHad channel systems. Without the flag, the network would be ignorant of this fact and may therefore mistake features of one channel for that of the other channel. For example, the p_{T} distributions of the

electrons and muons may be different, making it appropriate for the network to treat the channels differently.

D.2. Architecture

The choice of number of layers and number of nodes in each layer were chosen using general heuristics and the effect on performance after training. The network should be large enough to sufficiently learn the differences between signal and background processes. However, a network that is too large is prone to overfitting and can take an unacceptably long time to train. These concerns informed the final architecture of the network being eight layers total with nodes per layer of $N \times 2N \times 3N \times 3N \times 3N \times 3N \times 2N \times 1$ where N is the number of input features.

Dropout is a process that is intended to prevent overfitting during training. It masks certain nodes at random during training in order to force the network to adapt without them. This prevents the outputs from becoming too dependent on any given node, increasing the robustness of the network as a whole. When a discrepancy between train and test loss was noticed during training, dropout was added in between all layers of the network to prevent this. However, this caused strange clustering in the final S_{KNN} distribution. When the dropout was reduced to only being applied in a single layer, the clustering disappeared. The reduced dropout is used in the final network architecture.

The final network was trained over all signal samples along with $t\bar{t}$ and Z+jets samples. Because signal samples contribute events with truth label 0 due to non-truth-matched lepton systems and incorrect lepton channels, it was studied whether only including signal samples in training would be sufficient. Comparisons

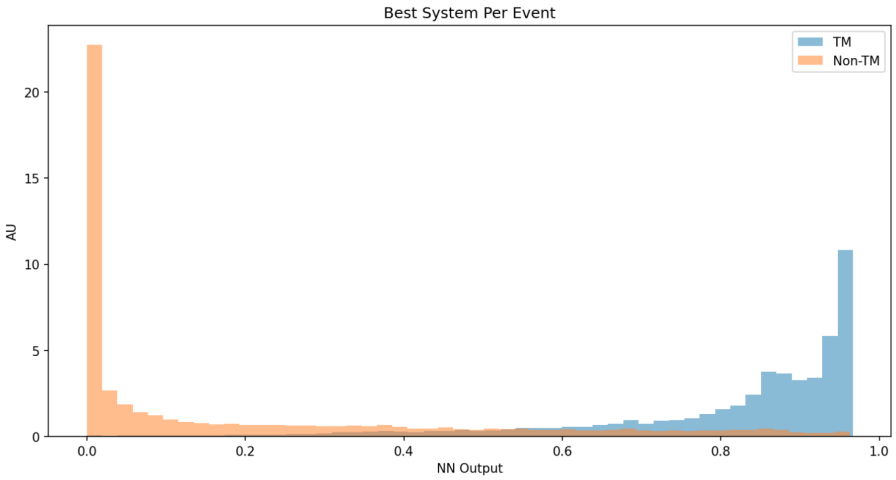
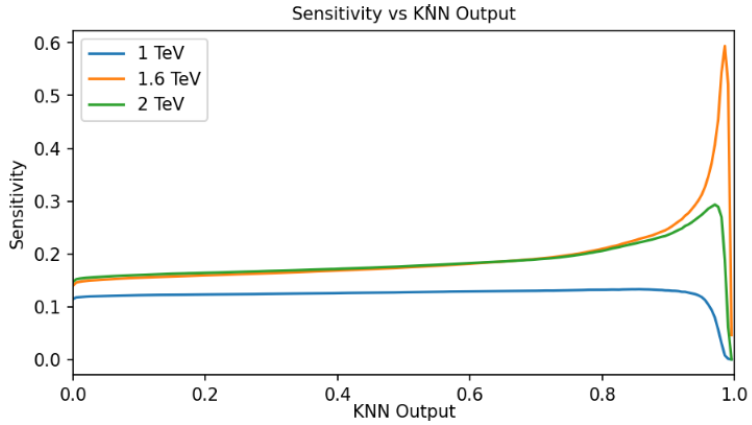


