

The Use of Energy Correlation Functions to Identify Jet Pairs from $H \rightarrow b\bar{b}$ in the ATLAS Experiment

Nicole Hartman

Mentor: Prof. Stephen Sekula

Southern Methodist University, Dallas, Texas 75275-0175 USA

E-mail: nhartman@smu.edu

ABSTRACT:

The typical proton-proton collision at the LHC produces about 20 “jets,” collimated sprays of particles from hadronization of quarks or gluons. The largest decay branching fraction of the Higgs boson is to a pair of bottom quarks, $H \rightarrow b\bar{b}$, resulting in a pair of jets; however, reconstructing this channel is made challenging by the sheer rate of QCD jet production in the LHC. This analysis used variables defined from Energy Correlation Functions to select and reconstruct the pair of bottom quark initiated jets most likely originating from Higgs decay in $H \rightarrow b\bar{b}$ simulation.

Contents

1	Introduction	3
1.1	Electroweak Symmetry Breaking	4
1.2	Large Hadron Collider	5
2	ATLAS Detector	5
2.1	ATLAS Coordinate System	5
2.2	Inner Detector	7
2.2.1	Pixel Tracking System	10
2.2.2	Silicon Microstrip Trackers	11
2.2.3	Transition Radiation Tracker	11
2.3	ECAL	12
2.3.1	Sampling vs. Homogenous Calorimeters	12
2.3.2	Liquid Argon Detector	13
2.4	HCAL	15
2.4.1	Tile Calorimeter	15
2.4.2	LAr hadronic end-cap calorimeter	15
2.4.3	LAr forward calorimeter	16
2.5	Muon Spectrometer	16
2.5.1	Toroidal Magnets	16
2.5.2	Monitored Drift Tubes	16
2.5.3	Cathode Strip Chambers	18
2.6	Forward Detectors	18
2.7	Data Acquisition Triggers	18
3	Jets	18
3.1	b tagging	19
4	Higgs Decay Modes	21
5	Energy Correlation Functions	24
5.1	C_1 Discriminating Variable	25
5.2	C_2 Discriminating Variable	26
6	Applying ECFs to ATLAS $pp \rightarrow HH \rightarrow b\bar{b}\gamma\gamma$ simulation	26
6.1	C_1 Discriminating Variable	26
6.2	C_2 Discriminating Variable	27
6.2.1	Optimization of β_{C_2}	28
6.2.2	Reconstructed Mass	29
7	Future Directions	33

8	Conclusions	33
9	Acknowledgements	33

List of Figures

1	The particles in the Standard Model. The Higgs is in the center, and the boson force carriers are shown in the circle right outside of it, where the lower half of the circle shows the carriers of the weak force and the upper half shows the photon and gluon, the carriers of the electromagnetic and strong forces, respectively. The outer circle has the fermions, which make up the matter of the universe. The upper half of the outer circle shows the quarks, while the lower half of shows the leptons [2]	3
2	The scalar potential used in spontaneous symmetry breaking [11]	4
3	Cut away of the ATLAS detector [17]	6
4	Illustration of the orientation of the subsystems inside the ATLAS Inner Detector [23]	7
5	List of the dimensions of the subsystems in the ID [17].	7
6	The structural elements in the ID at $\eta = 0.3$ [17].	8
7	The structural elements in the ID at $\eta = 1.4$ [17].	9
8	Schematic for the general idea for how a pixel detector tracks incident particles. [21]	10
9	Illustration of the active region of the pixel detector with the barrel and endcap layers [23]	11
10	Principle of operation for the TRT [24]	12
11	Schematic illustrating how a sampling calorimeter causes an incident particle to form a shower more quickly. [27]	13
12	Accordian sampling structure for calorimeter. [28]	14
13	Wedge showing the accordian structure for the liquid Argon portion of the calorimeter. [28]	14
14	Calorimeter system at ATLAS [17].	15
15	Main parameters for the Muon Spectrometer. The values in parentheses refer to the configuration from 2009 [17].	17
16	Monitored Drift Tubes: MDTs	17
17	Illustration for how jets resolved in their CM frame can become collimated when boosted into the lab frame [31].	19
18	“A secondary vertex with a significant decay length indicates the presence of a long-lived particle in the jet” [35]	20
19	The dominant decay mode for a b quark. The mediator, W^* is off-shell because the W is heavier than the b , and “ u ” stands for an “up-type-quark,” or a light quark. The b reconstruction resolution is impeded by neutrino in the final state.	20

- 20 ATLAS figures for the channels used for the 2012 Higgs boson discovery 21
- 21 Branching ratios for the various decay channels of the Higgs, where the orange 22 line at 125.9 GeV was added to emphasize the true branching ratios.[4] 22
- 22 Reconstruction of the di-jet mass of the two b quarks in the $ZH \rightarrow llb\bar{b}$ 22 search. The dominant backgrounds, and the expected signal (multiplied by a factor of 5) for $m_H = 120$ GeV are shown [39].
- 23 The cross sections for SM processes at the LHC, illustrating that weak and 23 Higgs production processes are orders of magnitude rarer than hadronic ones. [44].
- 24 C_1 distributions for quark vs. gluon discrimination in the boosted regime 25 where $\beta_{C_1} = 0.2$ [5].
- 25 C_1 overlay plot for $\beta_{C1} = 0.2$, which motivates a lower C_1 cut of 0.21. 27
- 26 Overlay the PDFs for the C_2 distributions, $\beta_{C_2} = 1.2$. 28
- 27 Likelihood functions for 2b events with $B_{C_2} = 1.2$ 29
- 28 ε_{0b} from Eq. (6.3), where the red (top) curve is for integrating L2(2b) at the 30 90% CL, the green (middle) curve integrated L2(2b) at a 85% CL, and the blue (bottom) curve integrated L2(2b) at a 80% CL. We want to choose a β_{C2} that minimizes ε_{0b} .
- 29 Reconstructed mass that maximized the likelihood that the di-jet pair originated from two bs (signal) and minimized the likelihood that the di-jet pair originated from two bs (control). 31
- 30 Breaking up the control curve from Fig. 29 to the three cases for the true 32 di-jet composition. We can see that some 2b di-jets sometimes have the min HDLR value, which contributes to the small peak at the Higgs mass for the control curve in Fig. 29.

1 Introduction

On July 4th, 2012, the ATLAS and CMS experiments jointly announced their discovery of a new resonance with mass 125 GeV,[†] [1] which was later determined to be highly consistent with the Standard Model (SM) Higgs boson. This was the culmination of the experimental verification of the particles hypothesized by the Standard Model (SM)[2]. Now that the Higgs has been discovered, it is essential to make precision measurements of its properties and couplings of the new boson. The Higgs self-coupling can be directly probed through $pp \rightarrow HH$ [3]. The most common decay mode for the Higgs is to bottom (b) quarks [4], but this channel can be difficult to distinguish from background processes due to generic processes from Quantum Chromodynamics (QCD). For this analysis, we were interested in investigating a new method to distinguish b -jets from jets initiated by gluons and other quarks; this same approach also has promise in distinguishing jets from resonant decays from jets arising from other sources. To do this, we exploited variables of interest from Energy Correlation Functions (ECF) [5], used by the authors to identify di-jet resonances in the boosted regime. We applied the ECFs to the resolved regime for SM Higgs decay.

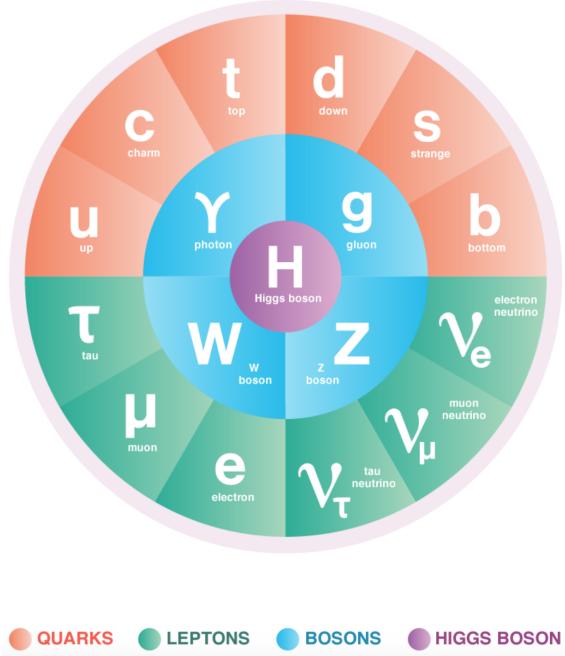


Figure 1: The particles in the Standard Model. The Higgs is in the center, and the boson force carriers are shown in the circle right outside of it, where the lower half of the circle shows the carriers of the weak force and the upper half shows the photon and gluon, the carriers of the electromagnetic and strong forces, respectively. The outer circle has the fermions, which make up the matter of the universe. The upper half of the outer circle shows the quarks, while the lower half shows the leptons [2]

[†]In particle physics, natural units are used, where $\hbar = c = k_B = 1$, which is why the resonance's mass is cited in MKS units of “energy.”

1.1 Electroweak Symmetry Breaking

Electroweak theory provided the unification of the electromagnetic and weak forces, and culminated with the formulation of the SM, which successfully described the fundamental particles and their interactions [6]. However, as originally formulated, the SM predicted massless force carriers [7]. The photon was known to be massless, but experiment had already concluded that the weak bosons were massive. This discrepancy was solved in the 1960s, when Robert Brout, François Englert, Peter Higgs, Gerald Guralnik, Richard Hagan, and Tom Kibble proposed spontaneous symmetry breaking to explain the discrepancy [8–10]. In Fig. 2 the proposed scalar potential field is shown [12]. The potential is even and symmetric about the origin, but the origin does not correspond to a local minimum. They proposed nature starts in the symmetric point, but random fluctuations cause a drop into a local minimum with an infinite set of equally valid vacuum.

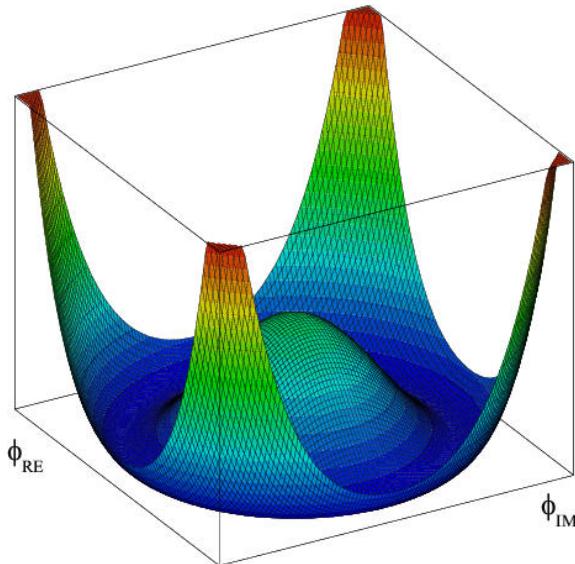


Figure 2: The scalar potential used in spontaneous symmetry breaking [11]

Therefore, what we observe today in nature does not look symmetric, but the symmetry of nature is replaced with a symmetry of possible solutions. As a consequence of this theory, three of the four of the original fields are “eaten-up” by the W^\pm and Z bosons, and the remaining field is the Higgs field. Since this mechanism gives mass to the vector bosons, the Higgs field can also be used to give the leptons mass so that particles that interact more strongly with the Higgs field gain more mass [13]. However, this does not lessen the number of inputs to the SM as the inputs of the particles’ masses are replaced with inputs of the particles’ couplings to the Higgs field.

As every field has an associated particle [14], the Large Hadron Collider (LHC) at CERN in Geneva, Switzerland was built to probe a high-energy regime to search for the Higgs particle, the physical manifestation of the Higgs field. Since the Higgs discovery announcement, Peter Higgs and François Englert won the 2013 Nobel Prize for physics for their contributions to the theory of spontaneous electroweak symmetry breaking [15].

1.2 Large Hadron Collider

The LHC is a 27 km circumference particle accelerator straddling the border between France and Switzerland [16]. It now collides bunches of 10^9 protons every 25 ns at a center of mass energy of 13 TeV [17]. Bunches of protons are used since the proton radius is about 1 fm, and thus the probability of a head-on collision between single protons is increased by increasing the number of colliding particles. We expect about 20 collisions per given bunch crossing at the LHC [18].

Since the proton is not an elementary particle, but rather a baryonic resonant QCD bound state, even a head-on collision will not have all of the available energy concentrated at a point. The proton is a composite particle made up of two up and one down valence quarks, along with a sea of virtual quarks and gluons that spontaneously come out of the vacuum because of the uncertainty relation $\Delta x \Delta p \geq \hbar/2$. The 6.5 TeV per proton is distributed among these partons[†]. Therefore, to look at the high-energy regime, we are interested in a “hard scatter” where a single quark or gluon carrying a large fraction of one proton’s momentum collides with a high energy constituent of the other proton, lending to the production of one or more Higgs bosons.

Because the Higgs has a short lifetime (on the order of 10^{-22} s [19]), the Higgs itself is not detected; rather its decay products are absorbed in the detectors at collision sites around the LHC ring. The two general-purpose detectors at the LHC are ATLAS (A Toroidal LHC ApparatuS) and CMS (Compact Muon Solenoid). SMU is a member of the ATLAS collaboration, which provided the Monte Carlo (MC) samples.

2 ATLAS Detector

A cutaway view of the ATLAS detector is shown in Fig. 3. The inner detector (ID) provides charged particle tracking information closer to the interaction point. The inner detector is immersed in a 2 T solenoidal magnetic field to bend the trajectories of charged particles and allow for momentum measurement. Energy measurement is made in two parts: with an electromagnetic and hadronic calorimeter. Finally, the outermost layer of the detector is the muon spectrometer, with a 4 T toroidal magnetic field.

