The Impact of Quark/Gluon Tagging on Searches for Dark Matter in Dijet Event Data with the ATLAS Detector

Lancaster University

Joe Peskett

Department of Physics,
Lancaster University

November 2020

Abstract

Differentiating between quark-and-gluon initiated events can be a useful tool for increasing the signal to noise ratio in searches of new physics. This thesis presents an analysis of the impact of quark-gluon tagging on the dijet invariant mass spectrum. Monte Carlo simulated data were compared with the ATLAS data from Run-2 of the LHC at 13 TeV. Truth information from the simulations was used to make a simple quark-gluon tagger based on the sum of the multiplicities of the two leading jets in a dijet event. A range of gluon-gluon selection efficiencies were applied to a H′ signal and 95% CL upper limits were found and compared between tagged and untagged samples. No improvements in the H′ upper limits were observed for any tested efficiency.

Declaration of Authorship

The work contained in this thesis is the result of research conducted by the author, however, would not be possible without the work of others from Lancaster University and the ATLAS collaboration. All figures and text that are not a result of the research of the author are appropriately cited.

Acknowledgements

Thank you to my supervisor, Prof. Bertram for all the help and advice he offered throughout my research. I would also like to express my gratitude to the others in the Particle Physics department who have been able to offer guidance and time to support my development and understanding.

Contents

1 The Standard Model

Taking its current form in the mid 1970s', the standard model of particle physics seeks to combine the theory of quantum chromodynamics (QCD) with that of the electroweak interaction, describing three of the four known fundamental forces of nature. So far, the model has withstood every experimental test. Since its conception, the standard model has successfully predicted the existence of the Z^0 and $W^{\pm}$ bosons (UA1 and UA2 Experiments, 1983)[1, 2], the top quark (Fermilab, 1995)[3, 4], the τ neutrino (Fermilab, 2000)[5] and perhaps most significantly the Higgs boson (LHC, 2012)[6, 7] along with many of their properties to a high precision.

Despite its success, there remain many problems which are not accounted for by the model. For instance, the theory does not include gravity, nor does it account for the observed matter-antimatter asymmetry in the universe. Importantly it does not provide a suitable candidate to explain the presence of dark matter, which constitutes 26.8% of the energy density of the universe [8].

The standard model is a quantum field theory which describes the interactions between all fundamental particles. Broadly these particles can be categorised into bosons and fermions shown in Figure 1. The elementary bosons are integer spin force carriers which obey Bose-Einstein statistics. This has the implication that there is no theoretical limit on the number of bosons that can occupy a single state. There are five known fundamental bosons, including four vector (spin 1) gauge bosons, the photon, gluon, $W^{\pm}$ and Z^0 bosons, and one scalar (spin 0) Higgs boson. In addition to this a tensor (spin 2) graviton is hypothesised to incorporate gravity into an extension of the standard model.

Fermions on the other hand are matter particles which follow Fermi-Dirac statistics and thus no more than one can exist in the same state. The fundamental fermions can be broken down into two groups of half integer spin (in units of Planck's constant $\hbar$) particles - quarks and leptons - each consisting of three generational pairs. Quarks, unlike leptons, possess colour charge and therefore interact via the strong nuclear force in addition to the other fundamental forces. This allows them to bind together to form composite particles
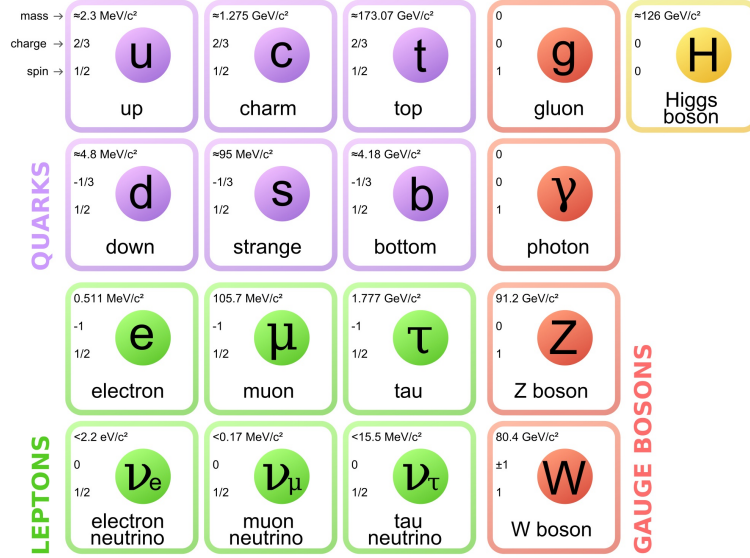
Figure (1): The structure of elementary particles that make up the standard model. Fermions are separated into leptons and quarks, each having three generations of increasing mass for the respective particle of a generation. The bosons are separated into the four gauge vector bosons and the Higgs scalar boson. [9].

in $q\bar{q}$ states called mesons, and qqq states called baryons, most commonly in the form of protons and neutrons. Leptons do not interact through the strong force and can be divided into charged leptons (the electron, muon and tau) and their corresponding neutral neutrinos.

2 Theoretical Construction

The standard model is a gauge theory of internal symmetry $SU(3) \times SU(2) \times U(1)$ and global Poincaré symmetry consisting of translational and rotational symmetries. Each particle in the theory is described in terms of a dynamical field that extends throughout space-time. All particles and their interactions are governed by the Electroweak, QCD, and Higgs sectors. The electroweak sector is a Yang-Mills gauge theory with the symmetry group $U(1) \times SU(2)_L$, combining electromagnetism with the weak interaction. The W$^\pm$ boson mediates the charged current (CC) weak interactions, with the symmetry subscript L referring to the fact that it couples to the left (chirality) component of quarks and leptons (and the right component of their antiparticles), while the Z^0 boson mediates the neutral current (NC).

The lagrangian of the electroweak sector does not contain a mass term and therefore

the Higgs field, a complex scalar of the group $SU(2)_L$, is introduced to cause spontaneous symmetry breaking during interactions due to its non zero vacuum expectation value. This results in the Brout-Englert-Higgs mechanism which generates the masses for the $W^\pm$ and Z^0 bosons through electroweak symmetry breaking, leaving the photon massless. The fermion masses on the other hand result from Yukawa coupling of the fermion field with the Higgs field [10, 11].

2.1   Quantum Chromodynamics

The QCD sector is a gauge theory of group SU(3) that defines the interactions between quarks and gluons. In many ways QCD is analogous to Quantum Electrodynamics (QED) however its field strength tensor (derived from the gluon field $\mathcal{A}_\alpha^A$) given in Eq. (1) contains a third "non-Abelian" term which allows triplet and quartic self-interactions between gluons, ultimately giving rise to the property of asymptotic freedom [12].

$$F_{\alpha\beta}^A = \partial_\alpha \mathcal{A}_\beta^A - \partial_\beta \mathcal{A}_\alpha^A - g f^{ABC} \mathcal{A}_\alpha^B \mathcal{A}_\beta^C \tag{1}$$

Here the superscripts A, B and C run over the eight colour degrees of freedom of the gluon field. As mentioned, all quarks have half integer spin, meaning spin 3/2 baryons (e.g. Δ^{++}) are symmetrical in their spin, space and SU(3) states. Total asymmetry of the wave function is required to obey Fermi-Dirac statistics which means it is necessary to introduce a colour degree of freedom, where each quark carries a colour charge, for example red, green or blue (A = 1, 2 or 3), in order to resolve this problem. Gluons on the other hand have a superposition of two colour states (one colour and one anti-colour) given by the Gell-Mann matrices. This results in eight possible linearly independent colour states for a gluon forming the nine effective states observed. Thus, the colour state of a quark can be rotated through the emission of a gluon i.e. $q_i \rightarrow g_{i\bar{j}} + q_j$. Experimental proof of colour degrees of freedom is seen in the decay $\pi \rightarrow \gamma\gamma$ and e^+e^- annihilation which both have reaction rates dependent on the number of colours [13].

No colour charged particle, such as a quark or gluon has ever been observed alone. Quarks are found only in bound states as colour neutral baryons or mesons. This property
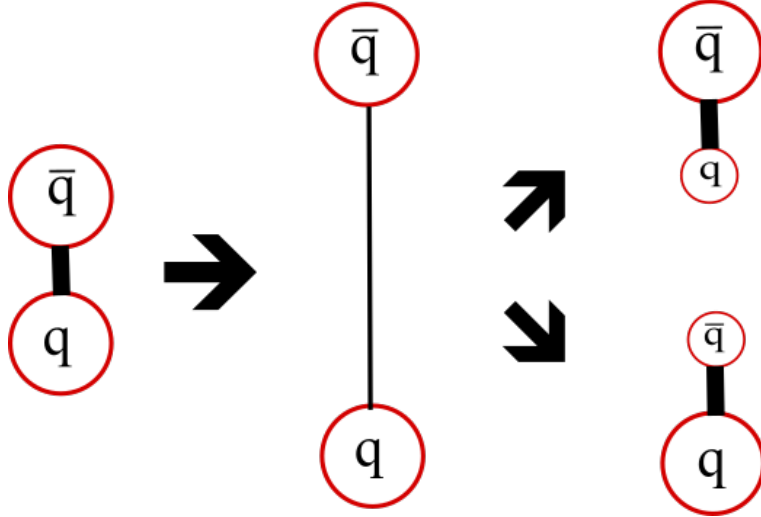
Figure (2): As two quarks are pulled a part, the potential between them becomes sufficient to create a new quark-antiquark pair. For this reason quarks can never be observed in isolation.

is known as confinement. There is no theoretical proof of confinement, however it is often described qualitatively as a string attached between quarks as shown in Figure 2. Whereas in QED the potential between two sources decreases with distance, the potential between two colour charged particles remains roughly constant as they are separated. As the separation between two sources becomes large enough, the string "snaps" as it becomes energetically favourable for a quark-antiquark pair to be created, forming two separate colour neutral bound states [14].

Inside a hadron, quarks are observed to possess asymptotic freedom. This is a property by which the interaction strength of the constituents of the colour neutral bound state gets asymptotically weaker. The QCD coupling constant as a function of the momentum transfer, q^2 between two interacting particles is [11]

$$\alpha_s(q^2) = \frac{\alpha_s(\mu^2)}{[1 + (\frac{11N - 2F}{3})\frac{\alpha_s(\mu^2)}{4\pi}\ln(\frac{q^2}{\mu^2})]}, \tag{2}$$

where μ^2 is some reference energy (for example the mass of the Z boson), N is the number of colours, and F is the number of flavours. It is seen that α_s decreases as Q^2 increases, and therefore as the distance between quarks decreases. Physically, this happens through a cloud of virtual gluons which carry colour charge causing anti-screening effects which result in the running coupling constant. This effect of asymptotic freedom is what
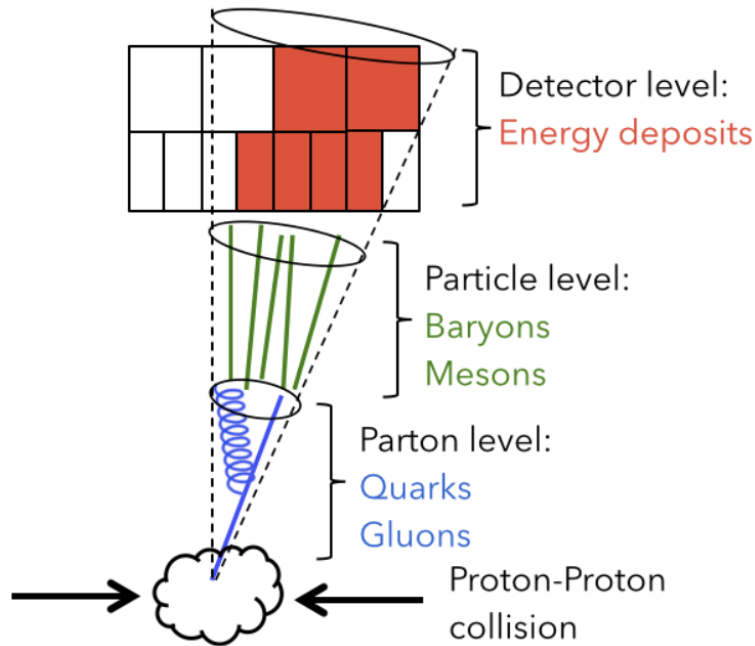
Figure (3): The formation of a jet from partons in an initial *pp* collision. High energy partons radiate gluons creating a collinear parton shower. [16].

allows the use of perturbation theory in the calculation of hard scattering processes in hadronic interactions.

2.2 Jet Phenomonology

At the Large Hadron Collider (LHC), bunches of protons collide inelastically with large amounts of energy being transferred between them. Within this collision, highly energetic parton-parton (i.e. quark or gluon) hard processes can be identified. This parton interaction can produce new elementary particles (e.g. a Higgs boson, a $t\bar{t}$ pair, new particles, etc.) with energy in the region of 100 GeV to several TeV. These elementary particles decay further to high-energy partons which radiate gluons, preferably with small energies collinear to their emitter. Radiated gluons can themselves create quark-antiquark pairs. The process continues to occur until the emitted parton energies become small (typically of the order 1 GeV), by which $\alpha_s(Q^2)$ becomes large and there is hadronisation between the partons. The partons are observed in the final state as a collimated bunch of hadrons (mostly pions and kaons) called a jet, which has properties proportional to the primary parton which formed the jet [15]. This process from the initial proton-proton collision to the final state hadrons observed in the detector is displayed in Figure 3.

Due to additional signals from other interactions overlapping with the final state hadrons which result from the primary parton, it is impossible to uniquely pinpoint the origin of the jet. Instead a jet algorithm must be introduced that determines which hadrons qualify as being included in the jet, this will be covered in more detail in Section 5.1. Once the jets are formed, the event can be characterised in terms of the resulting jets instead of the individual hadrons.

Jets observed from a large number of events can be used in the search for new hypothetical particles. If the theorised particle is expected to decay into other specific particles, events of this type are selected and the invariant mass of the expected decay particles is calculated. Resonances are then searched for with an invariant mass corresponding to the mass of the hypothesised particle, revealing itself as a local excess of events. If the new particle does not directly interact with the detector (eg. dark matter) the analysis may search for missing energy in the mass distribution or missing transverse momentum. Searches can broadly be split into theoretical-model-driven searches (eg. searches for SUSY particles, extra dimensional models, neutrino oscillations etc.) and phenomenological-model-driven searches (eg. dark matter particles, extended Higgs sector, higher generations of quarks, etc.). Theoretical model driven searches are intended to find particles which may provide evidence to support a new theoretical model including extensions of the standard model. Phenomenological model driven searches on the other hand are proposed to find new particles which could contribute to observations which are not currently described by theory alone [17, 18].

3 Dark matter

The particles of the standard model constitute only 5% of the total "cosmic energy budget" [19]. About 15% of the total baryonic matter consists of stars, the remaining mass exists in the form of the interstellar medium, composed of gas, dust and cosmic rays. However, cosmological evidence suggests that the total mass of most galaxy clusters is 10 - 100 times greater than that of the visible galaxies they contain.

This phenomenon was first identified by the velocity dispersion of stars orbiting the centre of the Milky Way. It was further noticed with galaxies in galaxy clusters, that the speed of the objects was far too small for the gravitational influence of the visible galaxies alone. More direct mass measurements from gravitational lensing, along with more precise mapping of the Comic Microwave Background (CMB) radiation have provided further evidence and insight into the 26% of the cosmic energy budget that consists of dark matter [19]. This evidence along with potential dark matter candidates and how they might be detected shall be discussed further in this section.

3.1 Rotation Curves of Spiral Galaxies

When studying the rotation curves of spiral galaxies, no Keplarian fall-off in velocities as a function of distance from the galactic nucleus is observed. The profiles of these curves, given in Figure 4, instead imply that the distribution of light coming from these galaxies does not match the corresponding distribution of mass. Instead, in an average galaxy, the ratio of mass to light increases with distance from the galactic centre by a factor of up to $\sim 10^3$. The most widely accepted explanation for this is the addition of an invisible mass component.

