A SEARCH FOR EMERGING JETS IN THE DIJET INVARIANT MASS SPECTRUM USING 139 FB^{-1} OF PROTON-PROTON COLLISION DATA AT A CENTER-OF-MASS ENERGY OF 13 TEV WITH THE ATLAS DETECTOR

by

AARON KILGALLON

A DISSERTATION

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

September 2022

DISSERTATION APPROVAL PAGE

Student: AARON KILGALLON

Title: A Search for Emerging Jets in the Dijet Invariant Mass Spectrum Using 139 fb^{-1} of Proton-Proton Collision Data at a Center-of-Mass Energy of 13 TeV With the ATLAS Detector

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

Eric Torrence	Chairperson
David Strom	Advisor
Davison Soper	Core Member
Michael Kellman	Institutional Representative

and

Krista Chronister	Vice Provost for Graduate Studies

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

Degree awarded September 2022

© 2022 AARON KILGALLON
This work is licensed under a Creative Commons
Attribution-NonCommercial-NoDerivs (United States) License.

DISSERTATION ABSTRACT

AARON KILGALLON

Doctor of Philosophy

Department of Physics

September 2022

Title: A Search for Emerging Jets in the Dijet Invariant Mass Spectrum Using 139 fb^{-1} of Proton-Proton Collision Data at a Center-of-Mass Energy of 13 TeV With the ATLAS Detector

A search for emerging jets in the dijet topology is presented here using $139\,\text{fb}^{-1}$ of $\sqrt{s} = 13\,\text{TeV}$ Run 2 ATLAS proton-proton collision data. Emerging jets constitute a class of dark jet models that have long-lived hadronization components, resulting in unique signatures within particle detectors. These jet signatures are the result of phenomenological considerations of self-interacting dark matter. These models provide an explanation for the baryon-antibaryon asymmetry as well as a well-motivated dark matter candidate particle, which make them particularly compelling. Due to the unusual nature of these jets containing high displaced track and displaced vertex multiplicities that vary significantly on the dark sector parameters, machine learning techniques such as unsupervised classification are ideal in the search for these types of models. A Classification Without Labels method known as the CWoLa method was used to extract limits

on heavy Beyond the Standard Model vector boson Z' particles that produce pairs of emerging jets in the large-R dijet topology. Limits were set on the cross-sections of these signatures and exclude Z' particles decaying to emerging jets from 10 fb to 2 fb between masses of 1.3 TeV and 4.0 TeV. Production of Z' particles with masses up to 3.1 TeV and a fixed 20 GeV width were excluded for dark sector couplings down to 0.015. Future considerations for emerging jets analyses are shown in the context of dedicated emerging jets triggers that were designed for use at ATLAS in Run 3.

This dissertation contains previously published and unpublished material and is not an official ATLAS result.

TABLE OF CONTENTS

Chapter Page

Chapter Page

LIST OF FIGURES

Figure Page

Figure Page

LIST OF TABLES

CHAPTER I

INTRODUCTION

1.1. Analysis Motivation

Despite the successes of the Standard Model (SM) [15] of particle physics in providing an excellent model and explanation for interactions of visible matter over a large energy scale, a number of unknowns in physics prompt the need for a more comprehensive model of the natural world. However, the nature and phenomenology of dark matter, the baryon / anti-baryon asymmetry seen in the Universe, and recent tensions in Standard Model measurements [16, 17] point to a large scope of unexplored particles that have yet to be directly measured. A number of searches are ongoing with the ATLAS experiment at the Large Hadron Collider (LHC) for these undiscovered particles; the theoretical phenomenology of their interactions range from Super-Symmetric models to those of exotic dark sectors that have couplings to SM particles.

Indirect evidence for dark matter makes searches for dark matter particles promising avenues of discovery for new physics, and a direct particle discovery would provide a massive boost in understanding of the likely interconnected problems currently observed in the Universe at a wide variety of energy scales. The dark matter hypothesis is a potential resolution for a generic set of energy density over-measurements in the Universe that we currently have no confirmed explanation for. Dark matter composes more than 80% of the matter density and more than 20% of the total observed energy density observed in the Universe [18]. It has no observed coupling with the photon and so it cannot be observed

directly in astronomical observations, but it does interact gravitationally and so its existence can be inferred from large-scale observations of galaxies and galaxy structures.

Dark matter has been traditionally considered in the context of Weakly Interacting Massive Particles (WIMPs) that are a relic from the Big Bang era [19], and while WIMP models are well-considered in the context of large-scale galactic structure, observations at smaller galaxy-sized scales tend to point towards models of self-interacting dark matter. The self-interacting aspect of these models is due to the wide mass hierarchy and gauge structure of the proposed dark sector where dark particles have strong couplings to each other, as compared to the isolated and inert particles seen in other WIMP-like models. The set of dark models that encompass these types of self-interacting dark sectors can also naturally incorporate an explanation for the baryon / anti-baryon asymmetry in the Universe when a portal from the dark sector to the visible matter sector is considered and when the energy scale of dark matter is similar to that of visible matter.

These two general ideas synergize into a unique topology of events that could be observed in proton-proton collisions at the LHC. Known as "emerging jets," this topology involves the creation of either a pair of bifundamental scalars from gluon fusion that decay into 4 fundamental particles that shower to form both dark and visible jets, or the production of a vector boson Z' that decays to pairs of dark quarks that hadronize within the dark sector to form pairs of dark jets. The small sector-to-sector coupling between the dark and visible sectors can cause dark jets to eventually manifest as visible jets, and the small coupling to the SM, in tandem with the similar energy scale between the two sectors, can make dark

mesons long-lived on a wide range of lifetime scales [20]. Therefore, as the progress of the dark shower continues from the production point, this small coupling from the GeV (giga electronvolts) to TeV (tera electronvolts) mass hierarchy in the dark sector creates a cascade of emerging energy from the dark particle decays, producing a set of secondary vertices and displaced tracks. This is known as an emerging jet, and its unexplored dijet topology will be investigated in this analysis. Figure 1.1 is a diagram showing pairs of emerging jets in the tracking and calorimeter regions of a particle detector.

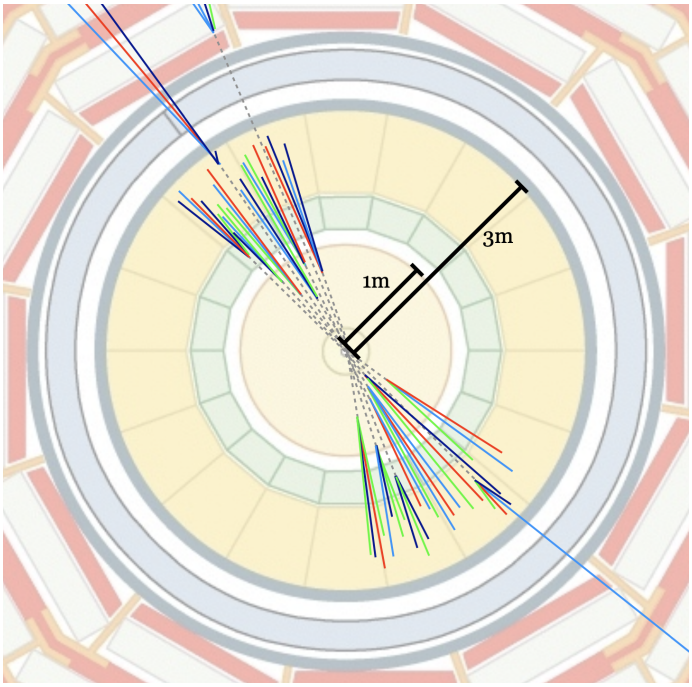


FIGURE 1.1. A schematic of pair production of jets from a pair of dark quarks that emerge within a detector [1].

The theory of these dark decays is well-motivated in the context of the many large-scale issues associated with dark matter. If the similarity between the energy density of dark matter and visible matter is not a coincidence, then the baryon / anti-baryon asymmetry that's seen in both the visible and dark sectors could

be due to the same fundamental asymmetry if there is a portal between these sectors. This requires the dark energy scale to be on the order of the visible sector (1 GeV - 10 GeV) if the number of dark and SM baryons produced during the Big Bang are the same. The TeV scale of the fundamental fields in the dark sector, in combination with this dark sector energy scale and the small SM coupling, ensures that dark mesons in the theory have long lifetimes. The emergent SM particles from decays of these dark meson are therefore produced at macroscopic (> 1 mm) distances from their production point [1].

Dark jet topologies have been explored in the general dark jet searches that have been ongoing at the LHC [21, 22], however, the unique topology produced by the long lifetimes of the dark mesons provides an interesting search case. Rather than relying on model-motivated jet substructure cuts or other specific event attributes, a combination of an invariant mass dijet search with a model-agnostic unsupervised machine learning technique allows a much more involved search for these dark jets with emerging qualities, with less model dependence and enhanced signal sensitivity. This thesis presents a first look at emerging jet signatures in the dijet topology as observed by the ATLAS detector at the LHC using 139 fb^{-1} of Run 2 data.

CHAPTER II

THEORY

2.1. The Dark Matter Problem and Opportunities for Discovery

There is overwhelming evidence for the existence of an abundance of gravitationally-interacting unseen matter in the Universe. Evidence ranging from rotation curves of galaxies, lensing effects, and the well-fitting nature of inflationary models containing dark matter all strongly point to an overabundance of matter that is not directly observable astronomically. The overall structure and effects of dark matter can be indirectly observed, but a specific particle detection and measurement would provide the jumping-off point to understanding the realm of dark particles that are currently unexplored. There is potential for these effects to be due to some other models such as modified gravity [23], however the focus on this analysis is the search for detector-observable models.

Models that explain dark matter vary significantly in their phenomenology and origin of the measured energy densities of visible and invisible matter in the Universe. Standard cosmological models assume the existence of dark matter as Cold Dark Matter (CDM) with some now-frozen relic density that is residual from earlier phases of the universe. These CDM models generally contain Weakly-Interacting Massive Particles (WIMPs); they are stable due to properties of their respective theories and match the frozen relic abundance.

The CDM assumption generally fits well within the standard cosmological model called ΛCDM [24], however, recent phenomenological studies indicate that dark matter might [25] have a self-interacting component. The overall density of

galaxy clusters and large-scale structure appear to be consistent with the ΛCDM model. N-body simulations of galaxy clustering [25] show good agreement between large-scale clusters simulated with CDM dark matter candidates and dark matter candidates that have a self-interacting component. However, discrepancies are noted between these classes of models in the central halo (envelope) densities of matter in galaxies, dwarf spheroid densities, and smaller-scale core substructure in galaxies.

Known as the core-cusp problem [26], simulations of galaxy structure formation indicate disparate core and halo densities from what is observed in a large range of dark parameter space. N-body simulations show a relatively over-dense core with rapidly dropping densities in the halo, however, observed low-mass galaxy rotational curves tend to be consistent with flat dark matter density profiles. In simulations incorporating dark matter with a self-interacting cross-section, this discrepancy can be resolved by this theory in the range up to a cross-section (interaction probability) per mass ratio of $\sigma/m < 1$ cm^2/g. This is comparable to the baryonic interaction cross-section - for protons interacting at 10 GeV, the proton-proton cross section is $\sigma/m \approx 0.023$ cm^2/g.

Additionally, the smaller-scale structures of galaxy clusters are significantly different from the structures seen in simulations. The observed density of dwarf galaxies and similar smaller-scale structures are several orders of magnitude smaller than those predicted by CDM models [27]. This seems to occur in a range of different star-forming regions, indicating that internal pressures such as outgassing from star-forming regions may not provide a good explanation for the structure-smearing effect seen in observations. In simulation with dark self-interactions in the range of $\sigma/m \approx 0.1$ cm^2/g, both these smaller-scale structures

and the core density profiles appear to agree with observational constraints. In order for kinetic energy from the core star-forming region to be flatly distributed, the average interaction rate $\Gamma(r) \approx \rho(r) \cdot (\sigma/m)\, v_{\mathrm{rms}}(r)$ must not exceed the age of the structure, and so the center-of-mass velocity of the collisions is accounted for in these cross-section per mass ratio estimates.

The number of baryons and anti-baryons produced during the Big Bang are expected to be equal due to symmetries of Standard Model interactions. However, the existence of the universe as we know it implies that some unknown mechanism produced a slight asymmetry in the annihilation of early baryon/anti-baryon pairs, leaving the visible matter as we see it now. No complete known dark-baryon / dark-anti-baryon annihilation is seen due to the actual existence of dark matter, so if equal dark baryon / dark-anti-baryon amounts were produced during the Big Bang, an additional dark asymmetry must exist. If the unknown asymmetry that produced the matter / anti-matter misbalance can be attributed to some physics involving a self-interacting dark sector, then an interaction at the same energy scale in the dark sector can explain the dark-matter asymmetry due to the bifundamental interactions in the portal between the sectors. To be consistent with the N-body simulated small-scale structure, the cross-section per mass ratios must satisfy $\sigma/m \approx 0.1$ cm^2/g. However, the dark baryon cross-section must be estimated from recent limits from direct detection experiments, and the mass depends on the dark sector model of choice. With the estimate that the upper limit of the dark matter cross-section is on the order of 1 b (1 barn = $1{\times}10^{-34}$ cm^2) [28] and if the dark sector baryon energy scale is similar to that of QCD (O(5 GeV)), then $\sigma/m = 0.112$ cm^2/g.

A recent set of phenomenology papers [1, 20] explore the unique model and phenomenology of self-interacting dark matter in the context of a hidden valley of dark particles. The gauge structure and properties of the confined dark sector were presented by the authors, and the phenomenology of the 4-jet and 2-jet final states were investigated for the LHC experiments. The gauge bosons associated with this hidden valley would be produced in proton-proton collisions due to this portal from the dark sector to the SM, and detection of these expected bosons at the mass scale of hundreds of GeV to several TeV is possible by reconstructing their resulting byproducts that are called "emerging jets."

2.2. Standard Model

The Standard Model is a remarkably accurate Quantum Field Theory that describes the nature and interactions of all known fundamental particles in the Universe up to the TeV energy scale. Built on generations of successive sets of experimental data and theoretical investigations, it has proven itself in its naturalness and predictive ability over orders of magnitudes in energy scales and cross-sections. With the sequence of particle discoveries made over the last 50 years and new theories to describe their interactions, it is a combined effort of many physicists to build this remarkable theory. It is a union of the ElectroWeak theory of Glashow, Weinberg, and Salam [29, 30] with the theory of the strong interaction [31]. The ElectroWeak aspect of this theory merges the interactions of photons and electrons in Quantum Electrodynamics (QED) with the dynamics of the Weak force. The incorporation of Quantum Chromodynamics within this theoretical framework brought the strong force and its color dynamics in to explain the structure and interactions of baryons such as the proton. The Higgs

mechanism and subsequent prediction and discovery of the Higgs boson gave the framework an explanation for the mass of the elementary particles. Finally, the ad-hoc incorporation of neutrino masses gave an explanation to flavor-mixing in the neutrino sector, and many further years of work have been spent constraining the various parameters of the SM couplings.

The forces in the Universe are mediated by Gauge bosons (photons, Z and W bosons, and gluons). The couplings of Gauge bosons define the type and strength of the forces that they correspond to, which are the electromagnetic, weak, and strong forces, however, the gravitational force is not incorporated into the SM. Matter seen in everyday life, such as protons and electrons, primarily interact via electroweak interactions. Nuclear forces are mediated by the weak and strong nuclear forces, and large-scale structure of the Universe is dictated by the gravitational force. The discrepancies from the Standard Model are seen at large scale variations such as Universe-ranging energy inconsistencies in matter densities and subatomic scale anomalies in measured flavor interactions. The wide range of these discrepancies point to not just a misunderstanding of one aspect or sector of interactions, but to a fundamentally-needed reshaping of our understanding of particle physics.

The elementary particles in the standard model are the 12 primary fermions (spin 1/2 particles), the 5 gauge bosons (spin 1 particles), and the scalar Higgs boson (a spin 0 scalar). The 6 fermions that are associated with the electroweak interaction are the electron, muon, tau, electron-neutrino, muon-neutrino, and tau-neutrino, and the 6 fermions associated with the strong force are the up, down, charm, strange, bottom, and top quarks. Each of these has their own corresponding anti-particle that carries the opposite quantum numbers. The

electroweak interactions are mediated by the photon, the Z boson, and the $W^{+/-}$ bosons, and a hypothetical graviton boson might mediate the gravitational force.

The discovery of the Higgs boson in 2012 effectively completed the theory, as the Higgs mechanism provides an elegant explanation for how elementary particles attain mass [32, 33]. The Higgs boson provides the mechanism that gives particles their mass via interaction with the Higgs field; this is a symmetry violation called the Higgs mechanism. A considerable validation of the model was made by the discovery of the Higgs boson by the ATLAS and CMS collaborations at the LHC, with validation of the predicted interactions of this boson within the Standard Model and a measured mass of 125 GeV [34, 35].

The strong force is, appropriately, the strongest of the forces. At a distance scale of the radius of the proton, $\approx 1 \times 10^{-15}$ m, it has a strength of $\alpha_s \approx 1$. The weak nuclear force mediates nuclear decay and fusion and has a greatly reduced strength, with an effective strength of $\approx 1 \times 10^{-6}$ and an effective range of $\approx 1 \times 10^{-17}$ m. The electromagnetic force is mediated by the photon, with a strength of $\alpha_{EM} \approx 1/137$.

With the exception of gravity, these fields were combined into a gauge-invariant quantum field theory that has been consistently verified across a large set of energy scales and interaction channels. The gauge invariance manifests the set of gauge bosons that mediate the four "distinct" forces (the Electromagnetic, Weak, Strong, and Gravitational forces) in the Universe, and the renormalizable quality of the Standard Model ensure that perturbative calculations can be done to accurately predict the probabilities of these mediated interactions.

The gauge structure of the Standard Model is

$$G_{SM} = SU(3) \times SU(2) \times U(1), \tag{2.1}$$

where SU(3) dictates the gauge structure of QCD, which is described by the set of interactions of 3 colors mediated by 8 gluons. $SU(2) \times U(1)$ is the combined gauge field theory describing the electroweak interactions that is spontaneously broken by the Higgs mechanism. At a high enough energy scale, QED and the Weak interactions combine under this structure. $U(1)$ describes the Electromagnetic force mediated by a single spin-1 boson with U(1) invariance, the photon.

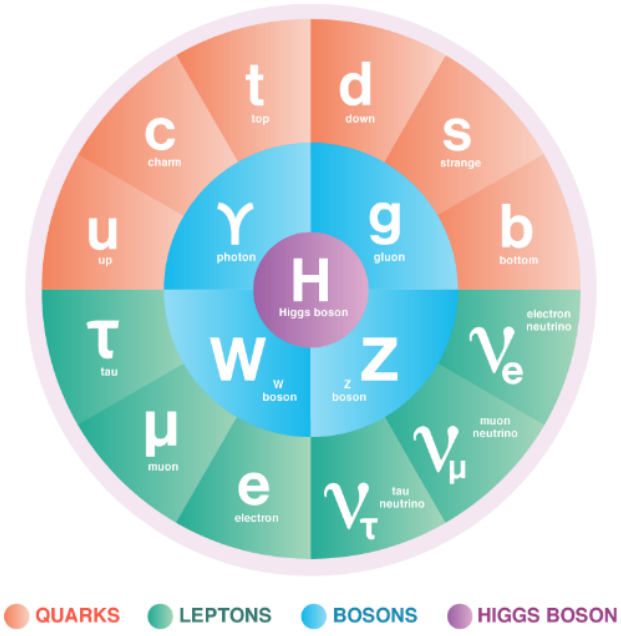


FIGURE 2.1. The Standard Model [2] of particle physics, showing all fundamental particles in the model.

Particles in this Quantum Field Theory are described as fluctuations in fields, and scalars are represented as fundamental fields with the spin-0 property. Fermions are represented by spin 1/2 fields, usually denoted by $\chi(x)$, and vector bosons are spin-1 fields denoted as $A_\mu(x)$.

Nature requires certain symmetries to produce gauge bosons in gauge theory. A local invariance of the Lagrangian density (L) of the theory inherently produces gauge fields that manifest with a certain set of interactions. These interactions are restricted by certain conditions such as whether they carry the charge relevant for each interaction, and relevant conservation laws must also be followed.

The Weak force involves interactions of particles containing two types of quantum numbers, weak isospin (T) and weak hypercharge (Y_W). Fermions in the SM have either right or left-handed chirality, which indicates whether their spins align or anti-align with the vector of their momentum. T = 0 for right-handed particles and T = 1/2 for left-handed particles.

The relationship between EM charge, Weak hypercharge, and weak isospin is $Q = T_3 + Y_W/2$, with T_3 the third component of weak isospin. Neither the Z nor the W bosons carry Weak hypercharge, but as the W boson carries EM charge, any interaction involving the W boson must inherently change T_3, indicating a change of flavor. As the Z boson doesn't carry EM charge or T_3, a much wider set of interactions is possible for the Z boson, for example the allowed couplings to particle/anti-particle pairs of the same flavor.

2.3. QCD

The SU(3) component of the Standard Model is formed by the set of interactions that mediate the strong force via color charges. Three colors are specified, known as red, blue, and green, and the color charge is carried by gauge bosons called gluons. Due to a property called color confinement [36], no free particles can contain a net color charge, so stable hadrons must contain no net color. Baryons are an overall zero-color state with three quarks and any number

of quark-antiquark pairs and of gluons. Gluons are the massless, neutral, and bicolored carriers of this color; they carry one unit of color charge and one unit of anti-color charge.

The representation of the gauge group can be reduced to a set of basis vectors (generators) that do not commute due to the non-Abelian nature of the group. The eight generators of the group correspond to the eight gluons seen in nature, and the generators can be expressed mathematically as the Gell-Mann matrices [31], $T_a = \frac{1}{2}\lambda_a$, $a = 1, ..., 8$.

The QCD Lagrangian therefore contains the set of couplings between fields of color-interacting particles, the 6 fermions (anti-fermions) from the 3 generations of quarks, and the gluon interactions. As the gluons carry color charge themselves, they have their own gluon coupling, and so the radiative processes of a gluon splitting can be difficult to compute in a showering process, especially in the non-perturbative regime.

The QCD Lagrangian can be written as

$$L_{\mathrm{QCD}} = -\frac{1}{4} G^a_{\mu\nu} G^{a,\mu\nu} + \sum_{q=1}^{6} \bar{\psi}_q \left(i\gamma^\mu D_\mu - m_q \right) \psi_q, \tag{2.2}$$

with $D_\mu = \partial_\mu + i g_s \frac{\lambda_a}{2} G^a_\mu$ the covariant derivative containing the fermion and gauge field interactions, and $G^a_{\mu\nu} = \partial_\mu G^a_\nu - \partial_\nu G^a_\mu - g_s f_{abc} G^b_\mu G^c_\nu$ with f_{abc} the commutators of the Gell-Mann matrices $i f_{abc} T_c = [T_a, T_b]$. This contains up to quartic terms of the gluon self-interacting, which is responsible for the color confinement properties consistently observed in a variety of hadronic experiments.