2.1 ATLAS Coordinate System

At the ATLAS detector, the z-axis is measured along the accelerator beam pipe, the x-axis points into the center of the ring, and y-axis is defined to point vertically up by the properties of a right-handed rectilinear coordinate system. Then a cylindrical coordinate system is used for the ATLAS detector. The azimuthal angle, $\phi = \arctan(\frac{y}{x})$, denotes orientation in the plane transverse to the beam, and the pseudorapidity, η , measures the polar angle inside the detector, where

$$\eta = -\ln \left(\tan \left(\frac{\theta}{2} \right) \right) \quad (2.1)$$

[†]A parton is a constituent of a hadron.

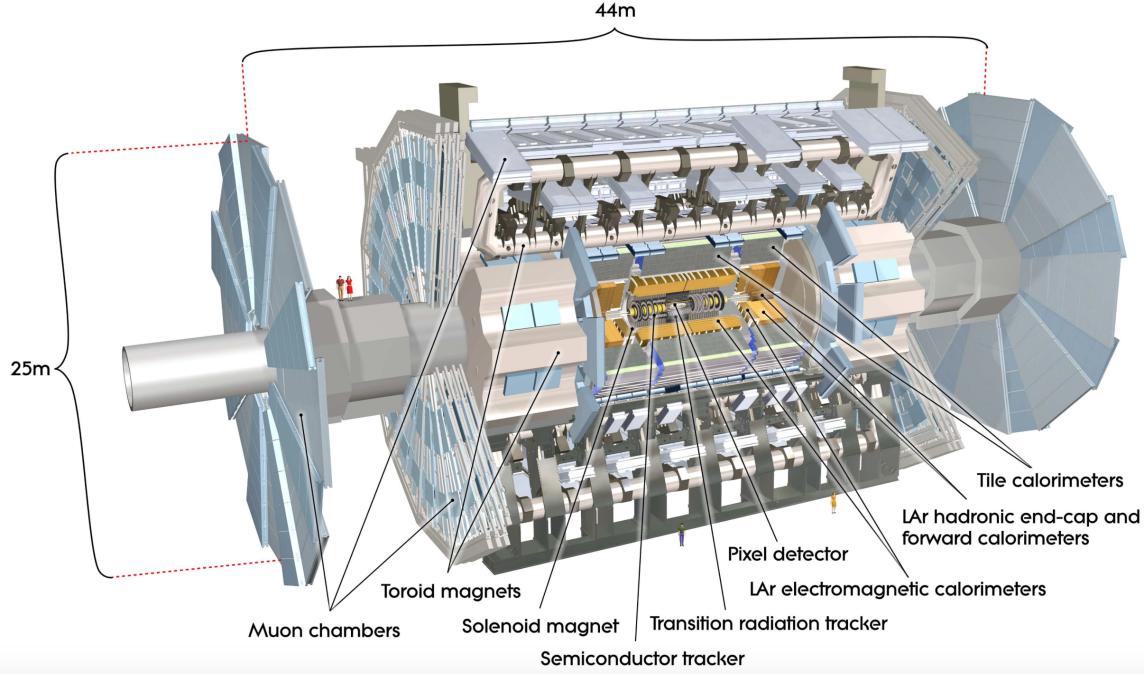


Figure 3: Cut away of the ATLAS detector [17]

and where $\theta \in [0, \pi]$ is the polar angle as measured from the z-axis. The ATLAS detector is forward-backward symmetric to maximize the detector coverage since colliding protons each have the same energy.

From these detector variables, we can get components in the four-momentum using the relations (derivation shown in Ref [20]).

$$(E, p_x, p_y, p_z) = (E, p_T \cos \phi, p_T \sin \phi, p_T \sinh \eta) \quad (2.2)$$

$$p = p_T \cosh \eta \quad (2.3)$$

We can also get a measurement of the momentum in the calorimeter for relativistic particles. In the relativistic limit, the energy and momentum are the same (in natural units). However, when the momentum is defined from a calorimeter measurement, E_T is used instead of p_T . Equations (2.2) and (2.3) still apply in this case.

2.2 Inner Detector

A cutaway view of the inner detector for ATLAS is shown in Fig. 4. With approximately

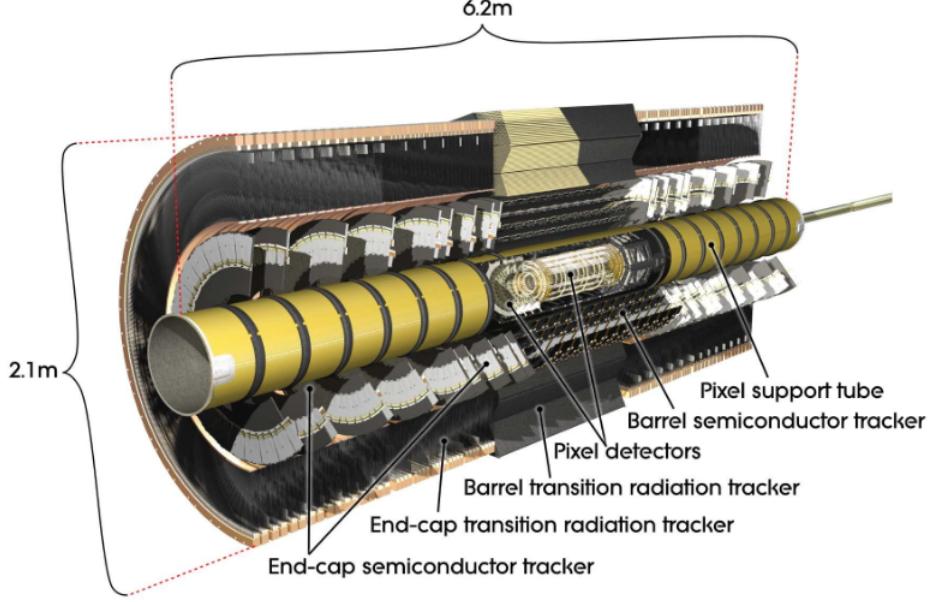


Figure 4: Illustration of the orientation of the subsystems inside the ATLAS Inner Detector [23]

Item		Radial extension (mm)	Length (mm)
Overall ID envelope		$0 < R < 1150$	$0 < z < 3512$
Beam-pipe		$29 < R < 36$	
Pixel	Overall envelope	$45.5 < R < 242$	$0 < z < 3092$
3 cylindrical layers	Sensitive barrel	$50.5 < R < 122.5$	$0 < z < 400.5$
2×3 disks	Sensitive end-cap	$88.8 < R < 149.6$	$495 < z < 650$
SCT	Overall envelope	$255 < R < 549$ (barrel) $251 < R < 610$ (end-cap)	$0 < z < 805$ $810 < z < 2797$
4 cylindrical layers	Sensitive barrel	$299 < R < 514$	$0 < z < 749$
2×9 disks	Sensitive end-cap	$275 < R < 560$	$839 < z < 2735$
TRT	Overall envelope	$554 < R < 1082$ (barrel) $617 < R < 1106$ (end-cap)	$0 < z < 780$ $827 < z < 2744$
73 straw planes	Sensitive barrel	$563 < R < 1066$	$0 < z < 712$
160 straw planes	Sensitive end-cap	$644 < R < 1004$	$848 < z < 2710$

Figure 5: List of the dimensions of the subsystems in the ID [17].

1000 tracks emerging from the interaction point every 25 ns, the ID is divided into three different regions to optimize the pattern recognition and momentum measurement algorithms [17]. The pixel detectors have the most precision, so this layer is closest to the interaction point. The pixel detector also has the highest cost, so the next least expensive tracking option is the silicon microstrip trackers (SCTs), at the next farthest region away from the interaction point. Finally, the Transition Radiation Trackers (TRTs) compose the outermost region of the inner detector. The ID uses pattern recognition to measure transverse momentum as low as 0.5 GeV, and provides electron identification for $|\eta| < 2.0$ for energies up to 150 GeV [17]. Displaced multi-particle vertexes can be resolved by the pixel detector and help identify long-lived B-hadrons.

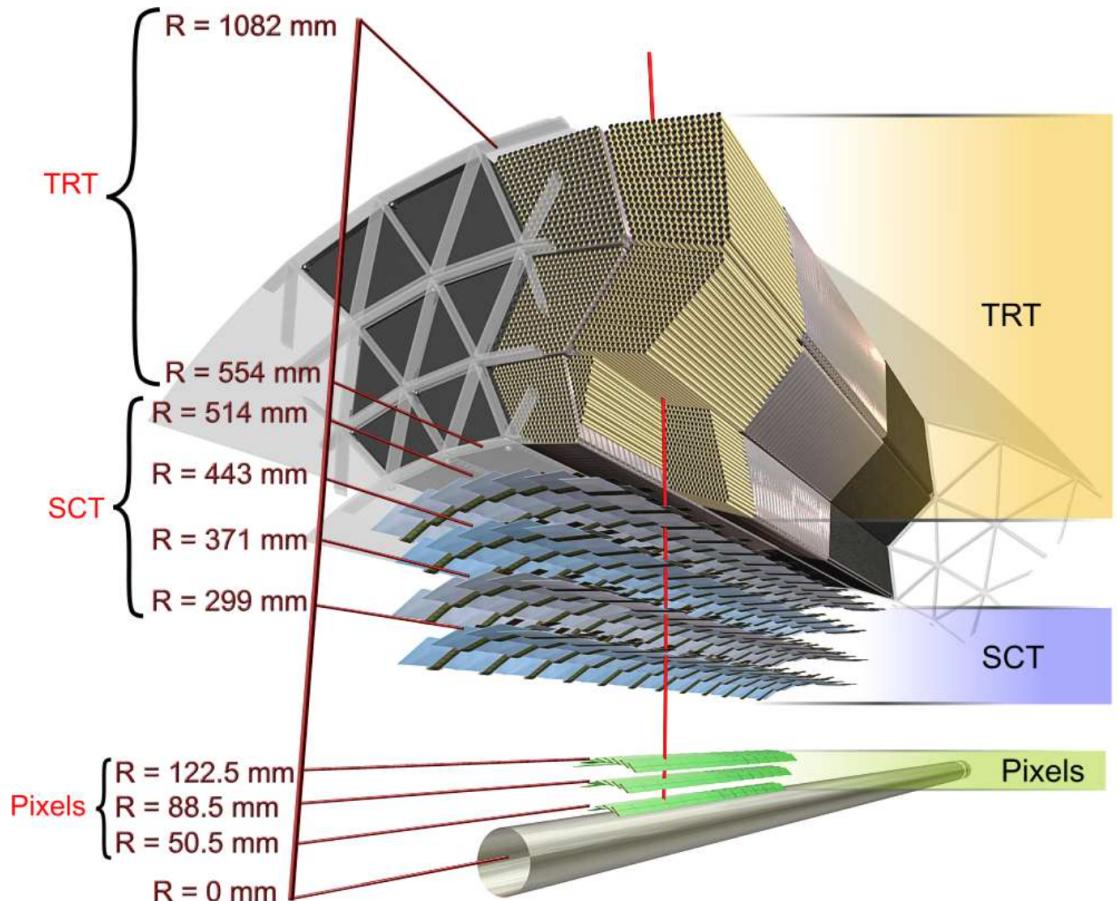


Figure 6: The structural elements in the ID at $\eta = 0.3$ [17].

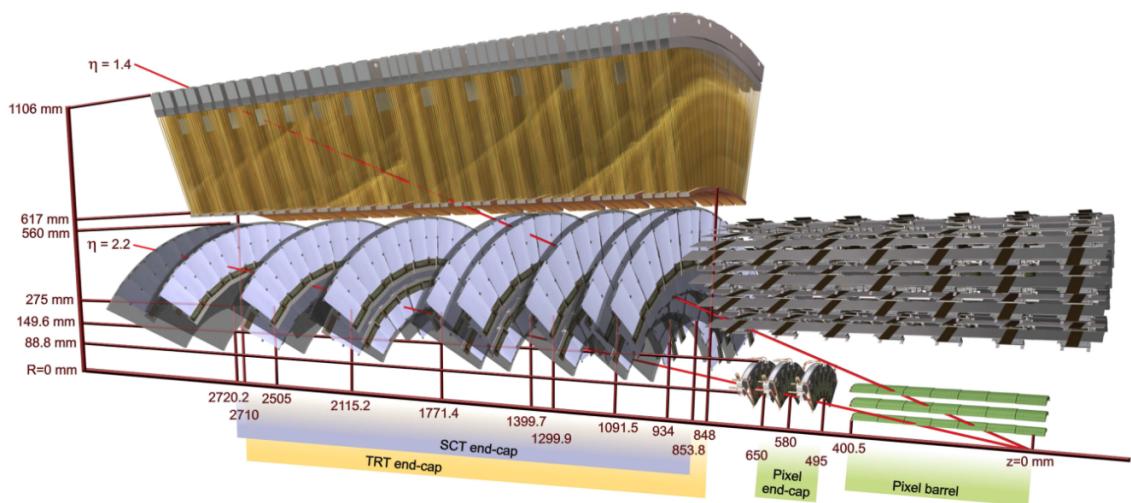


Figure 7: The structural elements in the ID at $\eta = 1.4$ [17].

2.2.1 Pixel Tracking System

ATLAS is often referred to as a “giant camera.” In a general sense, both a camera and the ATLAS detector provide ways to record specific events, but the parallels run deeper than their purposes, since both cameras and particle detectors can obtain information via pixels. In a camera an incident photon will go through a silicon diode and knock a valence electron free to generate a few electron hole pairs. An electric field is set up across the pixel and will then pull the electron and hole apart to the metal contacts where the charge can be read out [21]. The pixel detector at ATLAS works in a similar way, except the energy of the incident particles at ATLAS is orders of magnitude larger: while a photon of visible light has an energy of a few eV, typical “interesting” particles at the LHC are in the MeV to TeV range.

The pixel detectors do not “absorb” particles, but rather, as a particle passes through the pixel sensor, it creates electron-hole pairs that are separated by the electric field and read out by the electronics, as shown in Fig. 8.