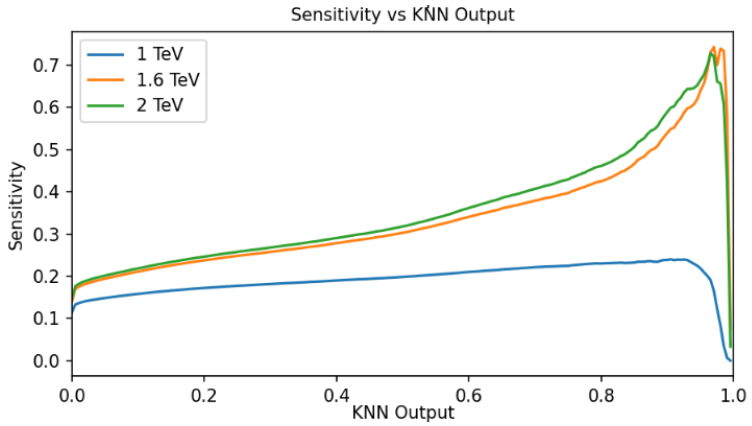
FIGURE D.1. Distribution of S_{KNN} variable with dropout applied between all layers.

of signal-only performance to signal+background performance are shown in Figure D.2. The network trained only with signal events does classify fairly well, particularly the 1.6 TeV signal events. However, adding in the background events to training improves the performance on 2 TeV events significantly and also slightly improves the 1 and 1.6 TeV events as measured by maximum sensitivity achieved.

After it was determined that the JZW background was significant in the signal region, it was tested whether the background sample events should also be added to the training of the network. This was attempted with results shown in Figure D.3. It is interesting to note that with JZW included in training, the 2 TeV signal events are more likely to be classified as background similarly to 1 TeV. Without JZW, they are classified as signal. This may be because highly boosted Higgs decay products in 2 TeV samples look similar to jets. The less boosted 1.6 TeV events may still retain distinctive properties that allow the network to distinguish them from jets.



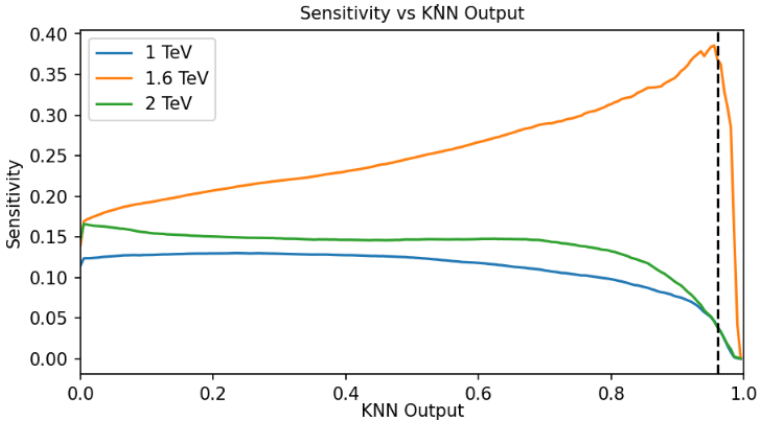
(a)



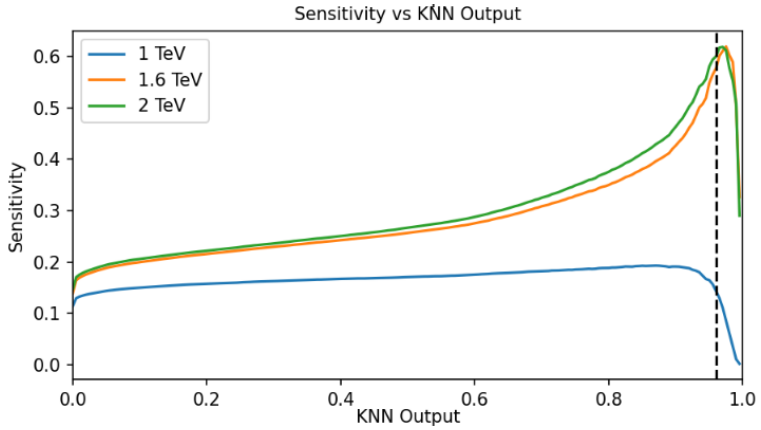
(b)

FIGURE D.2. Sensitivity as a function of S_{KNN} for network trained with only signal events (a) and with signal, $t\bar{t}$, and Z+jets events (b). Backgrounds used in sensitivity calculation are $t\bar{t}$ and Z+jets.