2.4. Hidden Valley Models

As motivated in Section 1.1, a natural extension to the SM gauge group involving a copy of QCD constrained in its own dark sector can provide an explanation for dark matter. This extension that is specified with a set number of dark flavors N_d adds to the set of gauge fields:

$$G_{SM} \times SU\left(N_d\right). \tag{2.3}$$

This extension is confined in the same manner as QCD, and the lightest baryon, the dark matter candidate, is stable due to conservation of a quantum number called "dark baryon number." The mesons do not carry this quantum number and are therefore allowed to decay within the available phase space.

Naturalness arguments point to the number of dark colors (N_d) being consistent between the SM and any hidden sector that has QCD couplings; therefore we consider models with $N_d = 3$ [1]. The connection of this dark sector to the visible sector is formed by interactions of two bosons, the complex scalar X_d and the vector boson Z_d^μ (referred to as a Z' from here on). These bifundamental particles carry both QCD colors and dark colors, thereby acting as a portal between the two sectors. The number of dark flavors n_{df} is set to seven to be consistent with arguments made by Schwaller et. el. - the number of dark flavors can be different from that of QCD, but must lie in the range $2 \leq n_{df} < 4 \cdot N_d$. This requirement is set to ensure that the sector behaved similarly as QCD; at above this $4 \cdot N_d$ threshold, the behavior of the dark sector is no longer asymptotically free.

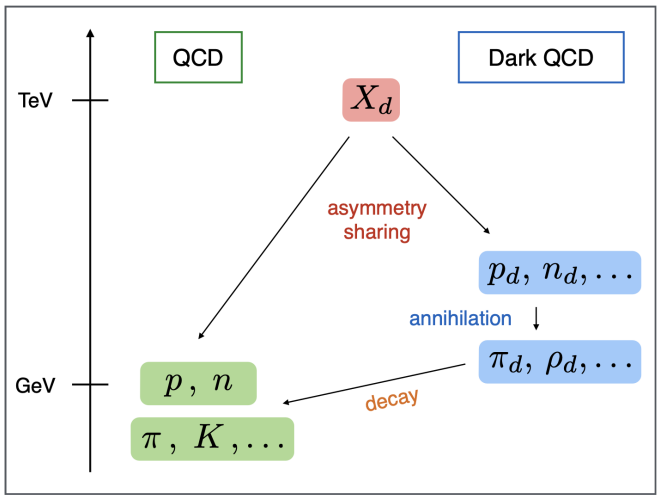


FIGURE 2.2. Graphical representation of the interface between the dark sector and the SM [1]. The dark energy scale Λ_D is roughly at the same energy scale as that of QCD.

The phenomenology of the analysis is concerned with only the hadronic production and dijet decays of the Z'. The Z' is assumed to be leptophobic (not coupling to either SM leptons or dark leptons) as flavor violations would otherwise by seen in the Standard Model measurements, so only the hadronic couplings to the SM and the dark sector are considered. The most simple extension to the Standard Model Lagrangian (L_{SM}) involving just dark quark (Q_d) and Z' hadronic interactions can therefore be written as [1, 37],

$$\begin{aligned} L \supset & L_{SM} \\ & + \bar{Q}_{d_i} \left(\not{D} - m_{d_i} \right) Q_{d_i} - \frac{1}{4} G_d^{\mu\nu} G_{\mu\nu,d} \\ & + \frac{1}{2} M^2 Z'^{\mu} Z'_{\mu} + Z'^{\mu} \left(g_q \bar{q} \gamma_\mu q + g_d \bar{Q} \gamma_\mu Q \right). \end{aligned}$$

The dark quarks are Dirac fermions and interact with the dark mediators with strength g_d, and the bifundamental Z' boson also has a small coupling to the SM with strength g_q. $G_d^{\mu\nu}$ are the dark gluon field tensors, and the couplings of

the dark fermions to the dark gauge fields are contained within their respective covariant derivatives. The set of terms in this Lagrangian define the field density of kinetic energy, the mass energy from self-interactions, gauge-field coupling, and cross interactions between the various fields in the theory. The vector boson Z' is assumed to couple vectorially to SM and dark quarks with coupling strengths g_q and g_d respectively; it is this coupling to the SM that leads to the visible final state of the hadronization byproducts.

Similarly to the QCD sector, the dark quarks can form quasi-stable dark mesons. In the simple dark sector structure considered for this analysis, only dark pions, dark quarks, and dark rhos were allowed to be formed in addition to the dark gluons. The Hidden Valley implementation in Pythia contains a set of partner particles to the SM quarks. These particles, named U_d, D_d, C_d, S_d, B_d, and T_d, are implemented in the sector as generically-name F_d particles and are charged under both the SM and dark colors. These decay to both sectors to a fermion f via $F_d \rightarrow f\, q_d$, with the q_d a heavy Hidden Valley particle in the fundamental representation of dark color [38]. Additional quantum numbers such as charge and SM color are carried by the SM fermion. An additional production mechanism for q_d particles is found from the interaction $Z' \rightarrow q_d\, \overline{q_d}$, which is the channel considered in this analysis and is specified directly in the Pythia configuration. The production of q_d particles is possible through gluon-gluon fusion via pairs of F_d particles, but this analysis is unlikely to be particularly sensitive to this signal due to its topology.

The authors of the Hidden Valley module assert that the coupling strength in the dark sector is relatively arbitrary [38]. Practically, it is difficult to separate effects from the mass hierarchy from effects related to the dark sector coupling α_d

and to its running. While the running of the coupling as a function of the energy scale was not implemented in the Hidden Valley initially due to this argument, it was later added to the sector in the Pythia 8.240 release. In the implementation, dark pseudoscalar mesons $\pi_{d,0}, \pi_d^{\pm}$ are produced and dark vector bosons $\rho_{d,0}, \rho_d^{\pm}$ also couple during the hadronization. The dark scalar mesons have higher mass than the pions and are therefore ignored within the hadronization due to their subdominant production [39]. The emission of a dark gluon from a dark quark $q_d \to q_d\, g_d$ is allowed, in addition to the coupling $g_d \to g_d\, g_d$, which are necessary for the dark jet production.

The dark pion and dark rho couplings were set manually in the production. The neutral dark pion was allowed to decay to pairs of down quarks with 100% branching ratio, and the neutral dark rho was given a 99% coupling to dark neutral pions. The remaining 1% was given to the coupling to a down quark pair to emulate the emerging aspect of the decays. The charged analog particles were given identical couplings but only to their charged dark sector analogs.

The hadronization continues in this Hidden Valley sector to first order until the dark mesons emerge into the visible sector due to their small coupling to the down quark. This can occur via a number of diagrams such as the Z' interaction or via a Yukawa coupling with a dark scalar X_d, although only the Z' interaction is considered in this analysis. The same symmetry-breaking mechanism that produces the SM baryon asymmetry is assumed to produce the DM asymmetry, so similar mass scales $O\,(1 - 10\ \text{GeV})$ are assigned to the dark sector energy scale in the model, which is how the dark pions are long-lived. The dark particles in the model are assumed to be on the order of 5 GeV as the dark matter mass density is known to be approximately 5$\times$ the visible matter density [18], so the mass of the

lightest baryons in the model must be on the order of 5× the mass of the proton if the baryon and the dark baryon abundances are approximately equal. This, in combination with the TeV scales for the dark mediators X_d and Z_d constraining their coupling to the SM, forces the dark mesons to have long lifetimes due to their small couplings and limited phase space available for decays due to the similar energy scales between the QCD and dark sectors. This creates the "emerging" aspect of the resulting jet, where the dark mesons will decay with exponentially-decreasing probability with the mean decay distance $d = \gamma\beta c\tau$, where γ is the standard Lorentz factor $\gamma = 1/\sqrt{1-\beta^2}$ and β is the particle velocity relative to the speed of light $\beta = v/c$.

Lower limits on the dark pion lifetime are set by experimental constraints; QCD lifetimes from the decay of B-hadrons limit analysis sensitivity below several hundred μm. Upper limits are set by Big-Bang Nucleosynthesis constraints, where the lifetime of the produced dark particles must be less than approximately 1 second [1]. It is within the experimental bounds of this lifetime range that this analysis aims to search within, with proper decay distances in the range 1 mm $< c\tau <$ 100 mm.

The lifetimes of the dark pions can be computed analytically as [37],

$$c\tau = 80 \text{ mm} \times \frac{1}{g_d^2 g_q^2} \times \left(\frac{2 \text{ GeV}}{f_{\pi_d}}\right)^2 \left(\frac{100 \text{ MeV}}{m_q}\right)^2 \left(\frac{2 \text{ GeV}}{m_{\pi_d}}\right) \left(\frac{M_{Z'}}{1 \text{ TeV}}\right)^4 \tag{2.4}$$

where g_d and g_q are the couplings of the dark quarks and SM quarks to the Z', f_{π_d} is a constant that is assumed to be on the same scale as the dark energy scale, and m_q, m_{π_d}, and $M_{Z'}$ are the masses of the dark quark, dark pion, and Z', respectively. For an α_s coupling of 0.1, an energy scale of 10 GeV, a quark

mass of 5 MeV, dark quark mass of 10 GeV, and Z' mass of 1 TeV, the $c\tau$ is approximately 3.6 m. These macroscopic decay lengths are significantly dependent on the couplings, the flavor of SM quark that the decay is associated with, and the mass of the Z'. Rather than varying the couplings and flavors of interactions, a highly simplified version of this model is considered where the dark pion lifetimes and Z' masses are set independently within an acceptable grid.

The dark baryon contribution to the jet is expected to be on the order of 10% when 3 dark colors are incorporated into the dark model [1], which is roughly the same fraction as that of QCD. These will escape the detector, but the loss in energy is roughly on the same scale as the Jet Energy Resolution (JER). This detector energy resolution will be discussed in further detail in Chapter III. So while the lifetime of the dark mesons is small enough to allow the decay byproducts to decay mostly within the detector volume, significant experimental sensitivity due to missing energy (MET) is not lost at reasonable lifetimes.

The five dark sector models considered in this analysis span the space of potential dark sector parameters that Schwaller, Stolarski, and Weiler [1] considered. These benchmark models, called Models A - E (Table 2.1), and the dark pion lifetimes were sampled 7 times over these dark sector models to produce $5 \times 7 = 35$ total models for each Z' truth mass. The dark sector meson mass range is motivated by the reconstruction capabilities of the ATLAS detector. However, the dark sector lifetimes are theory-motivated, and the Z' mass range is limited at the lower end (1.3 TeV) by data-acquisition threshold inefficiencies in the detector and at the upper end (4 TeV) by the analysis method. The lower end of the dark meson mass range (0.8 GeV) is motivated by the reconstruction capabilities of the tracker to reconstruct decays of this sort as a clean secondary

vertex. This is a relic of the adaptation of the existing 4-jet emerging jet analysis to this search, where greatly enhanced QCD background jet rejection is expected from secondary vertex counting. This lower range may need to be reconsidered for future analyses in the context of soft-bomb-like signatures, which are signatures containing a large multiplicity of very soft and relatively isotropically-decaying dark particles which do not produce a traditional jet signature after they decay into low-p_T SM particles. The upper end (40 GeV) represents a unique signature that is the limit of the emerging jets phenomelogy; this signature has most of the jet energy contained within a single dark meson. It is expected that an analysis such as the ongoing displaced jets ATLAS analysis will cover this region of the dark sector parameters with greater efficacy.

	Model A	Model B	Model C	Model D	Model E
Λ_d	10 GeV	4 GeV	20 GeV	40 GeV	1.6 GeV
m_{ρ_d}	20 GeV	8 GeV	40 GeV	80 GeV	3.2 GeV
m_{π_d}	5 GeV	2 GeV	10 GeV	20 GeV	0.8 GeV

TABLE 2.1. The dark models (A - E) considered by this analysis.

2.5. Parton Distribution Functions

The proton is a conglomerate sea of particles that are contained by the strong force. Three on-shell particles are considered the standard composition of the proton: two up quarks (q = $+\frac{2}{3}$) and one down quark (q = $-\frac{1}{3}$). However, more than half of the momentum of the proton is carried by various intermittent particles that include transient quark pairs and gluons.

The primary channel of consideration for this analysis is a leading-order s-channel interaction $f\bar{f} \rightarrow Z'_d \rightarrow Q\bar{Q}$ that forms from interactions of on- and off-shell fermion pairs within the colliding protons. Therefore, the simulations

that produce the signal samples needed for this analysis incorporate these PDFs into their simulations of the parton kinematics in the hard-scatter interaction. The kinematics of the resultant Z' are dependent on the carried parton momenta and the recoil of the proton remnants, and the resulting probability of the given hard-scatter (known as the matrix element) is calculated by a convolution of these PDFs over the cross-section of the interaction. The probability of finding a parton of a particular flavor i at an energy scale μ with a momentum fraction x of the proton p is incorporated into the PDF, $f_{i/p}\,(x, \mu^2)$.

2.6. Vector Boson Interactions

BSM vector bosons are generally produced close to their truth mass (on-shell); the probability of this depends on the flavors and energy scales of the production partons. With the assumption that the flavor conservation in the SM extends to the dark sector, standard processes that would produce a Z' involve pairs of fermions within the protons, such as $u\bar{u} \rightarrow Z'$ or $d\bar{d} \rightarrow Z'$.

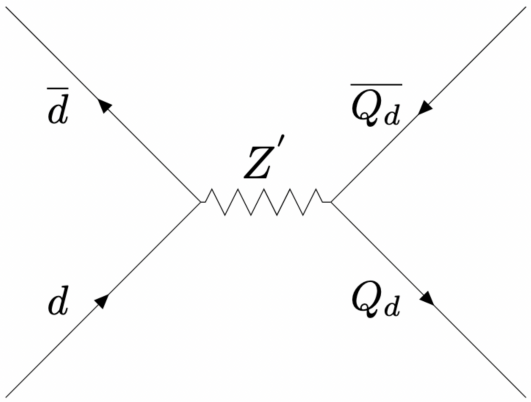


FIGURE 2.3. Feynman diagram of Z' production from a down/anti-down quark pair interaction. The dark quarks (Q_d) will hadronize in the dark sector and emerge in a detector due to their small coupling to the Standard Model.

The approximated Z' width in these hidden valley interactions was computed from perturbative calculations by Schwaller et. al. [37] and is

$$\Gamma\left(Z' \to X\bar{X}\right) \approx \frac{Nng_d^2 M_{Z'}}{24\pi}, \tag{2.5}$$

for a vectorial coupling of a generic particle X. This assumes a minimal contribution to the width from SM decays. With a Z' mass of 1 TeV, an assumed equality between the number of QCD and dark colors (N = 7), three flavors of dark quarks (n = 3), and a branching ratio to each quark flavor of 1/7, this width corresponds to approximately 10 GeV.

The model's full SU(3) dark sector is simplified into a phenomenological model that can reproduce a dark jet to first order. Z' bosons naturally have a coupling to dark quarks and are assumed to have couplings to SM quarks. This coupling is set to a very small number (0.001 in the nominal sample generation) - it is this small coupling that allows the jet to emerge. The dark quarks hadronize in the dark sector at a rate that is dependent on the α_S coupling, which is also dependent on the coupling that is set in the model.

Pythia is a well-known program in high-energy physics that is designed to produce Monte-Carlo simulated Standard Model and signal events [40]. The various Standard Model matrix elements can be computed by the program to a set order, and signal matrix elements needed in the signal generation are approximated by the computation of a set number of Feynman diagrams in the interaction. The program simulates the incoming beam effects, the parton kinematics via the PDFs, the hard scatter interaction via the Feynman diagrams specified in the production, and the final-state hadronization. The beam and parton simulation is a critical part of the program, and the matrix elements for

a simplified Hidden Valley of dark interactions can approximate dark or emerging jet processes.

Hadronization in both the Standard Model and the dark sector is computed first in the perturbative regime. However, the hadronization at energy scales closer to the dark energy scale is estimated via the Lund string model [41, 42] for dark scalars, dark vector bosons, dark mesons, and dark gluons. Dark baryon production is not included in the model, although this is estimated to contribute a maximum of 10% to the jet energy. The Lund string model is a description of the hadronization process that's consistent with lattice simulations and is consistently used as an accurate model of jet formation. The color gauge fields are constrained along tubes that propagate between pairs of colored particles. The energy density of the tubes increase as the separation increases, and at some point the stored energy is higher than the hadronization energy scale, and the string splits. The further cascade of color-charged particles progresses until the energy of the interactions is close to the energy scale of the group. The running of the coupling approaches ≈ 1 at these low energies and the non-pertubative dynamics continue until the final-state Hidden Valley bound states have completed their hadronization. Eventually, these quasi-stable states decay to the Standard Model and emerge within the detector.

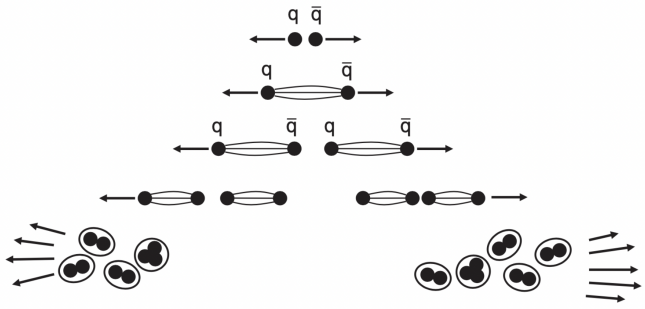


FIGURE 2.4. A graphical representation of how color confinement produces jets [3] in the non-perturbative regime of the Lund string model. A back-to-back pair of colored quarks separate in the center-of-mass frame and radiates gluons and pairs of quarks / anti-quarks - this process continues until around the energy scale of QCD. The attractive forces between the pair narrow the color flux into a tube and the stored energy increases linearly with the separation. At energies above the QCD energy scale, this tube splits into sets of hadrons that continue their own hadronization process until the average energy is below the energy scale of QCD.

The dynamics of these hidden valley sectors are incorporated in Pythia 8, but a limited set of dark sector interactions are considered as a simplified model. A dark quark Q_d is given a simplified set of couplings to dark pions and dark rhos. The dark pions are defined with the model's $c\tau$ and are given a small coupling to down quarks in the Standard Model. To allow the dark showers to progress kinematically, the dark quarks are given a mass of $2\times$ the dark pion mass. To be consistent with the SM, the dark rhos have a mass of $4\times$ the dark pion mass. The 6 dark particles considered in this simplified model of dark showers are in Table 2.2.

The events were generated within this limited fragmentation model for the grid of mass hierarchies and dark meson lifetimes. Variations on the dark sector mass hierarchy are used as systematics on the signal MC, resulting in systematic changes to the signal efficiencies, acceptances, and overall limits obtained.

Particle	PDG Id	Mass	$c\tau$
Z'	4900023	1500-5000 GeV	-
g_v	4900021	0	-
Dark Quark (Q_d)	4900101	$2{\cdot}m_d$	-
Dark Pion (π_d)	4900111	m_d	1-100 mm
Dark Rho (ρ_d)	4900113	$4{\cdot}m_d$	-
Off-diagonal Dark Pion	4900211	m_d	1-100 mm
Off-diagonal Dark Rho	4900213	$4{\cdot}m_d$	-

TABLE 2.2. The dark sector particles considered in these simplified dark shower models. Similarly to their analogous particles in the visible sector, dark gluson g_v, Z' bosons, dark quarks Q_d, and the dark rho particles decay almost immediately after production.

The cross-section measurements for these signals are dependent on the Z' coupling [37] and are estimated by Pythia by convolving the cross sections with the PDF distributions. The variation of these cross section measurements on the dark sector model were within the statistical uncertainty of the Pythia output; the hard-scatter cross section is expected to be independent of the hard sector hadronization at leading order. The decay cross-section of the leptophobic Z' then becomes a convolution over the parton PDFs ($f_i(x_1)$ and $f_i(x_2)$) and the available phase-space for this decay [37],

$$\sigma\left(pp \to Z' \to Q_d\bar{Q}_d\right) = \sum_{i=u,d} \int dx_1 f_i\left(x_1\right) \int dx_2 f_i\left(x_2\right) \times \frac{g_d^2 g_q^2}{72\pi} \left(\frac{x_1 x_2 s}{(x_1 x_2 s - M_{Z'}^2)^2 + \Gamma^2 M_{Z'}^2} \right). \tag{2.6}$$

The cross-sections from Pythia were extracted from the output event generation using the A14 tune of the NNPDF23LO PDF set, and they are listed at a fixed coupling of $g_d = 0.001$ in Table 2.2.

$M_{Z'}$ $[GeV]$	g_q	σ $[fb]$
1500	0.001	$1.714 \times 10^{0} \pm 7.809 \times 10^{-3}$
2000	0.001	$3.425 \times 10^{-1} \pm 1.448 \times 10^{-3}$
3000	0.001	$3.398 \times 10^{-2} \pm 1.049 \times 10^{-4}$
4000	0.001	$7.613 \times 10^{-3} \pm 1.837 \times 10^{-5}$
5000	0.001	$2.776 \times 10^{-3} \pm 5.375 \times 10^{-6}$

TABLE 2.3. The estimated cross sections from Pythia for each generated signal mass.

CHAPTER III

DETECTOR AND ACCELERATOR

The complexity in design and operation of the Large Hadron Collider (LHC) cannot be understated. As the world's most powerful accelerator, with four main experiments and numerous smaller experiments in the complex, it requires the effort of thousands of physicists and engineers to design and operate the machine as well as to analyze the enormous amount of data that is collected.

3.1. Accelerator

The LHC is a 27 km circumference particle accelerator that spans part of the French and Swiss border near Geneva. It is operated by CERN and is capable of performing proton-proton and heavy ion collisions at a number of collision points along the ring. The accelerator ring is located 100 meters underground to greatly reduce the background from cosmic rays, as well as to eliminate radiation leakage in populated areas. The accelerator collides its accelerants at four main points located at four main experiments around the ring by focusing the two beams using powerful superconducting quadrupole magnets.

The maximum center-of-mass (CoM) energy of the accelerator design is 14 TeV, and the accelerator has been approaching this design limit over the run years. The CoM energies were $\sqrt{s} = 8$ TeV during Run 1 (2010 - 2012), $\sqrt{s} = 13$ TeV during Run 2 (2015 - 2018), and anticipated to be $\sqrt{s} = 13.6$ TeV during Run 3 (2022 onward to 2025). This corresponds to a velocity of 99.996% of the speed of light for the colliding protons, increasing the sensitivity of the machine to high-mass exotic particles, Higgs bosons, and rare SM processes.