The pixel sensors in ATLAS consist of 80 million [22] rectangular $50 \mu\text{m} \times 400 \mu\text{m}$ “n⁺-in-” electrodes.[§] The pixel detector’s position with respect to the other components of the inner detector is shown in Figure 4. Zooming in on Fig. 4, the innermost part of the inner detector—the pixel detector—is shown in Fig. 9. The 80 million channels are divided into four cylindrical layers combined in a volume 1442 mm in length with 430 mm radius [23]. The pixel detector has a resolution of 15 microns, and this precision is limited by the size of the electronics.

The pixel detector starts 5 cm from the center of the beam pipe to gain as much information about the central reaction as possible, and has coverage for $|\eta| < 2.5$ [22]. Since the pixel detector is so close to the interaction point, its electronics must be able to withstand high radiation doses of up to 500 kGy. The p-n diodes can experience a leakage current when an electron-hole pair has enough energy to overcome the potential barrier. The detector minimizes the number of electron hole pairs spontaneously generated by cooling the system at -6°C .

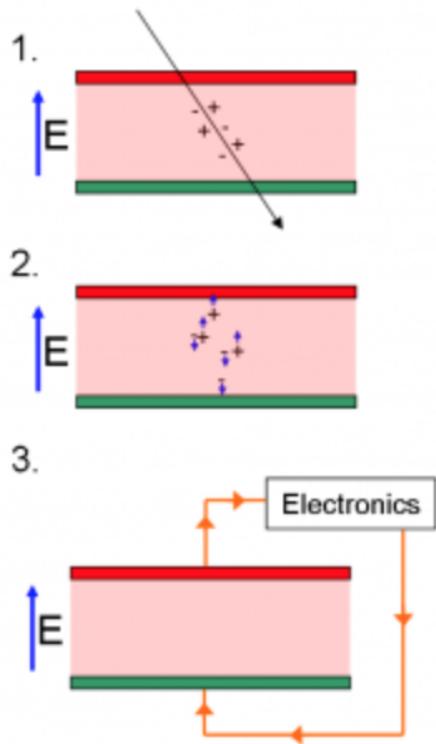


Figure 8: Schematic for the general idea for how a pixel detector tracks incident particles. [21]

[§]The “n” region is n-type, or doped with atoms that are electron donors; the “in” region is intrinsic, or undoped; and the “n⁺” region is also n-type, but doped more heavily than other n region.

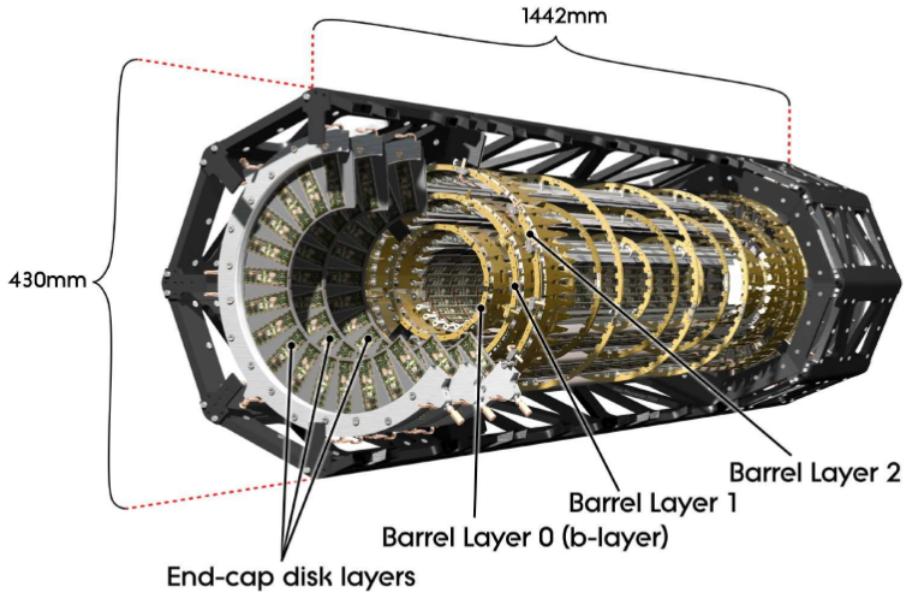


Figure 9: Illustration of the active region of the pixel detector with the barrel and endcap layers [23]

2.2.2 Silicon Microstrip Trackers

The Silicon Microstrip Trackers (SCTs) operate with a similar principle as the pixel detectors, but have an effective operating voltage determined by the effective doping level, and their leakage current also increases linearly with radiation dose. Initially the operating voltage was 150 V, but increased up to 250 to 350 V to compensate for the radiation dose after 10 years. The 15912 SCT sensors each have a thickness of $285 \pm 15 \mu\text{m}$ and a pitch of $80 \mu\text{m}$.

2.2.3 Transition Radiation Tracker

The Transition Radiation Trackers (TRT) make up the outermost region of the ID, and the dimensions of the TRT are shown in Fig. 5. “Transition Radiation” is the radiation emitted by a relativistic particle as it traverses an interface between two materials with different permittivities. Each TRT is a 4 mm polyimide drift tube, created by two $35 \mu\text{m}$ multi-layer films bonded back-to-back [17]. A $25 \mu\text{m}$ thick polyimide film has one side laminated with a $0.2 \mu\text{m}$ of Al with another $5\text{-}6 \mu\text{m}$ layer of graphite [17]. The inside of the tube is filled with a gas composed of 70% Xe, 27% CO₂, and 3% O₂. The anode is composed of $31 \mu\text{m}$ diameter cylinder of tungsten wire positioned at the center of the drift tube, and coated with $0.5\text{-}0.7 \mu\text{m}$ of gold. After fabrication, the tubes were cut to a 144 cm length for the barrel and 37 cm length for the end-cap region. The barrel straws are read out at each end of the tube, so the middle of the tube is insulated with a 6 mm glass layer which creates a 2 cm spot where the element is not sensitive to incident tracks. The

anode is grounded and connected to the front end electronics, while the cathodes are held at -1530V . Minimizing mechanical sag in the straw is crucial for accounting for the error in the position measurements, so the straws are mechanically supported by carbon fibers [17].

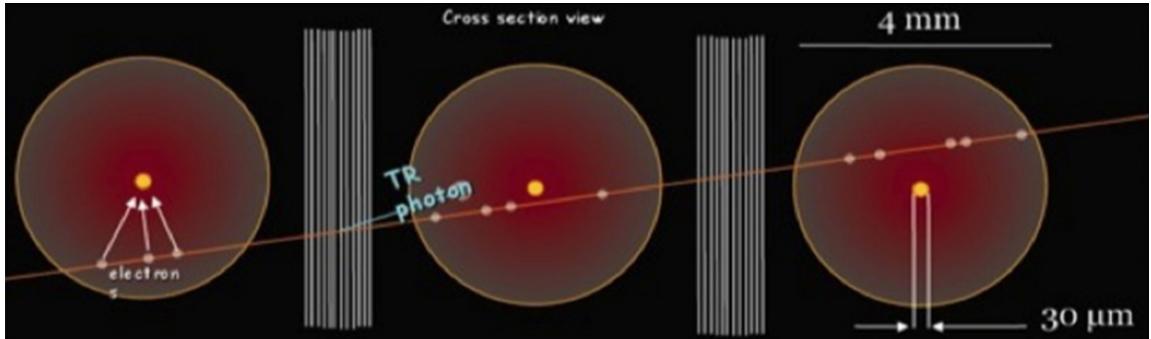


Figure 10: Principle of operation for the TRT [24]

An incident particle traversing a straw ionizes the gas molecules to free electrons and positive ions. The electrons accelerate through the electric field to the anode, and these electrons in turn will ionize other ions to create a gain of 2.5×10^4 . Since the anode is read out at both ends of the wire, this provides a drift time measurement from the difference of arrival times between the two ends of the wire. The time resolution is approximately a nanosecond, which corresponds to a spatial resolution of about 100 microns.

An outgoing particle will hit on average 36 of the TRT tubes; this improves the momentum measurement in the inner detector. During its lifetime, a straw detector can track 10^{15} particles, and a total integrated charge of 1000 C, which corresponds to about 20 years of the LHC’s operation. The ATLAS inner detector has 12,000 of these straws in the endcaps and 52,544 straws in the barrel, yielding a total of 351,000 read-out channels. Although the silicon trackers need to be cooled between -5 to -10°C , the TRTs operate at room temperature.

2.3 ECAL

Calorimetry literally means “heat measurement.” The incident particle interacts with the material in the detector to form a shower of particles which are decelerated and absorbed in the detector material to reconstruct the original particle’s energy. The ATLAS calorimeter is divided into two parts: the electromagnetic calorimeter (ECAL) and the hadronic calorimeter (HCAL). The ECAL is for electromagnetic interactions and has higher precision than the HCAL which reconstructs the particles interacting by the strong force.

2.3.1 Sampling vs. Homogenous Calorimeters

For an electromagnetic shower to develop, incident electrons (and positrons) will emit a photon through Bremsstrahlung approximately in through a distance characterized by the

radiation length

$$X_0 = \frac{716.4 \text{ g} \cdot \text{cm}^{-2} A}{Z(Z+1) \ln \frac{287}{\sqrt{Z}}} \quad (2.4)$$

where A is the number of nucleons and Z is the atomic number (number of protons) in the detector material [25]. The units of X_0 means that dividing by the density gives the actual distance traveled by a particle [26]. Then the photons will in turn pair-produce electrons and positrons when traveling a distance of $\frac{9}{7}X_0$. This showering phenomena will continue until the average particle energy decreases to below the critical energy, $E_c = \frac{610 \text{ MeV}}{Z+1.24}$. The hadronic showers are characterized by the interaction length, $\lambda = 37.8A^{0.312}\text{g} \cdot \text{cm}^{-2}$, the average distance that an incident particle travels before undergoing a nuclear interaction.

Two different calorimeter designs can be used to measure the incident particle's energy. In a sampling calorimeter the volume is divided into scattering and absorbing slabs, shown in Fig. 11. The scattering slabs use a high Z material to decrease the radiation length and allow the shower to develop more quickly. The energy deposited in the absorbing region of the calorimeter is measured, and the total energy is the energy deposited in the absorbing region divided by a scale factor, $f_{sampling} = E_{visible}/E_{deposited}$. A sampling calorimeter contains the shower in a smaller detector to help minimize the cost of the experiment. One of the downsides of this procedure is that the sampling method is not as precise because only a fraction of the energy deposited is measured, and fluctuations proportional to \sqrt{E} for Poisson statistics will introduce extra errors into the measurements.

A homogenous calorimeter circumvents this problem by using the whole calorimeter as an active volume. A homogenous calorimeter is only useful as an ECAL since hadronic interactions require more material to "contain," and it may not be monetarily feasible to construct such a large volume. To see this, we can look at the approximate formulas for the radiation length and interaction length, $X_0 \sim \frac{A}{Z^2}$, and $\lambda \sim A^{1/3}$. Since Z is approximately $\frac{1}{2}A$, this means $\frac{\lambda}{X_0} \sim A^{4/3}$, a number that can be as large as 30 for high Z materials such as lead, showing that hadronic showers have a larger extent.

2.3.2 Liquid Argon Detector

The ATLAS ECAL is a sampling calorimeter arranged in an accordion structure, as shown in Fig. 12. The high Z material ($Z = 82$) lead creates the shower, while the energy is measured in liquid argon (LAr), a low Z material ($Z = 18$). As illustrated in Fig. 12, the folding angle decreases as the the radius (measured out from the interaction point) increases. The folding angle varies between 90° and 67° to keep the LAr sampling region width approximately constant at 2.1mm between the absorbers.

The LAr scintillators are read out with wavelength-shifting photosensors. This type of a detector lends itself naturally to a tower structure with the modules forming wedges

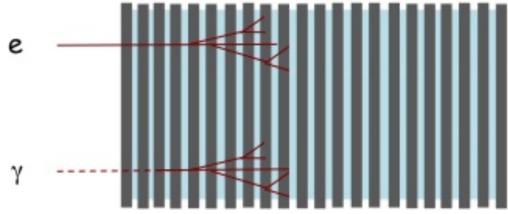


Figure 11: Schematic illustrating how a sampling calorimeter causes an incident particle to form a shower more quickly. [27]

– 13 –

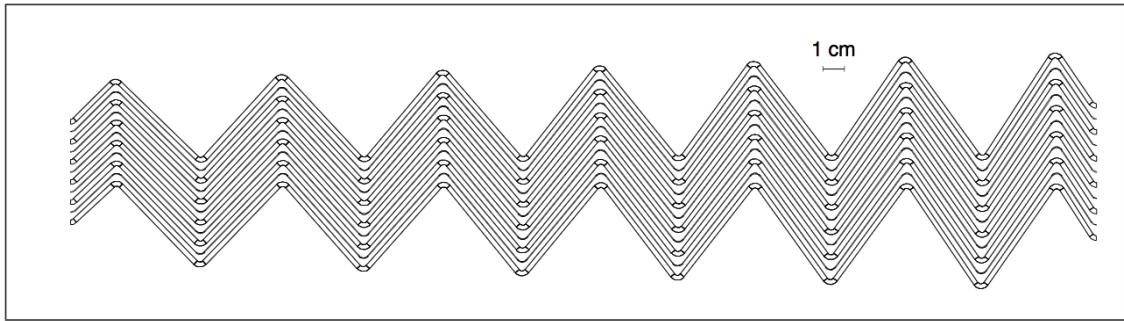


Figure 2-21 Transverse section through the active part of the barrel EM calorimeter (see also Figure 2-iii).

Figure 12: Accordian sampling structure for calorimeter. [28]

pointing back to the interaction point, shown in Fig. 13. This makes it easy for a particle's shower to be contained within a few modules or cells. The segmentation in $\Delta\eta \times \Delta\phi$ is 0.025×0.1 in the pre-sampler region, 0.0031×0.1 in the strips, 0.25×0.25 in the main region, and 0.5×0.025 in the back region. This yields an energy resolution of 10-12% $\text{GeV}^{-1/2}$.