As galactic luminosity decreases, the correlation between the light and the dynamical mass of the galaxy reduces. This can be seen in the "Universal Rotation Curve" of spirals, formulated as [20]

$$V(R) \simeq 200 \left(\frac{L}{L_*}\right)^\alpha [1 + f(L_B/L_B^*)(\frac{R}{R_M} - 1)] \qquad \text{km s}^{-1}, \tag{3}$$
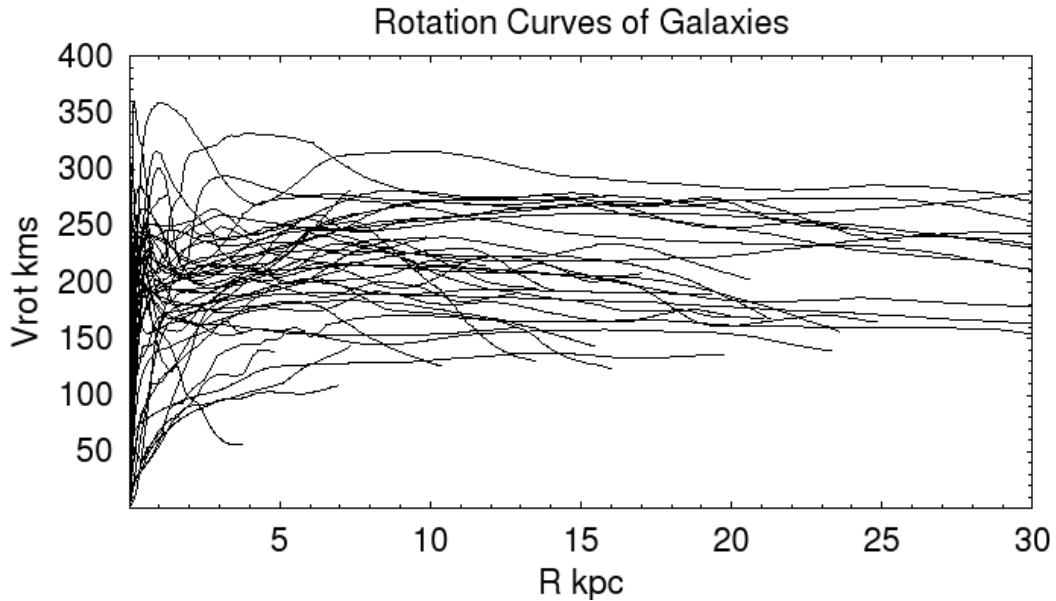
Figure (4): The rotation curves of spiral galaxies remain constant beyond their visible radius. This does not agree with Keplarian predictions in which the rotational velocity would fall off beyond this radius. The curves instead indicate that galaxies have significantly higher masses than that of the visible matter within them [22].

where L and L_B are the absolute luminosity and the absolute blue luminosity respectively with the asterisk denoting a reference luminosity. R is the radius and the reference radius, $R_M = 2.2$ disk length scales. This equation therefore gives the expected velocity of an object as a function of the luminosity and the radius from the galactic nucleus within the region $0.5 \leq R/R_{\mathrm{opt}} \leq 1$, where R_{opt} is the optical radius within which 83% of the I Band light is enclosed. It is important to note that there is a linear dependence on the radius and the rate of change is dependent on the luminosity, which holds within the optical radius of the galaxy.

Beyond the optical radius $(R_{\mathrm{opt}} \leq R \leq 2R_{\mathrm{opt}})$, the rotation velocity continues to increase and then asymptotically flattens out for low luminosity galaxies, whilst for high luminosity galaxies, there is a characteristic drop in rotational velocity just beyond the optical radius after which the curve becomes flat with $V_{\infty} \leq V(R_{\mathrm{opt}})$ [21]. This dependence of the rotational velocity on the luminosity of a galaxy implies that the structural properties of dark matter are connected to those of visible matter [23].

The use of the 21 cm radio waves emitted by hydrogen atoms has allowed observation of objects which are not energetic enough to be seen through visible wavelengths. These observations have shown that the rotational curves of spiral galaxies remain constant far

beyond their visible boundaries. The average visible radius of a spiral galaxy disk is $\sim 10\,\mathrm{kpc}$, however these discoveries show that the velocity of rotation remains almost constant up to a radius of 30 - 40 kpc. This corresponds to a galactic mass at least 10 times larger than that of the visible matter in the galaxies.

3.2 Virial Mass of Galaxy Clusters

There are several methods of estimating properties such as mass, density, and mass distribution of a galaxy cluster including the virial theorem, X-ray emissions, and gravitational lensing. In the virial theorem, galaxies are modelled as having random thermal motion relative to their galaxy cluster. Therefore, the average kinetic energy of a galaxy within the cluster is assumed to be proportional to the average gravitational potential energy binding the system, $KE_{av} = -\frac{1}{2}GPE_{av}$ [24]. This equation can be rearranged to get the mass of the galaxy cluster in terms of the average radial velocity of a galaxy within the cluster which can simply be obtained from their Doppler shift.

This approximation holds under the assumption that the cluster is in equilibrium, there is no bias in the sample of galaxies measured, the distribution of radial velocities is isotropic, and any difference in mass between bodies within the cluster are accounted for with "fudge factors". Applying this theorem, Fritz Zwicky became the first to notice the discrepancy in the ratio between the mass and the luminosity of the Coma galaxy cluster. The mass calculated from the average radial velocity was three orders of magnitude higher than that of the visible galaxies observed. More recent observations based on much more significant sample sizes suggest the average galaxy cluster is up to ~ 100 times more massive than suggested by its visible mass.

3.3 X-Ray Emission from Clusters

Most galaxy clusters have strong emissions in the X-ray band. This occurs primarily through the Bremsstrahlung radiation of high energy ($< K > \sim 10$ KeV) baryons in high temperature ($T \sim 10^8$ K) plasma. The mass of the cluster can be inferred from the gravitational potential that would be necessary to confine these high-energy particles
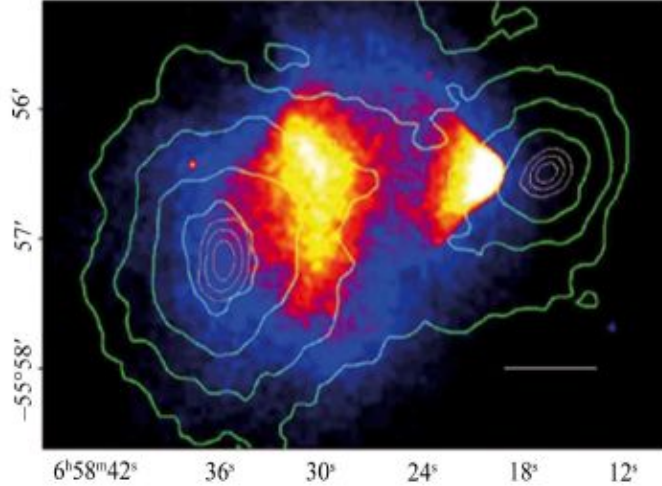
Figure (5): A collision between the Bullet cluster and another cluster obtained by the Magellan telescope at the Las Campanas Observatory. The contour lines give the convergence, κ, obtained through weak gravitational lensing, reconstructing the gravitational potential. The coloured heat map gives the X-ray emitting plasma temperature. For scale, the white bar corresponds to a distance of 200 kpc [27].

given the gas luminosity (in the X-ray band), the temperature of the plasma and the mass of the baryons. The X-ray emissions of clusters are extremely useful as they remain particularly consistent over time and the total luminosity has a simple relationship to the gas density. The X-ray emissions from two colliding galaxy clusters is shown in Figure 5.

Elliptical galaxy clusters have X-ray emitting halos which persist to a large radius from the clusters' centre. The total mass within the cluster can then be determined from the X-ray flux, and the temperature and density of the X-ray emitting plasma [25]

$$M_{tot}(\leq r) = -\frac{kT_{gas}(r)r}{\mu m_u G} \left(\frac{\partial \log T_{gas}}{\partial \log r} + \frac{\partial \log n_{gas}}{\partial \log r} \right), \qquad (4)$$

where $\mu = 0.62$ is the mean molecular weight and k is Boltzmann's constant. This holds when the gas is in hydrostatic equilibrium. Using the X-ray emission profiles, the total mass of an average cluster is measured to be $\sim 10^{15} M_\odot$ where $M_\odot$ is the solar mass. Visible stars provide only 2% of this mass, which approximately agrees with results from the virial theorem.

3.4 Gravitational Lensing

In more recent years, the main method of searching for dark matter (and dark energy) has been using gravitational lensing. Light rays from distant stellar objects do not travel in straight lines in a Euclidean frame under the presence of a large gravitational field produced by the likes of stars, galaxy clusters, or dark matter. Instead, the curvature of space-time around one of these objects causes the light to bend, shifting and manipulating the image of background objects. Measurements of the distortion or magnification of this image allow calculations of dark matter density relative to baryonic matter density, as well as information on properties of individual dark matter particles such as constraints on their size, mass and cross sections under the known fundamental forces. The field of gravitational lensing can be broken down into strong, weak and micro lensing along with flexion, however only strong and weak lensing will be covered as they have so far produced the most significant results related to dark matter.

Strong gravitational lensing is said to occur when an object warps space-time sufficiently, so that there are multiple paths that a ray of light from a background object can take around the lens. This is most easily observable around a particularly dense lens such as a galaxy cluster or the core of a galaxy. If the light source is perfectly behind a spherically symmetrical lens, light is able to travel around the object in every direction, forming what appears to the observer as an "Einstein ring", with a radius which is proportional to the square root of the mass of the lens.

In reality, the light source is rarely directly behind the lens, and the object forming the lens is unlikely to be perfectly symmetrical. This has the useful results that the image can appear in multiple locations when viewed from different angles and due to light from different parts of the source taking different paths through the lens, the image appears distorted. Instead of rings, the background source is observed as narrow arcs stretching around the lens (displayed in Figure 6), with the specific geometry of the arc being determined by factors such as the density of the source.

The paths to most distant objects in space do not pass by a strong gravitational lens on their way to earth, however they can still provide useful information through weak

Figure (6): A Hubble Space Telescope image of the Abell galaxy cluster. Weak gravitational lensing results in distant galaxies appearing as arcs around weak lenses. The combination of data from a large sample group of these objects can be used to understand the mass distribution in the region. [30].

lensing. In the weak lensing regime, the convolution of light sources (usually galaxies) as they pass through space is very small, however when taken over a sample of many sources the effects of weak lensing can give statistically significant results. As with strong lensing, the image is both distorted and magnified. Thus the transformation of the object can be split into convergent and shear terms for magnification and for the extent to which the image has been stretched respectively. Using these two terms, a map of the mass distribution about the path of the light rays can be formed. The relative mass density can be determined from patterns formed by images and their orientation. A region of "over-density" or "under-density" manifest themselves as E-modes. This method allows for a probe of the true distribution and structure of dark matter in the universe.

Observations show that individual galaxies consist mainly of a baryonic centre encompassed within a much larger halo of dark matter. Gravitational lensing is a useful tool in measuring this dark matter halo beyond where the visible mass of the galaxy lies. The use of weak lensing to combine signals from over 300 thousand galaxies in the SDSS survey resulted in an average galaxy with a stellar mass, $M_{stellar} = 6 \times 10^{10} M_{\odot}$ having a halo of mass $M_{halo} = 1.4 \times 10^{12} M_{\odot}$ [26]. More direct measurements taken in the Hubble Space Telescope SLACS survey of elliptical galaxies combined strong lensing and weak lensing to focus on the mass distribution within galaxies. This survey found that within

the central ~ 5 kpc of the galaxy, the mass is $27 \pm 4\%$ composed of baryonic matter, and beyond this radius, the amount falls off and the galaxy becomes much more dark matter dominated [26].

3.5   Dark Matter Candidates

Using the aforementioned methods, along with evidence from the Cosmic Microwave Background (CMB) radiation, various properties of dark matter candidates can be predicted. Based on the "free streaming length", which is a measure of the distance a dark matter particle (or object) would have travelled due to random motion in the early universe before slowing down due to cosmic expansion, dark matter can be categorised as hot, warm or cold. These correspond to a particle that travels a longer, similar, or shorter distance respectively than the radius of an average protogalaxy. This velocity is important as it determines the evolution of complex structures observed in the universe today.

Under a universe containing predominantly hot dark matter, early structures in the universe would be flat superclusters, which then collapse locally to form galaxies. If however dark matter is mostly cold, the first objects to form would be galaxies, which go on to produce galaxy clusters at a much later stage. The latter is what astronomers conclude to have occurred through deep field observations, with hot dark matter being excluded by high redshift early galaxies. The ΛCDM model of cosmology indicates the major component of dark matter to be non baryonic, however there may be a much smaller baryonic component ($\Omega_b \ll \Omega_m$) [29]. Some of the possible candidates are reviewed in Table 1.

Baryonic contributions to dark matter could exist in the form of brown dwarfs, neutron stars and black holes which do not interact with or emit enough light to be observable. These objects are collectively termed MACHOs (Massive Compact Halo Objects) and would exist in the galactic halo beyond the rest of the baryonic matter. These objects are partly observable through gravitational lensing due to their mass, however studies have provided the limit that no more than 8% of galactic halo mass comes from these MACHOs as the frequency of gravitational lensing from these objects is far too low [28].

18

	WIMPs	Sterile ν	Axions
Naturally Correct Ω	Yes	No	No
Production Mechanism	Freeze Out	Various	Various
Mass Range	GeV-TeV	keV	μeV - meV
Temperature	Cold	Warm	Cold
Collisional			
Early Universe			
Direct Detection	√√		√√
Indirect Detection	√√	√√	
Particle Colliders	√√		

Table (1): The table gives potential dark matter candidates along with their theorised properties and likely detection methods. The number of ticks corresponds to the likelihood of collision or detection [28].

The masses of the known neutrinos suggest the evidence that there could be right handed neutrinos called sterile neutrinos. Whether or not these particles are observable would be largely dependent on their masses. If one of their masses is below that of the $W^{\pm}$ boson, this sterile neutrino would be observable in certain phenomena and if their mass is much larger than the $W^{\pm}$ mass then they would be almost impossible to detect. However, if the mass is in the keV range, these particles could constitute a form of warm dark matter with a lifetime longer than the age of the universe and a possible decay to a photon and a lighter neutrino state. Even with a mass in this range, sterile neutrinos would not naturally account for 100% of the observed dark matter or for the structure of the universe.

Another candidate for dark matter is the axion, which was postulated to solve the strong CP problem but could also contribute to cold dark matter. The axion is very light ($10^{-6} \leq m_a \leq 10^{-3}$ eV) and thus there are very low limits on its interaction strengths. It can couple to two photons and thus when it travels through a strong electromagnetic field the particle will release a photon with energy equivalent to its mass.

The acronym WIMPs (Weakly Interacting Massive Particles) is a blanket term for dark matter candidates which only interact via the weak interaction. This includes supersymmetry (SUSY) particles such as a Majorana particle and the Kaluza Klein particle predicted from universal extra dimensions (UED) among others. One of the most likely

candidates for a dark matter particle is a neutralino, which if it exists in the mass range of 30 GeV $\sim$ a few TeV could give a dark matter density very close to that observed. If a WIMP exists as a stable particle, its current density may be simply explained as a relic from the big bang [28].

3.6   Detecting Dark Matter

Any interactions between dark matter and standard model particles are extremely rare and therefore notoriously difficult to detect. Dark matter is not directly visible through astrophysical observations, however it is possible that certain candidates could be indirectly observed through the small cross sections they have to standard model particles. For example, if large quantities of neutralinos exist in galactic halos or within the core of the Sun, they could pair annihilate to photons which could then be seen as anomalies in galaxy or solar spectra. However, for this to occur they must exist in significant quantities with sufficient cross sections.

Another method of detection is through the use of large underground detectors. WIMPs can become gravitationally bound within the Earth when they undergo scattering and no longer have sufficient energy to escape the Earth's gravitational well. These trapped particles accumulate beneath the Earth's surface and eventually annihilate, producing high energy neutrinos. Observatories such as IceCube [31], AMANDA [32] and Super-Kamiokande [33], are designed to detect these high energy neutrinos through their decay to muons when they interact with nuclei.

A more direct method of dark matter detection is through scintillation detectors which give off light due to the recoil of nuclei as dark matter passes through. A low temperature noble liquid, such as xenon or argon, interacts with a particle resulting in scintillation and ionisation which subsequently triggers a photomultiplier tube registering an event. Dark matter particles are identified as they scatter off nuclei, whereas background particles mostly scatter off electrons. So far, experiments of this type, such as ZEPLIN [34], XENON [35], etc., have found no significant evidence for dark matter, however they have been able to place upper limits on dark matter candidates cross sections with nucleons.

These methods all have the characteristic of waiting for a low probability process to occur. Depending on the mass and production mechanism of WIMPs, it may be possible for them to be produced in particle accelerators. Different WIMPs have different properties and thus there would be slightly different methods of searching for them. One of the most common methods is to search for missing energy in the final state radiation, tagged by a jet or a photon. Alternatively, if the dark matter candidate couples to a standard model particle, the dark matter may be inferred from a boost in the decay products of that particle.

3.7   H prime

If dark matter interacts non-gravitationally with standard model particles, there is a possibility that it would be produced in proton-proton collisions such as those at the ATLAS detector. Specifically, if it interacts through the strong nuclear force, a dark matter particle decay may lead to a dijet final state. There have been several proposed benchmark models for strongly interacting dark matter particles and other new physics scenarios which can be studied further through a dijet search [36]. This analysis focuses on the H' model.

New massive Higgs-boson-like particles are predicted in many Beyond Standard Model theories. Some of the simplest extensions of the Standard Model are the Two-Higgs-Doublet Models (2HDMs) which consist of adding two Higgs doublets to the SM. These have couplings that scale with the Yukawa couplings which results in the light scalar, h, the heavier, H, and the pseudoscalar, A, having suppressed decays to gluons and instead having dominant decays to quarks and to W and Z bosons [37].

Some extensions of the SM however predict new Higgs-boson-like particles with significant decays to gluons. For example, models which predict new heavy Dirac fermions lead to higher branching ratios to gluon and photon pairs through heavy fermion loops [39]. Other models, such as those which arose in response to the 750 GeV diphoton excess in 2015 predicted new particles with a $\gamma\gamma$ final state. In many of these models, strong decay rates to gg final states are also predicted, with the exact branching ratios being model

dependent [38]. In this analysis, the signal consists of a simulated SU(3) singlet scalar, H′, produced through gluon fusion and forced to decay to gluons.