To probe the high-energy regime of these potential exotic particles, as well as to allow greater sensitivity to the recently discovered Higgs particle, protons are a good candidate for the beam particles at the LHC. They are stable and relatively massive, which greatly decreases the power drain of the accelerator due to synchrotron radiation at these energy scales. The rate of synchrotron radiation is proportional to $1/m^4$ for a given particle of mass m, reducing the energy costs for the accelerator; however, this does increase the needed circumference of the beam ring with the same magnet technology. Lepton beams require smaller accelerator rings with less-powerful magnets, but lose substantially more energy due to radiation. While hadron (protons, anti-protons, etc.) accelerators allow for the probe of much higher energy scales than an electron accelerator, this comes at the cost of very contaminated and complex energy deposits in the detector due to elastic and inelastic pileup collisions. Pileup are the extra energy deposits from additional uninteresting collisions at the interaction point as well as energy deposits from beam effects. Each time a proton bunch crosses another bunch, any number of pairs of protons can produce an inelastic scattering. The typical number of proton pairs interacting per crossing at the LHC is between 15 and 60, as Figure 3.4 shows. This requires complex tagging algorithms to select objects associated to the hard scatter in each event, and the energy calibrations of the detector need to correct for these pileup energy deposits.

The proton beams are formed and clustered into bunches by the preliminary accelerators in the complex. First, hydrogen atoms are ionized by being passed through an electric field; this ionizes the atoms and allows the field to collect the resultant ions. These protons are then accelerated to 50 MeV through a linear series of magnets called LINAC2 [43]. The electromagnets in the series rapidly

oscillate their fields in a manner where the protons are attracted to the forward direction and repulsed from the back of the magnetic series. The protons then travel to the next magnet in the sequence where the same oscillation occurs, producing an incremental acceleration with each magnetic cycle.

These accelerated protons are then fed into the Proton Synchrotron Booster which gives a preliminary boost of the protons to a CoM energy of 1.4 GeV. They are then fed into the Proton Synchroton (PS) and then the Super Proton Synchrotron (SPS) rings. These accelerators are recycled from older programs and used as injection rings for the main accelerator. The final CoM energy of the protons before injection into the main ring is $\sqrt{s} = 540\,\text{GeV}$ after an acceleration to 25 GeV by the Proton Synchrotron.

The bunching of protons into clusters (bunches) is a critical aspect of the accelerator design. The timing separation between sequential bunches was reduced down to 25 nanoseconds (ns) in Run 2 from a spacing of 50 ns in Run 1. To accomplish this, 72 sequential radio frequency cycles of the PS are driven from the PS ring into the SPS ring by a kicker magnet system within a time of 200 ns. The SPS then uses another kicker system with a timing of 800 ns to inject the proton bunches into the main ring. The timing ratio of these injections mean that $72 \times 4 = 288$ bunches are injected into the main ring, but the timing separation requires each grouping of 72 bunches to be separated by 8 empty (200 ns / 25 ns) bunch spacings, and each group of 288 are separated by 32 empty bunch spacings (800 ns / 25 ns). The 2800 maximum total bunches are filled out with the exception of 100 empty bunches that are designed to synchronize with the abort system in the event of a beam dump. This operational configuration was used until

mid 2017, after which the bunch configuration was simplified to a sequence of 32 bunches followed by 8 empty bunches [44].

The main accelerator ring is then used to accelerate the protons up to their final CoM energy. The ring is composed of 1232 dipole magnets spaced around the beamline. These provide the Lorentz force to keep the protons contained inside the beampipe as well as the acceleration needed to bring the proton energies from 450 GeV up to 6.5 TeV. This smears the beams though, so quadrupole magnets placed before the collision points focus the beams from a width of 0.2 mm down to 16 nm. Due to the extremely high fields required, powerful superconducting magnets are needed, and liquid Helium-4 is used to keep them at 2 K to utilize their superconducting capabilities to produce a maximum field of 8.3 Tesla (T).

The beams intersect at four collision points along the circumference which correspond to the locations of the four main experiments at the LHC, ATLAS [6], CMS [45], ALICE [46], and LHCb [47]. The accelerator complex is shown in Fig. 3.1. ATLAS and CMS are general-purpose detectors, whereas LHCb focuses on B-physics and ALICE focuses on heavy ion physics during the dedicated heavy-ion data-taking periods of the LHC.

The field strength needed by the magnets is derived by first considering the effects of the Lorentz force law on protons moving through the ring,

$$\vec{F} = q\vec{v} \times \vec{B}, \tag{3.1}$$

with charge q, velocity vector $\vec{v}$, and magnetic field $\vec{B}$. When the velocity is perpendicular to the field, this simplifies to

$$R = \frac{|\vec{p}|}{qB} \approx \frac{E_{beam}}{qB}. \tag{3.2}$$

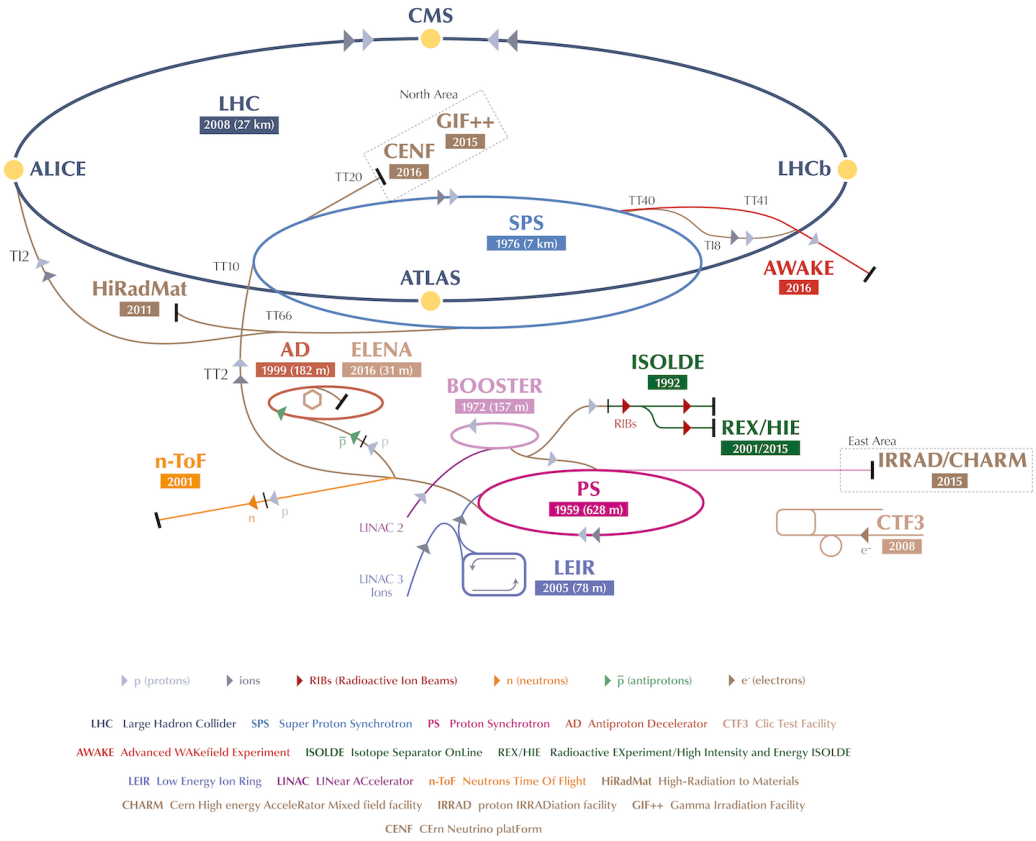


FIGURE 3.1. The CERN accelerator complex, showing the four main experiments on the ring, the linear accelerators and booster rings, and various smaller experiment [4].

For protons, this reduces to $E_{beam}[\text{TeV}] = 0.3 \cdot B \cdot R$, when B in is units of Tesla. For a 27 km circumference ring with radius 4.3 km at a CoM energy of 14 TeV, this corresponds to a minimal field strength of 8.33 T needed for the magnets.

3.2. Detector

ATLAS is a general-purpose particle detector designed for comprehensive particle detection at near-total volume coverage. The symmetric and cylindrical detector spans the 4π polar angle surrounding the transverse region from the

beamline, and pseudorapidity ($\eta = -\ln\tan(\theta/2)$) coverage for the calorimeter extends up to $|\eta| = 5.0$.

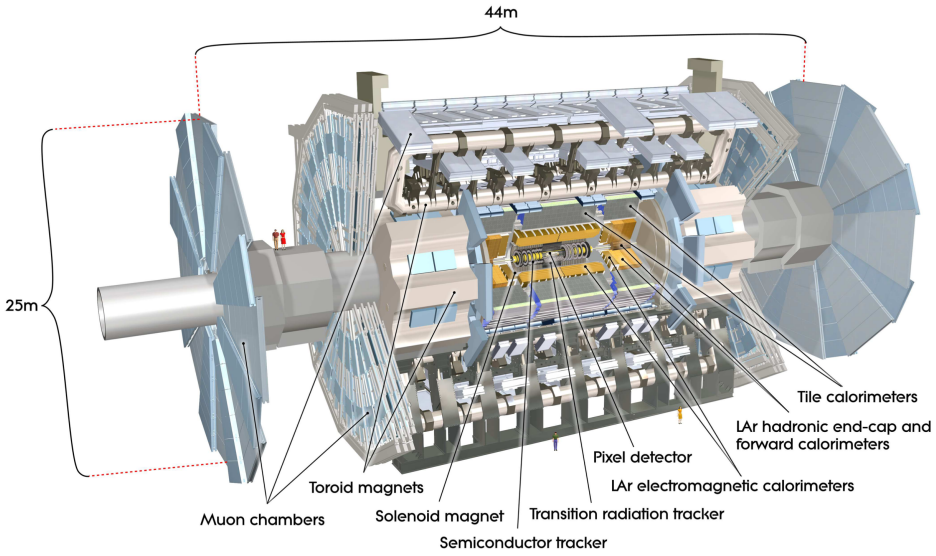


FIGURE 3.2. The ATLAS Detector layout showing the detector's critical subdetectors and magnet systems. Human figures are shown on the left to demonstrate the scale of the detector.

Figure 3.3 illustrates the coordinate systems used at ATLAS, showing the cartesian and cylindrical coordinates used in object geometric and kinematic labeling.

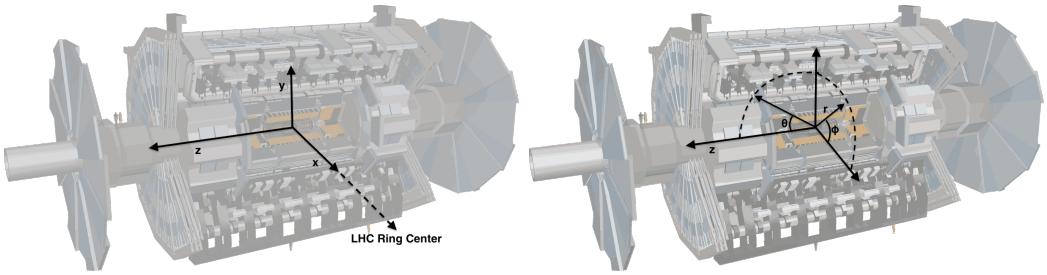


FIGURE 3.3. The ATLAS Coordinate systems.

The kinematic coordinate system at ATLAS uses the variable rapidity (y)

$$y = 2 \cdot \ln \frac{E + p_z}{E - p_z} \approx 2 \cdot \ln \left(\tan \left(\frac{\theta}{2} \right) \right) = \eta, \tag{3.3}$$

where y simplifies to η in the case of massless objects or when $|p| >> m$. η is generally preferred to y as geometric measurements have better resolution than kinematic measurements at ATLAS. In addition, differences in η are invariant to boosts up or down the beampipe, so distance measures using this difference are invariant across the detector for relatively massless objects. One such example is the distance measure ΔR, which characterizes the 2D separation of two objects,

$$\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}, \tag{3.4}$$

where for two objects with (η_0, ϕ_0) and (η_1, ϕ_1), $\Delta\eta = \eta_0 - \eta_1$ and $\Delta\phi = \phi_0 - \phi_1$.

The major detector components consist of the Inner Detector (ID), the Electromagnetic Calorimeter (ECAL), the Hadronic Calorimeter (HCAL), and the Muon Spectrometer (MS), as well as the associated magnet systems for each. An overview of the detector layout is shown in Figure 3.2; this image also shows the impressive scale of the detector, which weighs over 7000 tons.

A range of pileup (exteranous proton-proton collisions in a typical bunch crossing) conditions were present during the data-taking process in Run 2. At maximum design luminosity, approximately 1000 charged particles are expected to be produced within the tracking volume for each event, requiring a precise tracker in a radiation-proof inner detector.

During a bunch crossing, the vast majority of protons in a bunch will not interact with other protons in the opposing collision bunch. The beam focusing defines the resolution of the beamspot in the transverse direction, and the

bunch crossing is spread over a defined range along the beamline (z) to allow for discrimination between the primary collision in the event and the uninteresting pileup vertices. This greatly increases the physics capabilities of the machine, as soft inelastic interactions of the protons are not interesting relative to hard scatter events that have the capability of producing rare Standard Model (SM) processes and Beyond the Standard Model (BSM) signatures. The rate of these background events in the bunch collisions, known as pileup, depend on the beam focusing and the proton density in the bunch. After the initial injection of the 100+ billion protons into the accelerator, the density of the bunch starts to decrease. Proton-proton collisions reduce the number of protons available in subsequent crossings, and protons are also knocked out by beam effects. This reduces the amount of pileup in the events towards the end of a standard injection run, which presents the opportunity for lower threshold triggers to run during this lower-luminosity environment.

Cross-sections in most particle experiments are listed in unit of barns (1 b $= 1\times10^{-28}$ m^2) and overall integrated (time-summed) luminosity is listed in the inverse units, so the product of the cross-section and luminosity gives a number of expected events for that production channel over that given period of integrated data-taking,

$$\mathcal{L}_{\mathrm{b}} \times \sigma_{pp} = 1 \rightarrow \mathcal{L}_{\mathrm{b}} = \frac{1}{\sigma_{pp}}. \tag{3.5}$$

For N_1 and N_2 protons in each bunch of the bunch crossing, this instantaneous luminosity becomes

$$\mathcal{L}_{\mathrm{b}} = \frac{N_1 N_2}{\sigma_{pp}}. \tag{3.6}$$

The rate of pileup events affects the total luminosity captured by the experiment. These N_1 and N_2 protons circulate around the ring many times per second, so replacing this with an average approximation over the circulation frequency f and average number of interactions per bunch crossing $\langle\mu\rangle$ gives

$$\mathcal{L}_{\mathrm{b}} = \frac{\langle\mu\rangle f}{\sigma_{pp}}. \tag{3.7}$$

For N_b bunches within the ring, the instantaneous luminosity becomes

$$\mathcal{L}_{\mathrm{int}} = N_b \frac{\langle\mu\rangle f}{\sigma_{pp}}. \tag{3.8}$$

With a circular frequency of 11246 Hz, a peak μ of around 80 collisions per bunch crossing, a proton-proton cross-section of 68 millibarns, and a 32 filled / 8 empty bunch structure, the peak instantaneous luminosity becomes approximately $\mathcal{L}_{\mathrm{int}} = 2 \times 10^{-34}\ \mathrm{cm}^{-2}\,\mathrm{s}^{-1}$. The pileup profiles for the 4 data-taking years in Run 2 are shown in Figure 3.4.

3.3. Inner Detector

The Inner Detector of ATLAS is composed of a series of tracking layers that are designed to detect and compute the kinematics of charged particles coming from the collision point. The silicon technologies in the ID utilize ionization energy loss from the motion of charged particles through the material to measure their curvature and interaction point (IP) position. The detector lies within an axial magnetic field of strength 2 T produced by a thin superconducting solenoidal

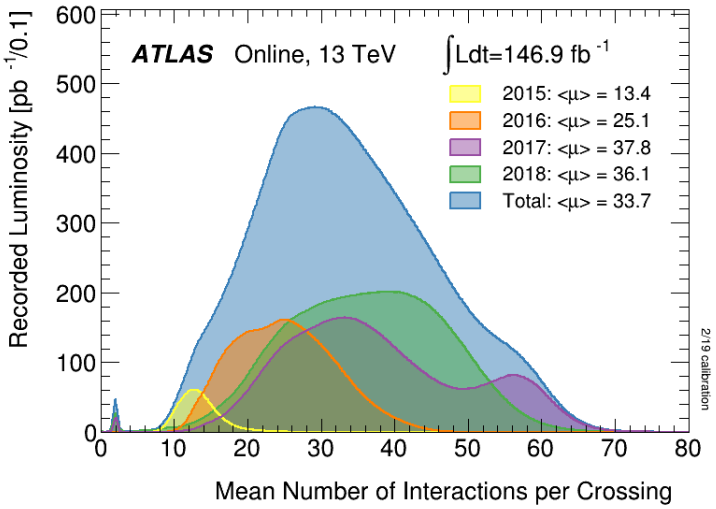


FIGURE 3.4. Average ATLAS μ profiles during data-taking years in Run 2 [5].

magnet; this field produces the curvature of the charged particles' tracks that is needed for momenta measurements.

The layout of the ID is shown in Figure 3.5, which shows the Pixel, SCT, and TRT subdetectors [6]. The ID is immersed inside an axial field produced by a 5.3 m long and 2.5 m diameter solenoidal magnet. The silicon layers in the ID span a range of $|\eta| < 2.5$ and are arranged coincentrically in the barrel region ($|\eta| < 1.7$) and in disks surrounding the beamline up to the maximum η acceptance ($1.7 < |\eta| < 2.5$).

3.3.1. Track and Vertex Reconstruction

The curvature of a charged particle in a magnetic field is dependent on the particle's charge, mass, momentum, and the strength of the field. Charged particles in an perpendicular field follow a helical path, with the orientation of the path dependent on the charge of the particle.

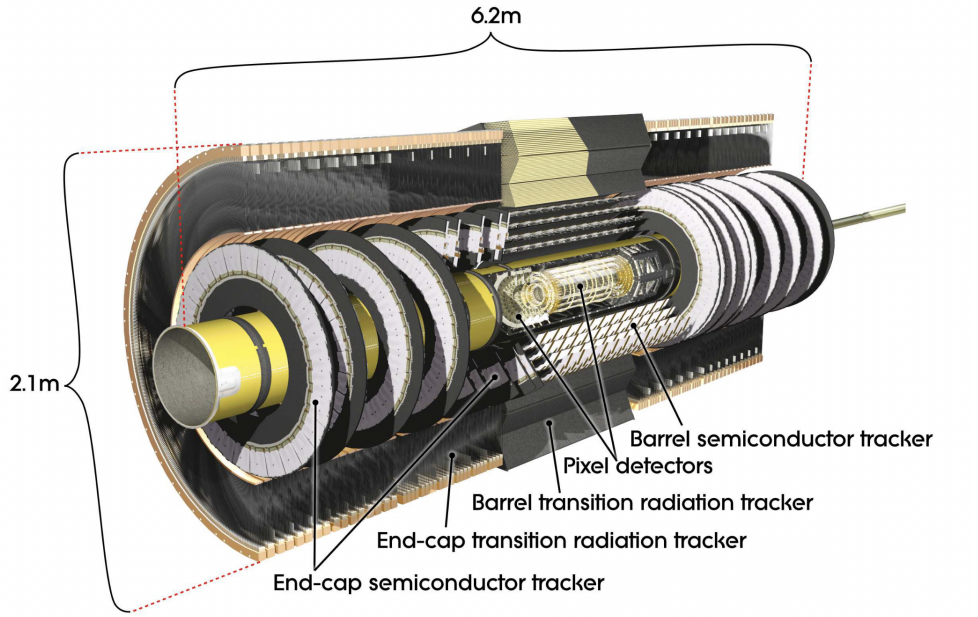


FIGURE 3.5. The ATLAS Inner Detector showing the various tracking layers and technologies used for charged particle identification [6].

Charged particle tracks are defined by 6 independent quantities that represent the particles' kinematics relative to either the relevant primary or secondary vertex in the event, $(q/p, \theta, \phi, d_0, z_0, m)$.

The trajectory of charged particles moving through the ID's magnetic field are curved by the Lorentz force law; the geometric location and curvature of the measured hit positions within the ID are used to construct the origin and kinematics of the charged particle producing this curved track. Analytic functional forms are fit to the layer hits and used to measure the track parameters such as the track's η and ϕ, which is critical for the ΔR association of the track to its corresponding jet.

The variable q/p is the ratio of the track's charge to its momentum, both of which are determined by the curvature of the track in the magnetic field. The curvature direction defines the charge, and for a particle of given mass m, the

track has a defined curvature within the measured magnetic field. θ is the polar angle of the track's initial vector relative to the beamline, and ϕ is the azimuthal angle of the track's initial vector. d_0 and z_0 are positional variables indicating the location of the track's perigee relative to a given vertex. d_0 is the signed quantity indicating the shortest distance between the track and the vertex in the azimuthal plane (x-y plane). Similarily, z_0 is the signed quantity indicating the perigee in the polar plane (r-z plane). A schematic of this coordinate system is shown in Figure 3.6.

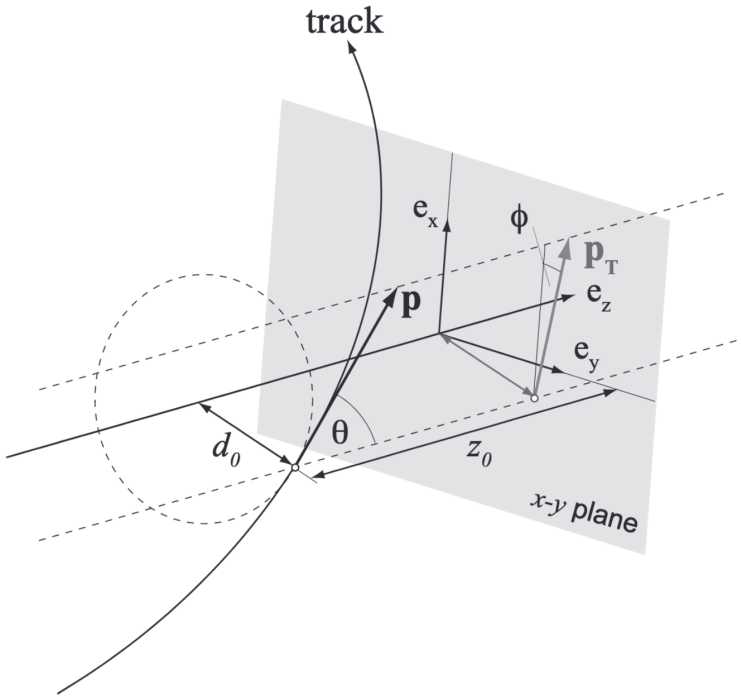


FIGURE 3.6. The track reconstruction geometry at ATLAS. Relative to a defined vertex, geometric descriptions of the closest approach and the kinematics are reduced to 6 variables [7]

.