As can be seen in plots of the performance metrics over time, toward the end of training the network continues to improve on training data while the performance on test data stays constant. This can be considered a mild form of overfitting, although a more common sign is that the performance on test samples decreases as performance on training samples continues to improve. The observed behavior motivates the use of 2-fold validation during training to ensure that results do not depend on applying a trained network to its own training samples.



(a)



(b)

FIGURE D.3. Sensitivity as a function of S_{KNN} for network trained with (a) and without (b) JZW events included in the training set. Backgrounds used in sensitivity calculation are $t\bar{t}$ and Z+jets.

As expected, performance decreased somewhat using the 2-fold training compared to training a single network with an 80/20 train/test split.

REFERENCES CITED

[1] Jannicke Pearkes. About the atlas inner tracker, 2022. URL https://www.symmetrymagazine.org/sites/default/files/images/hi-res/discoball_poster_sized.pdf.

[2] R. L. Workman and Others. Review of Particle Physics. *PTEP*, 2022:083C01, 2022. doi:10.1093/ptep/ptac097.

[3] Renilde Vanden Broeck. THE CERN ACCELERATOR COMPLEX. Complexe des accelerateurs du CERN. 2019. URL https://cds.cern.ch/record/2693837.

[4] L Arnaudon, P Baudrenghien, M Baylac, et al. Linac4 Technical Design Report. Technical report, CERN, Geneva, 2006. URL https://cds.cern.ch/record/1004186. revised version submitted on 2006-12-14 09:00:40.

[5] AC Team. Diagram of an LHC dipole magnet. Schéma d'un aimant dipôle du LHC. 1999. URL https://cds.cern.ch/record/40524.

[6] The ATLAS Collaboration, 2022. URL https://atlas.web.cern.ch/Atlas/GROUPS/DATAPREPARATION/PublicPlots/2022/DataSummary/figs/mu_run123.png.

[7] The ATLAS Collaboration. The ATLAS Experiment at the CERN Large Hadron Collider: A Description of the Detector Configuration for Run 3. In preparation.

[8] The ATLAS Collaboration, G Aad, E Abat, et al. The atlas experiment at the cern large hadron collider. *Journal of Instrumentation*, 3(08):S08003, aug 2008. doi:10.1088/1748-0221/3/08/S08003. URL https://dx.doi.org/10.1088/1748-0221/3/08/S08003.

[9] *ATLAS liquid-argon calorimeter: Technical Design Report*. Technical design report. ATLAS. CERN, Geneva, 1996. doi:10.17181/CERN.FWRW.FOOQ. URL https://cds.cern.ch/record/331061.

[10] Georges Aad et al. Technical Design Report for the Phase-I Upgrade of the ATLAS TDAQ System. 9 2013.

[11] Aranzazu Ruiz-Martinez, Javier Montejo Berlingen, Daniele Zanzi, Teng Jian Khoo, Mark Andrew Owen, and Savanna Shaw. Run 3 trigger menu design. Technical report, CERN, Geneva, 2019. URL https://cds.cern.ch/record/2683881.

[12] Stefan Höche. Introduction to parton-shower event generators, 2014. URL https://arxiv.org/abs/1411.4085.

[13] Joao Pequenao. Event Cross Section in a computer generated image of the ATLAS detector. 2008. URL https://cds.cern.ch/record/1096081.

[14] M. Aaboud et al. Electron reconstruction and identification in the ATLAS experiment using the 2015 and 2016 LHC proton–proton collision data at $$\sqrt{s} = 13$$ $$\text {TeV}$$. *The European Physical Journal C*, 79 (8), aug 2019. doi:10.1140/epjc/s10052-019-7140-6. URL https://doi.org/10.1140%2Fepjc%2Fs10052-019-7140-6.

[15] Georges Aad et al. Muon reconstruction and identification efficiency in ATLAS using the full Run 2 *pp* collision data set at $\sqrt{s} = 13$ TeV. *Eur. Phys. J., C*, 81:578, 2021. doi:10.1140/epjc/s10052-021-09233-2. URL https://cds.cern.ch/record/2746302. 64 pages in total, author list starting page 42, auxiliary material starting at page 59, 34 figures, 3 tables. All figures including auxiliary figures are available at https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PAPERS/MUON-2018-03/.

[16] G. Aad et al. Topological cell clustering in the ATLAS calorimeters and its performance in LHC run 1. *The European Physical Journal C*, 77(7), jul 2017. doi:10.1140/epjc/s10052-017-5004-5. URL https://doi.org/10.1140%2Fepjc%2Fs10052-017-5004-5.