4   The LHC and the ATLAS Detector

The Large Hadron Collider (LHC) is the world's largest particle collider, producing head-on collisions between protons (and sometimes heavy ions) currently at a record of 13 TeV centre-of-mass energy. Two beams of 2808 bunches, each consisting of 1.15×10^{11} protons, are accelerated in opposite directions and focused onto each other at four collision points, with collisions occurring every 25 ns. At each of these four points is a particle detector, two of which, ATLAS and CMS, are large general-purpose particle detectors, and the other two, ALICE and LHCb are more specialised for specific searches. The LHC was initially conceived to find evidence for the existence of the Higgs Boson, and to further probe for physics beyond the standard model such as SUSY particles and dark matter candidates. Following the conclusion of Run 2 of the LHC in 2018, the total integrated luminosity recorded by the ATLAS detector was $147 \, \mathrm{fb}^{-1}$ [41]. This large integrated luminosity has enabled higher precision measurements of particle masses and cross sections.

4.1   Acceleration Stages

The protons involved in the high-energy collisions are initially sourced from a bottle of hydrogen gas, each hydrogen atom consisting of one proton and one electron. Electrons are stripped from the hydrogen atoms using a strong electric field. The protons are then fed at a precise rate into the first of four intermediate accelerator stages, all shown in Figure 7, which each boost the energy of the protons before they are received into the LHC.

The first of these stages is Linac 2 (which will be replaced with Linac 4 for run 3), a 36 m long linear particle accelerator which accelerates the protons to 50 MeV, corresponding to 34.1 % the speed of light. It does this by using radio-frequency cavities to induce alternating charges on cylindrical conductors which exert a force on the protons, and smaller quadrupole magnets are used to confine protons to small beams.

These beams are injected from Linac 2 into the Proton Synchrotron Booster (PSB) which is the first of the circular accelerators. It is designed to accelerate the protons
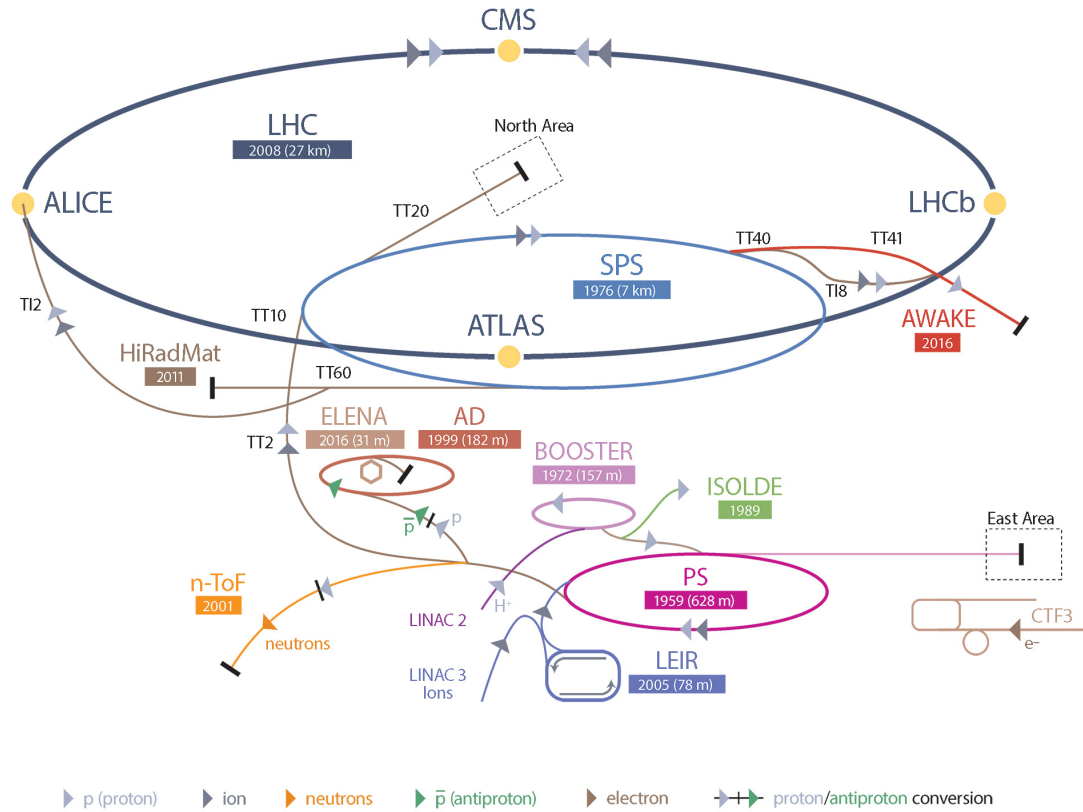
23

Figure (7): The various accelerators in the CERN accelerator complex are shown. A proton begins acceleration at Linac 2, and continues to accelerate in steps through PSB, PS, SPS, and finally the LHC before undergoing collisions at 13 TeV. The four major detectors in the LHC loop are CMS, ATLAS, ALICE and LHCb [40].

from 50 MeV to 1.4 GeV, which allows the next stage, called the Proton Synchrotron (PS) to accept more than 100 times the number of protons it would otherwise be able to accelerate. The PSB is made up of four stacked circular accelerators, each with a 25 m radius, allowing the proton package to be split into four to maximise the intensity of the beam. Each of these rings are split into 16 sections consisting of two dipole magnets, and three quadruple magnets are used to focus the beam. The four beams are recombined and passed to the PS which further accelerates the protons to 25 GeV, with a velocity 99 % the speed of light.

The final stage before the protons are injected into the LHC is the Super Proton Synchrotron (SPS). This synchrotron uses 1317 electromagnets to bend the protons around its 7 km circumference ring, bringing the proton energy up to 450 GeV. It splits the high-
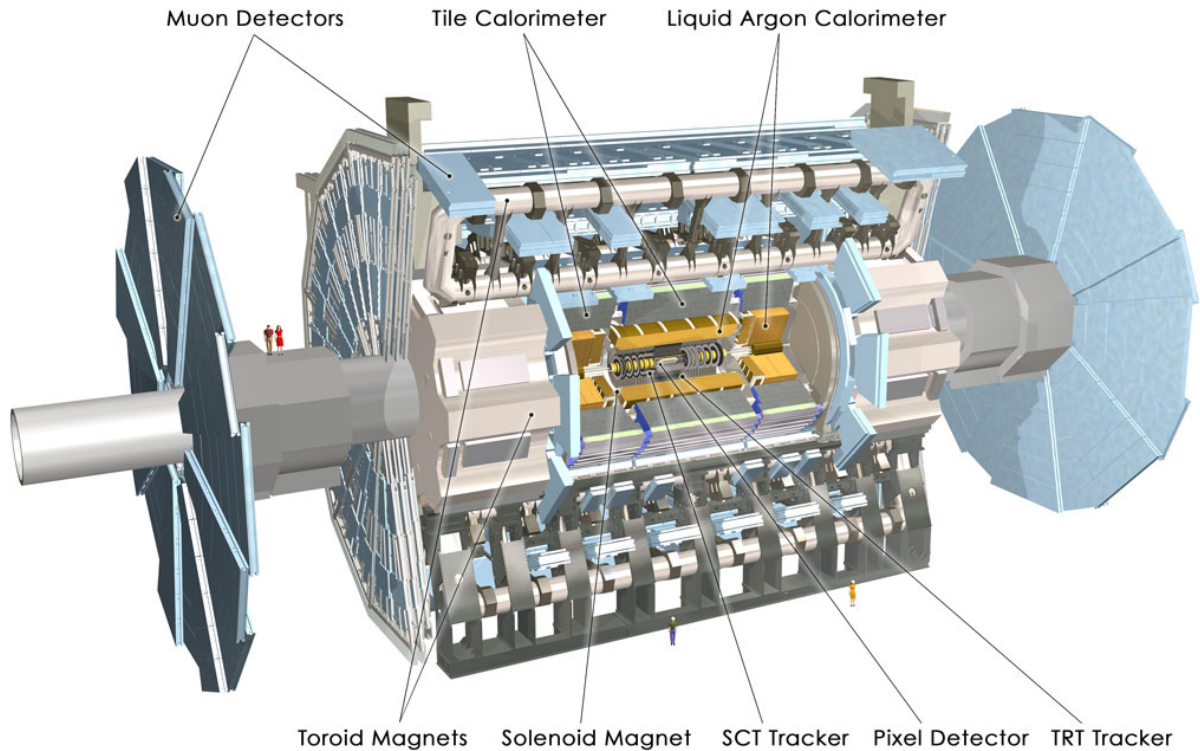
Figure (8): An artist rendition of a cross section of the ATLAS detector. [44].

energy beams between the LHC, the NA61 and NA62 experiments and the COMPASS experiment. The protons which are injected to the LHC are divided into clockwise and anticlockwise beams for the LHC's two pipes, where the protons will accelerate to their final collision energy of 6.5 TeV, taking approximately 20 minutes to reach this energy. The beams are then directed to their collision points by magnets at each of the four major detectors. Particle tracks resulting from these collisions are then recorded and analysed by the detector [42, 43].

4.2 The ATLAS Detector

The ATLAS detector is the LHC's largest general-purpose detector, covering almost the entire solid angle around the pp collision point with multiple layers of detector. The layout is shown in Figure 8, consisting of an inner detector, the calorimeters, the muon detectors and a magnet system. This combination of detectors allows ATLAS to record the trajectory, energy and momentum of individual high-energy particles around their collision point. The pseudo-rapidity, which is often used to define the coverage of a

detector is [45]

$$\eta = -\ln\left[\tan\left(\frac{\theta}{2}\right)\right], \tag{5}$$

where the origin of the coordinate system is based at the nominal collision point, and θ is defined as the angle from the beam axis. The central barrel of the ATLAS detector covers the region below $|\eta| \approx 1.2$ and the two end caps cover $1.2 \leq \eta \leq 2.5$.

The ATLAS magnet system consists of four large superconducting magnets. The magnetic field for the inner detector is produced by a central solenoid, aligned with the beam axis, which generates a 2 T field in the central region to allow momentum measurements of charged particles. Magnetic fields of 0.5 T in the central muon detectors and 1 T in the end cap muon detectors are obtained by a barrel toroid and two end cap toroids respectively.

4.3   The Inner Detector

The initial point of detection for particles resulting from a collision is the inner detector system which is intended to measure the momentum and vertex of charged particles and to identify electrons. The system is made up of three layers of sensors shown in Figure 9 and exists within the magnetic field of the magnet system, meaning the momentum of a charged particle can be determined from the degree of curvature through the magnetic field. The innermost layer of the system is the pixel detector, which combined with the next layer, the semiconductor tracker (SCT), enables high precision tracking. These are surrounded by the transition radiation tracker (TRT) which also has the purpose of distinguishing electrons from other more common particles.

The pixel detector contains approximately 80 million sensing elements shared between a central barrel and two end cap disks. The barrel is made up of three layers made from modules consisting of n-doped sensor tiles and read out chips. The end caps also have three layers meaning that every particle that passes through the pixel detector passes through three layers, each with a resolution of (14×115) μm^2, enabling precision vertex measurement. When a pixel is triggered by a particle, the signal produced is passed
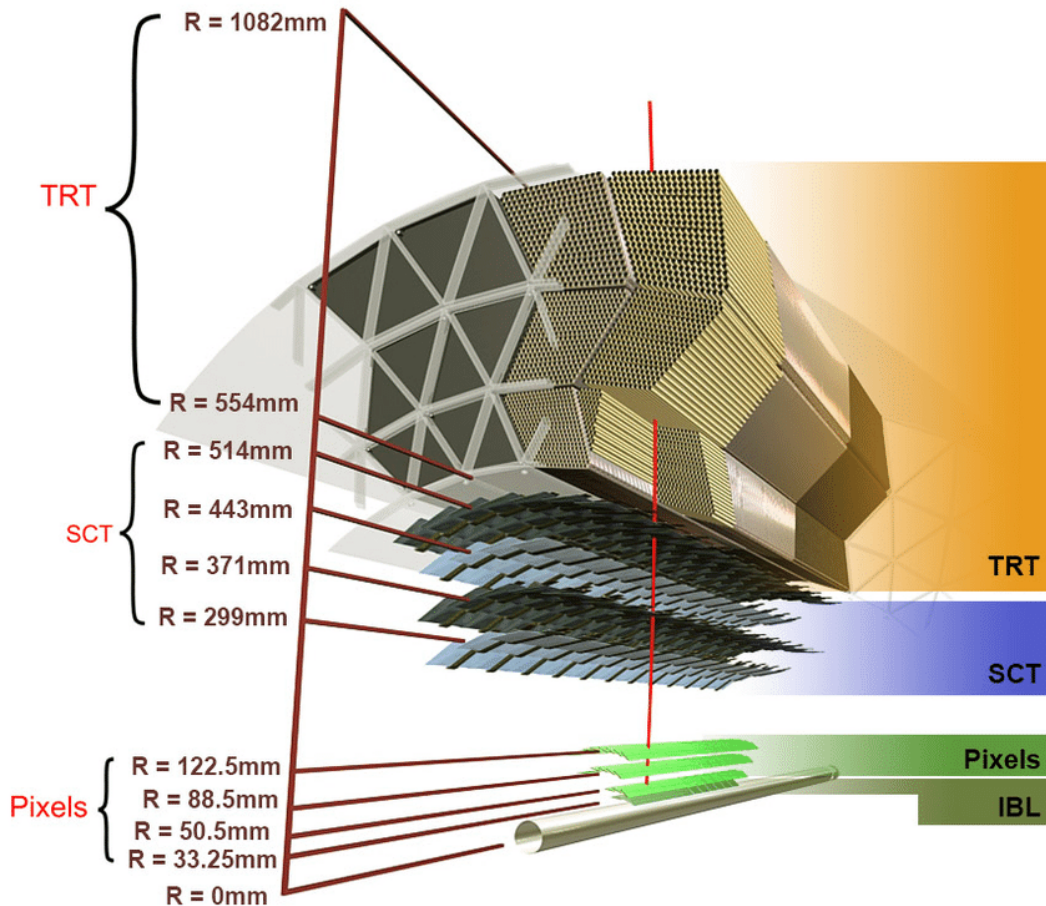
26

Figure (9): The three layers that make up the inner detector of ATLAS. The first point of interaction for a particle is the pixel detector, followed by the semiconductor tracker (SCT) and the transition radiation tracker (TRT) [46].

through an amplifier which is routed to a discriminator to determine whether the pulse is beyond a pre-determined threshold.

Surrounding the pixel detector is the SCT, made up of approximately 4088 sensor modules, each consisting of 285 μm thick silicon wafers with p-n micro-strip detectors attached. These sensor modules are split between four concentric barrel layers and two end caps each consisting of nine disks. Every module has 1536 readout channels, placed at an angle to avoid ghost hits and to maximise two track separation. These allow each layer of the detector to measure the location of a particle track with an accuracy of 17 μm. To ensure all particles pass through four layers in the region covered by the barrel, sensor modules from each layer overlap both in azimuth and in radius. There are, however, small gaps where particles may not be detected in the small region between the barrel and the end cap.

The final layer of the inner detector is the TRT detector, which detects radiation emitted by relativistic charged particles as they cross the boundary between two materials with different refractive indices. It does this using 372,032 drift tubes made from a gold plated tungsten hollow wire containing a gas which is a mixture of Xe (70%), CO_2 (27%) and O_2 (3%). Where the Xe is needed for efficient X-ray absorption, the CO_2 increases drift velocity, and the O_2 improves stability, and during operation this gas mixture is continuously circulated to maintain this mixture. The TRT again consists of a central barrel with the drift tubes arranged in three rings of increasing radius, and two end caps made up of wheels containing eight layers of drift tubes separated along the beam direction. Electrons and positrons can be distinguished from heavier leptons due to them producing stronger signals in the TRT, which is a consequence of lighter particles having a higher speed for a given particle energy.

4.4   Calorimeters

Calorimeters measure the energy deposited by a particle as it travels through the detector. They are intended to absorb most, or all of the energy of a particle produced in collisions with the exception of muons and neutrinos. The ATLAS detector contains two calorimeters, one electromagnetic calorimeter called the Liquid Argon (LAr) Calorimeter and the other is a hadronic calorimeter called the Tile Hadronic Calorimeter (TileCal).

The LAr Calorimeter is designed to measure the energy of electrons and photons among a large background of hadrons. It also gives information on the position and direction of photons which may not be achieved by the inner detector. The three LAr calorimeters are located in cryostats, one barrel calorimeter in the region $|\eta| \leq 1.475$, and two end caps in the regions covering $1.375 \leq |\eta| \leq 3.2$. The calorimeter itself is made of a stacked accordion structure consisting of layers of -183 degree liquid argon as an active medium and stainless-steel-clad lead absorber plates where the majority of a particle's energy loss occurs. These are joined to Copper and Kapton electrodes which are used for a capacitive readout.

TileCal, which is by volume the largest component of ATLAS, has more conventional

Iron scintillating tiles which are used as the active medium sandwiched between steel absorbers. The scintillation tiles are connected to photomultiplier tubes (PMTs), at the outermost radius of the module, by wavelength shifting fibres, allowing the amount of energy deposited to be measured. The purpose of the hadronic calorimeter is to absorb all the energy from hadronic showers so that any missing momentum can be identified. As it exists at a larger radius, the central calorimeter covers the region $|\eta| \leq 1.7$ and the end caps cover the pseudo-rapidity region $1.4 \leq |\eta| \leq 3.2$.

4.5   Muon Spectrometer

As mentioned, the inner detector and calorimeters are not sufficient to accurately measure the properties of muons. The ATLAS muon detector is therefore intended to measure the momentum of high energy muons over a pseudo-rapidity range $|\eta| \leq 2.7$ and provides a trigger in the $|\eta| \leq 2.4$ region. This is achieved with 4,000 muon chambers using a variety of technologies. These are immersed in an average 0.5 T toroidal magnetic field provided by the magnet system which means the bend of muon tracks through the detector is independent of their direction.