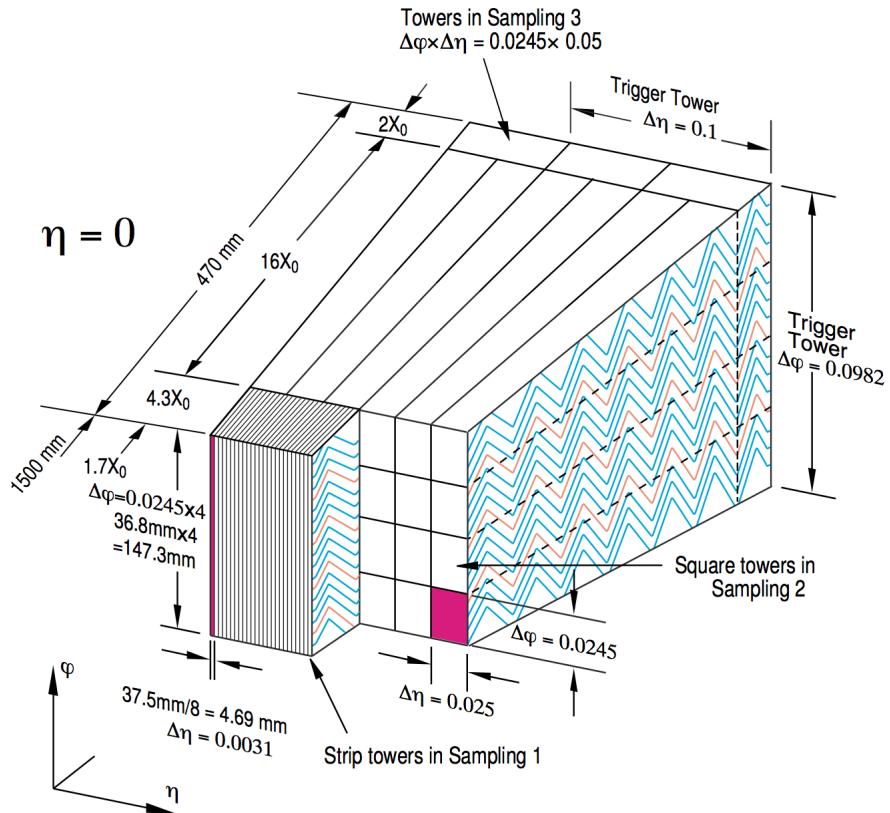


Figure 13: Wedge showing the accordian structure for the liquid Argon portion of the calorimeter. [28]

2.4 HCAL

The hadronic calorimeter (HCAL) is the layer just outside of the ECAL, and measures the energy of hadrons that traverse the ECAL without stopping, as well as minimum-ionizing particles like muons. Although the ECAL system could be used to measure the development of the hadronic showers as well, the HCAL system's coarser granularity decreases the monetary cost in covering this larger volume. It is divided into three parts: the tile calorimeter, the LAr hadronic end-cap calorimeter, and the LAr forward calorimeter, as shown in Fig. 14.

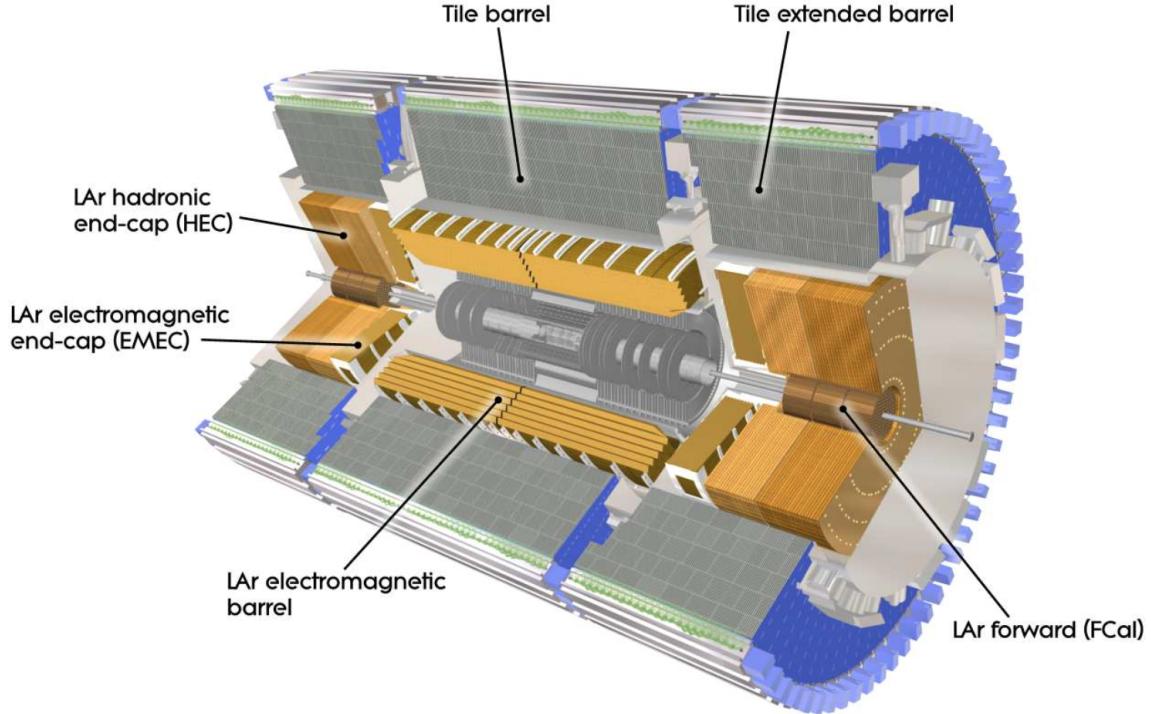


Figure 14: Calorimeter system at ATLAS [17].

2.4.1 Tile Calorimeter

The tile calorimeter is just outside ECAL and has an inner radius of 2.28 m and an outer radius 4.45 m. The barrel covers $|\eta| < 1.0$, while the extended barrels cover $0.8 < |\eta| < 1.7$. It is a sampling calorimeter with a steel absorber and scintillating tiles read out with wavelength shifting fibers and photomultiplier tubes. Azimuthally divided into 64 modules, the barrel is segmented into three regions at 1.5λ , 1.8λ , and 4.25λ , while the extended barrel has 1.5λ , 2.6λ , and 3.3λ depth segmentation. At $\eta = 0$, the HCAL has a 9.7λ depth [17].

2.4.2 LAr hadronic end-cap calorimeter

The LAr hadronic end-cap calorimeter has two wheels per endcap just behind the ECAL endcap calorimeter and housed within the same LAr cryostats. The wheels have two depth segments, 25 mm parallel copper plates for the wheels closest to the interaction point, and

50 mm plates for the wheels farther from the interaction point. The wheels are divided into 32 wedges, with inner and outer radii of 0.475m and 2.03m, respectively. The copper plates are filled with LAr as the active sampling volume [17].

2.4.3 LAr forward calorimeter

The forward calorimeter covers $\eta > 3.1$ and is approximately 10 interaction lengths deep. Each side has three modules, the first of copper for electronic measurements, and the second two of tungsten for hadronic measurements.

2.5 Muon Spectrometer

The tracking system determines the transverse momentum of a charged particle from its radius of curvature in a magnetic field since $R = p_T/(qB)$. A more massive object moving at the same velocity will have a larger transverse momentum, and therefore a larger radius of curvature. Since the muon is 200 times heavier than an electron, it can be difficult to have a tracking system that can accurately measure the momentum for relativistic muons and electrons, so the muon detection system is distinct to measure the p_T for a broad range of muon energies.

The muon detector is a gas detector like the TRT, and operates according to similar principles. The subsections below detail the components of the muon spectrometer, while the parameters for the coverage and number of channels for are listed in Fig. 15.

2.5.1 Toroidal Magnets

The muons are detected through the deflection of their tracks in a 4 T magnetic field. There are three large air–core toroids, and each of these 3 toroids has 8 coils. The particles of interest that reach this portion of the detector are muons and neutrinos, but only the muons will be detected because the electrically neutral neutrinos are not deflected by the magnetic field. Backgrounds for the muon spectrometer are photons and neutrons with energies below an MeV and 100 MeV, respectively [17]. The magnetic field is designed to be transverse to the muons’ flight direction to minimize multiple scattering. The magnets are housed within cryostats to keep them below the critical temperature for superconductivity. Inside these magnets are the muon chambers divided into three cylindrical chambers about the beam axis in the barrel region, while the transition and end–cap regions arranged as disks, also divided into three chambers [17].

2.5.2 Monitored Drift Tubes

The Monitored Drift Tubes (MDTs) are drift chambers that provide precision measurements. Each tube is made of aluminum with a 3 cm diameter and length between 0.9 and 6.2 m. The tube is filled with a 93% Ar, 7% CO₂ gas mixture, at a pressure of 3-bar [17]. It has a gain of 2×10^4 [17], similar to the TRT drift chambers (see Section 2.2.3). The precision for single muon events is 100 μm , while the precision for multi-muon events is 50 μm [17]. To control the precision, the sag of the wires is minimized by three kinematical mounts placed strategically to minimize distortion due to the support, as shown in Fig. 16b. There are a total of 1174 MDTs positioned in the barrel of the ATLAS detector ($|\eta| < 2$).

Monitored drift tubes	MDT
- Coverage	$ \eta < 2.7$ (innermost layer: $ \eta < 2.0$)
- Number of chambers	1088 (1150)
- Number of channels	339 000 (354 000)
- Function	Precision tracking
Cathode strip chambers	CSC
- Coverage	$2.0 < \eta < 2.7$
- Number of chambers	32
- Number of channels	31 000
- Function	Precision tracking
Resistive plate chambers	RPC
- Coverage	$ \eta < 1.05$
- Number of chambers	544 (606)
- Number of channels	359 000 (373 000)
- Function	Triggering, second coordinate
Thin gap chambers	TGC
- Coverage	$1.05 < \eta < 2.7$ (2.4 for triggering)
- Number of chambers	3588
- Number of channels	318 000
- Function	Triggering, second coordinate

Figure 15: Main parameters for the Muon Spectrometer. The values in parentheses refer to the configuration from 2009 [17].

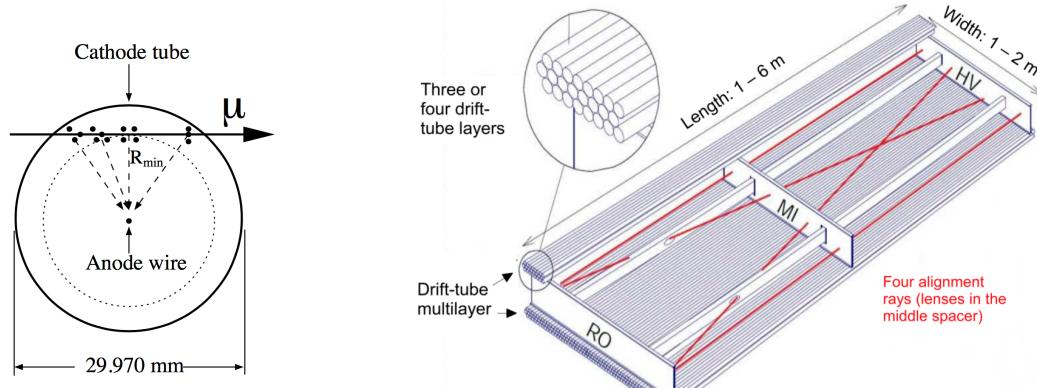


Figure 16: Monitored Drift Tubes: MDTs

2.5.3 Cathode Strip Chambers

For larger pseudorapidities ($2 < |\eta| < 2.7$), Cathode Strip Chambers (CSCs) are used for their higher granularity in the endcap region [17]. This is a multi-wire proportional chamber with strip read-out with a sense wire pitch of 2.54 mm and a read-out strip pitch of 5.08 mm resulting in a $60 \mu\text{m}$ track resolution [17].

2.6 Forward Detectors

The forward detector at ATLAS covers pseudorapidities $|\eta| > 8.2$, and has three subsystems: two to measure the luminosity and the last one for heavy-ion studies conducted when the collider accelerates with lead nuclei at $\sqrt{s} = 5 \text{ TeV}$. LUCID (LUminosity measurement using Cerenkov Integrating Detector) is located $\pm 17\text{m}$ from the interaction point, and is the primary online luminosity detector for ATLAS. ALFA (Absolute Luminosity For ATLAS) is located $\pm 240 \text{ m}$, and uses scintillating fibers as another measure for the luminosity, with these fibers extending as a millimeter from the beam pipe. Finally, ZDZ (Zero-Degree Calorimeter) lies ± 140 and has alternating layers of quartz rods and tungsten plates for detecting neutral particles at pseudorapities $|\eta| > 8.2$ [17].

2.7 Data Acquisition Triggers

The high fluency of particles prevents the ATLAS experiment from recording every proton-proton collision. There are three stages of triggers that discriminate which events should be written to storage. The Level I trigger is implemented in hardware before the information from an event even gets sent to a computing farm, and these Level I Triggering electronics need to be fast and efficient enough to pass on a manageable work load to the High Level Trigger [17]. The HLT records events at approximately 1000 Hz [29], while the beam collisions are at a rate of $4 \cdot 10^7$, so that the Level I needs to cut the rate by three orders of magnitude.

3 Jets

A given pp collision produces many quarks and gluons. However, because of the “confinement principle,” no free color charge can exist, which means no isolated quarks or gluons can exist in nature [30, 31]. It becomes energetically favorable for quark / anti-quark pairs to pop out of the vacuum to balance the color charge imbalance by forming color neutral hadrons. This sparks a chain reaction which produces a spray of particles in the detector.