[17] Matteo Cacciari, Gavin P Salam, and Gregory Soyez. The anti-k$_t$ jet clustering algorithm. *Journal of High Energy Physics*, 2008(04):063–063, apr 2008. doi:10.1088/1126-6708/2008/04/063. URL https://doi.org/10.1088%2F1126-6708%2F2008%2F04%2F063.

[18] M. Aaboud et al. Jet energy scale measurements and their systematic uncertainties in proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector. *Physical Review D*, 96(7), oct 2017. doi:10.1103/physrevd.96.072002. URL https://doi.org/10.1103%2Fphysrevd.96.072002.

[19] G. Aad et al. Performance of jet substructure techniques for large-r jets in proton-proton collisions at $ \sqrt{s}=7 $ TeV using the ATLAS detector. *Journal of High Energy Physics*, 2013(9), sep 2013. doi:10.1007/jhep09(2013)076. URL https://doi.org/10.1007%2Fjhep09%282013%29076.

[20] G. Aad et al. ATLAS b-jet identification performance and efficiency measurement with $$t{\bar{t}}$$ events in pp collisions at $$\sqrt{s}=13$$ TeV. *The European Physical Journal C*, 79(11), nov 2019. doi:10.1140/epjc/s10052-019-7450-8. URL https://doi.org/10.1140%2Fepjc%2Fs10052-019-7450-8.

[21] Reconstruction, Energy Calibration, and Identification of Hadronically Decaying Tau Leptons in the ATLAS Experiment for Run-2 of the LHC. Technical report, CERN, Geneva, 2015. URL https://cds.cern.ch/record/2064383. All figures including auxiliary figures are available at https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PUBNOTES/ATL-PHYS-PUB-2015-045.

[22] ATLAS Collaboration. Search for resonant and non-resonant higgs boson pair production in the $b\bar{b}\tau^+\tau^-$ decay channel using 13 tev *pp* collision data from the atlas detector, 2022.

[23] Georges Aad et al. Reconstruction and identification of boosted di-τ systems in a search for Higgs boson pairs using 13 TeV proton−proton collision data in ATLAS. *JHEP*, 2011:163, 2020. doi:10.1007/JHEP11(2020)163. URL https://cds.cern.ch/record/2725874. 46 pages in total, author list starting page 30, 13 figures, 3 tables, published in JHEP. All figures including auxiliary figures are available at https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PAPERS/HDBS-2019-22.

[24] Gustavo Carneiro, Jacinto Nascimento, and Andrew P. Bradley. Chapter 14 - deep learning models for classifying mammogram exams containing unregistered multi-view images and segmentation maps of lesions11this work is an extension of the paper published by the same authors at the medical image computing and computer-assisted intervention (miccai 2015) [1]. In S. Kevin Zhou, Hayit Greenspan, and Dinggang Shen, editors, *Deep Learning for Medical Image Analysis*, pages 321–339. Academic Press, 2017. ISBN 978-0-12-810408-8. doi:https://doi.org/10.1016/B978-0-12-810408-8.00019-5. URL https://www.sciencedirect.com/science/article/pii/B9780128104088000195.

[25] Deep Learning for Pion Identification and Energy Calibration with the ATLAS Detector. Technical report, CERN, Geneva, 2020. URL https://cds.cern.ch/record/2724632. All figures including auxiliary figures are available at https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PUBNOTES/ATL-PHYS-PUB-2020-018.

[26] John Baez. How many fundamental constants are there?, 2011. URL https://math.ucr.edu/home/baez/constants.html.

[27] Michael E. Peskin and Daniel V. Schroeder. *An Introduction to quantum field theory*. Addison-Wesley, Reading, USA, 1995. ISBN 978-0-201-50397-5.

[28] Peter W. Higgs. Broken symmetries and the masses of gauge bosons. *Phys. Rev. Lett.*, 13:508–509, Oct 1964. doi:10.1103/PhysRevLett.13.508. URL https://link.aps.org/doi/10.1103/PhysRevLett.13.508.