For every bunch crossing, 10 - 60 pairs of protons (on average) interact and produce a set of charged and neutral particles that can be associated back to their interaction points. The pointing of neutral particles such as photons is very

uncertain, so precise measurements of the tracks are relied on to locate the x, y, and z locations of these IPs. The interaction of interest, the hard-scatter event, is associated to a vertex by requiring at least two tracks with $p_T > 500\,\text{MeV}$ that originate from this vertex, and the vertex with the highest scalar sum of track p_T measurements associated to it is labeled the event's Primary Vertex (PV).

$$\text{PV} \leftarrow \max \sum_{\text{tracks}} p_{T_{\text{track}}} \left(p_{T_{\text{track}}} > 400 \text{ MeV}\right) \tag{3.9}$$

3.3.2. IBL

The Insertable B Layer (IBL) is the first set of tracking layer that charged particles coming from the IP will interact with. This subdetector consists of a series of silicon pixel sensors located 32.5 mm from the beamline. This layer was added in 2016 and was designed to improve tracking sensitivity in the ID, allowing for better primary and secondary vertexing resolution. The resolution of the pixels in this layer is $50\,\mu\text{m} \times 250\,\mu\text{m}$, adding an additional 12 million readout channels to the ID hardware.

A charged particle passing through a silicon pixel at a sufficient momentum is likely to produce least one electron-hole pair within the material. An applied voltage attracts the electrons towards electronics that register the electron accumulation as an energy hit within that pixel. This amplified electron signal that's collected is used to define an ionization energy deposit if this energy deposit is well-measured and sufficiently above noise fluctuations. The sensor detects an ionization hit if the accumulated charge passes defined "time-over-threshold" (TOT) thresholds that are calibrated to ensure consistent performance

of the pixels over the detector. The addition of this layer increases the d_0 and z_0 resolution of the ID by 40% [48].

3.3.3. Pixel Detector

The next three layers comprise the Pixel detector [49], which is a high-resolution set of tracking layers. The coverage of the pixel detector spans up to $|\eta| = 2.5$, and its coverage is composed of a central barrel ($|\eta| < 1.7$) and an endcap region ($1.7 \leq |\eta| < 2.5$. This subdetector is composed of a series of silicon pixels of size $40\,\mu\text{m} \times 500\,\mu\text{m}$ arranged in three layers. The first layer starts 50.5 mm from the beamline and the last layer extends out to 122.5 mm. The high granular resolution of the subdetector, especially in the ϕ coordinate, allows precision measurements of the curvature of the tracks that produce sufficient numbers of hits in the ID. The 92 million channels in the Pixel detector are fed to the readout system that stores the hit information if the event passes a trigger.

The high 3D spatial resolution of these layers are utilized in b-tagging, where it is necessary to have precision tracks that point to the secondary vertex created by long-lived B-hadrons in a jet.

3.3.4. SCT

The Semiconductor Tracker (SCT) is subdetector containing the next series of tracking layers [50]. The barrel region ($|\eta| < 1.4$) is composed of 4 layers of silicon strips arranged orthogonally to each other. This gives resolution along the two measured coordinates - along ϕ for each hit radius, and similarly for each z position. The forward regions ($1.4 \leq |\eta| < 2.5$) are covered by 9 disks with a similar orthogonal orientation of the strips.

This subdetector has slightly worse spatial resolution than the preceding layers due to the use of coarse strips instead of pixels, however, most tracks are well separated at that distance from the IP. So the 200 μm track resolution in this subdetector is sufficient for proper track reconstruction and provides a similar contribution to the track's momentum resolution as the precision layers closer to the IP. This subdetector adds an addition 6.3 million channels to the detector hardware readout.

3.3.5. TRT

The Transition Radiation Tracker (TRT) is a series of straw tubes filled with an argon and xenon gas mixture [51]. The 300,000 4 mm diameter tubes fill a substantial volume of the inner detector. The tubes act as cathodes and a wire running through the center of each tube acts as an anode, so the presence of charged particles produced in that volume induces a voltage through the tube that can be amplified and read out. Transition radiation is created by the boundaries between the layers in the tube structure, as well as the transitions from the gases inside and surrounding the tubes themselves. This radiation interacts with electrons in the tube's gas and accumulates them towards the readout electronics.

The induced voltage in the tube gives the geometric location of the hit in the ϕ coordinate. They are arranged parallel to the beamline in the barrel region and in rings surrounding the beamline in the forward regions, meaning that they only provide r - ϕ geometric hit locations. However, charged particle reconstruction is significantly enhanced by the addition of these high-radius hits, so the dramatically worse spatial resolution in this region complements the momentum measurements of the tracks, rather than relying on a costly set of silicon layers. The geometric

resolution is 130 μm in this region, although the large number of high-radius hits results in a roughly equal momentum resolution between the SCT and TRT regions.

3.4. Electromagnetic Calorimeter

The ID is intended to measure charged particles' properties while inducing a minimal effect on the kinematics of the particle itself. However, the calorimeter's purpose is to induce interactions from incoming particles to measure their energy. An electromagnetic or hadronic interaction with the incident particle will produce additional byproducts; these daughter particles will continue to cascade through the calorimeter layers. As each subsequent interaction produces additional interactions at a lower energy scale, the cascade will eventually complete its showering. The number of interaction lengths in the calorimeter is designed to ensure that little-to-no energy is lost out of the calorimeter. At some high energies, a sizable portion of the energy punches through all the detector layers, but this is accounted for by the jet calibration sequence. Other missing energy in the event is usually associated with low electromagnetic and hadronic cross-section SM particles - in particular the high SM background by neutrinos that pass readily through the detector.

The Liquid Argon Calorimeter (LAr) is the first detector in the calorimeter system [52]. It is designed to induce showering processes for electromagetically-interacting particles via Brehmsstrahlung with the highly dense liquid argon contained within the detector. Layers of absorbing lead allow greater absorption of the incoming energy by inducing more electromagnetic interactions, thereby greatly reducing the volume of cryogenically cooled liquid argon needed. LAr

is stable and relatively radiation-proof, allowing the detector resolution to stay constant over the lifetime of the detector.

LAr is composed of two half barrels, each with length 3.2 m, inner radius 2.8 m, and outer radius 4.0 m. It is designed with 3 main layers, with the highest granularity in the first layer to allow photon identification. The granularity decreases to the second layer and further on to the third layer. The final layer's purpose is to provide an energy measurement to correct for energy leakage due to punch-through [11].

The LAr system resides in the region of $|\eta| < 3.2$ and is composed of a barrel calorimeter with coverage up to $|\eta| < 1.475$ and two endcap wheels completing the remaining coverage. The granularity of this detector is η dependent and is the most granular in the first layers of the central region with cell dimensions of 0.003×0.1 in η - ϕ space. This granularity in the region of the ID allows for precision electron energy measurements as the particle track can be reliably extrapolated to the corresponding calorimeter cell hit. Less precision in the showering shape is needed, so the granularity of the second and third layers are 0.025×0.025 and 0.05×0.025.

The energy resolution of calorimeter systems can be approximated by the functional form

$$\frac{\sigma_E}{E} = \frac{\mathrm{a}}{\sqrt{E}} \oplus \frac{\mathrm{b}}{E} \oplus c. \tag{3.10}$$

The variable a parameterizes the energy resolution component that is caused by sampling errors. The b variable term parameterizes the resolution due to electronic noise, and c is a constant term associated to the innate resolution of the technology. In the case of the ATLAS LAr calorimeter, the noise contribution

is insignificant and the total resolution can be parameterized by Equation 3.11, which is the design resolution of the detector [53]

$$\frac{\sigma_E}{E} = \frac{10\%}{\sqrt{E/\text{GeV}}} \oplus 0.7\%. \tag{3.11}$$

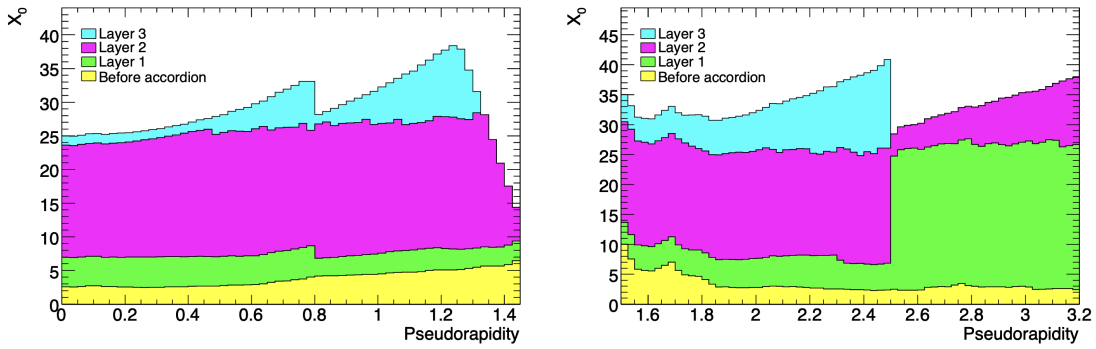


FIGURE 3.7. Electromagnetic interaction lengths as a function of η for the electromagnetic calorimeter system [6].

3.5. Hadronic Calorimeter

Interactions in the ECAL rely on Bremsstrahlung interactions with electron clouds within the detector material, whereas measurements of interactions in the HCAL rely on hadronic interactions with high-mass nuclei in the detector.

The HCAL sytem is composed of several sampling calorimeters with varying resolution and detector technologies. The barrel region has the highest design energy resolution and is composed of a scintillator tiles with steel absorbers in the region $|\eta| < 1.7$. LAr calorimeters comprise the region $1.7 < |\eta| < 3.2$ and also the forward regions $3.2 < |\eta| < 4.9$ near the beamline.

The HCAL barrel region has an inner radius of 2.28 m, an outer radius of 4.25 m, and is 5.8 m long. It is composed of three layers with interaction lengths λ_0 of 1.5, 4.1 and 1.8. The material budget decreases slightly in the forward region,

where the three layers have λ_0 values of 1.5, 2.5, and 3.3. The material budget for each of these regions is shown in Figure 3.8.

The layering of these tiles is accordion-shaped to allow relatively seamless transition between the modules and the central and forward regions. The central region is composed of 64 modules that contain samplers with resolution $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$ for the first two layers and 0.2×0.1 for the last layer [54].

The LAr technology is incorporated in the forward regions to allow for radiation hardening against the significant amount of pileup seen in those regions.

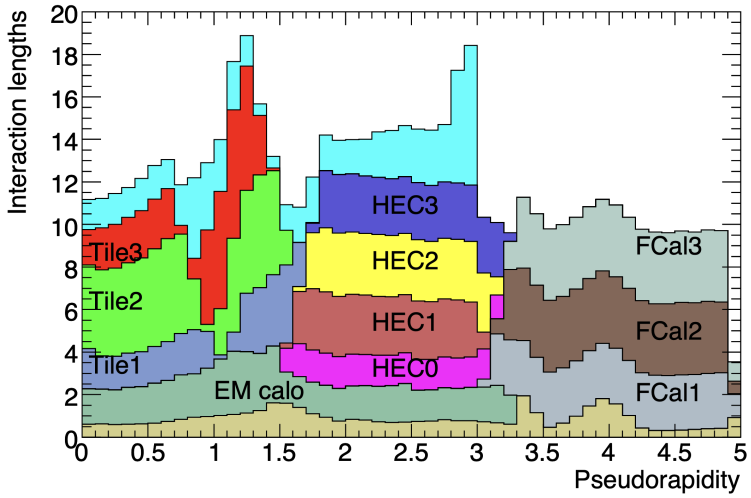


FIGURE 3.8. Hadronic interaction lengths as a function of η for the hadronic calorimeter system [6].

3.6. Muon Spectrometer

The MS is a tracking system that surrounds the calorimeter systems and is designed to precisely measure muon charges and momenta. Muons pass through the calorimeter systems with little energy loss due to the mass dependence of Bremsstrahlung processes. Muons can produce tracks in the ID, but their heavy mass induces very little curvature in the tracks, making the track p_T resolution

worse in that region. The high radius of the tracking layers in the MS creates an extension to the tracking system that is used for precision muon detection and measurement.

The MS is formed of a series of layered tracking technologies. The tracking measurements are done with high-precision tracking chambers that are composed of monitored drift tubes (MDTs) and cathode strip chambers (CSCs), with the triggering enabled by resistive plate chambers (RPCs) and thin gap chambers (TGCs). These systems are immersed in a $\approx 2\,\mathrm{T}$ magnetic field that is provided by barrel and end-cap air-core torroidal superconducting magnets.

3.7. Trigger System

With a bunch crossing rate of 40 MHz, approximately 1 billion proton-proton collisions occur per second at peak luminosity. The full raw detector information can be stored in 1.6 MB, so a disk-writing capability of 64 TB/s would be required to store all the data taken at the detector. As this is not possible with the currently available resources, a complex data acquisition system is needed to reduce this to a manageable rate.

The ATLAS trigger system, shown in Figure 3.9, is a sequence of hardware and software decisions used to record events usable for physics studies. The Level-1 (L1) system is a hardware-level set of trigger decisions used to reduce the raw trigger rate from 40 MHz to 100 kHz. These events are then sent to the High Level Trigger (HLT) system, which further reduces the data to a manageable rate of roughly 1 kHz. The trigger system is sensitive to a variety of objects at both L1 and HLT levels, such as muons, photons, and jets. The triggers can be prescaled,

meaning that a fraction of events are randomly discarded to reduce the output rate, or they can be unprescaled.

At L1, events are characterized by processors searching for energy deposits that are selected algorithmically for certain signatures. The L1Calo system will reconstruct calorimeter energy deposits and tag them using simple jet reconstruction algorithms. For example, a simple jet-finding algorithm is computed on topoclusters by considering calorimeter energy deposits in square 0.4×0.4 regions and summing them. Other systems, such as the photon trigger, look at isolated energy deposits within a small number of topoclusters that deposit most of their energy in the ECAL; these are indicative of a photon signature [55]. The regions surrounding these L1 objects are lumped into Regions of Interest (RoIs) and are passed to the next steps of the trigger system.

These hardware decisions and RoIs are passed to the Central Trigger Processor (CTP). This processor collects decisions and ROIs from the preceding systems and computes general event trigger decisions based on kinematics, trigger object separation, and multiplicities of the various L1 objects fed to it. The decisions incorporate the subdetector timing information to ensure that decisions from the trigger system are consistent with the collision timing. These decisions are made within $2.5\,\mu$s and objects for events that pass a trigger in the menu have their ROIs sent to the HLT system.

At the HLT trigger system, the ROIs are fed to a series of reconstruction algorithms that reconstruct the physics objects for each event within 300 ms. Software-level trigger decisions on these physics objects (referred to as "online" trigger objects) are made based on an HLT trigger menu which is run on a CPU farm with 40k cores. Approximately 1 kHz of the events passing these HLT

decisions are fed from the detector read-out system to permanent storage on disk. Later reprocessing on these raw events converts them to calibrated and further defined physics objects (referred to as "offline" reconstructed objects).

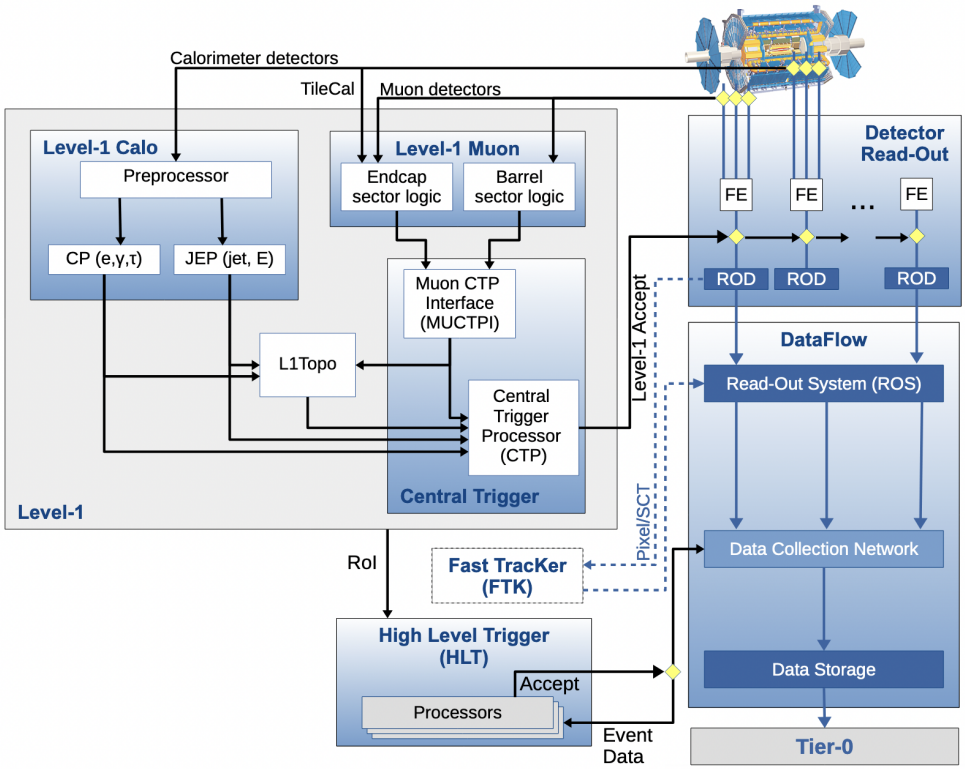


FIGURE 3.9. A flowchart representing the data flow of events through the data-acquisition system at the ATLAS detector. [8]

3.8. Data Reconstruction

A variety of data streams exist at ATLAS with applications that range from detector monitoring to the stream containing the critical physics data. The "physics_Main" stream is the general stream at ATLAS; this stream contains event that are output from the Physics trigger menu that are suited for general physics analysis. Other monitoring streams such as the monitoring stream outputs a small fraction of events to enable monitoring of detector conditions and the software

reconstruction. The Enhanced Bias stream is used to determine the average pileup and detector conditions by using minimally-biased triggers to select low-energy events that characterize an average event in the detector.

The data in the physics_Main stream are stored in RAW output format in preparation for reconstruction. Events are reconstructed by being passing through a series of algorithms sequenced by the ATLAS central software, Athena [56]. The detector objects such as the ID hits and the calorimeter energy deposits are sorted from the RAW file and the physics object reconstruction begins. For Run 2 data, tracks are reconstructed algorithmically from ID hits, and tracks that have $|d_0|$ values greater than 10 mm are discarded. However, in the special DAOD_RPVLL stream reconstruction applied to a filtered set of data for the 4-jet emerging jets analysis in Run 2, this range is extended out to 300 mm. In Run 3, this restriction will be lifted at both the trigger level and at the reconstruction level, so high d_0 tracks and high-radius displaced vertices will be reconstructed and available for the analysis. However, in the dataset available for this analysis, this information is not available and unfortunately cannot be used as a background discriminant. The use of this track reconstruction in the trigger will be discussed further in Chapter V.

3.9. Jets

Byproducts of standard hard-scattered quarks or gluons will undergo QCD hadronization, and the stable resultant particles will interact with the detector, unless they are particularly long-lived. Jets tend to consist of approximately 60% hadrons, with the majority of these being charged mesons such as pions which interact with the ID and then shower in the calorimeter. A smaller fraction (roughly 10%) are baryons such as protons and neutrons or mesons such as K_S or

K_L particles. Neutral pions will immediately decay into photons; in combination with additional photons from Final State Radiation (FSR) during hadronization, these constitute approximately 30% of the jet energy.

Energy deposits in the calorimeter can be clustered into jets with a variety of algorithms that vary in their resolution, choice of physics objects in the clustering, and clustering radius. The simplest jet algorithm will simply sum energy in a square surrounding the jet seed. This algorithm is used at the hardware level in the ATLAS trigger system; energy deposits in the calorimeter are formed into topoclusters and the topoclusters surrounding a 0.4×0.4 area from the jet seed are summed to form the jet. A slightly more complex algorithm would sum the energy in a cone of defined radius around the seed, however this increases the computational difficulty.

The final jets reconstructed at the HLT and offline levels in ATLAS use algorithms that are colinear and infrared (IR) safe. Due to the stochastic nature of jet fragmentation, jets can vary wildly in their clustering properties. Splittings of high-p_T incident particles and emissions of soft gluons during the hadronization can change the substructure, and colinear and IR safe algorithms are designed to be insensitive to these random processes within jets [57].

Topocluster energy reconstruction starts from the definition of the seed locations. Seeds are formed from cells that contain energy deposits that are greater than 4σ from the noise baseline, where σ is the standard deviation of the expected noise thresholds. Neighboring cells are then added to the topocluster sum if the energy deposit within the cell exceeds 2σ, and the process continues until there are no more cells within the topocluster region to add.

Calorimeter-only jets (EMTopo jets) use these topocluster energy deposits as inputs to the clustering algorithm. However, the p_T precision of the tracker motivates the use of tracks as inputs to the jet reconstruction. Tracks extrapolated to the topoclusters can provide a more accurate measurement of the contribution of the particles' momenta than the topocluster measurement. In the Particle-Flow (PFlow) algorithm, tracks associated to topocluster energy deposits are used within the clustering. If a track from the PV is associated to one or more topoclusters and the momentum measurement of the topoclusters is sufficiently close to the track's p_T, the associated topocluster(s) are removed and the track's kinematics is used in lieu of the topocluster(s). For emerging jet signals that are of interest here, the multiplicity of displaced tracks in the signature discourage the use of this algorithm, as a significant fraction of the energy in a signal jet will not be associated to a well-measured track. However, due to resource limitation in the TDAQ system, the Run 3 emerging jets trigger will be required to use this algorithm at the trigger level. Appendix A.1 and Chapter V contain further discussion of the trigger and the reconstruction's effects on the signal.

The large-R jets used in this analysis utilize the anti-Kt algorithm [10], and are reconstructed with a radius of 1.0 in η-ϕ space on the EM-scale topocluster energies (EMTopo). To reduce pileup and hadronization dependence of jet energy, Kt-reconstructed topoclusters that form each jet are discarded from the clustering if they contain less than 5% of the jet's total energy. These jets are referred to as trimmed anti-Kt R=1.0 LC jets, and they are designed to be highly resilient to the substantial pileup contributions that can be seen in large-R jets. The effects of this trimming relative to the standard pileup subtraction method are shown in Figure 3.10.

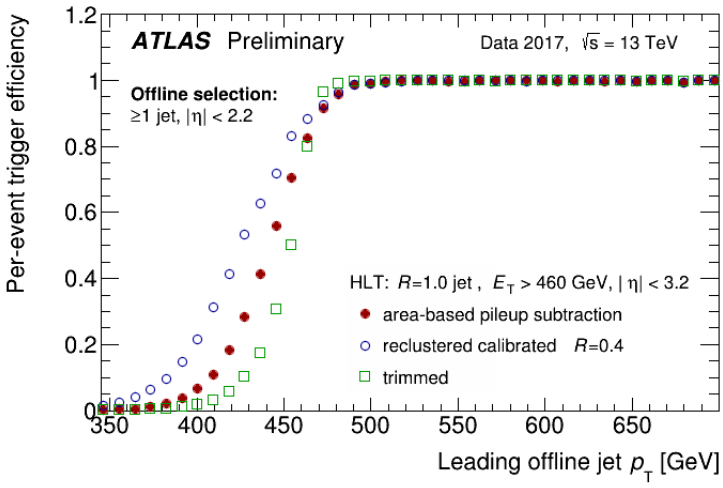


FIGURE 3.10. Trigger efficiency curves showing the effect of trimming on the large-R trigger turn on points. The effect of standard pileup subtraction methods is shown for reference [9].