The anti- k_T algorithm provides a way to cluster the energy deposited in the calorimeter to form a “jet,” and is the standard algorithm for defining jets at ATLAS [32]. First it defines two quantities,

$$d_{ij} = \min(p_{Ti}^{-2}, p_{Tj}^{-2}) \frac{\Delta R_{ij}}{R^2} \quad d_i = p_{Ti}^{-2} \quad (3.1)$$

where $\Delta R_{ij}^2 = (y_i - y_j)^2 + (\phi_i - \phi_j)^2$, and R is the jet-radius, defined by the user. To find the jets, the algorithm

- Calculates d_{ij} and d_i for all particles in the event, and lets d be the smallest d_i, d_{ij} .

- If $d = d_{ij}$, then it combines jet i and jet j .
 - If $d = d_i$, then it lets jet i be the final jet.
- It continues the previous steps until all the particles in the event have been accounted for.

This algorithm clusters high- p_T particles together if they fall within the jet’s radius.

The pp collisions at 13 TeV typically will give rise to about 20 moderate p_T jets in the detector, or around $\binom{20}{2} = 190$ viable di-jet candidates. So the challenge in elucidating the $H \rightarrow b\bar{b}$ or $HH \rightarrow 4b$ signals is accurately finding the “correct” pair(s) of jets.

Since the protons are not point-like, the colliding partons may not have equal p_T in the lab frame. By definition, the net momentum will be zero in the center of mass (CM) frame. A heavy (\sim TeV) resonance can be created at (or nearly at) rest in the CM frame, but when Lorentz boosting back into the lab frame the decay products can become highly collimated. When the decay products can no longer be resolved individually, we instead search for jets with a large radius parameter (“fat jets”), indicative of merged jets [31, 33]. The constituent clusters inside a fat jet are called *subjets*. Large R (about 1.0) correspond to fat jets, and R about 0.4 correspond to the standard resolved jets.

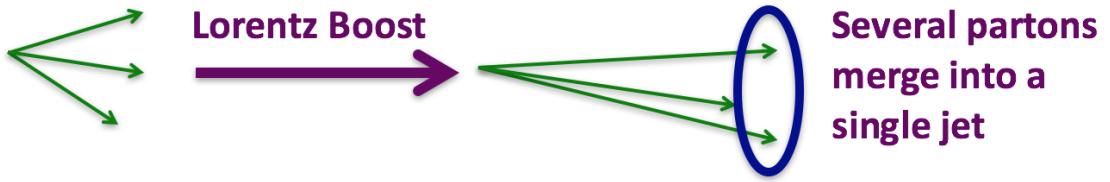


Figure 17: Illustration for how jets resolved in their CM frame can become collimated when boosted into the lab frame [31].

3.1 b tagging

A b -jet is a jet that has been identified as originating from a b -quark. Currently, b -jets are detected via “ b -tagging” [34–37]. The b quark tends to combine with other (mostly light) quarks to create a B hadron, which has a lifetime of about 1.5 ps [34, 36]. A relativistic hadron of this lifetime will travel about 1 mm in the detector. This displacement is long enough that it can be measured by the pixel detector if charged particles are produced. The signature for such an event is a vertex with a separation of approximately a few mm from another vertex, as shown in Fig. 18. Such a displaced vertex is identified by the impact parameter (IP) “the distance from the point of closest approach of the track to the interaction vertex” [36]. Jets originating from displaced vertices are used to identify b -quark initiated jets (b -jets) in ATLAS. This b -tagging algorithm can be employed either for resolved jets or for the subjets inside a fat jet. The b -tagging algorithm is only 70% efficient (at the most widely recommended working point) [37], which means identifying

two b-quarks this way is only 49% efficient, motivating the search for another method to find the right pair of jets from an $H \rightarrow b\bar{b}$ decay.

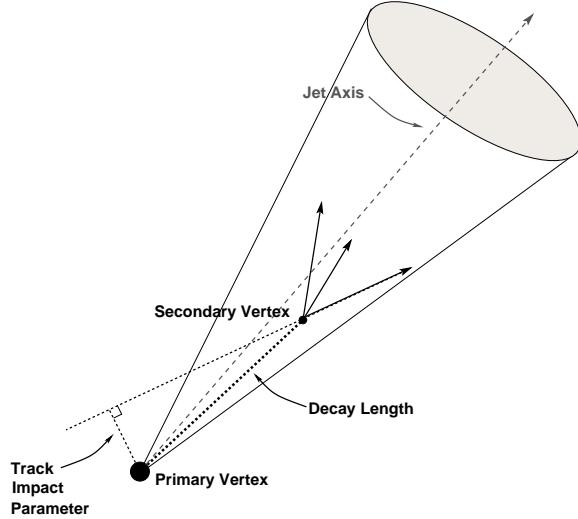


Figure 18: “A secondary vertex with a significant decay length indicates the presence of a long-lived particle in the jet” [35]

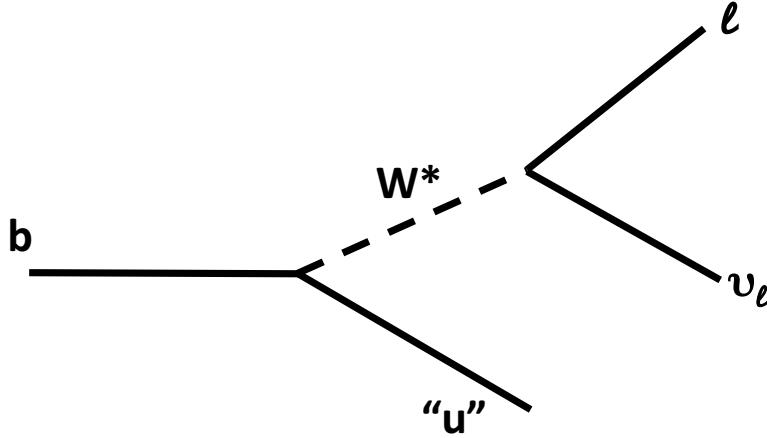


Figure 19: The dominant decay mode for a b quark. The mediator, W^* is off-shell because the W is heavier than the b , and “ u ” stands for an “up-type-quark,” or a light quark. The b reconstruction resolution is impeded by neutrino in the final state.

Since b is the lighter of the 3rd generation quarks, the b decays via a flavor changing current [38]. At tree level, the only way this reaction can occur is if it is mediated by

a W boson, i.e., by a flavor changing charged current, demonstrated in Fig. 19. The W^* then decays semileptonically, a channel which is difficult to reconstruct. This is because the final-state neutrino is not directly detected since its presence is induced by applying p_T conservation. Since conservation of transverse momentum is also affected by systematic uncertainties, the energy loss into neutrinos can never be reconstructed exactly. Because of the ubiquitousness of b final states in ATLAS, these processes have been studied extensively, and jet-correction algorithms exist which account for this energy loss.

4 Higgs Decay Modes

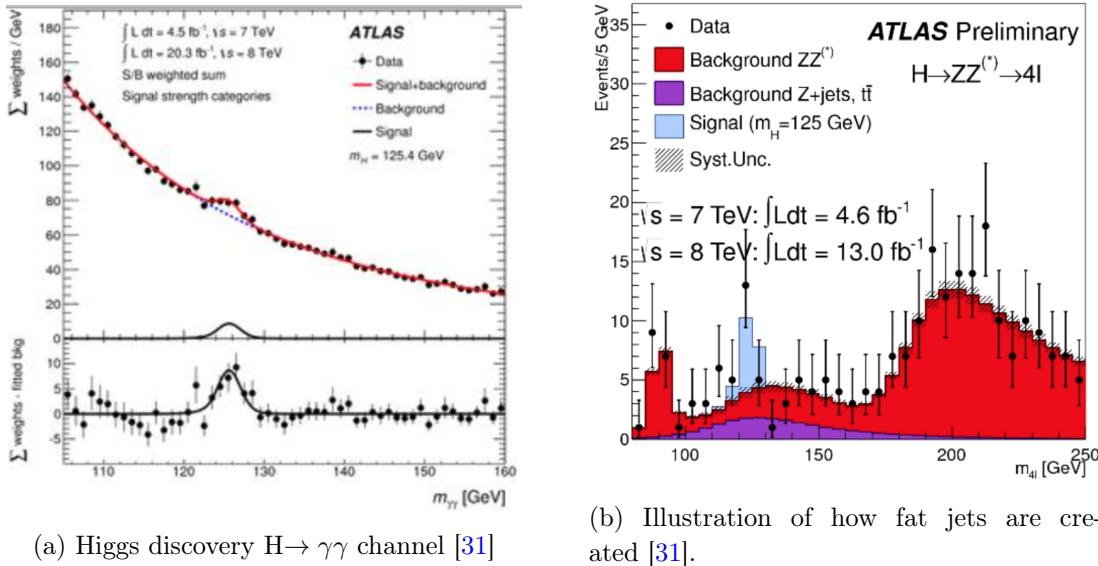


Figure 20: ATLAS figures for the channels used for the 2012 Higgs boson discovery

The Higgs discovery was made using the “clean” decay modes $H \rightarrow \gamma\gamma$ and the “golden channel” with a Higgs decaying to four leptons, displayed in Fig. 20 [31]. The branching ratios (BR) for the Higgs are given in Fig. 21. Note that the $H \rightarrow ZZ$ will have an additional suppression factor when each of the Zs is required to decay leptonically. Although the $H \rightarrow \gamma\gamma$ and $H \rightarrow ZZ$ modes have low branching ratios, they were used for the discovery mode because they involve particles that readily deposit most of their energy in the ECAL (and muon chambers for the leptonic muon final states) in a well-resolved way, and the background to these processes is low.

Studies have been done to try to observe the $H \rightarrow b\bar{b}$ process in association with a vector boson [39], in association with a hadronically decaying top [40], and in the di-Higgs to 4 b decay [41]. None of these searches were able to detect the signal because of the production rate of those channels (see Fig. 23) and the difficulty discriminating b -quarks from the large amount of QCD jet background at a hadronic collider. The data versus background and signal hypotheses plot for the $pp \rightarrow ZH \rightarrow llb\bar{b}$ search is shown in Fig. 22. The data did not peak significantly above the background at m_H .

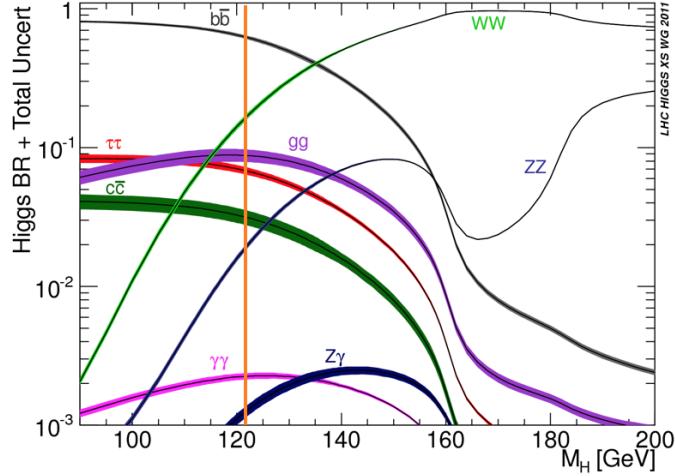


Figure 21: Branching ratios for the various decay channels of the Higgs, where the orange line at 125.9 GeV was added to emphasize the true branching ratios.[4]

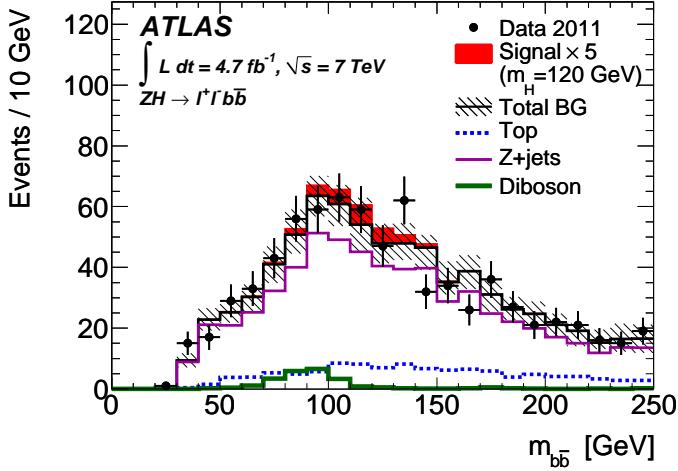


Figure 22: Reconstruction of the di-jet mass of the two b quarks in the $ZH \rightarrow ll\bar{b}\bar{b}$ search. The dominant backgrounds, and the expected signal (multiplied by a factor of 5) for $m_H = 120$ GeV are shown [39].

The current global measured mass of the Higgs is $m_H = 125.09$ GeV [42], and the predicted branching ratio of $H \rightarrow b\bar{b}$ is 0.568 ± 0.013 [43]. Investigating jet-substructure and novel algorithms for selecting or discarding jets can be helpful in sifting out this challenging SM signal.