[29] ATLAS Collaboration. Observation of a new particle in the search for the standard model higgs boson with the atlas detector at the lhc. *Physics Letters B*, 716(1):1–29, 2012. ISSN 0370-2693. doi:https://doi.org/10.1016/j.physletb.2012.08.020. URL https://www.sciencedirect.com/science/article/pii/S037026931200857X.

[30] CMS Collaboration. Observation of a new boson at a mass of 125 gev with the cms experiment at the lhc. *Physics Letters B*, 716(1):30–61, 2012. ISSN 0370-2693. doi:https://doi.org/10.1016/j.physletb.2012.08.021. URL https://www.sciencedirect.com/science/article/pii/S0370269312008581.

[31] F. Englert and R. Brout. Broken symmetry and the mass of gauge vector mesons. *Phys. Rev. Lett.*, 13:321–323, Aug 1964. doi:10.1103/PhysRevLett.13.321. URL https://link.aps.org/doi/10.1103/PhysRevLett.13.321.

[32] Vera C. Rubin, W. Kent Ford, Jr., and Norbert Thonnard. Extended rotation curves of high-luminosity spiral galaxies. IV. Systematic dynamical properties, Sa through Sc. *Astrophys. J. Lett.*, 225:L107–L111, 1978. doi:10.1086/182804.

[33] Douglas Clowe, Maruš a Bradač, Anthony H. Gonzalez, Maxim Markevitch, Scott W. Randall, Christine Jones, and Dennis Zaritsky. A direct empirical proof of the existence of dark matter. *The Astrophysical Journal*, 648(2): L109–L113, aug 2006. doi:10.1086/508162. URL https://doi.org/10.1086%2F508162.

[34] A. D. Sakharov. Violation of CP Invariance, C asymmetry, and baryon asymmetry of the universe. *Pisma Zh. Eksp. Teor. Fiz.*, 5:32–35, 1967. doi:10.1070/PU1991v034n05ABEH002497.

[35] Mark Trodden. Electroweak baryogenesis: A brief review, 1998. URL https://arxiv.org/abs/hep-ph/9805252.

[36] G.C. Branco, P.M. Ferreira, L. Lavoura, M.N. Rebelo, Marc Sher, and João P. Silva. Theory and phenomenology of two-higgs-doublet models. *Physics Reports*, 516(1-2):1–102, jul 2012. doi:10.1016/j.physrep.2012.02.002. URL https://doi.org/10.1016%2Fj.physrep.2012.02.002.

[37] Kaustubh Agashe, Hooman Davoudiasl, Gilad Perez, and Amarjit Soni. Warped gravitons at the CERN LHC and beyond. *Physical Review D*, 76(3), aug 2007. doi:10.1103/physrevd.76.036006. URL https://doi.org/10.1103%2Fphysrevd.76.036006.

[38] Georges Aad et al. Search for non-resonant Higgs boson pair production in the $bb\ell\nu\ell\nu$ final state with the ATLAS detector in pp collisions at $\sqrt{s} = 13$ TeV. *Phys. Lett. B*, 801:135145, 2020. doi:10.1016/j.physletb.2019.135145.

[39] Search for nonresonant pair production of Higgs bosons in the $b\bar{b}b\bar{b}$ final state in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector. 1 2023.

[40] Georges Aad et al. Search for resonant pair production of Higgs bosons in the $b\bar{b}b\bar{b}$ final state using pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector. *Phys. Rev. D*, 105(9):092002, 2022. doi:10.1103/PhysRevD.105.092002.

[41] Georges Aad et al. Search for Higgs boson pair production in the two bottom quarks plus two photons final state in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector. *Phys. Rev. D*, 106(5):052001, 2022. doi:10.1103/PhysRevD.106.052001.

[42] S Chatrchyan et al. The CMS experiment at the CERN LHC. The Compact Muon Solenoid experiment. *JINST*, 3:S08004, 2008. doi:10.1088/1748-0221/3/08/S08004. URL https://cds.cern.ch/record/1129810. Also published by CERN Geneva in 2010.

[43] The ALICE Collaboration et al. The ALICE experiment at the CERN LHC. *Journal of Instrumentation*, 3(08):S08002, aug 2008. doi:10.1088/1748-0221/3/08/S08002. URL https://dx.doi.org/10.1088/1748-0221/3/08/S08002.