The main barrel of the muon spectrometer consists of three layers of Monitored Drift Tubes (MDTs) which measure the curves of the tracks passing through the magnetic field in a radius of $5 - 10$ meters from the beam line. Within the middle of this tracking region there are two resistive plate chambers (RPCs) which are responsible for triggering and for second coordinate measurement in the central region, and there is a third triggering station on the outside tracking layer. The forward region is covered by large wheels at distances of approximately 7.4 m, 10.5 m, 14 m and 21.5 m from the interaction point. At the innermost of these layers, Cathode-Strip Chambers (CSCs) are used instead of MDTs as they perform at higher rates with a better time resolution. The other layers use MDTs and three layers of Thin Gap Chambers (TGCs) are used in the end caps for triggering due to higher background rates in this region.

4.6 Trigger and Data Acquisition System

Collisions occur at such a high rate that it is impossible to record every event that occurs in a collision. Therefore a trigger is used to identify which collisions may contain physically interesting interactions. These triggers use physical signatures that can be found in the final state particles which signal the data-acquisition (DAQ) system to log these events. The DAQ formats the data readout from the detectors and sends the selected events to permanent storage systems.

The trigger system is divided into a first-level trigger, a high-level trigger (HLT) and a final selection. The first-level trigger is based in the detector, and combines tracks from the calorimeters and muon spectrometer in search of particles that may have come from certain decays based off a set of selection criteria. It then gives the HLT information on the location of tracks which passed the first-level trigger and the HLT will obtain a more detailed readout from the detector to confirm whether the initial trigger was correct. The HLT will then access the information on the entire tracks from the events of interest and process the event data. This information will then go through a final selection where additional information such as missing neutrino momentum may be taken into account. Once an event has passed the final selection, the full data from the event will be passed on to off-line data storage systems. In the time the trigger system takes to make decisions, the information recorded by the detectors from all interactions is held in the data electronics [47, 48].

5   Analysis

The objective of this analysis was to determine and quantify the effect of quark-gluon-tagging on improving the sensitivity of a dijet resonance search. Specifically, this has been done by setting and comparing interaction cross section upper limits on a simulated H′ signal model with a simulated QCD background for tagged and untagged datasets. Both the H′ signal and the QCD background are generated using Pythia 8.1 with the PDF's, tunes, and selection criteria described in later in this section, with the QCD background intended to correspond to the data collected from Run 2 of the LHC.

The optimal variable to use to identify dijets as quark-quark, quark-gluon or gluon-gluon initiated was determined by manually applying cuts, as a function of dijet mass, to the Monte Carlo generated QCD background and using "truth" data to discover which variable provides the purest selection of gluon-gluon initiated jets. By also applying the gluon-gluon selection to the QCD Run 2 data, it could be verified that the simulated events matched the data. SWiFt, described in Section 5.6, was used to fit parameters to the untagged Pythia simulated background. Following this, a statistical analysis is applied to the QCD background with the H′ signal using HistFitter with the initial aim of rejecting the background only hypothesis, and in the absence of finding new physics, placing an upper limit on the H′ cross section. This process was repeated for the tagged background and signal, and the 95% CL upper limits were compared, where lower upper limits would indicate that quark-gluon tagging would allow for a higher sensitivity dijet resonance search.

5.1   Jet Reconstruction

There is no universal definition of a jet so jet reconstruction algorithms are used to approximate jets from particle tracks in the detector. The development of such algorithms faces problems due to cross over particle tracks from simultaneous events from separate collisions, which is termed "pile-up". The pile-up increases jet energy and can even contribute to "pile-up jets" which are additional jets in an event and is particularly a

problem at the LHC due to its high luminosity. The jet reconstruction and calibration algorithm should thus determine the energy scale and angular distribution of jets to a high resolution. To do this, cells around a signal in the ATLAS calorimeter are grouped into 3-dimensional "topo-clusters", shown in Figure 10, by a clustering algorithm and the resulting jets are cleaned and calibrated using a combination of Monte Carlo (MC) simulations. In situ calibration and information from the detector can be used to reduce pile-up.

As mentioned, a topological clustering algorithm is used to group energy deposited in the detector by particle tracks into jets. These algorithms find cells in the calorimeter with recorded absolute energy of E > 4σ, where σ is the total noise from electronics and pile-up. The cells which exceed this threshold are termed "seeds". Cells surrounding the seed, defined as those directly adjacent to it in every direction, must then pass the threshold of E > 2σ to be added to the topo-cluster and a final layer of E > 0 cells are included surrounding these cells. To distinguish between multiple overlapping topo-clusters, a splitting algorithm is used to identify local maxima in energy deposits. At ATLAS, the topological clustering algorithm is not intended to isolate energy deposits from individual particle tracks but instead separates individual electromagnetic or hadronic particle showers and provides their shape and energy [48].

The resulting topo-clusters along with their shape and energy parameters are used as the inputs to the anti-k_t algorithm. This is a sequential recombination algorithm, meaning the algorithm branches out from an individual topo-cluster, and has the advantage over other jet finder algorithms in that it exhibits IRC (both infra-red and collinear) safety [51]. At ATLAS, a distance parameter of R = 0.4 is used to specify the nominal angular width of the reconstructed jets and jets from both EM and LCW calibrations are used, corresponding to clusters calibrated to the electromagnetic scale and those where local calibration weighting has been applied respectively. Topo-clusters are sorted into jets using distance measurements, d_{ij}, which give the distance between two clusters i and j, and d_{iB}, which provides the distance between the cluster, i, and the beam, B. These properties are defined via [50]
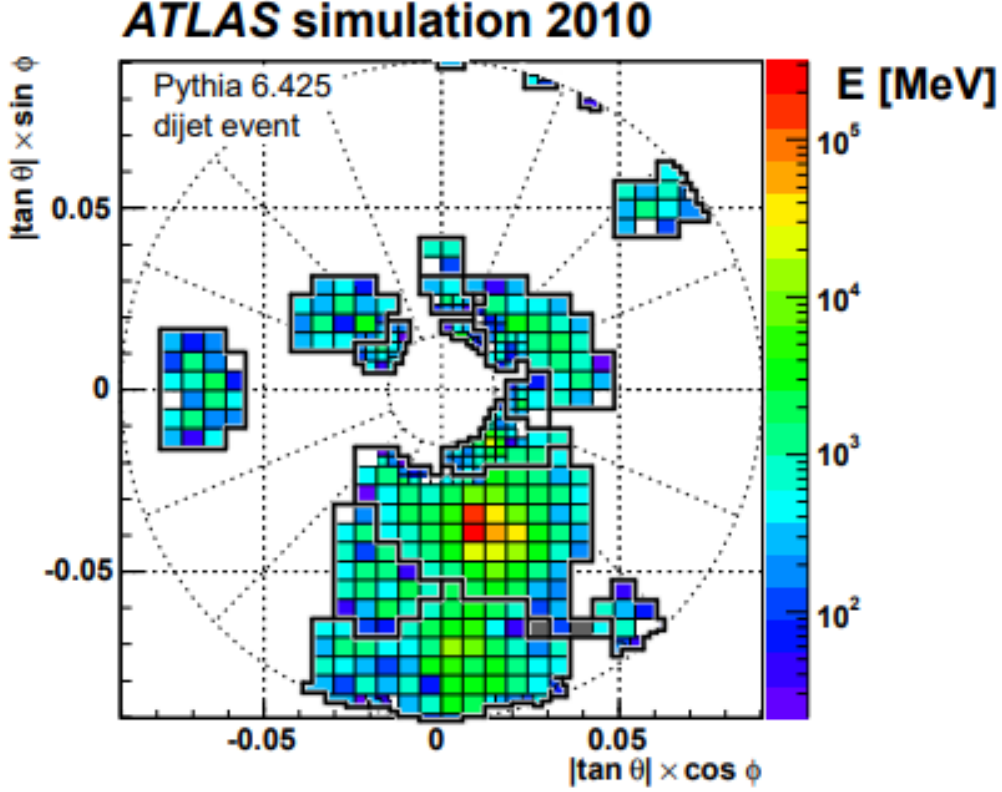
Figure (10): The figure shows a cluster identified by the topological clustering algorithm from the FCAL calorimeter. Clusters in this image can contain seeds in this module of the FCAL or in other layers as the cluster is three dimensional [49].

$$d_{ij} = min(k_{ti}^{2p}, k_{tj}^{2p})\frac{\Delta_{ij}^2}{R^2}, \tag{6}$$

$$d_{iB} = k_{ti}^{2p}, \tag{7}$$

where k_{ti} is the transverse momentum of particle i and $\Delta_{ij}^2 = (y_i - y_j)^2 + (\phi_i - \phi_j)^2$, where y_i and ϕ_i are the rapidity and azimuth respectively of particle i with respect to the beam axis. The parameter p determines whether the algorithm is dominated more by the energy as opposed to the geometrical factor, Δ_{ij}. For the anti-k_{ti} algorithm, $p = -1$ is used, which results in the geometry of the jet being largely dependent on the energy of particles within the jet compared with $p = 1$ algorithms such as the inclusive k_{ti} algorithm which is governed by the ordering between particles. Jets reconstructed using the anti-k_t algorithm are shown in Figure 11 in comparison with other reconstruction algorithms.

For a given particle, i, the algorithm determines the next particle, j, for which d_{ij}

33
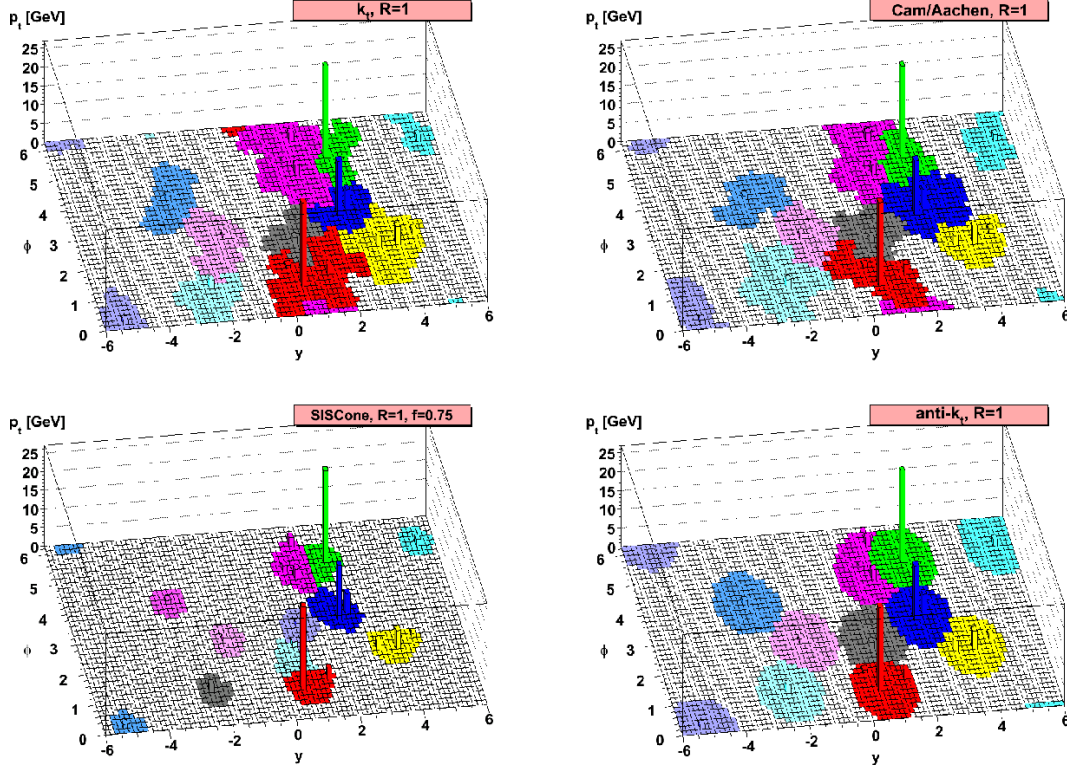
Figure (11): A comparison between clustering algorithms which determine jet shape using a parton-level event generated in HERWIG. Random soft ghost particles are clustered with hard particles resulting in hard jets [50].

is the smallest and recombines the two. If d_{iB} is the shortest distance, then object i is classified as a jet. This continues until all clusters are grouped into jets. The d_{ij} between a hard particle and a soft particle will be a product of the transverse momentum of the hard particle and their Δ_{ij}. The d_{ij} between two soft particles of the same separation however will be much larger due to their smaller transverse momentum. As a result of this all soft particles tend to cluster to a hard particle rather than to each other, and if there is only one hard particle in a $2R$ radius, all soft particles within a radius, R, will cluster to it in a perfect cone-shaped jet.

If two hard particles (1 and 2) exist within $R \leq \Delta_{12} \leq 2R$ of each other, they cannot each form a perfectly canonical jet and their shape instead depends on the respective magnitudes of k_{t1} and k_{t2}. If $k_{t1} \gg k_{t2}$, then the jet formed around particle 1 will be canonical and the jet around particle 2 will have the overlapping part cut off. However if the two particles are of equal transverse momentum, then soft particles in the overlapping region will split in a straight line between the two jets.

Two hard particles separated by $\Delta_{ij} \leq R$ will cluster together, forming a single jet. It
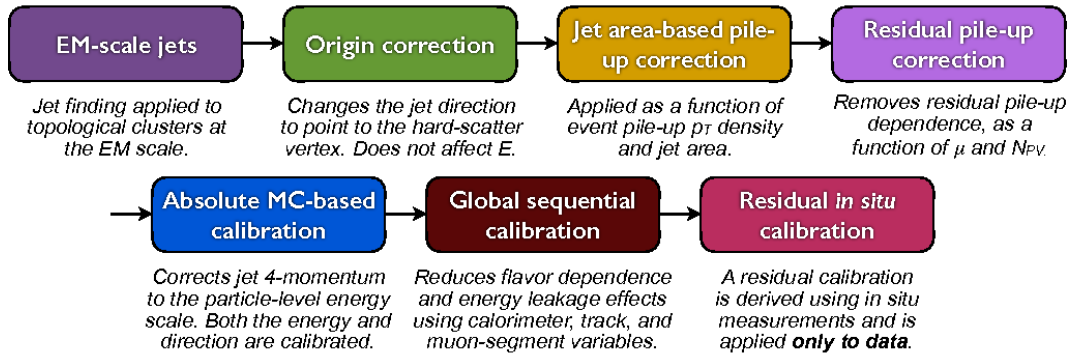
Figure (12): An overview of the 2015 ATLAS calibration scheme for EM-scale calorimeter jets. Each stage is applied to the jets' four momentum except the origin correction. [54].

is important to note that the soft particles do not influence the shape of jets in the anti-k_t algorithm, but instead cluster with the nearest hard particle. This feature is demonstrated in Figure 11 which displays the clustering algorithms dependence on the p value. As the anti-k_t algorithm is resilient with respect to soft radiation, the jet shapes are much less impacted by random "ghost" particles and instead all hard jets are mostly canonical and less complex [51, 52].

The energy of the reconstructed jets formed from the topo-clusters may not be equivalent to that of the primary parton which formed the jet. This can occur due to the effects of pile-up, dead material in the detector, energy escaping the calorimeters, and other factors. For this reason, reconstructed jets are calibrated following the ATLAS calibration scheme, outlined in Figure 12. A pile-up correction is applied to the jets through a p_t offset determined by MC simulations. Jet four-momentum and pseudo-rapidity are also calibrated by comparing reconstructed jets with MC simulated truth jets. Since the data do not entirely match with MC simulated data, a residual in-situ calibration is applied to account for miscalibration through multiplying jet p_t by the ratio of MC to data response rates.

5.2   Event Selection and Jet Cleaning

To ensure quality of data and maximise sensitivity to relevant physics, cuts are applied on the jets to be used in the analysis. This involves identifying problems at the detector level and also solving event-level issues using broader cuts on the properties of the event.

In the case of an issue that affects a significant section of a detector or sub-detector, entire runs and luminosity blocks can be flagged or removed from the data. Data quality (DQ) flags are issued for detectors, trigger slices, and physical objects (e.g. jets, electron, muon, etc.) and store information on the state of the hardware and the object reconstruction for each luminosity block. A DQ selection criteria is then applied to the block to form ATLAS' list of good runs and luminosity blocks. In this analysis, events are selected based on the following criteria in addition to the DQ selection to ensure sensitivity to the relevant physics:

- Leading jet $p_t > 420$ GeV,

- All other jet $p_t > 150$ GeV,

- $m_{jj} > 1100$ GeV,

- $|y^*| < 0.6$,

- Trigger: HLT_j420.

Here m_{jj} is the dijet invariant mass and $|y^*|$ is the Lorentz invariant rapidity difference between the two leading jets in an event. Selecting only central jets, confined by the $|y^*|$ cut off value, limits dominant background contributions from QCD processes such as gluon exchange, which is dominant at low polar angles. The trigger, HLT_j420, is a high-level trigger which cuts the minimal nominal jet p_t at 420 GeV and is chosen as it is the lowest single jet un-pre-scaled trigger available for the full Run 2 data. The dijet invariant mass cut off is chosen at the point where trigger is above 99.5% efficient, and the cuts on transverse momentum are intrinsic to the trigger.