Standard Model Total Production Cross Section Measurements

Status: Nov 2015

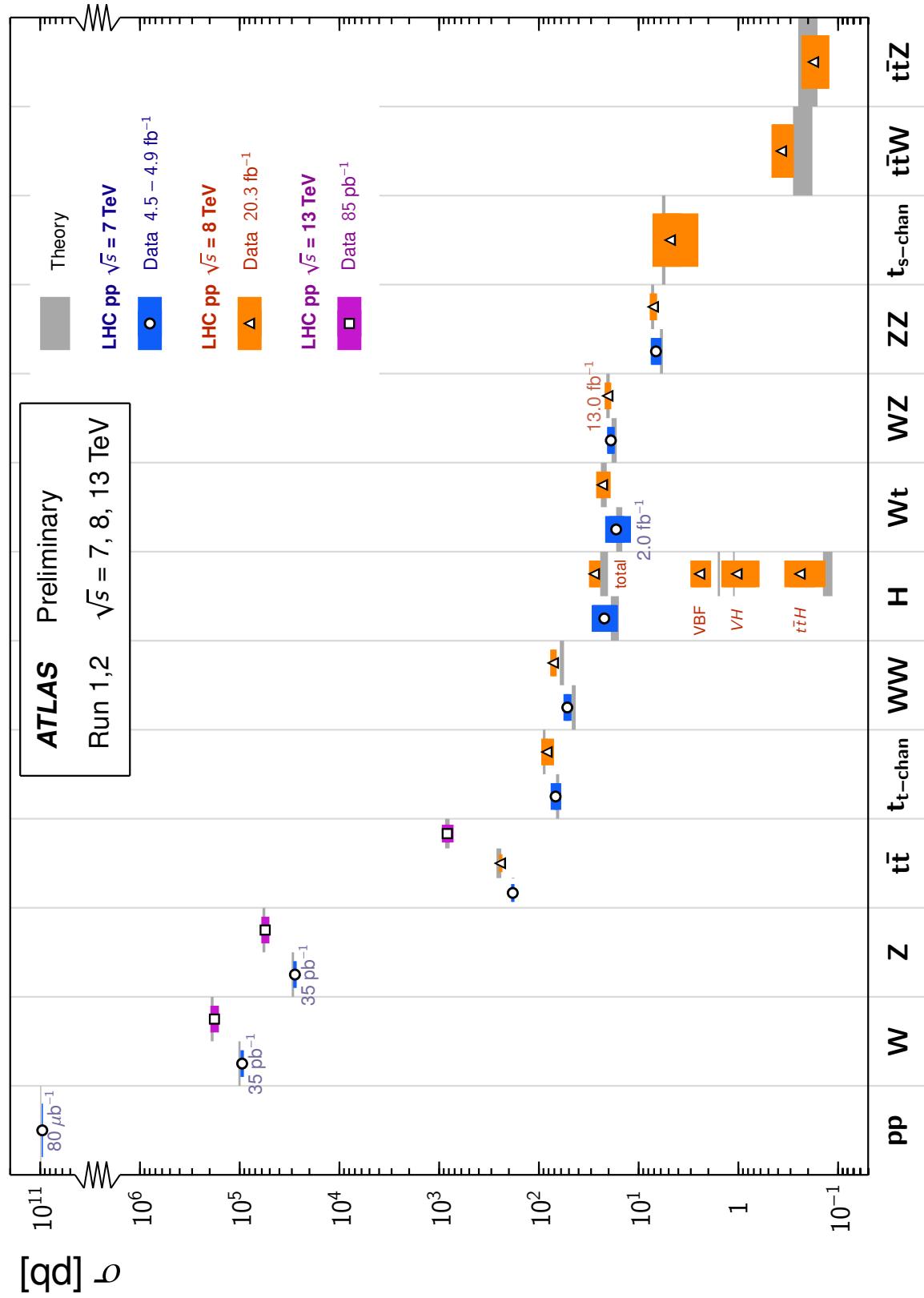


Figure 23: The cross sections for SM processes at the LHC, illustrating that weak and Higgs production processes are orders of magnitude rarer than hadronic ones. [44].

5 Energy Correlation Functions

In Ref. [5], the authors suggested the use of Energy Correlation Functions (ECF) – and in particular, two discriminating variables they call C_1 and C_2 – to help discriminate quark-initiated jets from each other and from gluon-initiated jets. Below, the relevant information from this paper is summarized.

The authors introduce generalized ECFs that can identify jets with N-subjets without needing a subjet finding feature [5]. The N-point ECF that they introduce is defined in Eq. (5.1).

$$ECF(N, \beta) = \sum_{i_1 < i_2 < \dots < i_N \in J} \left(\prod_{a=1}^N p_{T_{ia}} \right) \left(\prod_{b=1}^{N-1} \prod_{c=b+1}^N R_{i_b i_c} \right)^\beta \quad (5.1)$$

where J is the complete “fat jet” under consideration, i_j (with $j \in \{1, \dots, N\}$) labels the N subjets in J , and

$$R_{ij}^2 = (y_i - y_j)^2 + (\phi_i + \phi_j)^2, \quad (5.2)$$

$$y_i = \frac{1}{2} \ln \frac{E_i - p_{zi}}{E_i + p_{zi}}. \quad (5.3)$$

Since the rapidity y_i becomes the pseudo-rapidity in the relativistic limit (as is the case in LHC analyses), in this analysis I use pseudo-rapidity.

The first few N-values that are of interest to this analysis are:

$$ECF(0, \beta) = 1 \quad (5.4)$$

$$ECF(1, \beta) = \sum_{i \in J} p_{Ti} \quad (5.5)$$

$$ECF(2, \beta) = \sum_{i < j \in J} p_{Ti} p_{Tj} (R_{ij})^\beta \quad (5.6)$$

$$ECF(3, \beta) = \sum_{i < j < k \in J} p_{Ti} p_{Tj} p_{Tk} (R_{ij} R_{jk} R_{ik})^\beta. \quad (5.7)$$

When trying to construct the N^{th} ECF with less than N subjets, the ECF ideally would return 0. However, as the algorithm clusters the number of subjets in a fat jet, it will not be 100% efficient at identifying the true number of subjets, but realistically a fat jet with N -subjets should have $ECF(N+1, \beta)$ much smaller than $ECF(N, \beta)$. This motivates the use of $r_N^{(\beta)}$ defined as

$$r_N^{(\beta)} = \frac{ECF(N+1, \beta)}{ECF(N, \beta)} \quad (5.8)$$

If $r_N^{(\beta)} \ll 1$, then the jet is more likely to have N subjets than $N+1$. From there, the authors defined the test-statistic C_N as a ratio [5]

$$C_N^{(\beta)} = \frac{r_N^{(\beta)}}{r_{N-1}^{(\beta)}} = \frac{ECF(N+1, \beta) ECF(N-1, \beta)}{ECF(N, \beta)^2}. \quad (5.9)$$

For a jet with N subjets, C_N is expected to be significantly larger than C_{N-1} , and the authors note that “ C_N effectively measures higher-order radiation from leading-order (LO) substructure” [5].

5.1 C_1 Discriminating Variable

The authors suggest using C_1 for discriminating between quarks and gluons, where the fat-jet is expected to have only one subjet [5].

$$C_1^{(\beta)} = \frac{ECF(2, \beta)ECF(0, \beta)}{ECF(1, \beta)^2} = \frac{\sum_{i < j \in J} p_{Ti} p_{Tj} (R_{ij})^\beta}{\left(\sum_{i \in J} p_{Ti}\right)^2} \quad (5.10)$$

The C_1 distributions from Ref. [5] from with $\beta_{C_1} = 0.2$ for quarks and gluons are shown in Fig. 24. The gluon distribution is shifted toward larger C_1 values compared with the quark

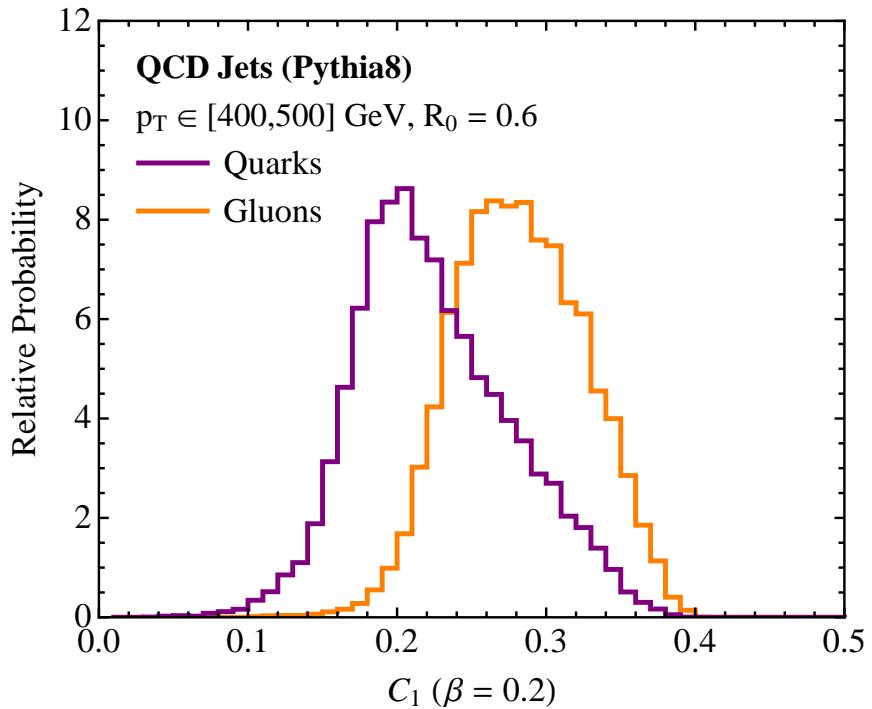


Figure 24: C_1 distributions for quark vs. gluon discrimination in the boosted regime where $\beta_{C_1} = 0.2$ [5].

distribution for very high p_T fat jets. In their energy regime (see Fig. 24) their discrimination power was optimized when $\beta_{C_1} = 0.2$, but they claimed that the discrimination would only depend on β beyond LO [5], implying LO analyses should not need to perform a β optimization for the C_1 test-statistic.

5.2 C_2 Discriminating Variable

The C_2 variable is useful for identifying two hard subjets in a fat jet and is applicable to analyses with boosted W s, Z s, and H s.

$$C_2^{(\beta)} = \frac{ECF(3, \beta)ECF(1, \beta)}{ECF(2, \beta)^2} = \frac{\left(\sum_{i < j < k \in J} p_{T_i} p_{T_j} p_{T_k} (R_{ij} R_{jk} R_{ik})^\beta \right) \left(\sum_{i \in J} p_{T_i} \right)}{\left(\sum_{i < j \in J} p_{T_i} p_{T_j} (R_{ij})^\beta \right)^2} \quad (5.11)$$

Unlike b-tagging, which requires the use of the tracking system, ECFs can be defined even when there is no tracking information (i.e., in the forward region).

6 Applying ECFs to ATLAS $pp \rightarrow HH \rightarrow b\bar{b}\gamma\gamma$ simulation

The authors suggested these discrimination variables be used with fat jets to identify substructure. However, we were interested in looking at these same variables in a different context, to see whether C_1 and C_2 could also help reject non b -quark jets and identify the correct di-jet pair coming from a SM Higgs decay even in high jet multiplicity environments. Since we are in a different regime than the authors considered, we need not expect our results to agree qualitatively with theirs as we are only concerned with the discrimination power of C_1 and C_2 .

For my analysis, I used a LO Monte Carlo (MC) sample $pp \rightarrow X \rightarrow HH \rightarrow b\bar{b}\gamma\gamma$, where the $m_X = 275$ GeV*. I used reconstructed anti- k_T truth jets in this study, and since the new resonances has low mass, the b -jets will still be resolved.

The reconstructed jets considered were required to have a $p_T > 25$ GeV and $|\eta| < 2.7$. In simulation, the jets will be labeled as coming from a b -quark if the jet has a b -hadron inside it. First I considered the comparison of the C_1 test-statistic for the b -quarks, versus the other jets in the sample, and used this distribution to implement a cut on C_1 .

We are interested in finding the di-jet pairs with two b -jets, so the $2b$ label corresponds to a di-jet pair having two b -jets, the $1b$ label corresponds to the di-jet with only one b -jet, and the $0b$ label corresponds to a di-jet not having any b -jets. Since the β_{C_2} parameter tunes the C_2 distribution, I then picked β_{C_2} to optimize the discrimination power of C_2 between the $2b$ signal and $0b$ background regions. Then I developed the Higgs Di-Jet Likelihood Ratio (HDLR) (explained in Section 6.2.2) to use the C_2 values for a given event to identify the most promising $H \rightarrow b\bar{b}$ jet pair candidate. Finally, I choose the least likely di-jet pair in the event to form a control region as a check on the HDLR method.

6.1 C_1 Discriminating Variable

First, for each jet I calculated the C_1 test-statistic. As the C_1 definition requires the $ECF(2, \beta)$ calculation, jets with fewer than 2 constituents (e.g., a neutral cluster of energy or a charged track) were discarded. Then I divided these variables into two categories:

*The MC Channel ID number for this sample is 341173

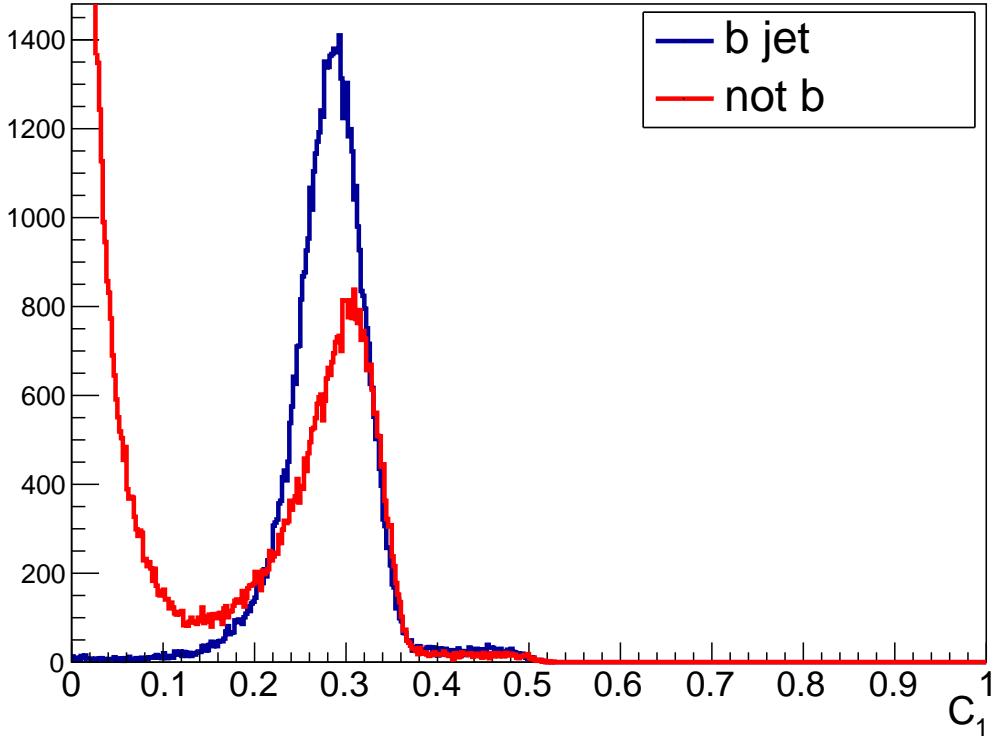


Figure 25: C_1 overlay plot for $\beta_{C1} = 0.2$, which motivates a lower C_1 cut of 0.21.