[44] A Augusto Alves et al. The LHCb Detector at the LHC. *JINST*, 3:S08005, 2008. doi:10.1088/1748-0221/3/08/S08005. URL https://cds.cern.ch/record/1129809. Also published by CERN Geneva in 2010.

[45] Jacques Gareyte. LHC main parameters. *Part. Accel.*, 50:61–68, 1995. URL https://cds.cern.ch/record/304825.

[46] CERN. Lhc run 3: physics at record energy starts tomorrow, 2022. URL https://atlas.cern/Updates/Press-Statement/LHC-Run3-Starts.

[47] Oliver Sim Brüning, Paul Collier, P Lebrun, Stephen Myers, Ranko Ostojic, John Poole, and Paul Proudlock. *LHC Design Report*. CERN Yellow Reports: Monographs. CERN, Geneva, 2004. doi:10.5170/CERN-2004-003-V-1. URL https://cds.cern.ch/record/782076.

[48] D Tommasini. The PS Booster, PS and SPS Magnets for the next 25 years. This note provides information and analysis on the present status of the magnets installed in the CERN Proton Synchrotron Booster (PSB), the Proton Synchrotron (PS) and the Super Proton Synchrotron (SPS) in view of their possible operation for the next 25 years. 2010. URL https://cds.cern.ch/record/1233948.

[49] Stephane Fartoukh, Sofia Kostoglou, Matteo Solfaroli Camillocci, et al. LHC Configuration and Operational Scenario for Run 3. Technical report, CERN, Geneva, 2021. URL https://cds.cern.ch/record/2790409.

[50] M. Aaboud, G. Aad, B. Abbott, et al. Measurement of the inelastic proton-proton cross section at $\sqrt{s} = 13$ TeV with the atlas detector at the lhc. *Phys. Rev. Lett.*, 117:182002, Oct 2016. doi:10.1103/PhysRevLett.117.182002. URL https://link.aps.org/doi/10.1103/PhysRevLett.117.182002.

[51] E. Daw. Lecture 7 - rapidity and pseudorapidity, 2012. URL https://www.hep.shef.ac.uk/edaw/PHY206/Site/2012_course_files/phy206rlec7.pdf.

[52] *ATLAS inner detector: Technical Design Report, 1*. Technical design report. ATLAS. CERN, Geneva, 1997. URL https://cds.cern.ch/record/331063.

[53] G Aad, M Ackers, F A Alberti, et al. ATLAS pixel detector electronics and sensors. *JINST*, 3:P07007, 2008. doi:10.1088/1748-0221/3/07/P07007. URL https://cds.cern.ch/record/1119279.

[54] M Moreno Llacer. ATLAS SemiConductor Tracker: operation and performance. Technical report, CERN, Geneva, 2011. URL https://cds.cern.ch/record/1399080.

[55] Vasiliki A. Mitsou. The ATLAS transition radiation tracker. In *8th International Conference on Advanced Technology and Particle Physics (ICATPP 2003): Astroparticle, Particle, Space Physics, Detectors and Medical Physics Applications*, pages 497–501, 11 2003. doi:10.1142/9789812702708_0073.

[56] M Capeans, G Darbo, K Einsweiller, M Elsing, T Flick, M Garcia-Sciveres, C Gemme, H Pernegger, O Rohne, and R Vuillermet. ATLAS Insertable B-Layer Technical Design Report. Technical report, 2010. URL https://cds.cern.ch/record/1291633.

[57] A Yamamoto, Y Makida, R Ruber, et al. The ATLAS central solenoid. *Nucl. Instrum. Methods Phys. Res., A*, 584(1):53–74, 2008. doi:10.1016/j.nima.2007.09.047. URL https://cds.cern.ch/record/1069672.

[58] *ATLAS tile calorimeter: Technical Design Report*. Technical design report. ATLAS. CERN, Geneva, 1996. doi:10.17181/CERN.JRBJ.7O28. URL https://cds.cern.ch/record/331062.

[59] *ATLAS muon spectrometer: Technical Design Report*. Technical design report. ATLAS. CERN, Geneva, 1997. URL https://cds.cern.ch/record/331068.

[60] T Kawamoto, S Vlachos, et al. New Small Wheel Technical Design Report. Technical report, 2013. URL https://cds.cern.ch/record/1552862. ATLAS New Small Wheel Technical Design Report.

[61] ATLAS level-1 trigger: Technical Design Report. 6 1998.