A "jet cleaning" process is needed to remove reconstructed jets which do not result from the intended proton-proton collision. Background processes often create so-called "fake" jets that can result from calorimeter noise, high-energy proton collisions with gas particles in the beam pipe, and particle showers from high energy cosmic rays. These events can be high-energy enough to be identified as jets by the jet reconstruction algorithm, however possess properties which allow them to be identified as fake jets. These

discriminating variables can be defined in terms of the quadratic difference between the actual and real electric pulse shapes in a given LAr calorimeter cell, Q_{cell}^{LAr}, where [55]

$$Q_{cell}^{LAr} = \sum_{j-1}^{4} \left(s_j - A \left(g_j - \tau g_j' \right) \right)^2,$$

(8)

Here s_j and g_j are the amplitude and normalised predicted ionisation shape respectively for each sample, j. A is the measured amplitude and τ is the time period of the signal. From this quantity assigned to each calorimeter cell, jet quality quantities can be defined. The average jet quality, $< Q >$, is the normalised mean quadratic difference, Q_{cell}^{LAr}, weighted by the squared energy detected. For a jet determined to have a poor signal shape quality ($Q_{cell}^{LAr} > 4000$), the fraction of its energy deposited in the LAr calorimeter cells and the hadronic calorimeter (HEC) cells are f_Q^{LAr} and f_Q^{HEC} respectively. A separate quantity E_{neg} can be defined by adding up the energy from all calorimeter cells which record negative energy from detector noise. The distributions of these quantities for good and fake jet enriched samples is displayed in Figure 13.

The quality selection used in this analysis is the LooseBad jet selection, intended to obtain a high efficiency in identifying good jets while rejecting as many fake jets as possible. Using the quantities outlined above, a jet will be classified as fake if it has any of the following properties [55]:

- $f_{HEC} > 0.5$ and $\left| f_Q^{HEC} \right| > 0.5$ and $\langle Q \rangle > 0.8$,

- $|E_{neg}| > 60$ GeV,

- $f_{EM} > 0.95$ and $\left| f_Q^{LAr} \right| > 0.8$ and $\langle Q \rangle > 0.8$ and $|\eta| < 2.8$,

- $f_{max} > 0.99$ and $|\eta| < 2$,

- $f_{EM} < 0.05$ and $f_{ch} < 0.05$ and $|\eta| < 2$,

- $f_{EM} < 0.05$ and $|\eta| \geq 2$,

Where, f_{HEC} and f_{EM} are the ratio of energy deposited in the HEC and EM calorimeters respectively to the total energy of the jet, f_{max} is the maximum possible energy that
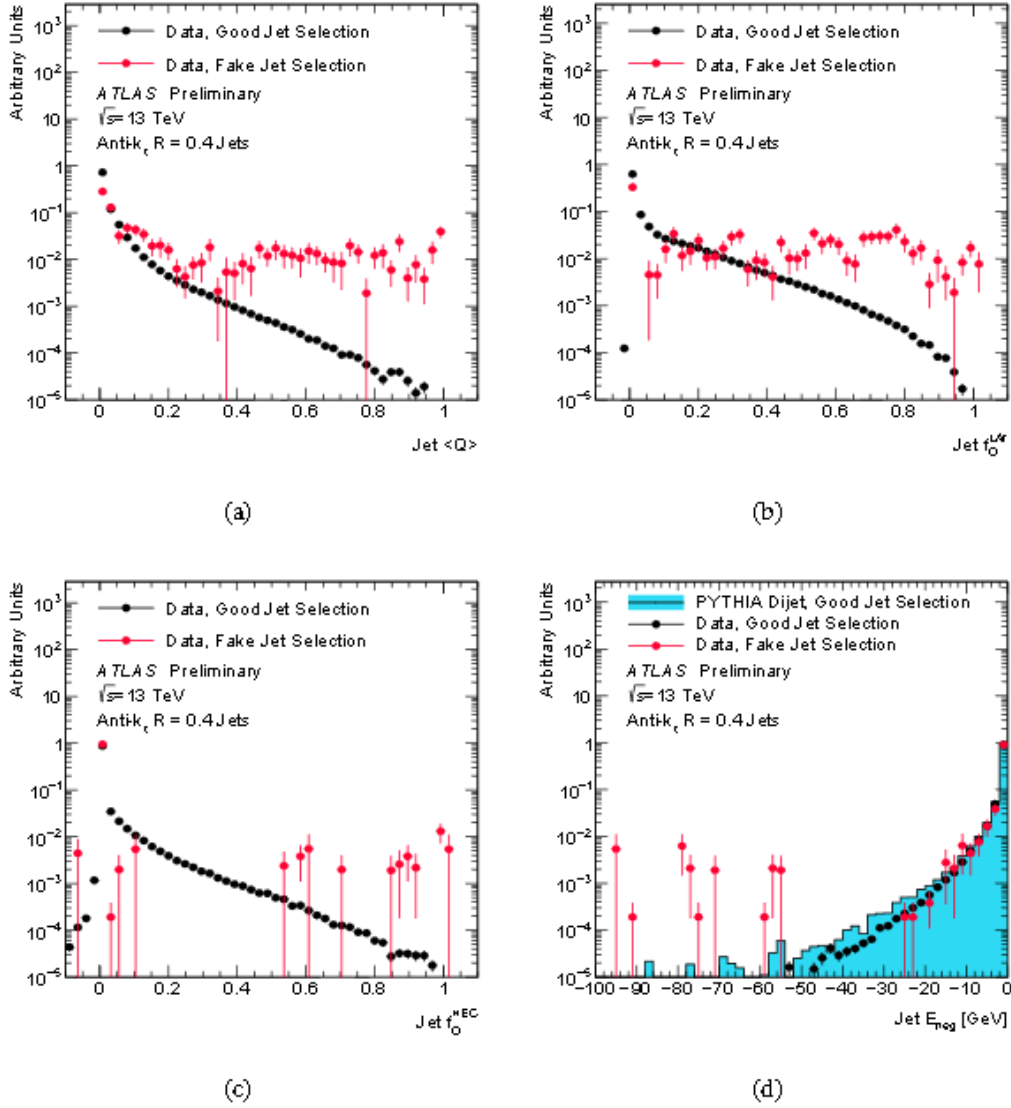
Figure (13): The distributions of (a) f_Q^{LAr}, (b) f_Q^{HEC}, (c) Q_{cell}^{LAr} and (d) E_{neg} for good (black) and fake (red) dominated samples. The blue histogram in (d) compares these with Pythia simulated data. [55].

can be deposited in a single layer of the calorimeter and f_{ch} is the charged jet fraction. Excluding jets that satisfy either of the first two points removes those caused by sporadic noise bursts in the HEC. Reconstructed jets that meet the criteria in the third point can be removed as they are likely due to noise in the EM calorimeter. The final three points are more broad, intending to filter fake jets which can result from cosmic rays, faulty hardware, and other non-proton-proton hard scattering events [55, 58].

5.3 Quark/Gluon Tagging

Jets resulting from high-energy particle collisions can broadly be split into quark or gluon jets based on their initiating parton. This is a useful property to distinguish data sets as it allows a small gluon (quark) dominated signal to be isolated from a large quark (gluon) background. In this analysis, the fact that the dominant decay of the H' is to two gluons is exploited to remove the dominant quark background in the dijet data. This allows for increased sensitivity in the search for new physics scenarios.

Differences between the final state partons in quark and gluon jets arise primarily due to the difference in colour charge between the two. As mentioned previously, while quarks carry only a single colour (or anti-colour) charge, gluons carry a combination of both a colour and anti-colour charge. As a result, the Altarelli-Parisi splitting functions contain a factor of C_A for a gluon radiating another gluon and a factor C_F for a quark radiating a gluon, where C_A and C_F are related by the Casimir Ratio $\frac{C_A}{C_F} = 9/4 \sim 2$. This results in gluon initiated jets containing more constituents and also being broader than quark jets as a quark jet can approximately be thought of as being dominated by its first gluon emission. The most commonly used identifiers are the jets' charged particle count, reconstructed from the tracks left in the ATLAS inner detector, and jet mass, however properties such as angular width [61, 62] and the types of hadrons found in the jet [63, 64] can also be used.

In this analysis, the parameter chosen to identify quark and gluon jets was the track multiplicity (nTracks), which is the number of charged tracks above 0.5 GeV in a jet. As can be seen in Figure 14, which shows truth quark and gluon jets in a MC simulated dijet dataset, gluon jets are dominant at high particle multiplicities and quark jets are dominant at low multiplicities. Therefore a threshold n_{gg} can be chosen to apply to the real dataset such that all jets with $n_{tracks} < n_{gg}$ are defined as quark initiated jets and those with $n_{tracks} \geq n_{gg}$ are gluon initiated. This allows n_{gg} to be varied to determine the purity of the gluon jet signal. The multiplicity of the jets also varies with the dijet invariant mass, and thus the multiplicity cut-off must be found in each m_{jj} bin to ensure the efficiency of the cut is approximately constant across the dijet mass range, enabling
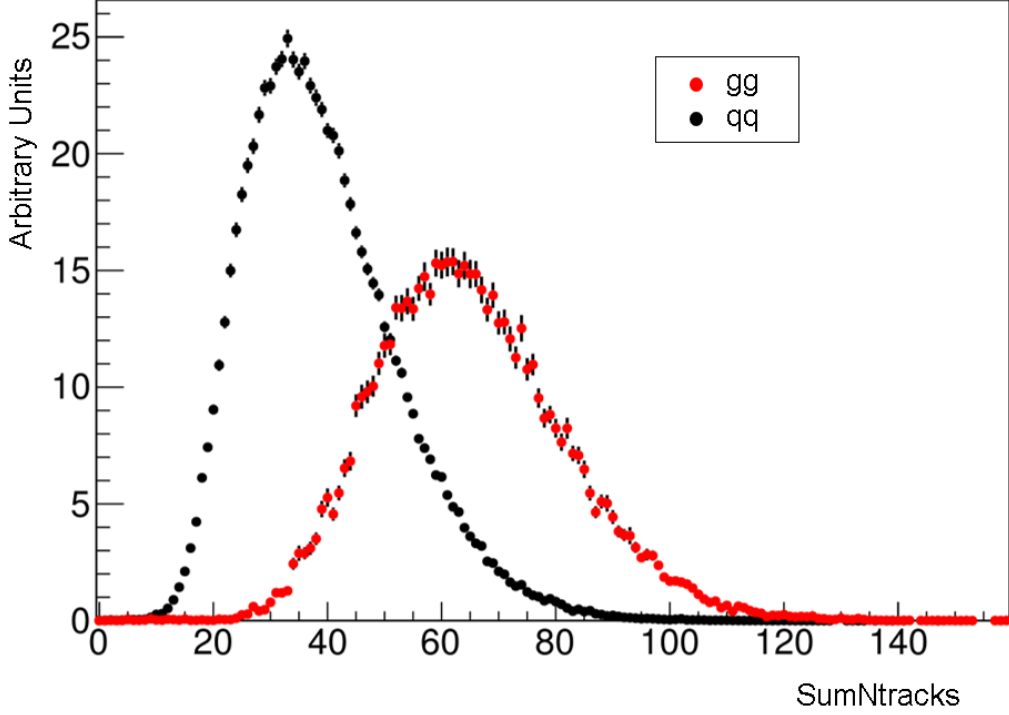
39

Figure (14): The histogram of the sum of the normalised number of tracks above 500 MeV in the dijet mass region $3040 \leq m_{jj} \leq 4270$ GeV shows that gluon-gluon initiated dijets tend to have higher multiplicities than quark-quark initiated dijets.

n_{gg} to be defined in terms of m_{jj}, or to ensure linear proportionality $\ln(m_{jj})$

$$n_{gg} = p_0 \times m_{jj} + p_1 \tag{9}$$

$$n_{gg} = p_0 \times \ln(m_{jj}) + p_1, \tag{10}$$

where p_0 and p_1 are constants determined from simulated data for a given efficiency; the best value for these parameters are those that maximize the significance ($\sigma \sim \frac{S}{\sqrt{B}}$). Using this equation for n_{gg}, each dijet sample in the data set is categorised as QQ, if both jets are quark initiated, GG, if both jets are gluon initiated, or QG if one jet is quark initiated and the other is gluon initiated. The cut off multiplicity is also regularly calculated as a function of the transverse momentum p_T, however m_{jj} is used in this analysis.

5.4   Monte-Carlo Simulations

To analyse the data reconstructed from the ATLAS detector, it is necessary to compare them with data simulated from theory. This involves simulation of the event, accounting for detector efficiencies, and applying all of the cuts which are placed on the reconstructed data. By comparing these sets of data, the extent to which data agrees with predictions by the standard model can be observed and deviations can be examined in the search for new physics. Due to the complexities in accurately predicting the complex evolution and hadronisation from initial collision to final state, a first principles approach to simulating high energy particle collisions is not possible, however close approximations can be made using Monte-Carlo event generators. A key benefit of MC simulated data in this analysis is that it provides "truth" information on whether jets are initiated by quarks or gluons, allowing them to be used to calculate the effectiveness of the gluon tagging algorithm used.

Parton shower MC event generators begin by calculating the cross sections to next to leading order (NLO) in perturbation theory. The branching from one parton to two partons is simulated, obeying local flavour and four momenta conservation. This means a single parton may either split to two partons or not, with a probability assigned to each outcome. The MC event generator used in this analysis is PYTHIA 8, which uses a shower evolution based on the standard (LO) DGLAP splitting kernels, $P(z)$ [68]:

$$P_{q \to qg}(z) = C_F \frac{1 + z^2}{1 - z}, \tag{11}$$

$$P_{g \to gg}(z) = C_F \frac{(1 - z(1 - z))^2}{z(1 - z)}, \tag{12}$$

$$P_{g \to q\bar{q}}(z) = T_R(z^2 + (1 - z)^2), \tag{13}$$

where the Casimir factors $C_F = \frac{4}{3}$ and $C_A = N_C = 3$, $T_R = \frac{1}{2}$ and z is the energy sharing fraction between daughter partons. Both of the algorithms for the initial state radiation (ISR) and final state radiation (FSR) are based on these splitting kernels. The hadronisation of the FSR is based on the Lund String fragmentation framework which

is a manifestation of the string model of linear confinement discussed previously. Here a colour flux tube is stretched between two colour charged particles, such as two quarks, at a typical hadronic scale (~ 1 fm) and the tension of the string is $\kappa \approx 1$ GeV/fm. Depending on the invariant mass of the two particles, the string can either snap forming further hadrons or the colour singlet hadron can remain intact. Every event generated contains all the particles from an interaction (with a primary vertex at the origin of the interaction) and is assigned an event number.

To account for interaction with the detector, the MC events generated by PYTHIA are passed to the GEANT 4 toolkit which simulates the passage of particles through matter. The simulated data are then reconstructed from the simulated detector response and the cuts specific to the required physics are applied. The "A14" (ATLAS 14) [69] tune is used in this analysis which determines the input parameters to PYTHIA from previous data. This is used in conjunction with the NNPDF23LO leading order parton density function. A comparison of the performance of this tune with the AU2 [70] and Monash [71] tunes for ATLAS dijet observables is shown in Figure 15 [68, 75, 76]. As can be observed, the A14 tune provides a closer fit to the real ATLAS data for dijet specific variables.
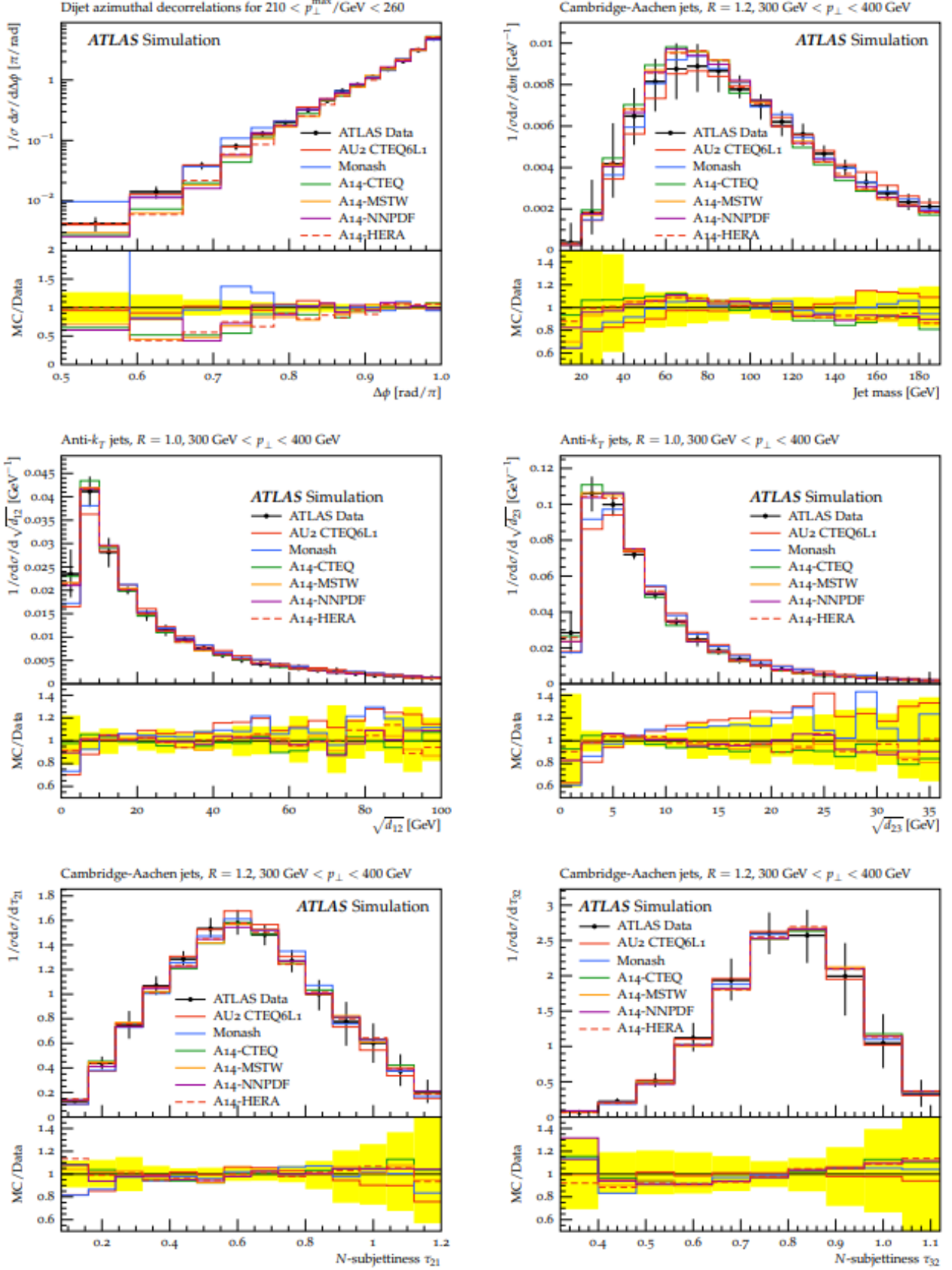
Figure (15): Predictions for ATLAS dijet decorrelation and jet substructure observables using the AU2, Monash, and A14 tunes with a range of parton distribution functions [69].