3.9.1. Anti-Kt Algorithm

The Anti-Kt algorithm belongs to a class of jet reconstruction algorithms that include the Kt, Cambridge-Aachen, and Anti-Kt algorithms [10]. These algorithms utilize a subjet weighting scheme to recluster subjets in a self-consistent method. The variables considered within this algorithm are:

$$d_{ij} = \min\left(k_{ti}^{2p}, k_{tj}^{2p}\right) \frac{\Delta_{ij}^2}{R^2} \text{ for } i \neq j \tag{3.12}$$

$$d_{i,i} = k_{ti}^{2p} \tag{3.13}$$

where $\Delta_{ij}^2 = (y_i - y_j)^2 + (\phi_i - \phi_j)^2$ is the rapidity-ϕ distance between a pair of subjets and k_{ti} is the transverse momentum of the ith subjet. The variable p is the exponent that defines the algorithm type: $p = 1$ for the Kt algorithm, $p = 0$ for the Cambridge-Aachen algorithm, and $p = -1$ for the anti-Kt algorithm.

The algorithm first computes d_{ij} and d_{ii} for each pair of subjets and finds the minimum over all the sets of d_{ij} and d_{ii}. If a minimum d_{ij} exists, the corresponding ith and jth subjets are combined in energy and position and replace the original pair of subjets and all the $d_{ii,ij}$ variables are recomputed. If the minimum is one of the d_{ii}, this subjet is considered a fully reconstructed jet and is removed from the list of subjets.

The algorithm is designed to cluster the soft radiation around an energetic energy deposit into a full jet. The highest k_t subjets will cluster first and incorporate the surrounding radiation into a roughly conical collection of energy. This algorithm also handles situations with isolated splashes of energy in the detector; isolated subjets with no other weighted k_t within the jet radius will be considered full jets. This is shown in Figure 3.11 where the highest p_T subjet generally resides in the center of the jet, with soft radiation within the jet radius typically clustered close to the hardest subprocess. It handles overlaps in energy deposits by associating overlap energy to the high-$\mathrm{p_T}$ neighboring jet.

Standard jet algorithms that are IR safe are preferred, meaning that they are invariant to any additional emission of radiation during the fragmentation process. The hardest subprocess in the jet region will tend to form the seed of a jet, with additional soft radiation in the jet radius tending to be clustered alongside the hardest subprocesses. This makes the measured jet energy relatively invariant to the hadronization process or model for the jet, as changes to coupling parameters such as α_s will generally change the contribution of soft radiation, which will be absorbed by the Kt algorithms.

Typical QCD-originated jets can be readily clustered within $\mathrm{R} = 0.4$ jets, so this is the typical jet radius used at ATLAS at 13 TeV. However, many

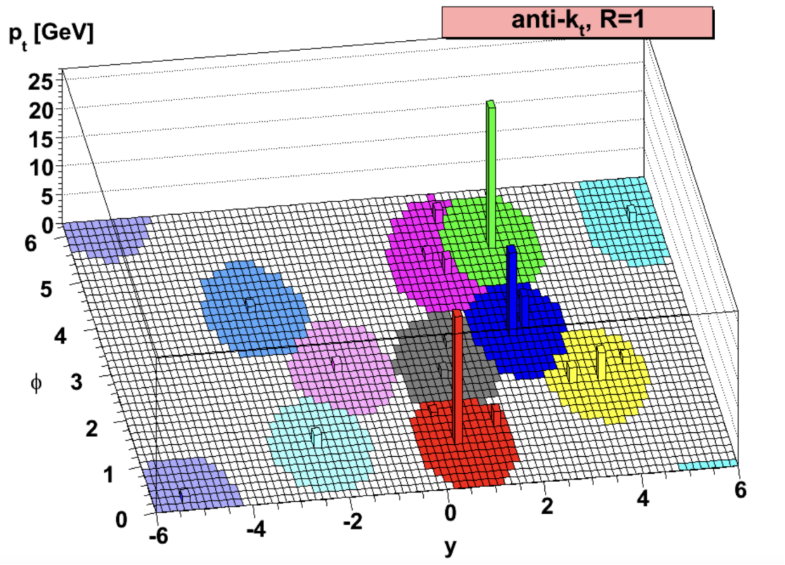


FIGURE 3.11. A demonstration of the Anti-Kt algorithm for a simulated event that shows the typical clustering of energy deposits for R=1.0 jets [10].

boosted topologies produce much larger jets and the typical jet radius for the reconstruction of a top quark is R = 1.0. As Figure 4.7 shows, emerging jet signatures can contain large amount of soft and wide radiation, making the R = 1.0 algorithm optimal for this analysis. The same jet radius was used by the original ATLAS CWoLa analysis [58] that is a precursor to this analysis. That previous search was searching for heavy boosted decays producing a pair of R = 1.0 jets; further details on this will be discussed in Chapter IV.

Truth jets are reclustered from final-state truth particles, and represent a collection of the final states of a hadronization process. Truth jets are used in the context of the emerging jet signal generation, and are simply formed via anti-Kt reclustering of the Pythia final-state particles, with the final-state label determined by the particle's stored barcode that is produced during the generation. Standard truth jets are clustered with a radius of R = 0.4.

3.10. Jet Calibration

Jet calibration is an intricate process designed to shift the energy scale of the jets in the data to be consistent with the overall energy scale of the MC and to increase the p_T resolution of the jets. The data and simulated events are known to have quite differing responses in the detector, and so these energy scales must be calibrated to be consistent, in addition to correcting energy measurements from problematic regions of the detector. Jets are reconstructed with a baseline energy scale that is defined algorithmically, potentially followed by trimming, and then followed by a defined set of calibration steps.

The calibration steps depend on the jet algorithm and radius that are used in the reconstruction. Energy resolution is a critical component of $R = 0.4$ jets, and so effort is made to increase the energy resolution using pileup residual corrections and sequential calibrations for high-p_T jets. Mass resolution is a critical aspect of large-R jets as large-R jets are typically used to reconstruct heavy objects such as top quarks, W-bosons, and other boosted objects that have a crucial measured mass component.

The first calibration step for $R = 0.4$ jets start at the electromagnetic scale, so energy measurements in the topoclusters are made consistent with the energy that would be seen in the ECAL. Jets are origin-corrected to have η values that are consistent with the location of the PV in the event, instead of the online definition - at the center of the detector. This improves the η resolution of the calibrated jets as the momentum vector of the jet more closely aligns with truth particles from the PV. A pileup correction is then applied to the jets that corrects the jet's response to the in- and out-of-time pileup environments from the various data-taking conditions. The overall jet energy scale (JES) is then corrected to be

consistent with MC events to ensure that energy measurements are meaningful for physics analysis. The higher-order response of the detector to jets with different track multiplicity and substructures is then corrected in the Global Sequential Calibration (GSC) step. Then, data-driven in-situ methods are applied to correct the varying response as a function of p_T and η, in addition to correcting effects from troublesome regions in the detector.

In the case of the large-R jets utilized in this analysis, they are set to a baseline energy scale at the hadronic energy scale using a local weighting scheme called LCW [11]. The previously mentioned η correction is applied and is followed by the JES calibration step. In a very similar vein to the JES calibration, the jet masses are also corrected in the Jet Mass Scale (JMS) calibration step [59]. Insitu methods utilizing the balance of a large-R jet to a well-measured set of small-R jets, a well-measured photon, or a well-reconstructed $Z \rightarrow \mu^+\mu^-$ decay correct the JES and JMS of the barrel region. An η-intercalibration then corrects the forward region JES and JMS relative to the barrel region.

3.10.1. Pileup Correction

Pileup contributes a relatively isotropic distribution of energy to the detector on average. Therefore, to reduce the expected pileup count (μ) and event Number of Primary Vertices (NPV) dependence of the subsequent calibration steps, an average pileup correction is derived and applied to R = 0.4 jets. The first step is an area-based correction that removes the median contribution of in and out-of-time pileup contributions to the jet p_T that distort the true jet energy.

$$p_T^{\text{area}} = \rho \times A \tag{3.14}$$

The computation of the median ρ energy contribution must be sensitive to low p_T contributions from soft particles in the event. Therefore, the pileup contribution is computed from R = 0.4 Kt reconstructed jets within the $|\eta| < 2.5$ domain. The Kt algorithm is more sensitive to soft radiation further from the jet seed, and ρ is derived from the median energy deposit within these jet areas.

The area A is computed by ghost track association. Many ghost tracks with infinitesimal p_T are randomly added to the jet cone and are clustered along with the topoclusters used in the clustering algorithm. The small p_T values do not affect the clustering or jet measurement itself, but the average fraction of clustered ghosts over the total associated jet cone area becomes the jet area.

This area correction is only derived from jets within the central detector and is insensitive to effects that are dependent on NPV. The contributions to the jet energy from in and out-of-time pileup can be approximated by examining differences between truth and reconstructed jet p_T measurements as a function of NPV and μ, thereby representing effects of both in and out-of-time pileup contributions. The second step of the pileup correction corrects these contributions using MC simulation of hard-scatter QCD events to estimate the jet exposure to pileup. Truth and reconstructed jets are matched with a $\Delta R < 0.3$ and the reconstructed p_T difference between the truth and reconstructed jets is measured as a function of η_{det}, the central detector η. The difference is histogrammed as a function of μ and NPV and a functional form is fit to the difference, allowing for a sequential removal of the pileup dependence.

The correction uses this measured functional form in the calculation across both NPV and μ,

$$\alpha\left(\eta_{deta}\right) = \frac{\partial p_T}{\partial NPV}\left(\eta_{det}\right) \tag{3.15}$$

$$\beta\left(\eta_{det}\right) = \frac{\partial p_T}{\partial \mu}\left(\eta_{det}\right) \tag{3.16}$$

The total correction from this calibration step starts from the EM-scale reconstructed jet, subtracts the average pileup contribution over the jet, then applies the ρ correction in sequence.

$$p_T^{residual} = \alpha \times (NPV - 1) + \beta \times \mu. \tag{3.17}$$

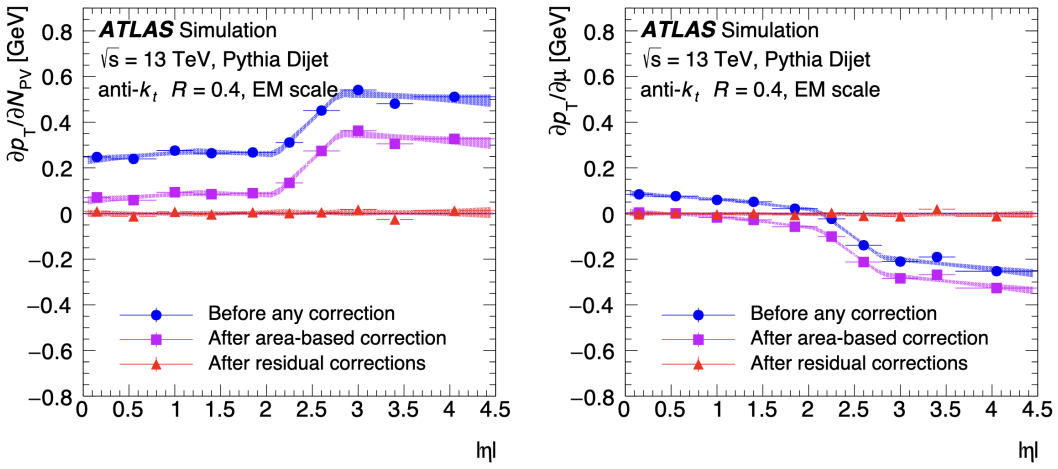
The corrected jet $\mathrm{p_T}$ then becomes

$$p_T^{corr} = p_T^{reco} - p_T^{area} - p_T^{residual} = p_T^{reco} - \rho \times A - \alpha \times (NPV - 1) - \beta \times \mu. \tag{3.18}$$

This calibration was performed on Run 2 ATLAS data and is shown in Figure 3.12.

3.10.2. JES Correction

The stable components of MC truth jets have a well-defined response in the simulated detector, and so the four-momenta of simulated reconstructed jets can be corrected on average to match the true energy distributions seen in the calorimeters. This relies on having a fairly accurate detector model to use in the simulation, however mismodeling of the detector can be corrected for in the in-situ methods described in the following sections.



(a) In-time pile-up dependence

(b) Out-of-time pile-up dependence

FIGURE 3.12. A demonstration of the pileup correction as applied to reconstructed jets matched to truth jets with $p_T = 25\,\text{GeV}$ [11] as a function of (a) NPV and (b) $\langle\mu\rangle$.

Simulated reconstructed jets are matched and corrected to their corresponding truth jets within a radius cut of $\Delta R < 0.3$ ($\Delta R < 0.6$ for large-R jets). Events are classified into various bins of E_{truth} and η_{det}. The distributions of E_{reco}/E_{truth} are formed for each bin; the means of Gaussian fits to the cores of these distributions indicate the average energy scale to be corrected to. The actual response of E_{reco} at each E_{truth} and η_{det} bin is derived from a numerical matrix inversion, and the overall correction factor is the inverse of this response and is applied to each jet [11].

The η correction is an additional correction applied to the JES after the energy correction. A similar approach is applied as in the energy correction. Jets are placed into a series of E_{truth} and η_{det} bins, and an average response is determined from the η_{det}/η_{truth} distributions. This is a smaller-order correction to the energy vector of the jets, but helps increase the jet energy and reconstructed invariant mass resolution.

3.10.3. Jet Mass Scale Correction

The Jet Mass Scale (JMS) correction is a set of correction factors derived as a function of η and the mass that is applied to the uncalibrated jet's mass. They are derived in an identical manner to the JES factors, only binned and applied to the jet mass instead of the jet p_T. MC samples are examined to determine the ratio of the truth to the reconstructed jet mass, and the ratio becomes a response factor that can be inverted to derive a calibration factor for the data. This method is validated by examining the jet mass shape in $t\bar{t}$ (top/anti-top quark event) data. The JMS mass resolution can be seen in Figure 3.13.

3.10.4. GSC Correction

The Global Sequential Calibration (GSC) corrects higher-order jet energy from p_T-dependent and parton-level effects [11]. Jets from different initiating partons have varying responses in the detector. Jets originating from the hadronization of a quark tend to have a lower multiplicity of tracks associated to the jet, and these tracks tend to have higher p_T distributions. Gluon initiated jets tend to have a wider and softer set of energy deposits that do not penetrate as far into the detector.

Since this series of sequential calibrations utilizes energy deposits in various detector layers to account for the penetration depth of the energy deposits, it is natural to consider the effect of high p_T jets that have significant missing energy in the context of this calibration step. High p_T jets can stochastically have a significant fraction of their true energy lost out of the detector due to punch-through of the energy through all the calorimeter layers. This energy will be

visible as a set of energy deposits in the muon system geometrically close to the energy deposits of these high p_T jets.

Similarly to the previous procedure, a series of calibration curves that correct the data to the MC energy scale were derived. These corrections are functions of the energy fractions in the first layer of the HCAL, to account for higher p_T effects from, in particular, high p_T quark-initiated jets. Variables such as the number of jet-associated tracks and the number of topoclusters are also incorporated into the correct to account for the quark/gluon jet differences. These derived response curves sequentially account for detector response differences to these various hadronization and punch-through effects.

3.10.5. In-Situ Corrections

The η-intercalibration is an in-situ calibration method that is applied to jets after the GSC stage. This method is a data-driven method to perform a higher-order correction to the JES to account for problematic regions of the detector and to correct the overall scale differences between the forward and central regions. Transition regions in the detector such as shifts between regions of different granularities or detector technologies create both symmetric and asymmetric response differences for jet measurements. The η dependence of these response difference will not be captured by the JES correction step, and so this higher-order correction is applied afterwards.

Well measured dijet events have a back-to-back pair of jets that have minimal contributions from pileup jets or out-of-cone radiation from leading jets in the event. To leading order in QCD, these events are expected to be p_T balanced at the truth level and also at the reconstruction level, if the events are well-

measured and calibrated properly. This is utilized in the intercalibration; a shift in the average p_T balance of these dijet events is attributed to detector-level miscalibrations from previous calibration steps and is corrected for in probe bins of η and p_T. Events are selected from data and from MC QCD events that have leading and subleading jets that are back-to-back in ϕ with a $|\Delta\phi| > 2.4$ cut. To reduce asymmetries in the dijet pair caused by third jets in the events, the third jet is required to have no more than 40% of the average p_T of the leading and subleading jets. To reduce the effect of pileup contamination of the event-level asymmetry, the Jet Vertex Tagger (JVT) is applied with a "Tight" threshold to reject jets that are not from the PV in the event. This is a hybrid machine-learning method that tags R = 0.4 jets as originating from the PV or from pileup vertices using jet and jet-associated tracks, however, this is not derived for R = 1.0 jets. Events are then sorted into their respective calibration bins of p_T and η with event performance trigger selections which ensure that events are selected with at least a 99%-efficient trigger.

The p_T asymmetry within each of these bins is computed, where the asymmetry of two jets is defined as

$$A = \frac{p_{T,0} - p_{T,1}}{p_{T,avg}}. \tag{3.19}$$

These distributions are computed for every pair of bins and reflect the asymmetry shift for every probe bin relative to every other reference bin. Gaussian fits to the core of these asymmetry distributions determine the mean of this asymmetry shift for that particular set of bins, and well-fitting distributions are taken as appropriate bins to use in the calibration. In the traditional "central" method, events in the barrel ($|\eta| \leq 0.8$) region are considered well measured due to the

flat response of the JES in this barrel region, and events in the forward ($|\eta|$ ¿ 0.8) barrel region are corrected relative to the central region. However, this method is statistics-limited and a matrix method was developed midway through Run 2. In this method, every probe bin is defined relative to every other reference bin in the detector. An overall calibration loss function is defined via product of the asymmetry matrix of the bins and parameterized response factors in a similar bin matrix. The product of these matrices is minimized via matrix inversion and the simultaneous set of response coefficient equations corresponding to the correction factors for each bin are solved for.

The ratio of these response coefficients of the MC and the data then become the calibration factors needed to push the data's JES scale to that of the MC's for each bin. Spline fits are then applied to the calibration factors to smooth the detector response and to allow interpolation for any jet p_T and η.

An additional in-situ method using the event asymmetry of either a Z+γ event or a $Z \to \mu\mu$ event provides a higher-order correction to the eta-intercalibration. Barrel-region photons are precisely calibrated, so the absolute JES of the recoil jet can be determined, and an overall correction factor can be derived. Similarly, a recoil jet off a dimuon pair can have its absolute energy scale corrected for [60].

The combination of these calibration methods gives very precise Jet Energy Resolution (JER) for jet measurements, in particular at low p_T. For $R = 0.4$ jets, the final jet energy resolution at 300 GeV is roughly 6% [61], and for $R = 1.0$ jets, it is roughly 12% at 400 GeV [12]. Figure 3.13 shows precision of the JES and JMS corrections for $R = 1.0$ jets

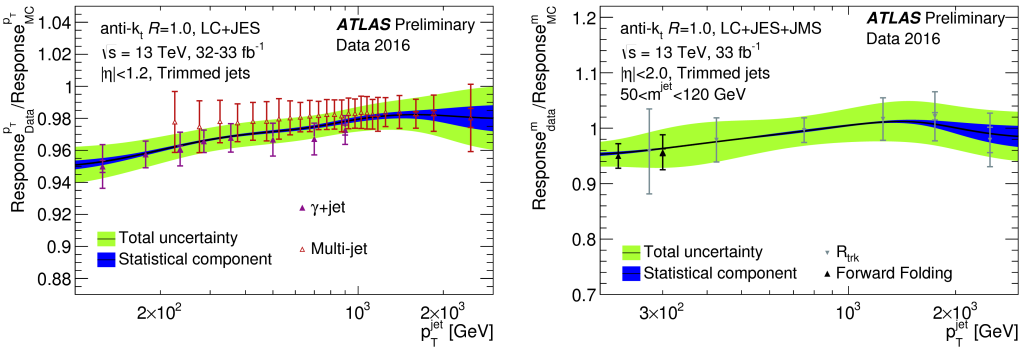


FIGURE 3.13. In-situ methods such as γ+jet measurements are used to validate the previous calibration steps [12]. The JES response is shown to be calibrated within 2%, and the JMS response is consistent within error bounds.

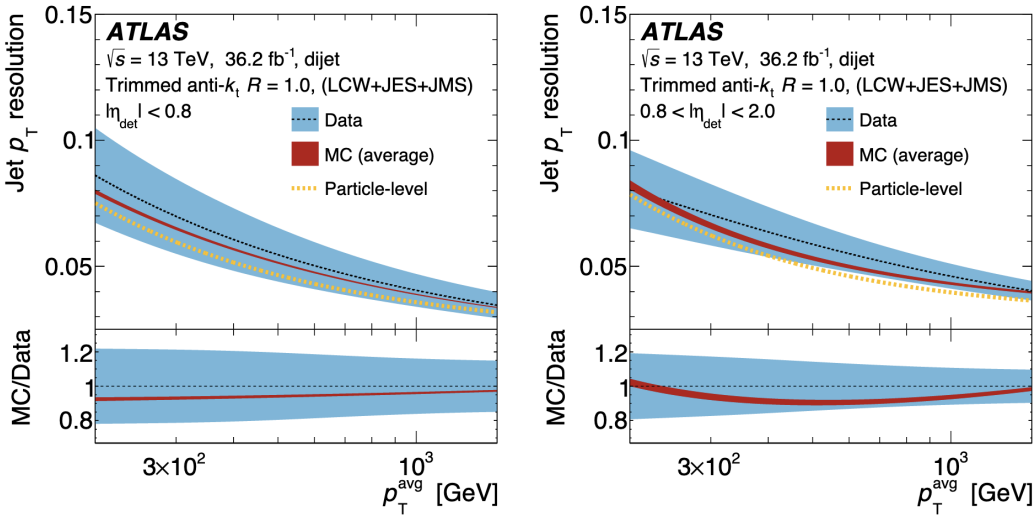


FIGURE 3.14. Calculated JER for R = 1.0 jets after application of the JES and JMS calibrations [12]. An analytic fit to the MC JER was used in the systematic evaluation for this analysis - f(p_T) = $\sqrt{a^2/p_T^2 + b^2/p_T + c^2}$ with a=5.018×10^{-4}, b=1.092, and c=-2.090×10^{-2}.

3.11. Dijets

This analysis focuses on the dijet topology, where balanced pairs of jets are expected to lie relatively back-to-back in the transverse plane. An example of one such event is shown in Figure 3.15.