[62] W Buttinger. The ATLAS Level-1 Trigger System. Technical report, CERN, Geneva, 2012. URL https://cds.cern.ch/record/1456546.

[63] Federico Meloni. Primary vertex reconstruction with the ATLAS detector. Technical Report 12, CERN, Geneva, 2016. URL https://cds.cern.ch/record/2222390.

[64] Nicoletta Garelli. Performance of the ATLAS Detector in Run-2. *EPJ Web Conf.*, 164:01021, 2017. doi:10.1051/epjconf/201716401021. URL https://cds.cern.ch/record/2310929.

[65] G. Aad et al. Electron and photon energy calibration with the ATLAS detector using LHC run 1 data. *The European Physical Journal C*, 74(10), oct 2014. doi:10.1140/epjc/s10052-014-3071-4. URL https://doi.org/10.1140%2Fepjc%2Fs10052-014-3071-4.

[66] David Krohn, Jesse Thaler, and Lian-Tao Wang. Jet trimming. *Journal of High Energy Physics*, 2010(2), feb 2010. doi:10.1007/jhep02(2010)084. URL https://doi.org/10.1007%2Fjhep02%282010%29084.

[67] Optimisation and performance studies of the ATLAS b-tagging algorithms for the 2017-18 LHC run. Technical report, CERN, Geneva, 2017. URL https://cds.cern.ch/record/2273281. All figures including auxiliary figures are available at https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PUBNOTES/ATL-PHYS-PUB-2017-013.

[68] David Krohn, Jesse Thaler, and Lian-Tao Wang. Jets with variable r. *Journal of High Energy Physics*, 2009(06):059–059, jun 2009. doi:10.1088/1126-6708/2009/06/059. URL https://doi.org/10.1088%2F1126-6708%2F2009%2F06%2F059.

[69] Measurement of the tau lepton reconstruction and identification performance in the ATLAS experiment using pp collisions at $\sqrt{s} = 13$ TeV. Technical report, CERN, Geneva, 2017. URL https://cds.cern.ch/record/2261772. All figures including auxiliary figures are available at https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/CONFNOTES/ATLAS-CONF-2017-029.

[70] William Keaton Balunas et al. Combination of searches for Higgs boson pairs in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector supporting note. Technical report, CERN, Geneva, 2021. URL https://cds.cern.ch/record/2775165.

[71] Thomas H. Kittelmann, Vakhtang Tsulaia, Joseph Boudreau, and Edward Moyse. The virtual point 1 event display for the ATLAS experiment. *J. Phys. Conf. Ser.*, 219:032012, 2010. doi:10.1088/1742-6596/219/3/032012.

[72] J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, O. Mattelaer, H.-S. Shao, T. Stelzer, P. Torrielli, and M. Zaro. The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations. *Journal of High Energy Physics*, 2014(7), jul 2014. doi:10.1007/jhep07(2014)079. URL https://doi.org/10.1007%2Fjhep07%282014%29079.

[73] Johan Alwall, Michel Herquet, Fabio Maltoni, Olivier Mattelaer, and Tim Stelzer. MadGraph 5: going beyond. *Journal of High Energy Physics*, 2011 (6), jun 2011. doi:10.1007/jhep06(2011)128. URL https://doi.org/10.1007%2Fjhep06%282011%29128.

[74] Torbjörn Sjöstrand, Stefan Ask, Jesper R. Christiansen, Richard Corke, Nishita Desai, Philip Ilten, Stephen Mrenna, Stefan Prestel, Christine O. Rasmussen, and Peter Z. Skands. An introduction to PYTHIA 8.2. *Computer Physics Communications*, 191:159–177, jun 2015. doi:10.1016/j.cpc.2015.01.024. URL https://doi.org/10.1016%2Fj.cpc.2015.01.024.

[75] Johannes Bellm, Stefan Gieseke, David Grellscheid, Simon Plätzer, Michael Rauch, Christian Reuschle, Peter Richardson, Peter Schichtel, Michael H. Seymour, Andrzej Siódmok, Alexandra Wilcock, Nadine Fischer, Marco A. Harrendorf, Graeme Nail, Andreas Papaefstathiou, and Daniel Rauch. Herwig 7.0/herwig++ 3.0 release note. *The European Physical Journal C*, 76 (4), apr 2016. doi:10.1140/epjc/s10052-016-4018-8. URL https://doi.org/10.1140%2Fepjc%2Fs10052-016-4018-8.