1. Jets originating from a b-quark
2. Jet that did not originate from a b-quark.

This is a different application of C_1 than the way the authors of this paper had motivated it. We were interested specifically in identifying b quarks for this analysis, and Fig. 25 shows the distinction of these two C_1 distributions. Clearly, the non-b-jet C_1 distribution contains a large number of small C_1 values, motivating a lower cut of $C_1 > 0.21$, the C_1 value for which the two distributions cross.

6.2 C_2 Discriminating Variable

Next, I wanted to use C_2 to isolate the di-jet pairs most likely to originate from $H \rightarrow b\bar{b}$, then formed all possible unique jet pairs and divided them into three separate distributions:

1. Both of the jets originated from b quarks, “2b”
2. Only one of the jets originated from a b quark, “1b”
3. Neither of the jets originated from a b quark, “0b”.

We use the ratio of the probability density distributions to help exaggerate the discrepancies between the variables. Since my dominant background for 2b is going to be the 0b

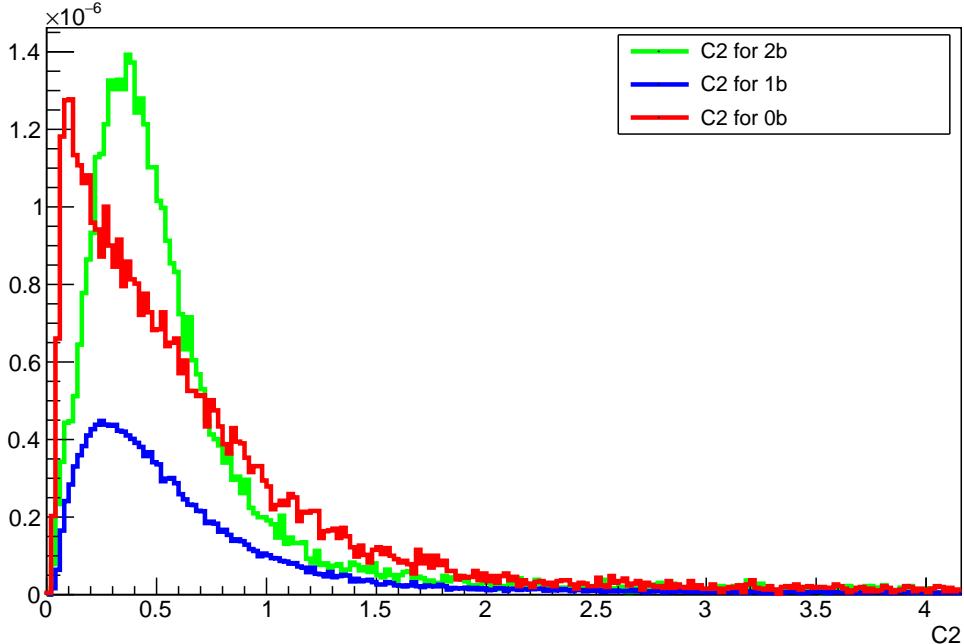


Figure 26: Overlay the PDFs for the C_2 distributions, $\beta_{C_2} = 1.2$.

events, to optimize β_{C_2} , I simplified the problem to a binary one by just considering the shape of the $0b$ and $2b$ C_2 distributions, called $P_0(C_2)$ and $P_2(C_2)$, respectively. Then I considered the distributions $L_2(x)$ and $L_0(x)$ defined as

$$L_2(C_2) = \frac{P_2(C_2)}{P_0(C_2) + P_2(C_2)} \quad L_0(C_2) = \frac{P_0(C_2)}{P_0(C_2) + P_2(C_2)} \quad (6.1)$$

The variable L_2 is called the Higgs Di-jet Likelihood Ratio (HDLR). In an ideal case where $P_2(C_2)$ and $P_0(C_2)$ do not overlap, a $2b$ di-jet pair would have an $L_2 = 1$ and $L_0 = 0$, and a $0b$ di-jet pair would have an $L_2 = 0$ and $L_0 = 1$. When creating the L_2 and L_0 distributions, I used the flavors in the di-jet pair to divide L_2 and L_0 into sub-cases. L_2^{2b} corresponded to an L_2 calculated for the di-jet pair of two real b -jets, and L_0^{2b} corresponded an L_2 calculated for a di-jet pair where neither jet was a true b -jet. A similar definition held for L_0^{2b} and L_0^{0b} .

6.2.1 Optimization of β_{C_2}

After the distributions $P_2(C_2)$ and $P_0(C_2)$ are obtained, C_2 was calculated for each di-jet pair, and $P_0(C_2)$ and $P_2(C_2)$ were used to calculate L_2 and L_0 . Then to quantify the difference between these distributions, I used the L_2^{2b} and L_0^{2b} distributions, shown for $\beta_{C_2} = 1.2$ in Fig. 27. From their definitions, the L_2^{2b} is expected to be shifted to the right of L_0^{2b} . First I found the value, a , such that

$$\int_a^1 N(L_2^{2b}) dL_2^{2b} = 0.9 \quad (6.2)$$

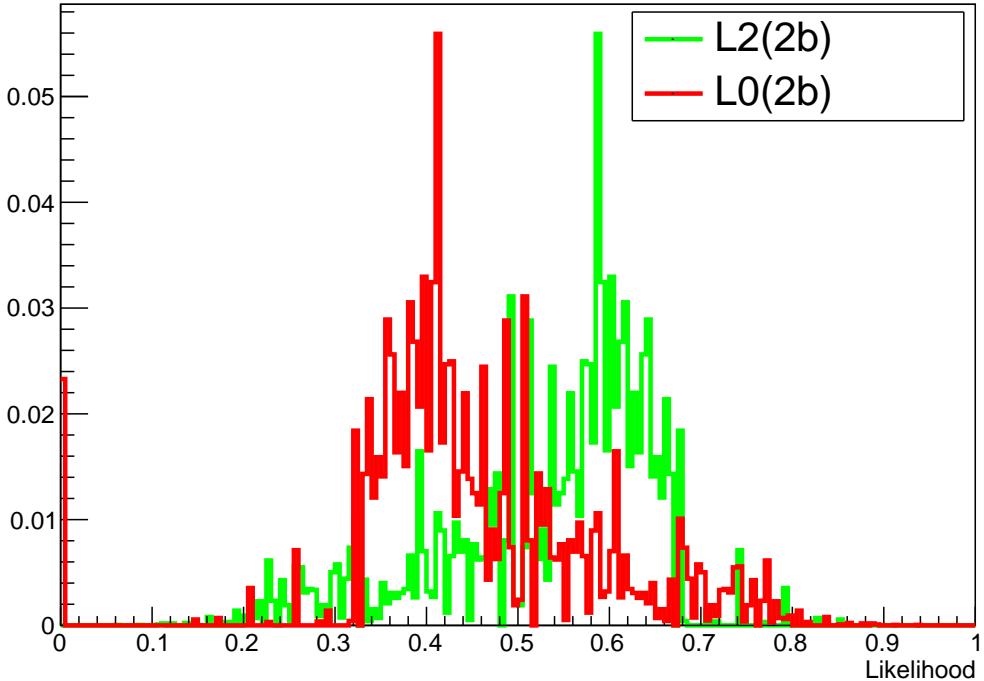


Figure 27: Likelihood functions for $2b$ events with $B_{C_2} = 1.2$

where a is the lower limit on the integral that encloses 90% of the L_2^{2b} curve. Then to quantify the amount of $0b$ background incorporated in such an L_2 cut, I calculated the “efficiency,” ε_{0b}

$$\varepsilon_{0b} = \int_a^1 N(L_2^{0b}) dL_2^{2b} \quad (6.3)$$

where the lower limit a was calculated from Eq. (6.3). Then I found ε_{0b} for $\beta_{C_2} = 0.0 - 2.2$ in increments of 0.2. To minimize the background, I wanted to minimize the $0b$ efficiency. In Fig. 28, the efficiency plateaus for β_{C_2} values greater than 0.8, so any β_{C_2} value larger than 0.8 will have comparable discrimination power. Therefore, I chose $\beta_{C_2} = 1.2$.

6.2.2 Reconstructed Mass

Finally, I used the HDLR to pick the most likely di-jet pairs to originate from two b -quarks. Because this is a $HH \rightarrow b\bar{b}\gamma\gamma$ sample, every event must have a true di-jet $b\bar{b}$ pair in it originating from a Higgs, if the b -jets are within the acceptance of the ATLAS detector.

In each event, I selected the jet pair with the maximal value of L_2 such that $L_2 > a$ (from Eq. (6.2)), and designated this the “ $H \rightarrow b\bar{b}$ candidate” I then plotted the invariant mass of this pair in Fig. 29. As a control, I also selected the pair with the minimum value of L_2 such that $L_2 < a$. For $\beta_{C_2} = 1.2$, $a = 0.3575$. Cutting L_2 based on a ensures that a di-jet pair cannot contribute to both the signal and control distributions, and eliminates a significant proportion of the background as the area under the control

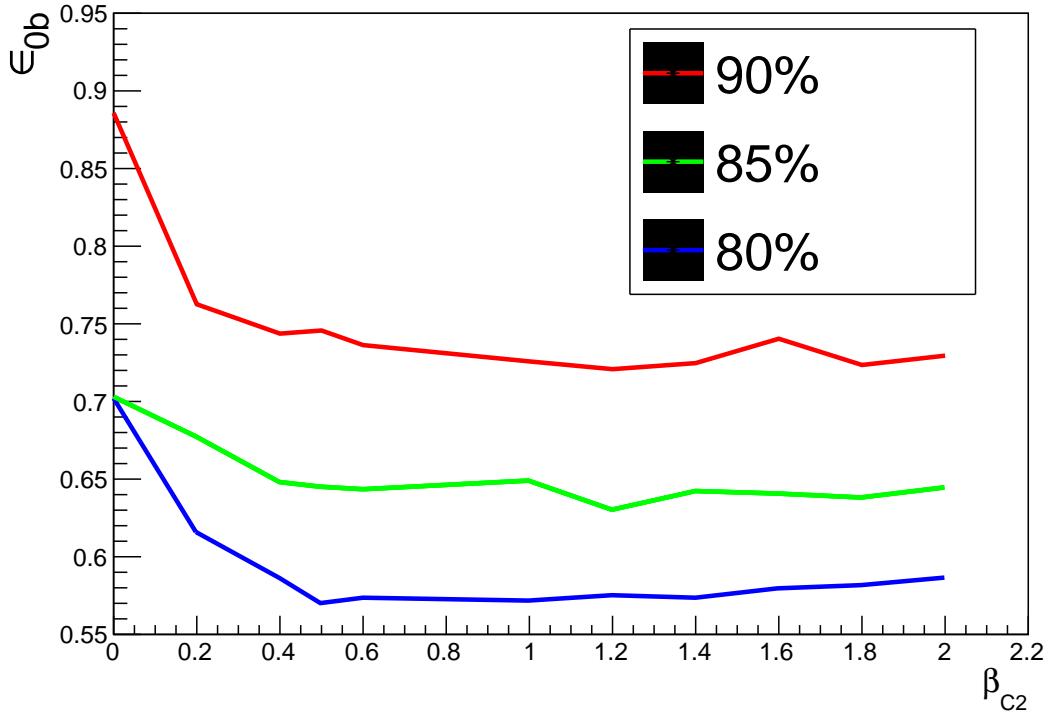


Figure 28: ε_{0b} from Eq. (6.3), where the red (top) curve is for integrating L2(2b) at the 90% CL, the green (middle) curve integrated L2(2b) at a 85% CL, and the blue (bottom) curve integrated L2(2b) at a 80% CL. We want to choose a β_{C2} that minimizes ε_{0b} .

(red) curve is much less than the area under the signal (green) curve. The control sample is mostly unstructured, although a small bump at the Higgs mass is evident. The average value for the signal shape is 119.6 GeV, a little less than the mass of the Higgs, for reasons stated in Section 3.1

To investigate the bump at the Higgs mass further, I decomposed the control region from Fig. 29 into the flavor tags of the di-jet pairs, shown in Fig. 30. We can see that sometimes the 2b events do have the smallest HDLR value, although the control region is mostly composed of 1b and 0b events.