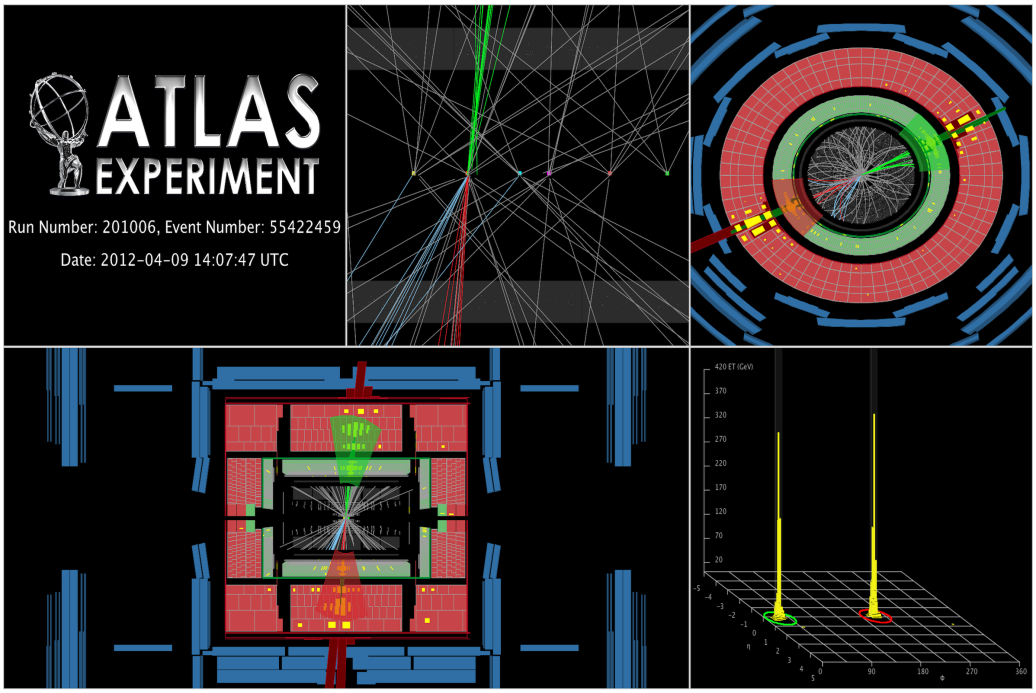


FIGURE 3.15. A high-mass dijet data event with $m_{jj} = 3.91$ TeV illustrating a back-to-back jet pair originating from a single vertex with 5 reconstructed pileup vertices. [13]

During trigger development, the resulting jets from simulated Hidden Valley samples were examined for their ability to properly reconstruct Z's that were generated with mass 600 GeV. It was found that R = 0.4 jets that are standard for properly reconstructing QCD jets are insufficient for containing most of the energy of these dark jets, as Figure 4.7 shows. The total invariant mass loss away from the truth Z' mass was reduced from up to 40% with the R = 0.4 jets to 10% with the R = 1.0 reconstruction. This motivated both the trigger strategy and the analysis methodology to focus on pairs of large-R jet events.

The next chapter shows the analysis method wherein these physics objects that were reconstructed from the detector are analyzed for the presence of signal. A set of cuts and machine-learning taggers separate out the signal-like events and the resulting invariant mass spectra are then analyzed statistically. The Search

Phase and Limit-setting phases are detailed, indicating the level to which this analysis can conclude on the presence of signal (Search Phase), and the cross-section limits that can be placed on excluded signal (Limit-Setting Phase).

CHAPTER IV

ANALYSIS

4.1. Dijet Bump Hunt Searches

Searches for exotic particles within the dijet and multijet spectra can be hampered by the statistical dominance of QCD events, even after the application of optimized cuts designed to reduce the background contamination and enhance the signal sensitivity. However, the dominance of these background events can be taken advantage of by providing a background-rich set of data for training autoencoders or machine-learning classifiers, and these analyses can use aspects of the background dominance to utilize data-driven background estimates that rely on significant statistics in the data set. This analysis takes advantage of both methods using a modified weak classification method - a modified Classification Without Labels (CWoLa) method.

4.2. Data-Driven Background Estimate

Due to the predominance of these QCD events, the dijet invariant mass spectrum fits well to a series of analytic functional forms. The smooth response of QCD cross-sections with PDFs, with a relatively smooth jet calibration, can make these spectra fit to a variety of power-law functions, and the global or local applicability of these fit functions depends dramatically on the luminosity. The relatively small width of any resonances compared to the overall acceptable fit range of the spectrum means that these resonances will appear as a small bump in the invariant mass (m_{jj}) spectrum. The invariant mass in natural units is

$$m_{jj}^2 = (E_1 + E_2)^2 - \left|\vec{p}_1 + \vec{p}_2\right|^2, \tag{4.1}$$

but in collider experiments is usually computed as

$$m_{jj}^2 = 2 \cdot p_1 \cdot p_2 \ (1 - \cos\theta), \tag{4.2}$$

where θ is the angle between the momenta vectors. Depending on the analysis, the masses of reconstructed objects can be small enough to be insignificant, in which case the invariant mass becomes

$$m_{jj}^2 = 2 \cdot p_{T1} \cdot p_{T2} \ \left(\cosh(\eta_1 - \eta_2) - \cos(\phi_1 - \phi_2)\right). \tag{4.3}$$

The analytic functions that are fit to invariant mass spectra have varied in their functional form, number of fit parameters, and range of applicability to increasing luminosities. Care needs to be taken that the choice of fit function does not under- or over-fit the data, that the range of the fit does not obscure any signal, or that the fit itself is not significantly affected by the presence of detectable signal.

At small enough luminosities, these analytic functions typically fit globally across the entire range of the spectrum. However, with an increase in statistics in the spectrum, either by an increase in luminosity or a reduction in the lower m_{jj} limit of the data, these fit functions can fit within a narrower range of the spectrum due to the decreased statistical uncertainty no longer masking higher-order effects on the JES. The Sliding Window Fit (SWiFt) method was developed to accommodate this.

Most power-law functions that do not globally fit to the dijet spectra will fit appropriately within some narrower range in m_{jj}. So in the SWiFt method, to derive a background estimate for each m_{jj} bin, a chosen functional form is fit to the bin of interest and the previous and subsequent N bins. N varies on the choice of fit function and the statistics within the spectrum itself, but is generally chosen to be as large as possible without under-fitting the data. An evaluation of the χ^2/ndf can indicate if either under or overfitting occurred in the fit. A visualization of this is shown in Figure 4.1. At the lower end of the spectrum, the background fit is more uncertain as it can only fit the upper N/2 bins from the central bin. As the window slides down the center of the spectrum, the background fits the previous and following N/2 bins. At the high edge of the spectrum, only the previous N/2 bins can be fit.

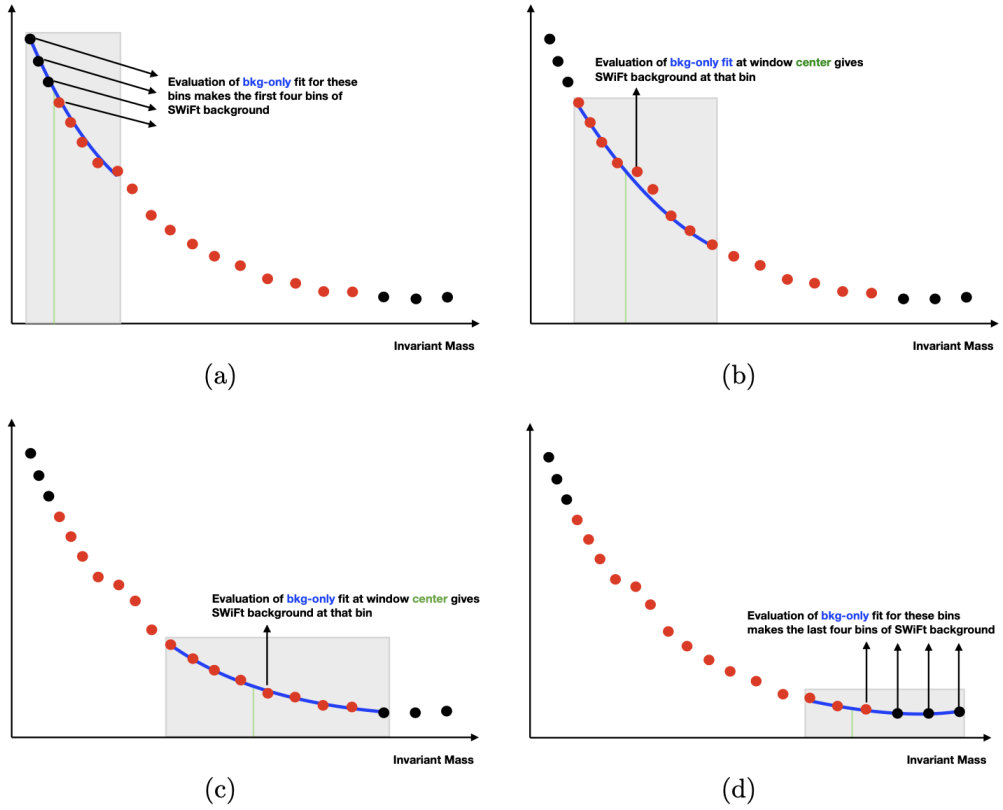


FIGURE 4.1. A visual demonstration of the SWiFt fitting mechanism for a given window width [14].

4.3. CWoLa Method

The CWoLa method is a Weak Classification Machine Learning method that is used to reject uninteresting events from the background-dominated m_{jj} spectrum and to enhance the signal contribution to the spectrum. The method relies on classification theorems regarding optimal classifiers for mixed samples [62]. Given two samples, both of which are dominated by background-type events, in the presence of signal, one of the mixed samples might contain different distributions in the feature space considered. Any classifier that learns the optimal

classification of these mixed samples is the optimal classifier for distinguishing individual signal events from background events [58].

Figure 4.2 demonstrates the classification procedure, where neural networks are trained to classify events from the central signal bin (labeled as '1' events) from events in the sidebands (labeled as '0' events). The same window width is then evaluated for the next signal bin, thereby sliding the window down the spectrum.

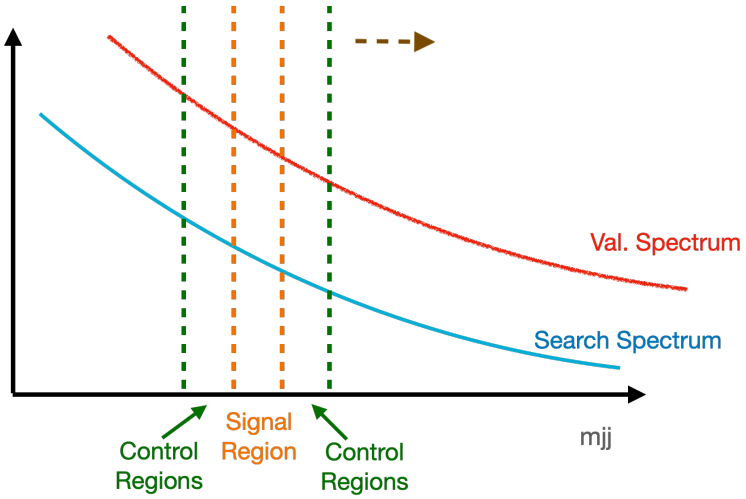


FIGURE 4.2. Visual demonstration of the CWoLa method combined with a SWiFt fit.

The advantages of this are apparent as a specified signal model is not required and generic signals with enhancement in physics-motivated feature space can be searched for. The decision on the set of features to train the classifier remains a choice that is motivated by the physics of interest. However, the classifier will be naturally invariant to the signal model, making this method ideal for discovery when there is a large parameter space to search.

This method has been used previously at ATLAS in one preparatory analysis [58], hereafter referred to as the "original CWoLa (ATLAS) analysis". This analysis searched for heavy pairs of large-R jets produced by a heavy resonance. This analysis was intended to validate the use of unsupervised learning at ATLAS, and so the input feature space that was used just contained the leading and subleading jet masses. In practice, as long as the response of the networks across the input feature space is relatively invariant across m_{jj}, the feature set can be expanded by the physics motivation of the analysis. This method could potentially be extended as a deep learning problem, where the entire jet image (calorimeter image) is processed with a Convolutional Neural Network and the associated tracks are processed with Recurrent Neural Networks (RNNs). However, a more thorough analysis method would be required to ensure this doesn't induce bumps in the spectra.

4.3.1. Feature Space

The authors of the original ATLAS CWoLa analysis [58] intended for their analysis to be a proof-of-concept of the application of this method in dijet searches, and so the only features trained on were the leading and subleading jet masses in the event. This analysis searched for heavy LLP decays of the type $A \rightarrow B + C$ with an agnostic approach to the coupling mechanics or byproducts, with the exception of the assumption of central s-channel production of heavy jets. However, the jet masses of emerging jet signals will have similar jet mass spectra as in QCD as the initiator quarks have O(1 - 10 GeV) masses. In addition, the overall jet mass spectrum is very model dependent due to the wide hierarchy of particle masses in the dark sector decays, as shown in Section 4.9. So the feature

space was expanded by a number of variables that are motivated by the physics of these events.

- The leading and subleading jet masses.
- The leading and subleading jet promptTrackFrac (defined in Equation 4.4).
- The number of jets in the event.
- $|y^*| = \left|\frac{y_0 - y_1}{2}\right|$ using the leading and subleading jets.
- $|\Delta\phi| = |\phi_0 - \phi_1|$ is the ϕ separation between the leading and subleading jets.

The macroscopic lifetimes of the dark pions in the dark showers produce a set of displaced tracks with high $|d_0|$ values. These non-prompt tracks can be associated to their resulting jet and used as a measure of the prompt and non-prompt components of the jet. A new jet variable called promptTrackFrac (PTF) is therefore introduced,

$$\text{promptTrackFrac} = \frac{1}{p_{\text{T jet}}} \sum_{i \in tracks} p_{T_i} \left(|d_0| < 2.5 \; \sigma_{d_{0,i}} \left(p_{T_i} \right) \right). \tag{4.4}$$

Here, σ_{d_0} is the resolution of the track d_0 measurement, and is very dependent on the track properties. The 2.5 multiplicative factor was determined to be the optimal factor via signal studies on R-Hadron LLP samples; more details on this are in Appendix A.1. The track measurement error has a series of complex dependencies that are interrelated with the ID efficiency, the track's p_T and other kinematics, and the PV / beamspot uncertainty. However, a generalized parameterization of this significance is applied in this analysis to be consistent with the events that will be selected by the Run 3 emerging jets trigger.

However, a more complex parameterization would include higher-order terms in the power-law series used to parameterize σ_{d_0}, or an individualized track-by-track computation would take the beamspot (and therefore the PV) geometric uncertainties into account. This track-based uncertainty treatment is available at the offline level in ATLAS, but to be consistent with the trigger selection, the analytical fit is currently preferred.

The chosen functional form of σ_{d_0} was determined via a study of Z bosons in "Enhanced Bias" data. Enhanced bias data is a separate data stream [8] in ATLAS that is used to characterize standard bunch-crossings and pileup effects within the detector, in addition to its used in the calibration of low-p_T objects. The stream uses a set of minimally-biased and highly-prescaled hit triggers to select events just above the detection threshold of the detector. These so-called "spacepoint" triggers search for a minimal energy-producing jet in the event or a minimal threshold of energy that's detected in the pixel system. This stream is ideal for these sorts of performance studies requiring low-p_T reconstructed Z bosons with a standard pileup environment.

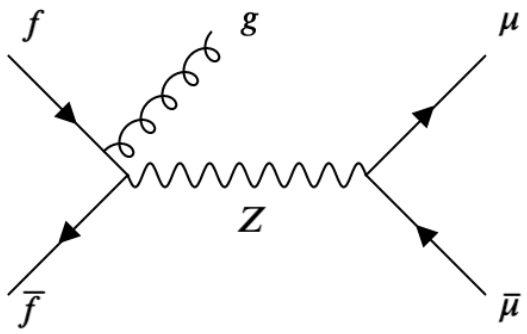


FIGURE 4.3. A Feynman diagram of the Z boson + photon process used to determine the associated track d_0 distributions from QCD ISR jets.

To ensure minimal contamination from pileup tracks on the σ_{d_0} measurements, recoil jets off an s-channel produced Z boson (shown in Figure 4.3) were used to derive the d_0 resolution. The selected events were required to have two opposite-sign muons with $|\eta| < 2.4$ and $70 < m_{jj} < 110\,\text{GeV}$. A recoil jet with $|\Delta\,\phi\,(Z, jet)\,| > 2.0$ was required to have at least 50% of the Z boson's p_T. Tracks were selected within a $\Delta R < 0.6$ region from the jet and required to have $\text{p}_\text{T} > 1\,\text{GeV}$ and $|PV_z - z_0| < 10\,\text{mm}$. The d_0 distributions of these tracks were binned in track p_T in bins of width 0.5 GeV, and Gaussians were fit to the central peaks of the d_0 distributions to determine the width of the d_0 distributions.

A power-law function was fit to the d_0 resolution histogram for tracks with $1.0\,\text{GeV} \leq p_T < 8\,\text{GeV}$, as shown in Figure 4.4, and defines the resolution function used in the PTF variable's definition. The functional fit was $\sigma_{d_0} = 0.0436\ /\ p_T\,[\text{GeV}] + 0.0195$, with the resolution function σ_{d_0} defined in mm.

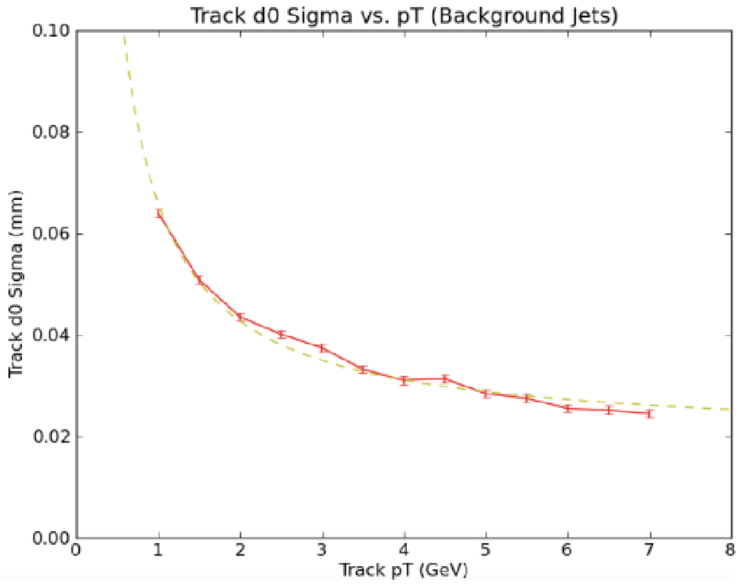


FIGURE 4.4. The d_0 resolution function from Z boson recoil jet studies and its analytical functional fit.

The PTF variable was originally designed for use as an LLP selector, so its sensitivity was tested on R-Hadron samples with lifetimes 0.1 ns and 10 ns, and the optimal multiplicative factor found was 2.5. PTF was later determined to be an ideal variable of choice for separating emerging jets signals from the QCD background. Figure 4.5 demonstrates PTF values for a variety of signal samples shown relative to the background-dominated data. A clear separation exists between these distributions, allowing both simple cut-based and machine learning methods to enhance signal sensitivity in the analysis. Further information on the derivation of this variable can be found in Appendix A.1.

For this analysis, the PTF definition was modified slightly from the original studies. The ΔR association was changed to a radius of 1.0 to match the radius of the large-R jets, and the track association z_0 cut was changed to $|PV_z - z_0| < 8\,\text{mm}$ to be consistent with the configuration used in the Run 3 emerging jets triggers. These triggers and their configurations are elaborated on more in Chapter V.

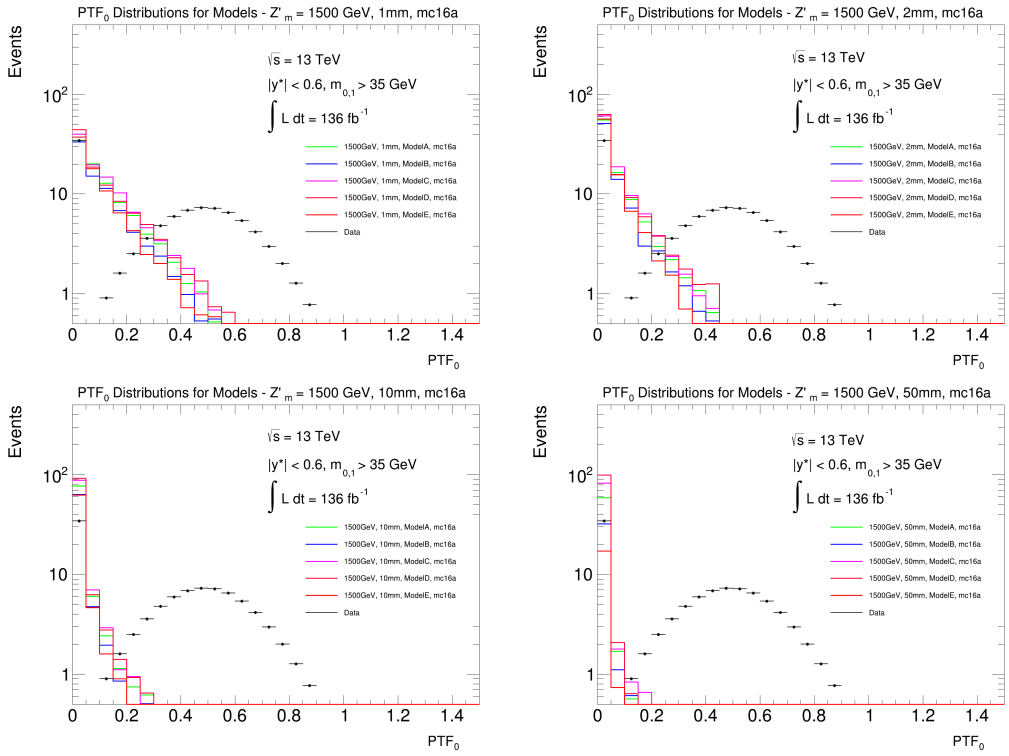


FIGURE 4.5. Z' samples with mass 1500 GeV and data leading jet PTF distributions for a variety of lifetimes and dark models. The data are scaled by an appropriate factor to fit the plotting scale.

The event variable $|y^*|$ is used in a variety of dijet analyses searching for s-channel resonances. Most bump-hunt searches consider resonances such as BSM Z' particles with vector or scalar s-channel resonances over the t-channel dominated QCD background. The $|y^*|$ distribution peaks at high values for t-channel QCD processes and at low values for s-channel processes. The authors of the original ATALS CWoLa paper [58] estimated that the signal fraction of most heavy resonances in the search region were enhanced by at least a factor of 5 by this cut. This is demonstrated in Figure 4.6, where the peak of each QCD channel's cross section for an incoming parton is at high $|y^*|$ values.