[76] ATLAS Pythia 8 tunes to 7 TeV data. Technical report, CERN, Geneva, 2014. URL https://cds.cern.ch/record/1966419. All figures including auxiliary figures are available at https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PUBNOTES/ATL-PHYS-PUB-2014-021.

[77] Stefan Gieseke, Christian Röhr, and Andrzej Siódmok. Colour reconnections in herwig++. *The European Physical Journal C*, 72(11), nov 2012. doi:10.1140/epjc/s10052-012-2225-5. URL https://doi.org/10.1140%2Fepjc%2Fs10052-012-2225-5.

[78] D. J. Lange. The EvtGen particle decay simulation package. *Nucl. Instrum. Meth. A*, 462:152–155, 2001. doi:10.1016/S0168-9002(01)00089-4.

[79] Enrico Bothmann, Gurpreet Singh Chahal, Stefan Höche, Johannes Krause, Frank Krauss, Silvan Kuttimalai, Sebastian Liebschner, Davide Napoletano, Marek Schönherr, Holger Schulz, Steffen Schumann, and Frank Siegert. Event generation with sherpa 2.2. *SciPost Physics*, 7(3), sep 2019. doi:10.21468/scipostphys.7.3.034. URL https://doi.org/10.21468%2Fscipostphys.7.3.034.

[80] Richard D. Ball, , Valerio Bertone, Stefano Carrazza, Christopher S. Deans, Luigi Del Debbio, Stefano Forte, Alberto Guffanti, Nathan P. Hartland, José I. Latorre, Juan Rojo, and Maria Ubiali. Parton distributions for the LHC run II. *Journal of High Energy Physics*, 2015(4), apr 2015. doi:10.1007/jhep04(2015)040. URL https://doi.org/10.1007%2Fjhep04%282015%29040.

[81] Torbjörn Sjöstrand, Stephen Mrenna, and Peter Skands. A brief introduction to PYTHIA 8.1. *Computer Physics Communications*, 178(11):852–867, jun 2008. doi:10.1016/j.cpc.2008.01.036. URL https://doi.org/10.1016%2Fj.cpc.2008.01.036.

[82] G. Aad et al. The ATLAS simulation infrastructure. *The European Physical Journal C*, 70(3):823–874, sep 2010. doi:10.1140/epjc/s10052-010-1429-9. URL https://doi.org/10.1140%2Fepjc%2Fs10052-010-1429-9.

[83] S. Agostinelli et al. Geant4—a simulation toolkit. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 506(3):250–303, 2003. ISSN 0168-9002. doi:https://doi.org/10.1016/S0168-9002(03)01368-8. URL `https://www.sciencedirect.com/science/article/pii/S0168900203013688`.

[84] Christian Wiel. *Reconstruction and Identification of Boosted Tau Pairs in the Semi-Leptonic Decay Channel at ATLAS*. PhD thesis, Dresden, 2018.

[85] Adam Paszke, Sam Gross, Francisco Massa, Adam Lerer, James Bradbury, Gregory Chanan, Trevor Killeen, Zeming Lin, Natalia Gimelshein, Luca Antiga, Alban Desmaison, Andreas Kopf, Edward Yang, Zachary DeVito, Martin Raison, Alykhan Tejani, Sasank Chilamkurthy, Benoit Steiner, Lu Fang, Junjie Bai, and Soumith Chintala. Pytorch: An imperative style, high-performance deep learning library. In *Advances in Neural Information Processing Systems 32*, pages 8024–8035. Curran Associates, Inc., 2019. URL `http://papers.neurips.cc/paper/9015-pytorch-an-imperative-style-high-performance-deep-learning-library.pdf`.

[86] Sebastian Ruder. An overview of gradient descent optimization algorithms. *arXiv preprint arXiv:1609.04747*, 2016.

[87] Identification of Boosted Higgs Bosons Decaying Into $b\bar{b}$ With Neural Networks and Variable Radius Subjets in ATLAS. Technical report, CERN, Geneva, 2020. URL `http://cds.cern.ch/record/2724739`. All figures including auxiliary figures are available at https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PUBNOTES/ATL-PHYS-PUB-2020-019.