43

5.5   MC Comparison with Data

With the Q/G tagging cut applied to both the simulated QCD background and real reconstructed dijet events, a comparison is made for a range of parameters in the dijet resonance samples with the baseline selections applied to check for differences in the data. As can be seen in the plots below (Figures 16-18), there is good agreement between the data sets for all variables examined.
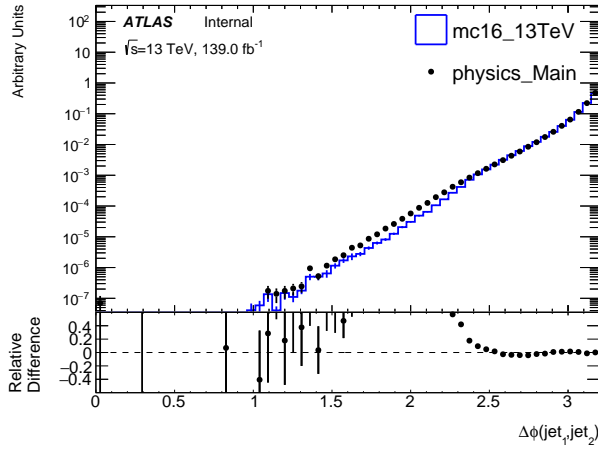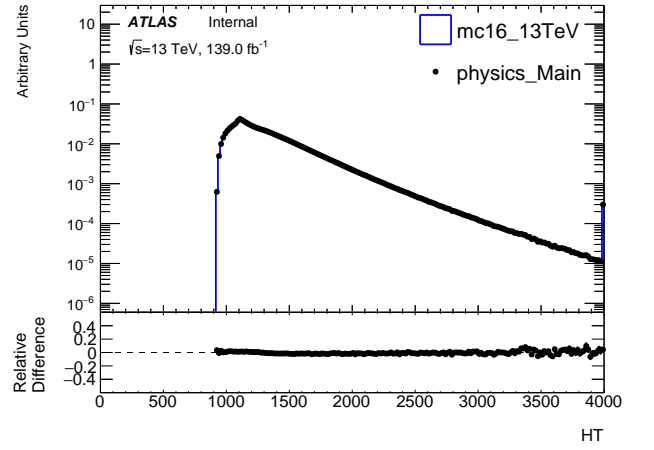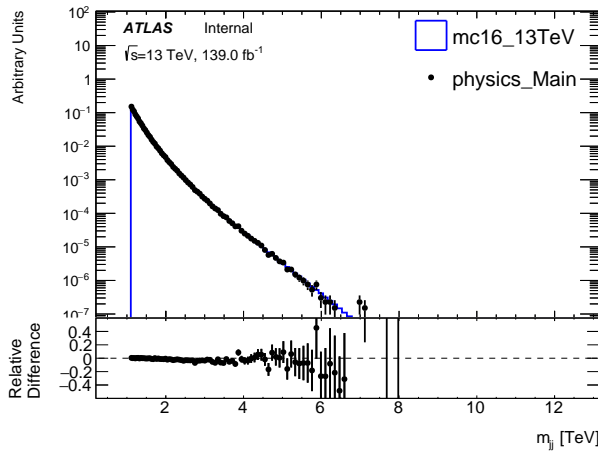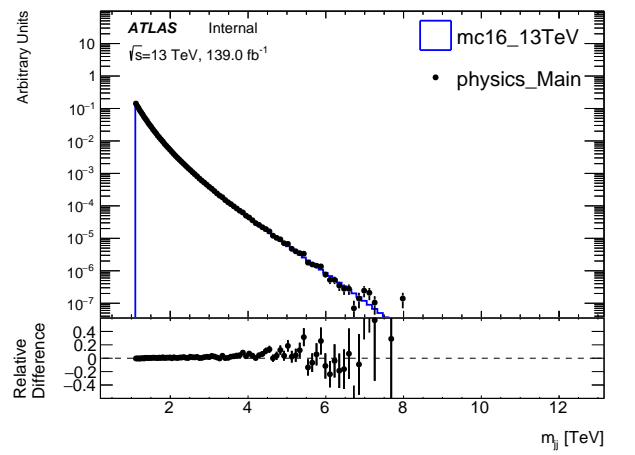
Figure (16): Comparison of MC data (blue) with ATLAS data (black) for (a) the angle between the jets ($\Delta\phi$), (b) the scalar sum of the jet momentum (H_T), (c) and (d) the dijet invariant mass spectrums for GG jets and all jets respectively.
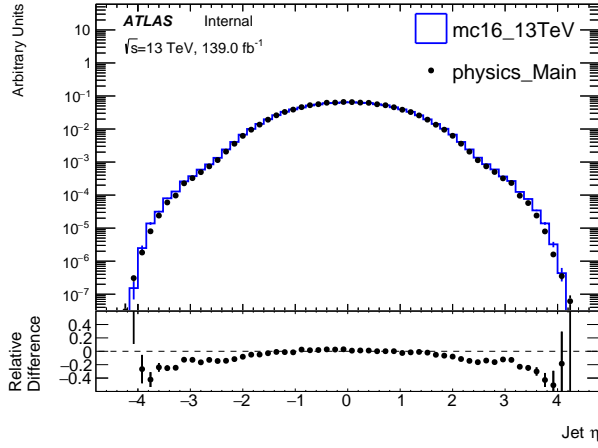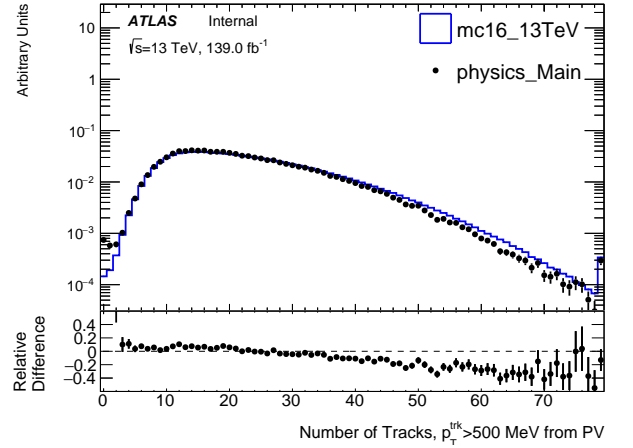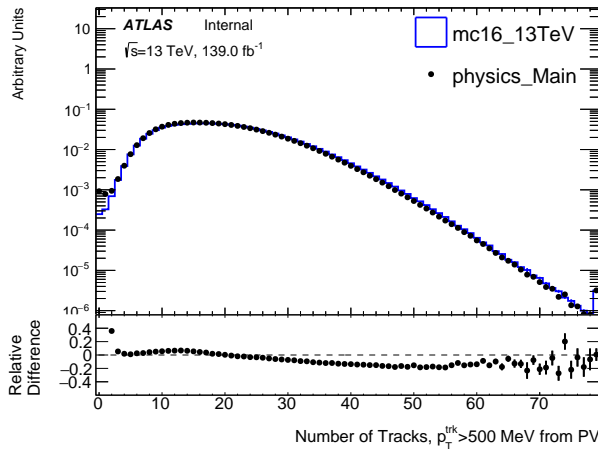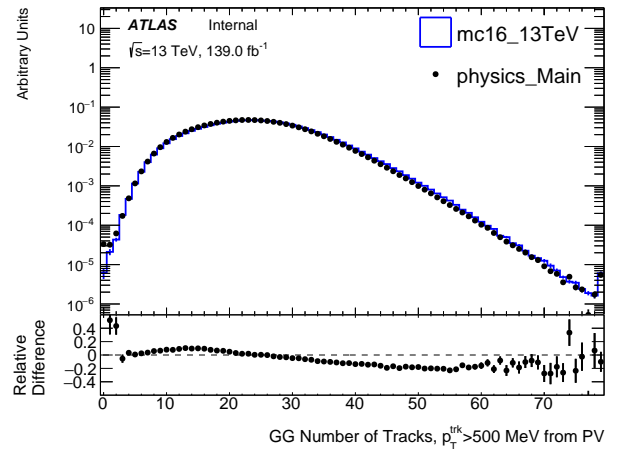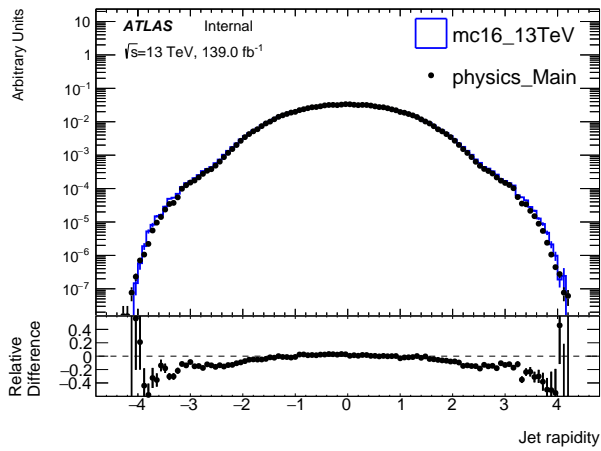
Figure (17): Comparison of MC data (blue) with ATLAS data (black) for (a) the angle of the reconstructed jets (η), (b) the multiplicity of dijets in the region $3100 \leq m_{jj} \leq 4070$, (c) the multiplicity of jets across the entire m_{jj} spectrum, and (d) the multiplicity for GG tagged events.
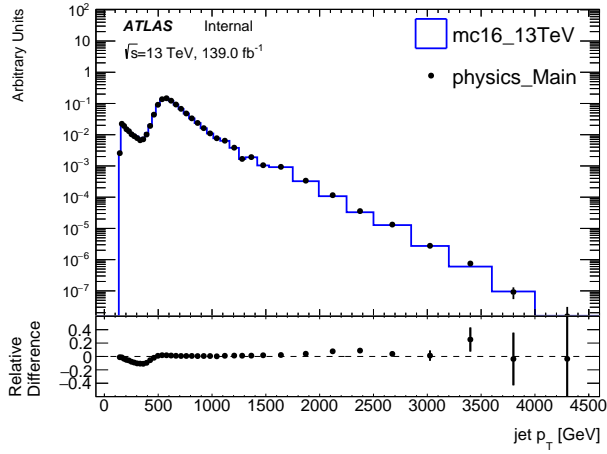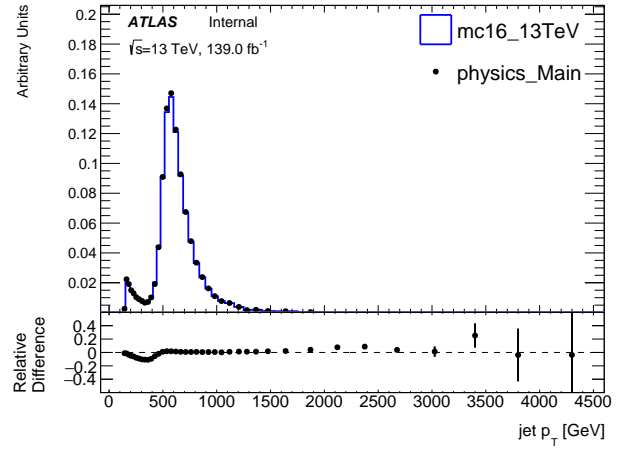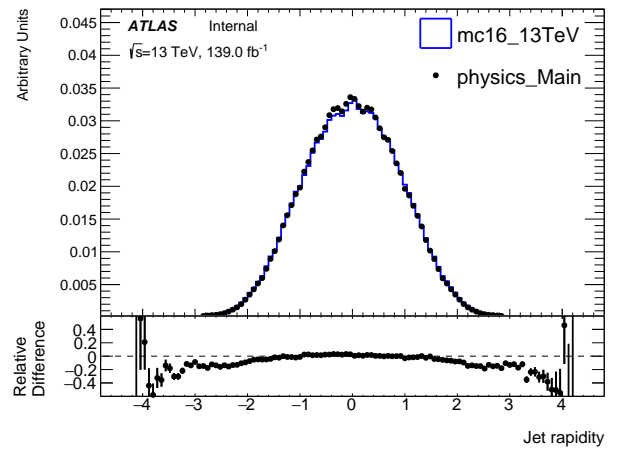
Figure (18): Comparison of MC data (blue) with ATLAS data (black) for (a) and (b) jet p_T, (c) and (d) jet rapidity with both linear and logarithmic axes.

5.6 Background with SWiFt

Searches for new physics related to dijet events are usually conducted through searching for a local excess in events among a smooth background. Traditionally this has been achieved through fitting an ad-hoc high parameter function to the dijet mass spectrum, however with the high luminosities achieved at ATLAS, this method becomes statistically inaccurate. A solution to this is to use a sliding window fit, where background functions are applied to localised subsections of the dijet mass distribution and the background estimations are combined across the entire data range. This is a good solution as potential resonances are expected to be highly localised and therefore a single fit across the entire spectrum is not necessary.

SWiFt [79] is the technique used in this analysis to fit parameters to the background. Using a fixed window size, it applies a fit to this data window to obtain an estimate of the background value for the central bin. This process is repeated, sliding across one bin at a time and combining together the background values of each bin to provide the background distribution. For bins at each end of the spectrum, which do not have enough bins either side of them to be a central bin, SWiFt takes the first and last windows of the spectrum and will use the background fit of each of these to calculate background values for all of the bins below and above respectively. In this study, a window size of 24 bins is used as it has been found to provide a balance between not over fitting and providing good sensitivity. The fit used to estimate the background for the dijet mass distribution is the "dijet fit function"

$$f(x) = p_1(1-x)^{p_2} x^{p_3 + p_4 \ln x + p_5 (\ln x)^2}, \tag{14}$$

where $x = m_{jj}/\sqrt{s}$ and p_i are the fit parameters whose values are determined by minimising the log likelihood function for each window. The function can be fit with four ($p_5 = 0$) or five parameters, or more parameters can be added including higher logarithmic terms of x. To determine which number of parameters has the best fit, both the four and five parameter fits are plotted and their χ^2 values are compared. Both fits can be seen in
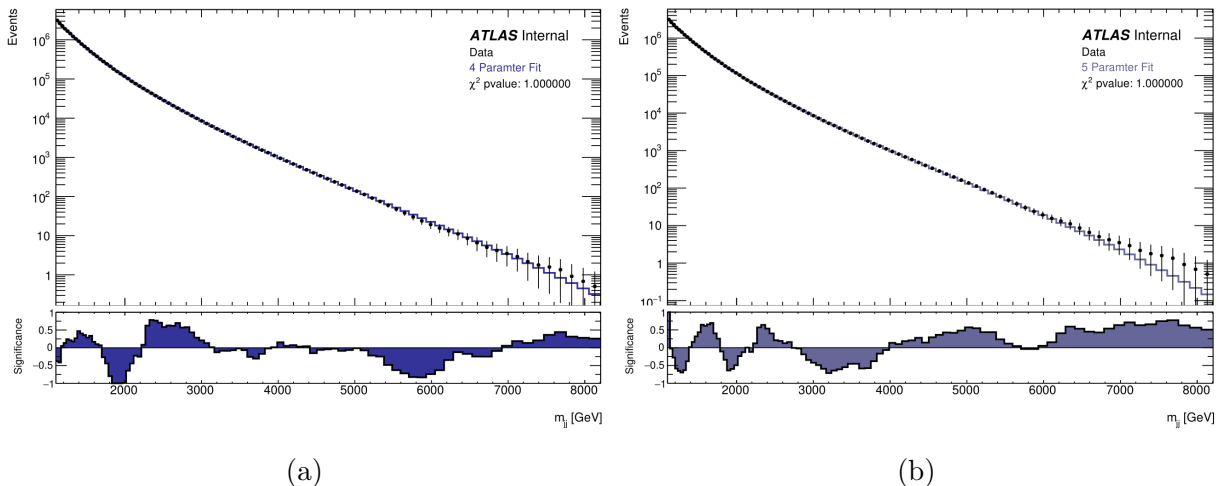
48

Figure (19): SWiFt background fits using (a) 4 parameters and (b) 5 parameters over the dijet mass spectrum.