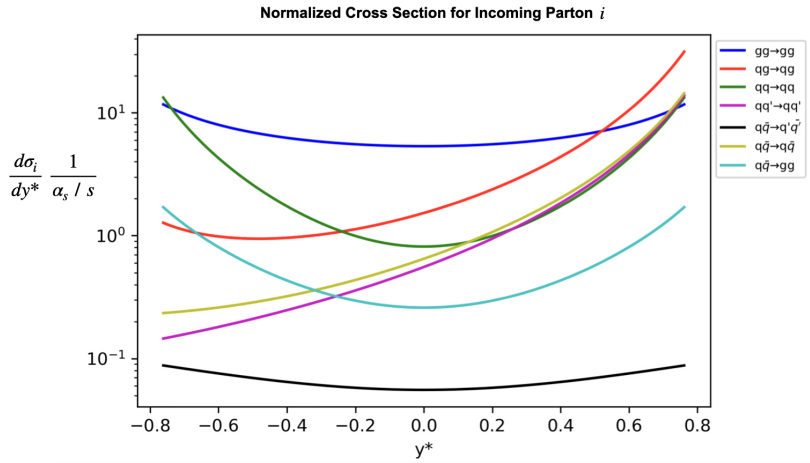


FIGURE 4.6. QCD y^* distributions for a given interacting parton at an energy scale s. Forward scattering cross-sections tend to dominate in QCD due to its t-channel predominance.

The inclusion of $|\Delta\phi|$ as an input feature was intended to account for events with hard radiation away from the leading or subleading jets. Effects from this third-jet radiation are likely correlated with the number of jets, but will add additional information into the NN that is correlated with how clustered jets are along the back-to-back axis of the Z' decay's byproducts.

The number of jets in the event was used as a feature as it has been shown to be a good discriminant in model-dependent machine-learning methods used in the ongoing ATLAS 4-jet emerging jet analysis. QCD events tend to contain smaller numbers of jets than dark jet events; at a p_T scale of 100 - 200 GeV, the small-R jet multiplicity is generally 4 or 5. Emerging jet models tend to have a large tail in the number-of-jets distribution as the wide decays tend to have substantial out-of-cone radiation. However, this effect is likely to be somewhat diminished in this analysis as R = 1.0 jets are used instead of the standard R = 0.4

jets. This is demonstrated in Figure 4.7 where most of the decay energy tends to be captured by R = 1.0 jets and where R = 0.4 poorly reconstruct the Z'.

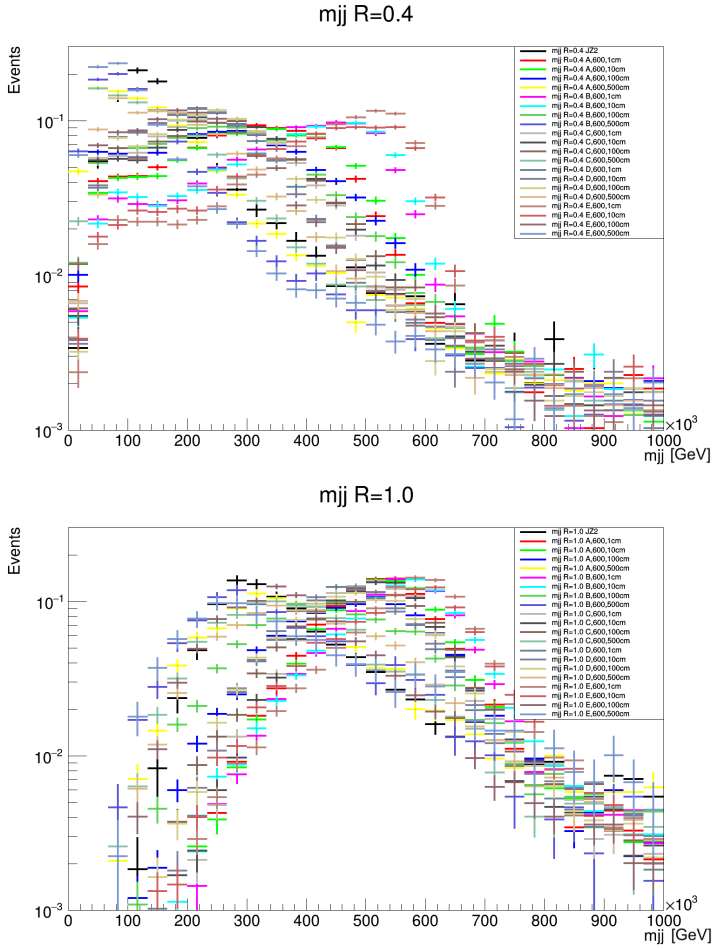


FIGURE 4.7. Invariant mass distributions for R=0.4 jets (top) and R=1.0 jets (bottom) for a variety of 600 GeV Z' samples and low-p_T simulated QCD background (labeled as JZ2). Significant model and lifetime dependence is seen; this illustrates that at high dark meson lifetimes, the signal follows the background distribution due to the missing energy in the events.

Events passing the previously listed cuts were sorted into an m_{jj} histogram with binning that is consistent with the jet resolution. Then, events in each bin and its sliding window were trained on to classify the events in the signal bin from events in the sidebands using the input features. A cutoff threshold was defined

by examining the neural network output scores for events in the sideband regions, and the top 10% or 1% quantile becomes the threshold that was extrapolated to the signal bin. The events with scores exceeding that threshold as evaluated by the trained neural network on the test events in that signal bin were selected and placed in a filtered histogram, and the BumpHunter algorithm was run on the spectra after the background fits.

4.4. Monte Carlo Events

The analysis method employed in this analysis can determine limits on generic signals, particularly Gaussian-shaped signals with a defined width and strength. However, knowledge of how the emerging jet signatures appear within the detector is critical for extracting the signal efficiency of these processes and for extracting fiducial cross-section limits. To model the effects on these signals, emerging jet Z' events were simulated with Pythia 8.245 using the Hidden-Valley module. The simplified model's structure is specified in Chapter II, and the Pythia configuration is shown in Appendix A. The PDF set used to simulate the nominal samples used in this analysis was the NNPDF 2.3 LO set from the NNPDF collaboration [63].

The subsequent decay of the Z' into pairs of dark quarks is dependent on the properties of the dark sector, where the coupling is set by the model and the decay is further constrained by the kinematic phase space available to the dark quark pair. The perturbative hadronization in the dark sector is performed by Leading Order (LO) calculations in the Hidden Valley model of the cross-sections and kinematics in the dark color sector. The small coupling to the Standard Model down quark creates the long-lived nature of the dark pions and defines the Z'

cross-sections. The lifetime specifies the emerging aspect of the dark jets such as the displaced vertices and tracks that are produced after their hadronization has completed.

The authors of the emerging jets phenomenology paper [37] calculated approximations to the width of the Z' decay - $\Gamma(M_{Z'}) \approx M_{Z'}/100$. This corresponds to a width of roughly 15 - 40 GeV in the mass range considered in this analysis. The width was set to 20 GeV in the sample generation to be consistent with these approximations. After simulation of the generated events was complete, including the hadronization, events were filtered to those with two anti-Kt R = 0.4 truth jets in the central detector region with ($|\eta| < 2.4$) and $p_T > 125$ GeV. As detector simulation and reconstruction of events is resource-intensive, this filter was set to remove events that are not likely to be reconstructed within the detector or readily used in the analysis. The Pythia configuration and software tags used in the Athena sample generation process are listed in Appendix A.1.

The response of the detector to the generated particles in these events was simulated using GEANT4 [64]. This program uses the estimated detector model to simulate the response of each region of the detector to the generated truth particles produced by Pythia. An alternate and less resource-intensive program, the ATLAS Fast (AF2) simulation program was found to be insufficient for the modeling of the signal samples as it uses parameterized jet modeling and doesn't take the lifetime of BSM truth particles into account during the simulation.

The samples were then run through the standard ATLAS reconstruction software, Athena [56]. Athena digitizes the simulated detector hits, overlays simulated pileup onto the events with the data-appropriate pileup (μ) distributions, and then outputs a software-readable file containing all the

reconstructable objects from the detector, including large-R jets, tracks, tagged electrons and photons, b-tagged jets, and muons. The overlaid pileup distributions were generated using Pythia with the A3 tune of the NNPDF 2.3 LO PDF. These samples were then reduced by the application of a derivation filter called EXOT3. Derivation filters reduce the file size through a removal of unneeded objects, apply the needed trigger decisions, and apply event-level vetoes. The cuts in the EXOT3 derivation are $p_T > 120\,\text{GeV}$, $\text{m} > 30\,\text{GeV}$ applied at the trigger level to the leading and subleading jets in the event, and application of all available single and multi-jet unprescaled large-R triggers. Figures 4.8 - 4.10 are virtual displays of some of these simulated and reconstructed signal events using visualization tools incorporated within Athena.

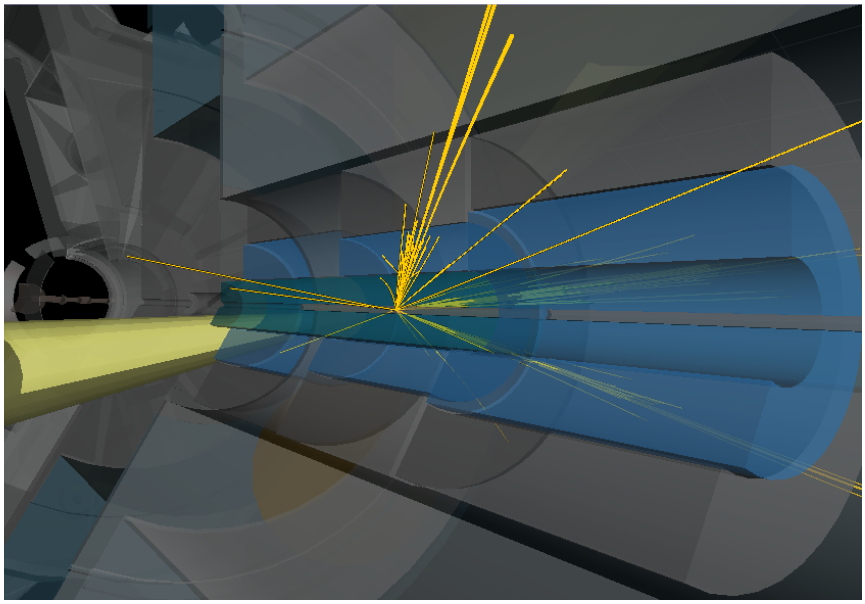


FIGURE 4.8. A 1500 GeV Z', $c\tau = 1\,\text{mm}$, Model A event showing the leading central jet in the event and tracks with p_T greater than 2 GeV, showing the lack of reconstructed tracks associated to the leading jet in the event.

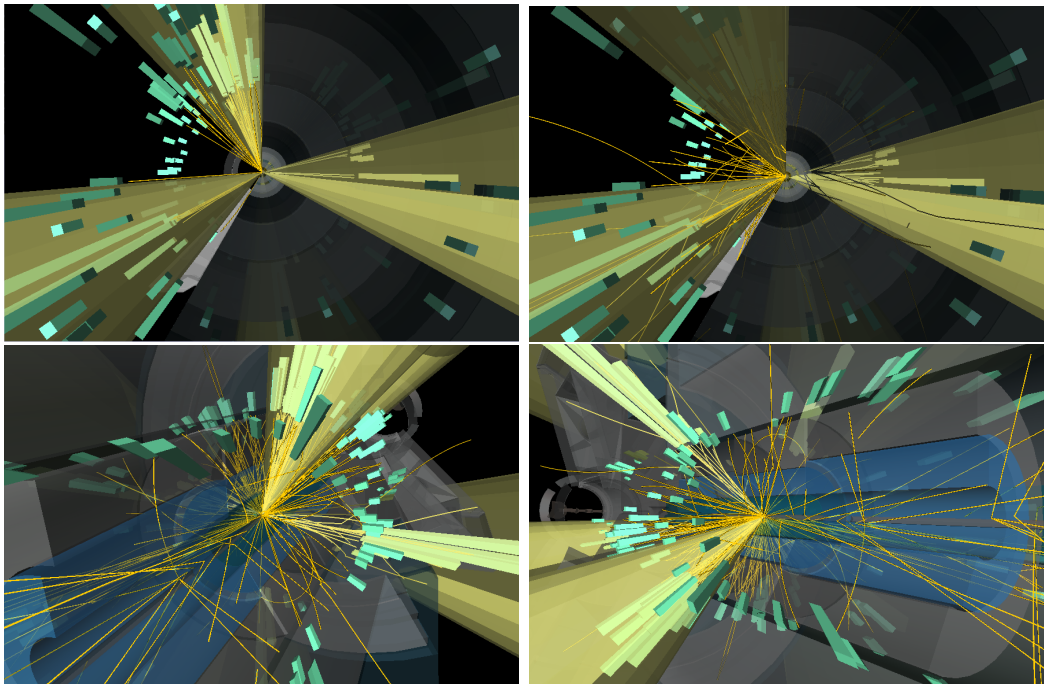


FIGURE 4.9. A 1500 GeV Z', $c\tau = 5$ mm lifetime, Model B event. Top left - beampipe-transverse view of the event with $p_T > 50$ GeV ($R = 1.0$) jets and $p_T > 2$ GeV tracks. Top right - same as the left plot but showing all tracks associated to the PV. Bottom row, oblique projections showing calorimeter energy deposits, jets, and all the tracks in the event.

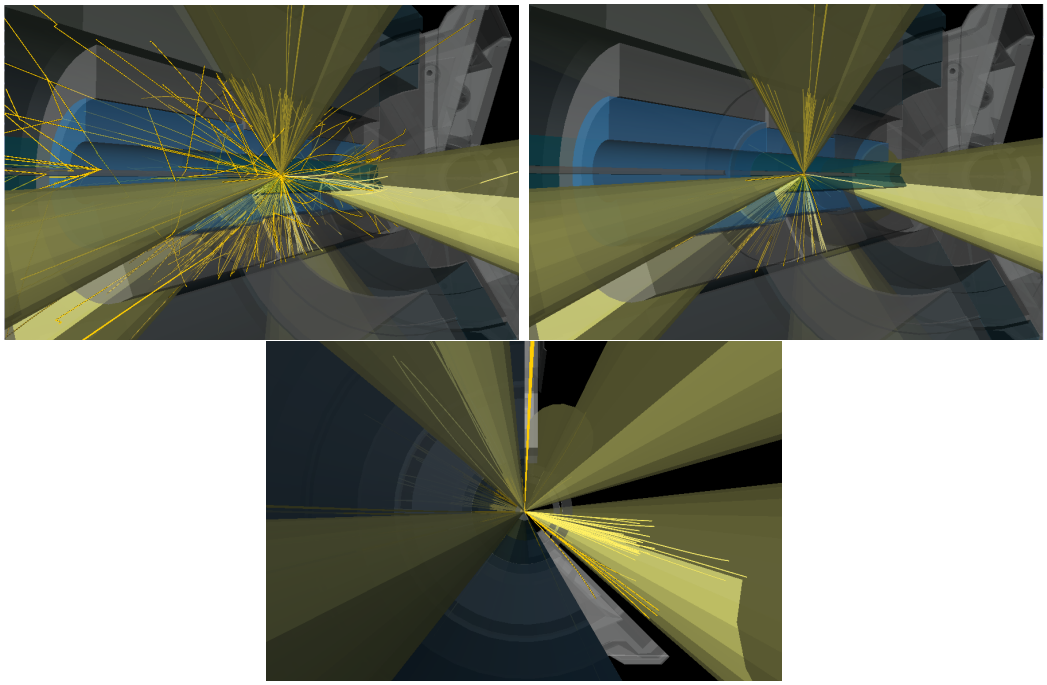


FIGURE 4.10. A 1500 GeV Z', $c\tau = 10$ mm, Model B event with $p_T > 50$ GeV (R = 1.0) jets. Right and bottom plots - oblique and beamline transverse plots with $p_T > 2$ GeV tracks associated to the PV.

4.5. Data Selection

Data were taken from Run 2 of the LHC as recorded in the "physics_Main" stream of ATLAS. Events were filtered to remove events that were not taken during stable data-taking periods, such as those during unstable beam conditions or during detector issues. The list of acceptable data-taking periods are listed in Good Run Lists (GRLs), and the applied GRLs are shown in Appendix A.1.

These events lie in a range of data-taking conditions, such as during different pileup conditions, trigger menu settings, and bunch-crossing configurations. Events were selected that passed the lowest unprescaled single-jet hardware trigger, known as L1_J100. The L1 trigger system performs square clustering on the coarse topocluster calorimeter trigger-tower segments, and events not meeting the

p_T cutoff of 100 GeV (at the L1 energy scale) are discarded. These events represent a large fraction of the total number of events triggered by ATLAS at L1.

Events were then filtered into the EXOT3 derivation format. The EXOT3 derivation is a filtered set of events and physics objects for specific analysis use, particularly for high jet-mass events with decays to large-R jets. This derivation of the data required "an OR combination of single large-R jet and HT triggers", as well as "at least two online jets with mass $>$ 30 GeV and $p_T > 100$ GeV, or two jets with $p_T > 1$ TeV."

The leading jet masses for the MC are plotted in Figure 4.17. Data with reconstructed large-R (R = 1.0) jets are known to be efficient above a jet mass cut of 35 GeV due to the studies shown in Section 4.5.1. The jet mass distribution for QCD jets is known to peak close to 0 GeV, so this cut reduces a large quantity of background with little hit to signal efficiency.

The efficiencies of these signal samples after passing through the DAOD step are listed in Appendix A. The table shows some reduction in signal selection (30% - 60%), however these cuts are very efficient in removing background events; the background-dominated data had a selection efficiency of 0.177% after these analysis-level cuts. The fraction of events passing the sequential p_T and η cuts, jet mass cuts, $|y^*|$ cut, and PTF cuts are 14.8%, 6.9%, 4.2%, and 0.177% respectively.

To ensure that the truth-level cuts in the generation were not significantly affecting the efficiencies at the reconstruction level, the ntuple-level efficiencies were determined for one signal sample (2000 GeV, $c\tau = 1$ mm, Model A) without the presence of any truth-level cuts. From the 1k events originally generated, 755 passed the analysis-level (DAOD+cleaning+event-level) cuts, and 738 passed the

truth-level cuts for R = 1.0, while 694 passed based on cuts on the R = 0.4 truth jets.

4.5.1. Trigger selection

The trigger menu changed significantly over the run-taking and varied over the data-taking years to account for the expected pileup distributions in each year. The primary triggers implemented for each year are listed in Table 4.1, which represent the entirety of the unprescaled primary large-R dijet triggers that collected data during this period of time. The backup triggers are not listed but have insignificant luminosities and vary significantly on the pileup profile. Events were filtered into the final selection using these triggers, and the efficiency point for the highest-p_T threshold trigger was used to derive a conservative measurement for the m_{jj} efficiency point for all the Run 2 data.

Year	Lowest Unprescaled Triggers	Online p_T Cut	Offline m_{jj} Cut
2015	`HLT_j360_a10_lcw_sub_L1J100`	360 GeV	1.3 TeV
2016	`HLT_j420_a10_lcw_L1J100`	420 GeV	1.3 TeV
2017	`HLT_j460_a10_lcw_subjes_L1J100`	460 GeV	1.3 TeV
	`HLT_j480_a10_lcw_subjes_L1J100`	480 GeV	1.3 TeV
	`HLT_j460_a10t_lcw_jes_L1J100`	460 GeV	1.3 TeV
	`HLT_j480_a10t_lcw_jes_L1J100`	480 GeV	1.3 TeV
2018	`HLT_j480_a10t_lcw_jes_L1J100`	480 GeV	1.3 TeV

TABLE 4.1. The lowest p_T unprescaled primary large-R jet triggers operating during Run 2 data-taking periods.

Due to differences between reconstruction at the trigger level and offline calibrated scale, the mass and p_T cuts that are placed online are not fully efficient offline at their specified online thresholds. Therefore, to apply an appropriate set of cuts to ensure that the data is efficiently selected from the output stream,

trigger efficiency plots were created. The efficiency of events selected by the probe trigger relative to a lower-threshold reference trigger were plotted as a function of the variable of interest, after appropriate cuts and calibration steps have been applied.

Events are generally required to be at least 99% efficient in m_{jj} in order to appropriately apply a background fit, although this threshold is luminosity dependent as higher luminosities increase the analysis' sensitivity to these trigger effects. Since the derivation specifies an online mass cut, the efficiency of the probe triggers used through the data-taking years was plotted, and the mass cut was chosen at the point where the trigger is 99% efficient.

An example of this method is shown in Figure 4.11 for the variable $\mathrm{p_T}$, from which the m_{jj} efficiency point is estimated. A conservative estimate is to assume that the m_{jj} turn on point is twice the $\mathrm{p_T}$ turn on point, although in practice it can be slightly lower than this. To ensure that the selected jet mass cut (35 GeV) is trigger efficient, trigger efficiency plots were created as a function of the jet mass. Events were required, at the offline level, to have jet cuts of $p_T > 500\,\mathrm{GeV}$, $|\eta| < 2.0$, and $\mathrm{m} > 35\,\mathrm{GeV}$, in addition to the event-level cut $|y^*| < 0.6$.

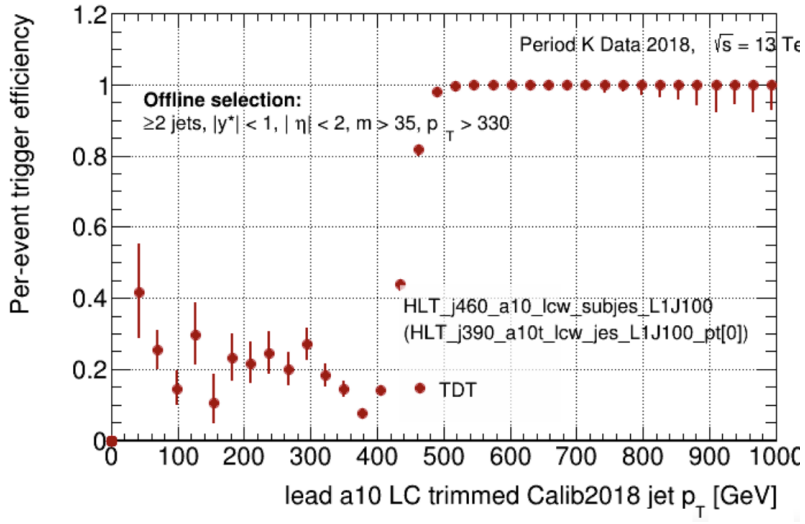


FIGURE 4.11. Offline p_T trigger efficiency plot for HLT_j460_a10_lcw_subjes_L1J100 triggered events relative to HLT_j390_a10t_lcw_jes_L1J100 events for 2018 Period K data. Events are fully efficient down to the $m_{jj} = 1300\,\text{GeV}$ cut with the conservative assumption of the m_{jj} turn on point equaling $2\times$ the p_T turn on point. This trigger is the highest p_T threshold trigger for the unprescaled primary triggers, so the m_{jj} efficiency point here is applicable to the other primary triggers in the dataset selection.