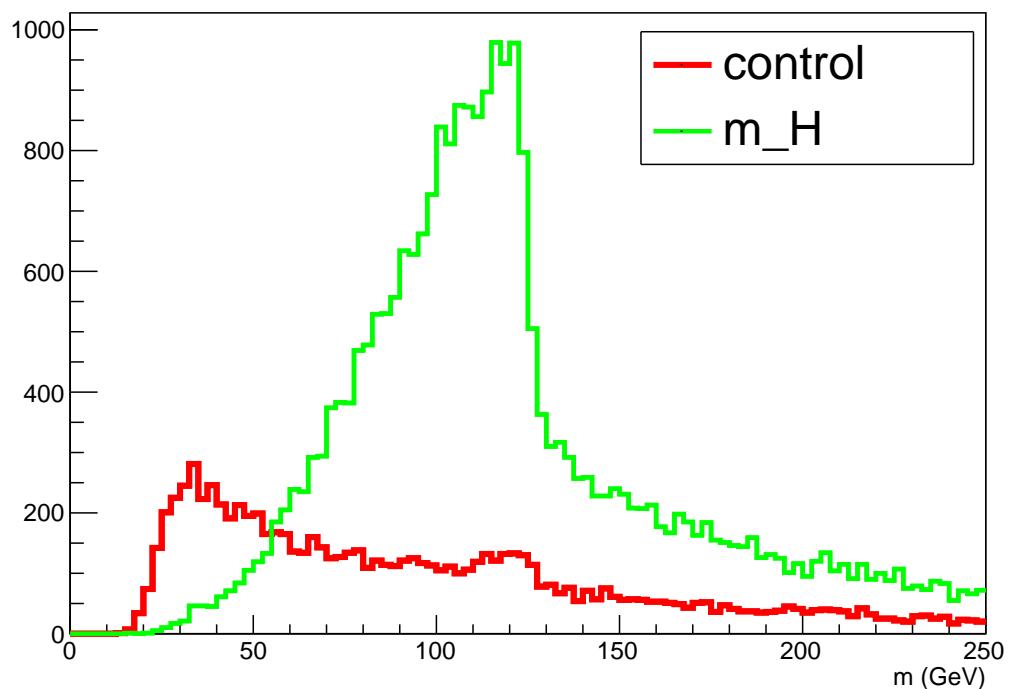


Figure 29: Reconstructed mass that maximized the likelihood that the di-jet pair originated from two bs (signal) and minimized the likelihood that the di-jet pair originated from two bs (control).

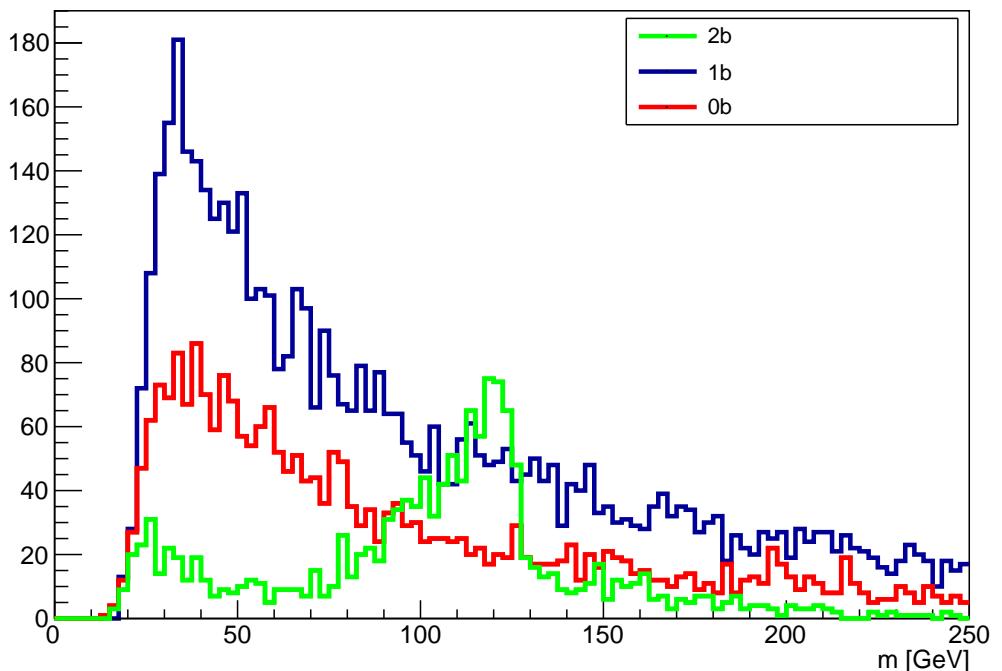


Figure 30: Breaking up the control curve from Fig. 29 to the three cases for the true di-jet composition. We can see that some $2b$ di-jets sometimes have the min HDLR value, which contributes to the small peak at the Higgs mass for the control curve in Fig. 29.

7 Future Directions

If time permitted, below are the next steps.

- I would apply the ATLAS jet energy corrections software package to shift the reconstructed mass closer to the true Higgs mass, based on effects mentioned in Section 3.1. Then I would run on an independent MC sample, and make sure that the HDLR algorithm still reconstructs the invariant mass of the Higgs. A final algorithm test would be to run over ATLAS’s 13 TeV data set, although it is not likely that the algorithm will be able to detect the signal over the background at this stage.
- One of the motivations for using ECFs was that without b -tagging, we would not be restricted to make an $|\eta| < 2.4$ cut since we are no longer relying on b -tagging to identify b -quarks. It would be interesting to loosen the η cut and look in the forward region of the detector since we have calorimetry out to $\eta = 4.9$.
- On a more technical note, I saved the C_2 variables in a histogram that could hold values up to 10.0, and discarded C_2 values that were larger than 10.0. From Fig. 26, it is clear that the C_2 distributions fall off rapidly for larger C_2 values, so this “cut” is probably insignificant. However, it would be more accurate to save the C_2 values as NTuples, so that one does not need to know ahead of time how large to set the bounds on the C_2 histogram.
- Finally, just as we ran over a range of β_{C_2} values to optimize the algorithm, β_{C_1} could be optimized as well instead of just accepting the value 0.2 recommended by the authors, since they were considering a boosted regime different than our SM analysis.

8 Conclusions

We developed a new variable, the HDLR, based on ideas in Ref. [5]. We used the HDLR variable to identify the di-jet pair in an event to have originated from a Higgs. This method allowed us to find a peak at the Higgs mass in simulation without significantly shaping the background. This result shows promise, and has the added benefit of allowing the detection of b -jet in the forward region of the detector where traditional b -tagging algorithms fail. We suggest further avenues for investigating this idea.

9 Acknowledgements

Multitudes to thank Dr. Sekula for mentoring me for the past three years and teaching me how to be a physicist! I’d like to also thank grad students Jeff Hetherly and Matthew Feickert for helping me with code and concepts and modeling what being a successful grad student looks like. Thanks to Maddie McKay for her inspiration through starting this project by applying the ECF Paper [5] to this analysis and giving me her code when I started in the fall. Finally, I’d also like to thank Dr. Cooley for being a great teacher and showing me that being a physicist was an actual career path. Also thank you to the Hamilton and Engaged Learning programs for funding this project.

References

- [1] "The Higgs boson." CERN, <http://home.cern/topics/higgs-boson>, accessed 11 April 2016.
- [2] "The Standard Model of Particle Physics." *Symmetry Magazine*.
<http://www.symmetrymagazine.org/standard-model/>, accessed 11 April 2016.
- [3] Matthew J. Dolan, Christoph Englert, Michael Spannowsky. "Higgs self-coupling measurements at the LHC" 10.1007/JHEP10(2012)112.
- [4] John F. Gunion, Howard E. Haber, Gordon Kane, and Sally Dawson. *The Higgs Hunter's Guide* Front. Phys, Vol 80 (2000).
- [5] Andrew J. Larkoski, Gavin P. Salam, and Jesse Thaler, "Energy Correlation Functions for Jet Substructure." arXiv:1305.0007v3 [hep-ph] 10 Jul 2013.
- [6] A. Salam, in Elementary Particle Theory, p. 367. Almqvist and Wiksell, Stockholm, 1968.
- [7] P. W. Higgs, "Spontaneous symmetry breakdown without massless bosons," Phys. Rev. 145 (1966) 1156D1163.
- [8] Englert, François; Brout, Robert (1964). "Broken Symmetry and the Mass of Gauge Vector Mesons." Physical Review Letters 13 (9): 321-23.
- [9] P. W. Higgs, "Broken symmetries and the masses of gauge bosons," Phys. Rev. Lett. 13 (1964) 508-509.
- [10] Guralnik, Gerald; Hagen, C. R.; Kibble, T. W. B. (1964). "Global Conservation Laws and Massless Particles." Physical Review Letters 13 (20): 585-587.
- [11] wikipedia image, "Higgs Mechanism." https://en.wikipedia.org/wiki/Higgs_mechanism Accessed 15 April 2016.
- [12] S. Dawson, "Introduction to electroweak symmetry breaking," arXiv:hep-ph/9901280 [hep-ph].
- [13] S. Weinberg, "A model of leptons," Phys. Rev. Lett. 19 (1967) no. 21, 1264 - 1266.
- [14] Y. Grossman. "Introduction to the Standard Model." *CERN Summer Student Lectures* 4 July 2014. <https://indico.cern.ch/category/345/>, Accessed April 16 2016.
- [15] "The Nobel Prize in Physics 2013."
http://www.nobelprize.org/nobel_prizes/physics/laureates/2013/ Accessed 25 Dec 2015.
- [16] "The Large Hadron Collider." CERN, <http://home.cern/topics/large-hadron-collider>. Accessed 11 April 2016
- [17] G. Aad, E. Abat, J. Abdallah, *et al.*, "The ATLAS Experiment at the CERN Large Hadron Collider." Journal of Instrumentation, Vol. 3, August 2008.
- [18] "LHC Collisions - CERN." cern.ch/lhc-machine-outreach/collisions.htm. Accessed 15 April 2015.
- [19] CMS Collaboration. "Limits on the Higgs boson lifetime and width from its decay to four charged leptons." arXiv:1507.06656v2.
- [20] "Pseudorapidity." wikipedia article. Accessed 15 April 2016.
- [21] S. Zenz "How a Pixel Detector Works." Imperial College London Blog.
<http://www.quantumdiaries.org/2008/07/25/how-a-pixel-detector-works>. Accessed 4 Oct. 2015.

- [22] S. Stapnes. “Detector Challenges at the LHC.” *Nature*, Vol. 448. 19 July 2007.
- [23] G. L. Aad, M. Ackers, F. A. Aleppo, *et al.*, “ATLAS Pixel Detector Electronics and Sensors,” *Jinst*. 3 (2008) P07007.
- [24] Romaniouk, Anatoli. “ATLAS TRT (Transition Radiation Tracker).” <http://particle.mephi.ru/en/research/atlas/trt/>. Accessed 23 Dec 2015.
- [25] James E. Brau, John A. Jaros, and Hong Ma. “Advances in Calorimetry.” PH-EP-Internal-Technical-Note-2010-013.
- [26] Makund Gupta. “Calculation of Radiation Length in materials.” *Annual Reviews*, Vol 41 (1991).
- [27] Halkiadakis, Eva. “Introduction to the LHC Experiments.” arXiv:1004.5564.
- [28] ATLAS Collaboration. *Technical Design Report: Liquid Argon*, Ch 1 and Ch 6, 1996.
- [29] ATLAS Experiment - Public Results. “Trigger Operation Public Results.” https://twiki.cern.ch/twiki/bin/view/AtlasPublic/TriggerOperationPublicResults#2015_pp_at_13_TeV. Accessed 14 April 2016.
- [30] T. Shears. “The Particle World: an introduction to particle physics.” *CERN Summer Student Lectures* 1 July 2014. <https://indico.cern.ch/category/345/>, Accessed April 16 2016.
- [31] P. Mattig. “Standard Model at Hadron Colliders: IV Higgs Boson.” *CERN Summer Student Lectures* 24 July 2014. <https://indico.cern.ch/category/345/>, Accessed April 16 2016.
- [32] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. “The anti-k_T jet clustering algorithm.” *JHEP* 0804:063, 2008.
- [33] ATLAS Collaboration, “Performance of jet substructure techniques for large-R jets in proton-proton collisions at $\sqrt{s} = 7$ TeV using the ATLAS detector,” *JHEP* 09 (2013) 076 , arXiv:1306.4945 [hep-ex]
- [34] The LHCb Collaboration, “b-hadron lifetime measurements with exclusive $b \rightarrow J/\Psi X$ decays reconstructed in the 2010 data.” *LHCb-ANA-2011-001*.
- [35] “Performance of the ATLAS Secondary Vertex b-tagging Algorithm in 7 TeV Collision Data,” Tech. Rep. ATLAS-CONF-2010-042, CERN, Geneva, June 2010.
- [36] ATLAS Collaboration, T. Gopfert, *et. al.* “Tagging b-jets in ATLAS.” Prepared for 24th International Symposium on Lepton-Photon Interactions at High Energy (LP09), Hamburg, Germany, 17-22 Aug 2009.
- [37] Nagai, Yoshikazu. “b-tagging in ATLAS.” CPPM Marseille.
- [38] Particle Data Group Collaboration, Z. Chai, *et al.* *Review of Particle Physics*, Vol 38, No. 9, 2014.
- [39] ATLAS Collaboration. “Search for the Standard Model Higgs boson produced in association with a vector boson to a b-quark pair with the ATLAS detector.” *Phys. Lett. B* 718 (2012).
- [40] ATLAS Collaboration. “Search for the Standard Model Higgs boson decaying to $b\bar{b}$ produced in association with top quarks decaying hadronically in pp collisions at $\sqrt{(s)} = 8$ TeV with the ATLAS detector.” *Eur. Phys. J. C* (2015) 75:349.
- [41] ATLAS Collaboration. “Searches for Higgs boson pair production in the $hh \rightarrow bb\tau\tau, \gamma\gamma WW^*, \gamma\gamma bb, bbbb$ channels with the ATLAS detector.” *Phys. Rev. D* 92, 092004 (2015).

- [42] The ATLAS and CMS Collaborations. “Combined Measurement of the Higgs Boson Mass in pp collisions at $\sqrt{s} = 7$ and 8 TeV with the ATLAS and CM Experiments” Phys. Rev. Lett. 114, 191803 (2015).
- [43] CERN twiki. “SM Higgs Branching Ratios and Total Decay Width (update in CERN Report4 2016)”
<https://twiki.cern.ch/twiki/bin/view/LHCPhysics/CERNYellowReportPageBR> Accessed 14 April 2016.
- [44] ATLAS Collaboration. “Summary plots from the ATLAS Standard Model physics groups.”
<https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/CombinedSummaryPlots/SM/>
Accessed 15 April 2016.