Figure 19 which shows the SWiFt background fits for the tagged dijet samples.

5.7   Limit Setting

To determine any improvement of quark/gluon tagged data over the untagged dijet samples, the limits placed on their calculated cross sections can be compared. This analysis uses the HistFitter statistical data analysis framework [82] to obtain the limits based on the SWiFt background and dijet signal histograms. These histograms are combined by the HistFactory package to form probability density functions (PDFs) for the signal and background based on the normalised signal (or background) strength per bin. The PDFs produced by HistFactory include information on the signal rates, factors of normalisation for background signals, and nuisance parameters, $\boldsymbol{\theta} = (\boldsymbol{\theta}_s, \boldsymbol{\theta}_b, b_{tot})$, which give a measure of systematic uncertainties.

From the PDFs the likelihood function is calculated as the product of Poisson probabilities for all bins in the histograms [80]

$$L(\mu, \boldsymbol{\theta}) = \prod_{j=1}^{N} \frac{(\mu s_j + b_j)^{n_j}}{n_j!} e^{-(\mu s_j + b_j)} \prod_{k=1}^{M} \frac{u_k^{m_k}}{m_k!} e^{-u_k}, \tag{15}$$

where μ determines the strength of the signal process, i.e. for the background only hypothesis, H_0, $\mu = 0$ and for the nominal signal hypothesis, H_1, $\mu = 1$. The parameters s_j and b_j are the mean number of entries in the jth bin from the signal and background

respectively calculated by integrating the PDFs. n_j is the number of entries in the jth bin of the histogram whose expectation value is $E[n_j] = \mu s_j + b_j$, and the product on the right hand side is intended to take into account the nuisance parameters. This allows a measure of the compatibility of a particular hypothesised value of μ to be defined, called the profile likelihood ratio

$$\lambda(\mu) = \frac{L(\mu, \hat{\hat{\boldsymbol{\theta}}})}{L(\hat{\mu}, \hat{\boldsymbol{\theta}})}, \tag{16}$$

where $\hat{\hat{\boldsymbol{\theta}}}$ is the set of nuisance parameters which maximise L for the given μ value, and the denominator, $L(\hat{\mu}, \hat{\boldsymbol{\theta}})$ is the maximum possible likelihood function where $\hat{\mu}$ and $\hat{\boldsymbol{\theta}}$ are their maximum likelihood estimators. Using the profile likelihood ratio of the signal, HistFitter constructs a test statistic, $q_{\mu_{sig}}$

$$q_{\mu_{sig}} = -2 \ln \lambda(\mu_{sig}). \tag{17}$$

Through running multiple pseudo experiments, a distribution of the test statistic, $f(q_{\mu_{sig}} | \mu_{sig}, \theta)$, is used to calculate the p-value, p_μ for a test hypothesis. HistFitter uses the CL_s method to obtain a confidence level for the signal. This approach calculates the confidence level of the signal as a ratio of the confidence levels of the signal plus background and background only hypotheses [83]

$$CL_s = \frac{CL_{s+b}}{CL_b}. \tag{18}$$

CL_{s+b} and CL_b are determined by finding the probability that the test statistic could be less than or equal to the observed value. The signal hypothesis is said to be excluded at the confidence level when $1 - CL_s \leq CL$. The HistFitter software applies the CL_s method with a confidence level of 95% to set the upper limit on the given dijet sample [81, 83].

6   Results

The purpose of this study was to produce a quark-gluon tagging algorithm to be applied to the dijet mass spectrum which obtains a high efficiency of selecting gluon-gluon dijets, e_{gg}, while maintaining a high exclusion rate of truth quark-quark dijets, $(1 - e_{qq})$, and quark-gluon dijets, $(1 - e_{qg})$. Determining the optimal property for discriminating between jets was thus done by comparing the QQ and QG jet efficiencies for a range of these discriminating properties with a fixed e_{gg}. Using MC simulated truth data the 65%, 75%, 85%, and 95% e_{gg} cut off values for DiffnTrack (the absolute difference in the multiplicity of the two leading jets in an event), and SumNtracks (the sum of the multiplicity of the two leading jets) were found. For each of these values, the percentage of QQ and QG truth jets which passed the cuts were found. Whilst both tagging properties achieved similar exclusion rates for QQ and QG jets, using SumNtracks resulted in consistently higher exclusion rates. The cut off values found were also more linear as a function of $ln(m_{jj})$ and therefore easier to maintain a constant efficiency across the dijet mass spectrum when implementing the tagging cut on the data. For these reasons the sum of the multiplicities is used as the tagging variable for the remainder of this study.

As mentioned previously, the cut off multiplicity, n_{gg}, above which a dijet is identified as gluon-gluon initiated, varies across the m_{jj} range as multiplicities are higher for larger dijet masses. To determine the parameters for the cut off value as a function of $ln(m_{jj})$, n_{gg} was found for a given e_{gg} in each dijet mass data subset. The weighted mean dijet mass was determined for each m_{jj} subset and n_{gg} was plotted against $ln(m_{jj})$ as shown in Figure 20. A linear fit was applied across the dijet mass spectrum to determine the cut off multiplicity parameters.

For the H' signal, both the tagged and untagged dijet samples were passed through SWiFt to obtain the background estimations. The SWiFt backgrounds for the untagged and $e_{gg} = 0.85$ MC simulated sample are shown in Figure 21. It can be observed in both cases that the fit is suitable and consistent with no large overestimations or underestimations outside the low mass region.
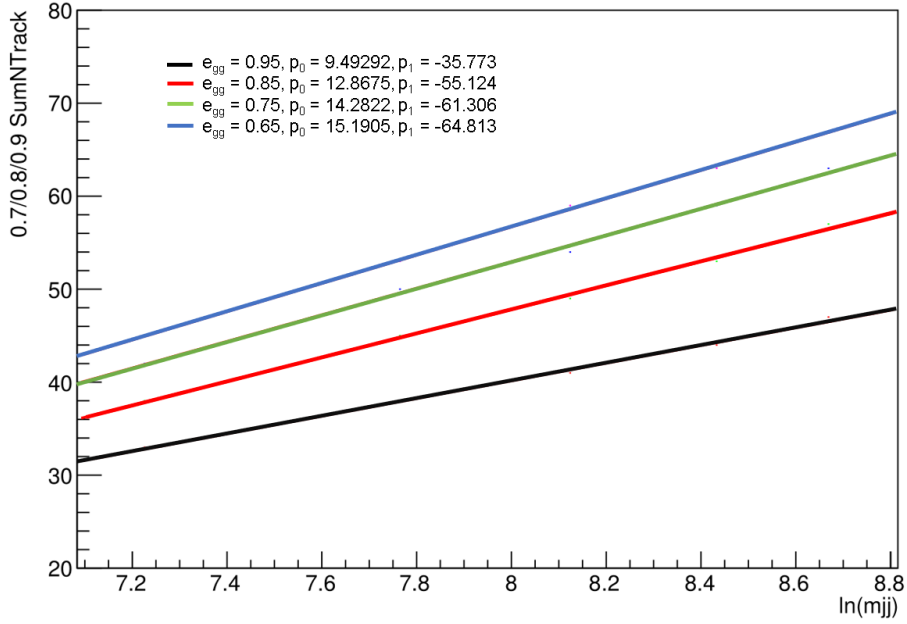
Figure (20): A linear fit was applied to the sum of the multiplicity of the two leading jets for dijet events as a function of m_{jj} to obtain the parameters needed to apply quark/gluon tagging to the events.

The SWiFT backgrounds along with the signal and data histograms are used as inputs to HistFitter which finds the 95% CL_s upper limits for the H' sample for each e_{gg} value. These upper limits were compared with those obtained using no quark-gluon tagging to determine the impact of the cut. These results are shown in Figures 22 and 23, which give the upper limits on the cross section times acceptance times branching ratio for each generated H′ mass where the green and yellow bands represent the tagged simulated data and the solid black line represents the untagged simulated data, along with the ratio between the rates for each cut. For each tested gluon selection efficiency, no improvement in upper limits across the dijet mass region was observed, with the upper limit increasing for lower selection efficiencies. Better results may be obtained by tagging each jet in a dijet event individually based on their multiplicities and determining the discriminant as a function of p_T. This may allow for a higher rejection rate of QQ and QG dijets for a given GG selection efficiency. Other discriminating variables, such as those discussed in the analysis section, may also provide better limits for the H' signal.

A significant limitation of this analysis is the lack of systematic uncertainties applied at each stage of calculation. In attempt to account for this the preceding steps were
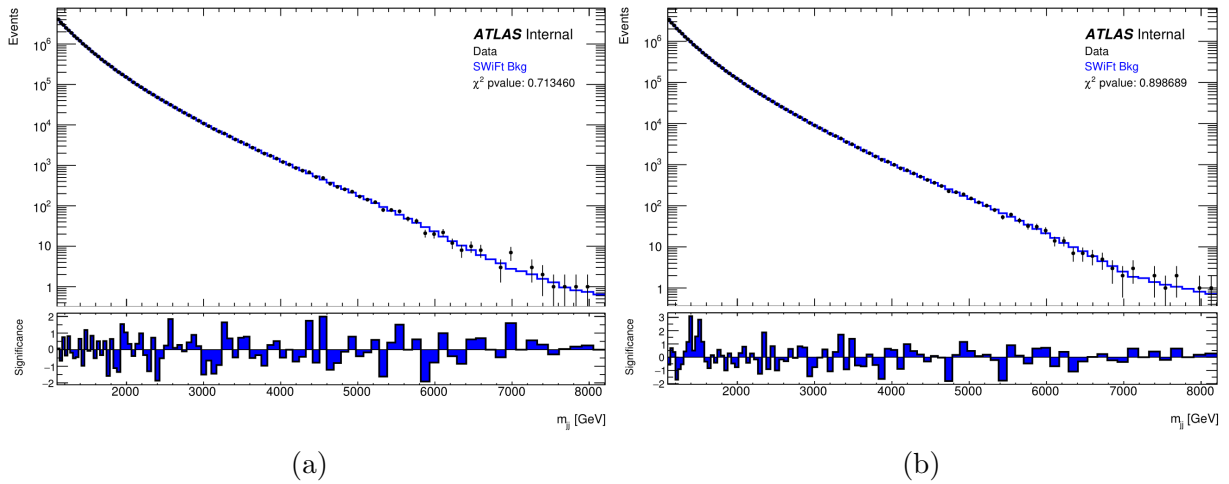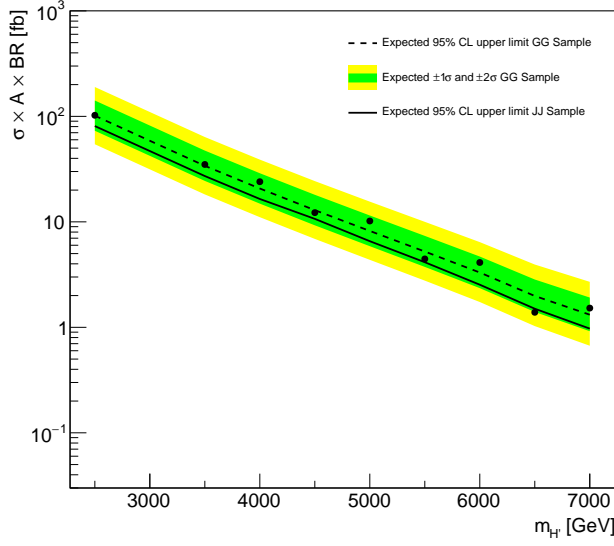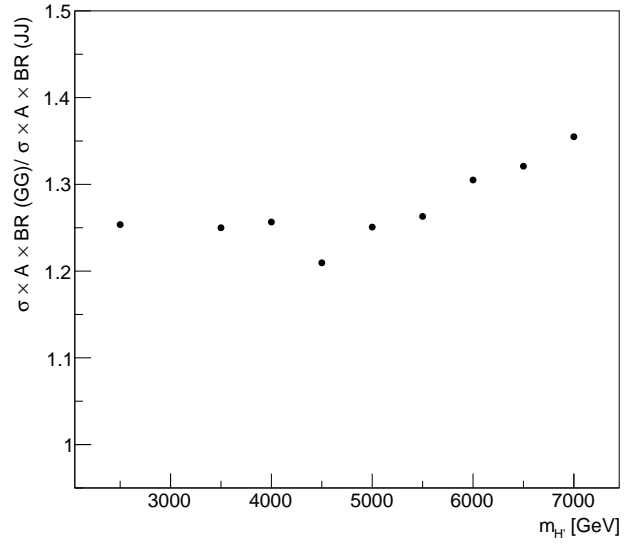
Figure (21): SWiFt background fits for (a) no quark/gluon tagging and (b) $e_{gg} = 0.85$ over the dijet mass spectrum.
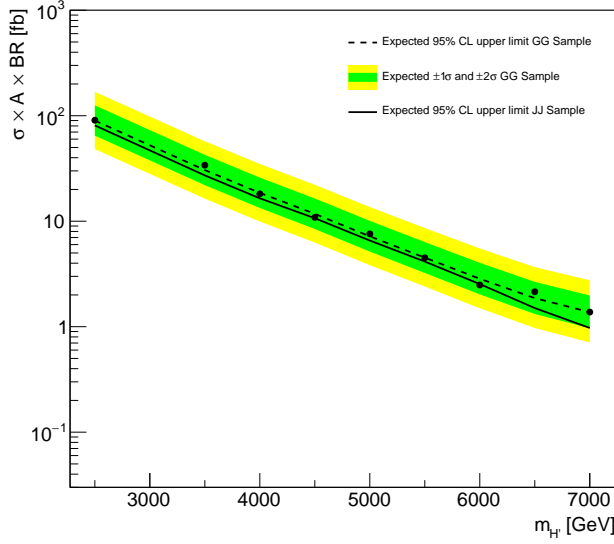
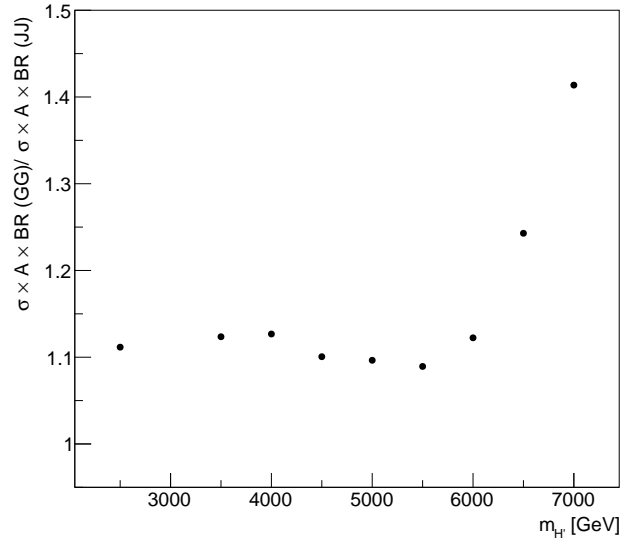repeated four times, each time showing similar upper limits, therefore only one set of results is displayed.
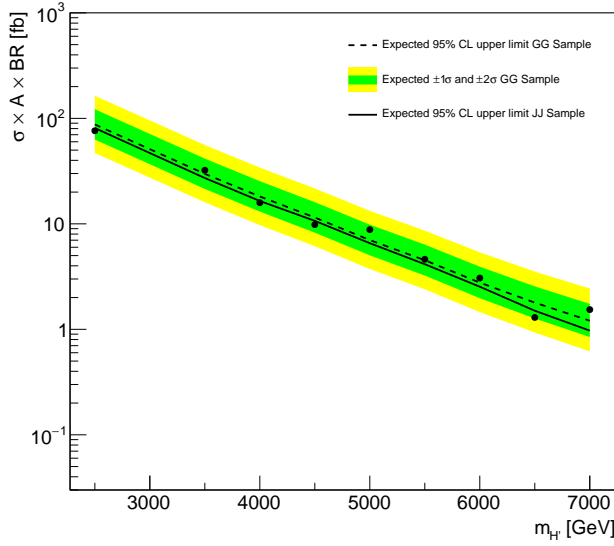
(a)

(b)

(c)

(d)

Figure (22): The 95% CL upper limits for the tagged dijet data (solid line) and untagged dijet data (dotted line and green and yellow bands) samples along with the ratios between their cross sections × acceptance × branching ratio for (a) & (b) $e_{gg} = 0.65$, and (c) & (d) $e_{gg} = 0.75$.
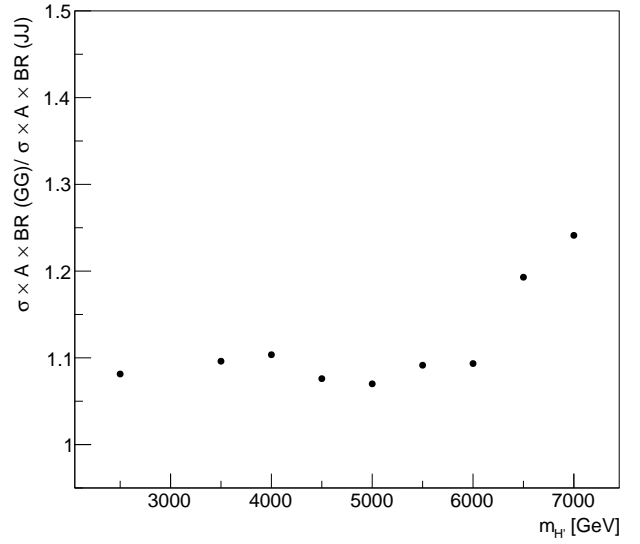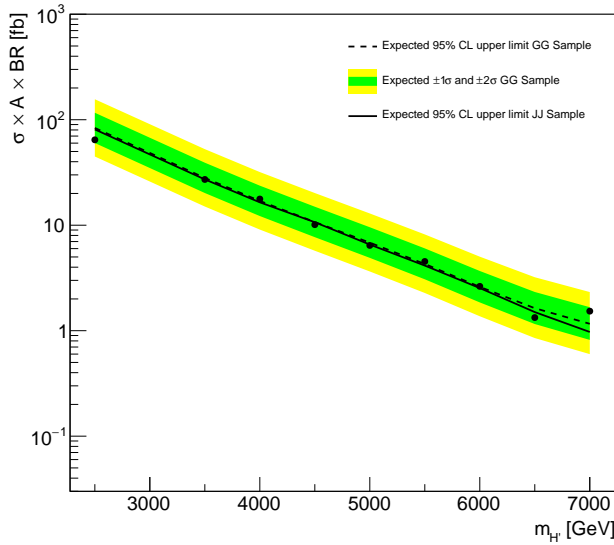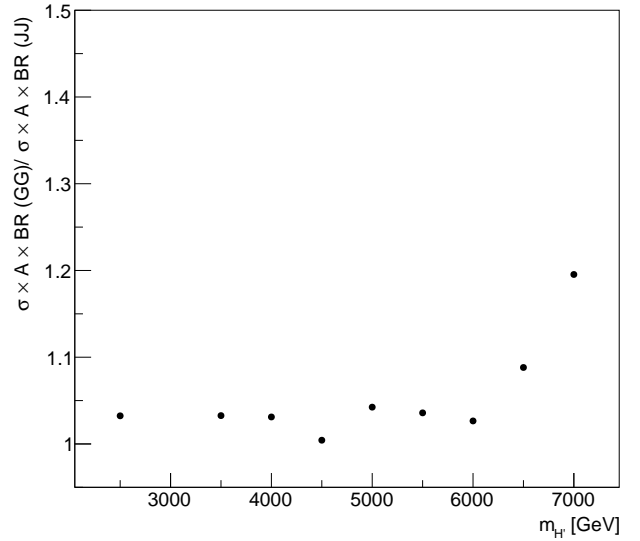
(a)

(b)

(c)

(d)

Figure (23): The 95% CL upper limits for the tagged dijet data (solid line) and untagged dijet data (dotted line and green and yellow bands) samples along with the ratios between their cross sections $\times$ acceptance $\times$ branching ratio for (a) & (b) $e_{gg} = 0.85$, and (c) & (d) $e_{gg} = 0.95$.

7 Conclusion

The purpose of this thesis was to examine the impact of quark/gluon tagging, specifically on a GG dijet signal among a strong QQ and QG background, to determine whether quark/gluon tagging may be a useful tool for identifying dijet resonances which may be indicative of new physics. Specifically, this was done by applying a simple GG dijet cut on an MC simulated H' signal based on the 13 TeV data collected during Run 2 of the LHC at the ATLAS detector. The upper limits were calculated and used to compare tagged and untagged samples. Dijets were identified as GG initiated based on the sum of the multiplicity of the two leading jets in the dijet. The cut off multiplicity was found using a fixed GG selection efficiency using truth jets from PYTHIA 8 simulated dijet data. A background estimation was modelled using SWiFt and 95% CL upper limits on the cross sections across the dijet mass spectrum were calculated with HistFitter.

The resulting upper limits showed no improvement compared to the untagged samples. This could in part be due to the QQ and QG rejection efficiencies not being sufficiently high to reduce the background signal. However, the use of other methods to discriminate between quark and gluon initiated jets such as tagging based on the multiplicity of individual jets, using a more complex algorithm which uses multiple parameters, or fine tuning the selection efficiency for the cut used in this thesis may produce improvements in the calculated upper limits. Additionally, accurate calculation of the systematic uncertainties involved in each stage of this analysis would be necessary to definitively prove any improvements in upper limits for future studies.

References

[1] G. Arnison *et al.*, Experimental observation of large transverse energy electrons with associated missing energy at $\sqrt{s} = 540$ GeV, Physics Letters B **122**, 103-106 (1983).

[2] P. Bagnaia *et al.*, Evidence for $Z^0 \rightarrow e^+e^-$ at the CERN pp collider, Physics Letters B **129**, 130-140 (1983).

[3] S. Abachi *et al.*, Observation of the top quark, Phys. Rev. Lett. **74**, 2632-2637 (1995).

[4] F. Abe *et al.*, Observation of Top Quark Production in $\bar{p}p$ Collisions with the Collider Detector at Fermilab, Phys. Rev. Lett. **74**, 2626 (1995).

[5] T. Patzak, First direct observation of the tau neutrino, European Physical Society, (2000).

[6] G. Aad *et al.*, 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**, 1-29 (2012).

[7] S. Chatrchyan *et al.*, Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC, Physics Letters B **716**, 30-61 (2012).

[8] Plank Mission Brings Universe Into Sharp Focus, `https://www.nasa.gov/mission_pages/planck/news/planck20130321.html` (2013) (accessed 29/07/2020).

[9] Standard Model of Particle Physics image, `https://www.quantumdiaries.org/wp-content/uploads/2014/03/2000px-Standard_Model_of_Elementary_Particles.svg_.jpg` (accessed 29/07/2020).

[10] A. Bettini, Introduction to Elementary Particle Physics, Cambridge University Press. pp 304-346 (2008).

[11] R. Mann, An Introduction to Particle Physics and the Standard Model, CRC Press, pp 9-20, 46-61, 163, 346-352, 483-510 (2010).

[12] R. Ellis, W. Stirling, B. R Webber, QCD and Collider Physics, Cambridge University Press, pp 1-7, 22-39, 63-71, 86-268, 483-510 (1996).

[13] T. Muta, Foundations of Quantum Chromodynamics - An Introduction to Pertubative Methods in Gauge Theories. World Scientific Lecture Notes in Physics -**78**, pp 10-24 (2010).

[14] A. Smilga, Lectrures on Quantum Chromodynamics. World Scientific, pp 292-295 (2001).

[15] S. Marzani, G. Soyez, M. Spannowsky, Looking inside jets: an introduction to jet substructure and boosted-object phenomenology, arXiv:1901.10342v3 [hep-ph] (2020).

[16] Structure and formation of a Jet from a PP collision `https://scholar.harvard.edu/sprince/home` (accessed 30/07/2020).

[17] A. Banfi, Hadronic Jets, University of Sussex, Brighton, UK (2016).

[18] D. Kar, Experimental Particle Physics - Understanding the measurements and searches at the Large Hadron Collider, School of Physics, IOP Publishing, (2019).

[19] Y. Nagashima, Beyond the Standard model of Elementary Particle Physics, John Wiley & Sons, pp 475-512 (2014).

[20] M. Persic, P. Salucci, The Universal Galaxy Rotation Curve, The Royal Astrophysical Journal, **368** 60-65 (1991).

[21] P. Salucci, C. Frenk, The effect on the disk on the rotation curves of spiral galaxies, Monthly Notices of The Royal Astronomical Society, **237**(1), 247-256 (1989).

[22] Y. Sofue *et al.*, Central Rotation Curves of Spiral Galaxies, University of Tokyo Astro, (1999).

[23] M. Persic, P. Salucci, F. Stel, The Universal Rotation Curve of Spiral Galaxies: I. The Dark Matter Connection, Mon. Not. R. Astron. Soc (2008).

[24] Using the Virial theorem: the mass of a cluster of galaxies, `http://spiff.rit.edu/classes/phys440/lectures/gal_clus/gal_clus.html` (accessed 13/08/2020).

[25] S. Ettori *et al.*, Mass Profiles of Galaxy Clusters form X-ray Analysis, Space Science Reviews, **177** 119-154 (2013).

[26] R. Massey, T. Kitching, The Dark Matter of Gravitational Lensing, Rep. Prog. Phys, (2010).

[27] Merging galaxy clusters image, `https://ned.ipac.caltech.edu/level5/March15/Roos/Roos15.html` (accessed 14/08/2020).

[28] J. Feng, Dark Matter Candidates from Particle Physics and Methods of Detection, Department of Physics and Astronomy, University of California, Irvine, California, (2010).

[29] P. A. R. Ade *et al.*, Planck2015 results. XIII. Cosmological parameters, Astromomy and Astrophysics manuscript, astro-ph (2016).

[30] Hubble Space Telescope image of the Abell cluster, `http://www.cosmostat.org/research-topics/weak-lensing` (accessed 15/08/2020).

[31] M. G. Aartsen *et al.*, First Observation of PeV-Energy Neutrinos with IceCube, Phys. Rev. Lett. **111**, 021103 (2013).

[32] E. Andres *et al.*, The AMANDA neutrino telescope: principle of operation and first results, **13** (2000).

[33] Y. Ashie *et al.*, Measurement of atmospheric neutrino oscillation parameters by Super-Kamiokande I, Phys. Rev. D **71** 112005 (2005).

[34] D. Yu. Akimov *et al.*, The ZEPLIN-III dark matter detector: Instrument design, manufacture and commissioning, astro-ph **27** 1 (2007).

[35] E. Aprile *et al.*, The XENON100 dark matter experiment, **35** 9 (2012).

[36] D. Abercrombie *et al.*, Dark matter benchmark models for early LHC Run-2 searches: Reportof the ATLAS/CMS Dark Matter Forum, arXiv:1507.00966 (2015).

[37] R. Harlander *et al.*, Interim recommendations for the evaluation of Higgs production cross sections and branching ratios at the LHC in the Two-Higgs-Doublet Model, arXiv:1312.5571 [hep-ph] (2013).

[38] Y. Nakai1, R. Sato, K. Tobioka, Footprints of New Strong Dynamics via Anomaly and the 750 GeV Diphoton, arXiv:1512.04924 [hep-ph] (2016).

[39] D. Curtin *et al.*, Exotic Decays of the 125 GeV Higgs Boson, arXiv:1312.4992 [hep-ph] (2017).

[40] CERNs accelerator complex, `https://stfc.ukri.org/research/` `particle-physics-and-particle-astrophysics/large-hadron-collider/` `cern-accelerator-complex/`, (accessed 19/08/2020).

[41] ATLAS Multiple year collision plots `https://twiki.cern.ch/twiki/bin/view/` `AtlasPublic/LuminosityPublicResultsRun2` (accessed 18/08/2020).

[42] CERN's accelerator complex, `https://home.cern/science/accelerators/` `accelerator-complex` (accessed 18/08/2020)

[43] L. Robertom, F. Castillo, The Search and Discovery of the Higgs Boson, The Chinese University of Hong Kong, pp. 29-50 (2015)

[44] Image of the ATLAS detector, `https://uonews.uoregon.edu/archive/news-release/` `2010/3/uo-physicists-thrilled-lhcs-opening-new-frontier` (accessed 20/08/2020).

[45] G. Aad, Measurement of the centrality dependence of the charged-particle pseudorapidity distribution in proton-lead collisions at $\sqrt{S_{NN}} = 5.02$ TeV with the ATLAS detector, The European Physics Journal C, **76**, 199 (2016).

[46] R. Stefan, Measurements of ZZ production with the ATLAS detector and simulation of loop-induced processes with the HERWIG event generator, UCL, London, (2017).

[47] The ATLAS Collaboration, ATLAS: A 25-Year Insider Story of the LHC Experiment, World Scientific Publishing, pp 31-58 (2018).

[48] D. Green, At the Leading Edge: The ATLAS and CMS LHC Experiments, pp 69-95, 129-155, 233-259, 329-351 (2010).

[49] The ATLAS Collaboration, Topological cell clustering in the ATLAS calorimeters and its performance in LHC Run 1, European Physics Journal, C77 490, (2017).

[50] M. Cacciari, G. Salam, The anti-k_t jet clustering algorithm, JHEP, (2008).

[51] P. Berta, ATLAS jet and missing ET reconstruction, calibration, and performance, Nuclear and Particle Physics Proceedings, 273-275, (2016).

[52] The ATLAS Collaboration, Jet reconstruction and performance using particle flow with the ATLAS Detector, The European Physics Journal C, **77** 466 (2017).

[53] The ATLAS Collaboration, Performance of the ATLAS track reconstruction algorithms in dense environments in LHC Run 2, The European Physics Journal C, **77** 673 (2017).

[54] The ATLAS Collaboration, Jet energy scale measurements and their systematic uncertainties in proton–proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector, Phys. Rev. D. 38 (2017).

[55] The ATLAS Collaboration, Selection of jets produced in 13 TeV proton-proton collisions withthe ATLAS detector, ATLAS-CONF_029 (2015).

[56] N. Anjos, The ATLAS Jet Trigger for LHC Run 2, Proceedings of Science, The European Physical Society on High Energy Physics (2015).

[57] The ATLAS Collaboration, Data-Quality Requirements and Event Cleaning for Jets and Missing Transverse Energy Reconstruction with the ATLAS Detector in Proton-Proton Collisions at a Center-of-Mass Energy of $\sqrt{s} = 7$TeV, ATLAS-CONF_038 (2010).

[58] The ATLAS Collaboration, Performance of the ATLAS track reconstruction algorithms in dense environments in LHC Run 2, Eur. Phys. J. C, **77** 673 (2017).

[59] The ATLAS Collaboration, The new Fast Calorimeter Simulation in ATLAS, ATL-SOFT-PUB_002 (2018).

[60] Z. Hubacek, Upgrading the ATLAS fast calorimeter simulation, ATL-SOFT-PROC-002 (2016).

[61] B. Abbottet *et al.*, A Comparative Study of the Decays $B \rightarrow (K, K^2)\ell^+\ell^-$ in Standard Model and Supersymmetric Theories Phys. Rev. Lett. **83**, 1908 (1999).

[62] T. Aaltonenet *et al.*, Measurement of the Top Quark Mass and $p\bar{p} \rightarrow t\bar{t}$ Cross Section in the All-Hadronic Mode with the CDFII Detector, Phys. Rev. D **81**, 052011 (2010).

[63] P. Abreuet *et al.*, The ZEUS Leading Proton Spectrometer and its use in the measurement of elastic ρ^0 photoproduction at HERA, Phys. Lett. B **401**, 181 (1997).

[64] K. Ackerstaffet *et al.*, Production of K_S^0 and Lambda in Quark and Gluon Jets from Z^0 Decay, Eur. Phys. J. C8, **241** (1999).

[65] J. Gallicchio, M. Schwartz, Quark and Gluon Jet Substructure, JHEP (2013).

[66] J. Gallicchio, M. Schwartz, Quark and Gluon Tagging at the LHC, Phys. Rev. Lett **107** (2011).

[67] The ATLAS Collaboration, Quark versus Gluon Jet Tagging Using Charged-Particle Constituent Multiplicity with the ATLAS Detector, ATL-PHYS-PUB-009 (2017).

[68] T. Sjostrand *et al*, An Introduction to PYTHIA 8.2, SLAC-PUB-16122 (2014).

[69] The ATLAS Collaboration, ATLAS Pythia 8 tunes to 7 TeV data, ATL-PHYS-PUB-021, (2014).

[70] The ATLAS Collaboration, Summary of ATLAS Pythia 8 tunes, ATL-PHYS-PUB-2012-003, (2012).

[71] P. Skands, S. Carrazza, J. Rojo, Tuning PYTHIA 8.1, the Monash 2013 tune, Eur. Phys. J. C 74:3024, (2014).

[72] The ATLAS Collaboration, The ATLAS Simulation Infrastructure, Eur. Phys. J. C70 pp 823-874, (2010).

[73] S. Hoche, Introduction to parton-shower event generators, SLAC-PUB-16160, (2014).

[74] S. Marzani, G. Soyez, M. Spannowsky, Looking inside jets: an introduction to jet substructure and boosted-object phenomenology, Lecture Notes in Physics, **958** (2019).

[75] T. Sjostrand *et al*, A Brief Introduction to PYTHIA 8.1, FERMILAB-PUB-512 (2007).

[76] S.Agostinelli *et al*, Geant4—a simulation toolkit, Nuclear Instruments and Methods in Physics Research Section A, **506** iss.3 (2003).

[77] T. Sjostrand *et al*, Monte Carlo Generators, CERN-LCGAPP-06 (2006).

[78] B. Allbrooke *et al*, Search for new phenomena in dijet mass and angular distributions from pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector, Physics Letters B, **754**, 302-322, (2016).

[79] K. Sekhon *et al*, SWiFt: Sliding Window Fit Method for Resonance Searches, ATL-COM-PHYS-2018-161, (2018).

[80] G. Cowan *et al*, Asymptotic formulae for likelihood-based tests of new physics, Eur. Phys. J. C71, (2011).

[81] J. Lorenz *et al*, HistFitter - A flexible framework for statistical data analysis, J. Phys: Conf. Ser. **608**, (2015).

[82] M. Baak *et al*, HistFitter software framework for statistical data analysis, Eur. Phys. J. C 75.153, (2015).

[83] A. Read, Presentation of search results: $theCL_s$ technique, J. Phys. G: Nucl. Part. Phys, **28**, 2693-2704, (2002).