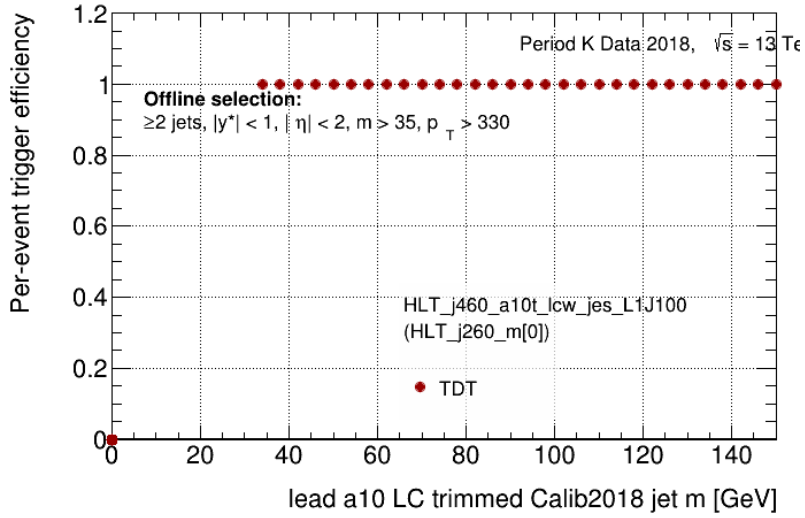


FIGURE 4.12. Leading offline jet mass trigger efficiency plot for HLT_j460_a10t_lcw_jes_L1J100 triggered event relative to HLT_j260_L1J75 events for 2018 Period K data. Triggered events are shown here to be fully efficient down to the 35 GeV offline mass cut.

Selection	Search Region Cut	Validation Region Cut
Trigger	Unprescaled R = 1.0 Jet Triggers	-
Offline Jet p_{T_0}	$> 500\,\text{GeV}$	-
Offline Jet p_{T_1}	$> 200\,\text{GeV}$	-
Online Jet p_{T_0}, p_{T_1}	$> 125\,\text{GeV}$	-
Offline Jet $\lvert\eta_0\rvert, \lvert\eta_1\rvert$	< 2.0	-
Offline m_0, m_1	$> 35\,\text{GeV}$ and $< 300\,\text{GeV}$	-
Online m_0, m_1	$> 30\,\text{GeV}$	-
$\lvert y* \rvert$	< 0.6	-
$\text{PTF}_0, \text{PTF}_1$	< 0.3	≥ 0.3 and < 0.5
m_{jj}	$> 1300\,\text{GeV}$	-

TABLE 4.2. Search and Validation region cuts applied to the data. The Validation region cuts are identical unless specified otherwise.

Due to trial-and-error experimentation during training, it was found that the NNs tended to learn to prioritize information about the highest jet-mass events, and were not necessarily selected due to their low PTF values. To reduce the effect of these spuriously-selected high-mass events, a tighter set of PTF and jet-mass cuts were applied; the jet masses were constrained to $35 \leq \text{m} < 300\,\text{GeV}$ and the PTF cuts were reduced to $\text{PTF} < 0.3$. This greatly improved the training stability and signal selection of the networks, and the PTF cut greatly reduced the overall training time due to the reduced number of events. This did come at the expense of statistics available in the tail-end of the spectra, where the method had issues extrapolating from the quantile NN score cuts in the sidebands to the signal region. This pushed the upper mass bound of the limits down to 4088 GeV and 3629 GeV for the spectra with the 10% and 1% neural network score cuts, respectively.

4.6. Bump Hunt Analyses

The general analysis procedure is: events matching the analysis cuts were filtered into datasets to form the spectra and training sets after the calibration sequence is applied. For each m_{jj} bin, events in the neighboring 4 bins on each side were used for the sideband training. The spectrum was filtered into separate spectra for events that pass each quantile cut over the sliding windows. The SWiFt method was then applied to the NN filtered m_{jj} spectra for background estimation using fits with a window width of 13 bins. The set of functions used to fit to the spectrum are the three parameter dijet function and the UA1 fit function, with the dijet function taking the form

$$f\left(x = \frac{m_{jj}}{\sqrt{s}}\right) = p_0\ x^{p_1}\ \ (1-x)^{-p_2-p_3\ln(x)-p_4\ln(x)^2}\,. \tag{4.5}$$

The three parameter function is Equation 4.5 with $p_3 = p_4 = 0$. In dijet (R = 0.4 jet) searches, the window width that is used for the fit is generally 11 - 25× the jet m_{jj} resolution, but it is adjusted to ensure that the function does not overfit or is under-constrained given the degrees of freedom available to the fit. The binning was chosen to be consistent with other dijet searches at ATLAS [14], where the binning also matches the m_{jj} resolution. The evolution of the binning starts at around 30 GeV in width for the first bins at around 1 TeV and increases to over 100 GeV at the 5 TeV scale, and is listed in Appendix A.1. The UA1 function is shown in Section 4.13 and was used as a fit systematic in the Search and Limit-Setting procedures.

A resonance within the data will appear as a bump in the m_{jj} spectrum that must be evaluated in a statistical manner. The BumpHunter algorithm

paper [65] elaborates on this algorithm that was developed to evaluate spectra for localized resonances. A large number of pseudoexperiments are formed to test the background hypothesis against the data. In each pseudoexperiment, the background estimate is Poisson-fluctuated, the starting fit parameters are varied, the fit function is varied, and a local test statistic is evaluated for each possible signal interval along the spectrum. The interval widths are allowed to vary from 2 to 5 maximum bins and define the range of potential bumps that can be statistically evaluated by the algorithm. The local p-value is evaluated in each interval by computing the Poisson probability of the data given that pseudoexperiment's fluctuated background estimate. A global test-statistic (t) is then computed by taking the minimum of the local p-values: $t = -\log\left(p_{\text{min,intervals}}\right)$ over all the evaluated intervals. The overall distribution of these test statistics is formed and also evaluated on the data, and the fraction of pseudoexperiments lying below the data's global test-statistic value is the overall spectrum's global p-value.

The interpretation of the global p-value is that any similar cut-based method used to rule out the null hypothesis would be incorrect at the p-value's level. For example, a p-value of 0.7 indicates that any attempt to refute the null hypothesis with an alternate model, using the same test statistic, would be incorrect 70% of the time. Therefore, the 0.01 cut that was placed on the p-value for the discovery threshold in this analysis is a 99% indicator of the likelihood that an alternate model (possibly a signal plus background model) exists that is more consistent with the data.

4.7. Neural Networks

In similar fashion to the behavior of neurons within biological systems, artificial neural networks (NNs) are a series of layered and weighted nodal decisions that are trained to match a desired output. While vastly less complex than biological systems such as the brain, artificial neural networks are seeing rapidly increasing use in many modern applications, and their complexity and use are widely varying in application. They are used to solve a variety of problems - common examples include reading handwritten digits and image classification. Traditionally in particle physics, NNs are applied to separate events by their background-like or signal-like qualities to provide a relatively accurate classification. Other applications in particle physics include the use of Recurrent Neural Networks (RNNs) to tag jets as b-quark initiated jets, the use of NNs in jet calibration, and the use of NNs to classify complex event types that have difficult-to-distinguish final-state topologies.

Preprocessed data are fed through the series of layers within a NN, with normalization and weighting applied to the computed information that's fed forward to each layer. The final layer's outputs are a set of numbers for each output layer; these can give classification probabilities if the model act as a multi-feature classifier. The action of this normalization, the individual and group weighting of nodes in the layer, the learning algorithm, and the structure of the network affects the overall accuracy and response of the network to its inputs. The output of these neural networks in the case of a binary classification problem is generally a set of numbers in the range from 0 to 1, with "1" indicating events with the most signal-like qualities, and the network can then define a probability of a particular output belonging to a particular class.

The features that are fed into a NN are generally preprocessed in advance for accuracy and numerical stability of the model. Normalization of the input data features and expected outputs in the general range of $[-1, 1]$ is preferred to increase stability of the trained NN to both outlier events and overtraining. In this analysis, event features were normalized by the subtraction and division of the feature distributions' approximate means. Further details on this normalization are in Appendix A.1. This approximately centered the distributions at the central values and normalized the features to their minima and maxima.

To allow for comparable results between this analysis and the original CWoLa analysis [58], a similar NN structure was chosen, $7 \times 64 \times 32 \times 8 \times 2$. The input feature space was increased to 7 input variables from the 2 (jet masses) that were previously used, but the same number and dimension of hidden layers was chosen, with 64 nodes in the first hidden layer, 32 nodes in the second hidden layer, and 8 nodes in the last hidden layer. A dropout layer between the first and second hidden layers in the network was set with a dropout fraction of 10%. If a network becomes too reliant on the information provided by a small subset of nodes, it can learn a model that's applicable only to the training data and can be not transferable to the test or other similar data. This is referred to as overtraining. By applying a dropout layer, a fraction of the nodes are dropped randomly from the model during the training and the learned model is more generalizable to additional data such as the validation data. A diagram of the network is shown in Figure 4.13. The output has two classification nodes, and the activation function is a softmax function to approximate the probability of each output being from either output class. The softmax function is a generalization

of the logistic function that allows the output values to represent approximated normalized probabilities of each output class. Analytically, it can be expressed as,

$$s(z_i) = e^{z_i} / \sum_{j=1}^{K} e^{z_j} \tag{4.6}$$

for each nodal output z_i.

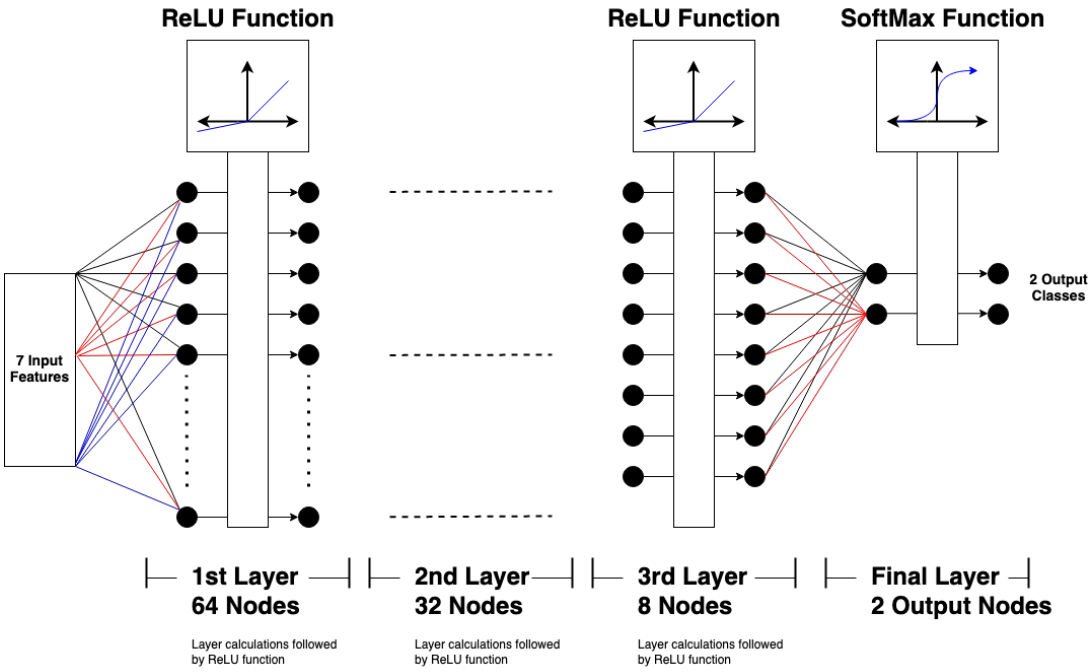


FIGURE 4.13. A visualization of the network architecture used for this analysis. Note that the dropout layer is not shown.

Networks were implemented in the Tensorflow framework [66] using the Keras backend [67]. Optimizer parameters were chosen to be consistent with the previous CWoLa analysis; the learning rate for the networks was 1×10^{-3} and the networks were trained for a maximum of 1000 epochs on a GPU. To help prevent overtraining, a scheduler reduced the learning rate by 0.25% for each epoch after the first 100. 20% of the input events were randomly selected as the validation set, and output plots of the AUC and the the validation loss were used to check for

over- and under-training. The loss function defines the overall inaccuracy of the network, and the learning algorithm attempts to minimize the overall loss function on the training data. The weighted Binary Cross-Entropy loss function [68] was used; it is a well-used loss function for binary classification problems. The analytic form of this loss function (with no event or class weighting) is

$$L(\vec{y}) = -\frac{1}{N}\sum_{i=1}^{N} y_i \cdot \log\left(p(y_i)\right) + (1-y_i) \cdot \log\left(1-p(y_i)\right), \tag{4.7}$$

with the true class label y_i (for binary separation problems, this is 0 or 1) and $p(y_i)$ is the evaluated probability of that label given the model, all for N evaluated events. As the training problem is highly imbalanced, with up to a factor of $10\times$ more sideband events than signal region events, class weighting was used to prevent the trained models from getting stuck in a local minimum, such as the situation where a false positive classification of all events as sideband events is the most optimal model configuration. The class weighting applied to this loss function effectively changes the representation of the number of events in each classification. For example, the class weighting of the signal region events w_s, with population N_s is

$$w_s = \frac{1}{\sum_i w_i(y_i = 1)} \frac{\sum_i^{N_s} w_i}{2}, \tag{4.8}$$

for event $i's$ weight w_i. With this weighting scheme, the value of the loss function becomes meaningful; for a properly reweighted set of signal and sidebands with identical distributions, if the relevant NN is unable to differentiate the mixed samples, then the loss should approach $\ln(2) = 0.693$.

The AdaM (Adaptive Moment) optimizer [69] was used to minimize this loss function due to its prevalent use in similar binary classification problems,

including the original CWoLa analysis. This algorithm uses a momentum term in the adaptation of the model parameters to help prevent the loss function from getting stuck in a local minimum. The network parameters are adjusted for each step of the algorithm proportionally to the first derivative of the loss function and inversely proportional to the second derivative, and the momentum has been shown to ensure quick and effective minimization of loss functions [70] across a variety of machine learning applications. Additional AdaM optimizer settings that are set are $\beta_1 = 0.8$, $\beta_2 = 0.99$, $\epsilon = 1 \times 10^{-8}$, and $r = 0.0025$, which correspond to the exponential constant of the first and second moment calculations, a divisor for numerical stability, and the learning rate decay constant, respectively. The batch size was set to 1% of the total training events to be consistent with the original analysis. In addition to the class weighting, individual sample weights were applied to each event that were trained on. When signal events are injected into various sidebands, they have a weight that depends on their estimated cross-section, the number of events in the signal sample, and the overall strength of the injection - this appropriately weights the events to their expected cross-sections. The event weighting w_e for signal events is therefore

$$w_e = \sigma_{\text{fid}} \cdot S \ / \ N, \tag{4.9}$$

for N generated events with a fiducial cross section of σ_{fid} at a signal strength of S. When the weighting of a particular event was well over unity, training instability was found to be problematic, as highly weighted events in smaller training batches have large effects on the computed loss function's gradients. So, in the case that a signal event's weight was above 2, the weightings were replaced with 1's and

the signal events were over-sampled many times to match the expected signal contribution for that particular weighting.

Outputs from the NN can have spurious values if the hyperparameters of the network allow the node weightings to change dramatically, or when the inputs have spurious values away from the normalized ranges from the training set. To reduce sensitivity to this effect and ensure that information fed into the next layer is appropriately normalized, each layer's outputs can be passed through an activation layer. Activation layers map output distributions to values generally between -1 and 1 by use of an analytical function. The choice of activation functions depends on the use case of the NN; some examples are the tanh function, the sigmoid function, and the rectified linear function.

The sigmoid function is the first layer's activation function; this allows the networks to be less sensitive to outlier features that might be overprominent in the training after the width-normalized preprocessing. The suppression of the outliers is due to the exponential powers in the sigmoid function. The three hidden layers in the network have the leaky rectified linear (ReLU) function applied in sequence. The use of these ReLU functions is a way to introduce non-linear effects into NNs and their use is standard for many binary classification problems in the data science field. ReLU is effectively a filter on information that suppresses negative layer information at each layer in the network; this information leads to a zeroed output and a negligible effect on the model's training gradients. The discontinuity and wider range of feature space that can be omitted from the NN's learning greatly improve the accuracy and resistance to overtraining, and introduces non-linearities into the model [71]. The leaky ReLU function allows for negative layer

information to still exert an influence on the NN's training, albeit with a user-set reduced scale ($\alpha = 0.3$) for negative contributions. The function is defined as,

$$\mathrm{ReLU}_L(x) = \{x \text{ for } x >= 0;\ \alpha \cdot x \text{ for } x < 0\}. \tag{4.10}$$

To account for situations where the network's initialization state was a local minimum that the solver could not adapt to, five independent networks were trained for each signal bin. The samples were randomly shuffled before the sample's training and test set split. The network with the minimum loss function was chosen as the network implemented within the analysis for each particular bin.

Since the highly-imbalanced nature of the datasets prevent the accuracy metric from being particularly useful, the Area Under the Curve is considered. It is the area under the true positive and true negative score ROC curve, and should reach 1.0 for perfect classification. Some examples of the training loss and AUC values from particular sidebands are shown in Figure 4.14.

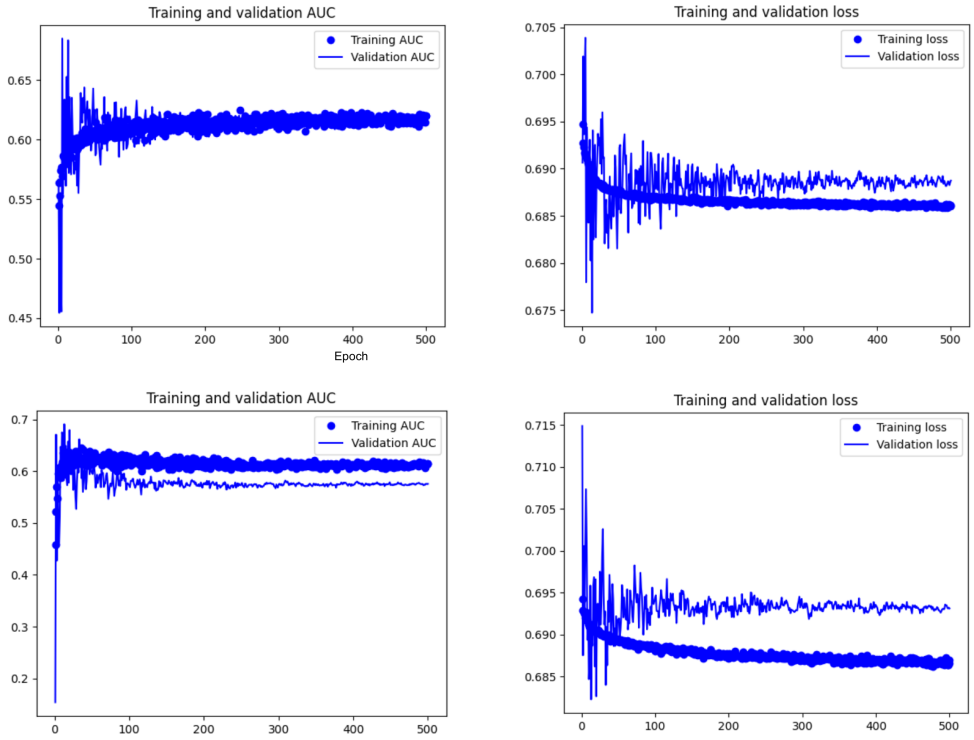


FIGURE 4.14. Sideband 8's (upper row) and 14's (lower row) evolution of the AUC (left) and loss function (right) over the training epochs. A small amount overfitting is demonstrated by these networks, however generalized hyperparameters that prevented overfitting for every bin were difficult to find.

4.8. CWoLa Method Validation

The efficacy of the CWoLa method in selecting signals of interest was determined via signal injection studies. The CWoLa method can act in two general capacities for enhancing the signal sensitivity - as an autoencoder and as a mixed classifier. Signal events are not guaranteed to be evaluated as "signal-like" by the trained networks, however their NN scores are not necessarily consistent with the "background-like" score distributions in each sideband. Therefore, the NNs can act as autoencoders as the more variate signal scores are more likely to lie in the upper 10% / 1% quantile of signal region scores. Of course, the signal events

can be enhanced by the central intent of the CWoLa method - to enhance the signal-like component of events in the signal regions as compared to events in the sidebands. The performance of the CWoLa method as an autoencoder can be tested by examining the signal efficiency of selection of injected signal for NNs that were trained without the presence of injected signal. The CWoLa performance as a signal classifier can be tested by examining the selection efficiency for NNs that were trained with signal injected at a variety of cross-section strengths.

One unfortunate quirk of the CWoLa method is that if few signal events appear in the signal region and if many appear in the sideband regions, the network learns to anti-tag signal events. This results in an all-or-nothing type enhancement of signal in this method; the signal sensitivity is sharply pointed towards the peak of the signal distribution and tends to select fewer signal events closer to the tails of the signal distributions. This greatly reduces the overall signal width, which is shown in Table 4.5.

The original CWoLa analysis used a different bin configuration for the construction of their signal regions and sidebands. Wide (1 - 2x signal width) regions were selected for the signal region, which often corresponded to regions of 5 - 10 bins. The same-sized neighboring regions were selected as the sidebands.

To determine the effects of the single-bin choice for the signal region in this analysis, the signal region was varied by increasing numbers of bins in the plus and minus directions. The signal efficiency for each signal region selection was evaluated for an injected signal with the signal region width varying from 1 to 7 total bins. A dependency of the overall signal efficiency was found on the number of bins, but this test case illustrates that the peak signal sensitivity is greatest when 1 bin is selected as the signal region. This is likely due to the fact that this

signal region definition contains the highest proportional fraction of signal, and the classifiers learn to distinguish these mixed samples more readily. This effect is illustrated in Figure 4.15.

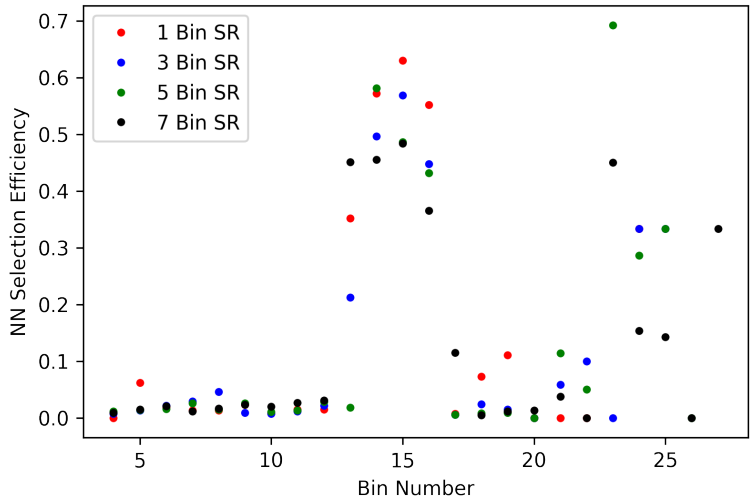


FIGURE 4.15. Effects from varying the signal region width used in training. The injected signal was a 2000 GeV Z' with cτ=1 mm and Model A. The signal peak at around bin 15 is enhanced the most by the 1 bin signal region. The spuriously-selected events in the high mass bins away from the peak are likely selected due to the autoencoder effect of this method.

4.9. Signal Evaluation

The 175 signal points that were generated for this analysis and in preparation for future Run 3 analyses are composed of exotic Z' particles produced with masses 1500, 2000, 3000, 4000, and 5000 GeV. Models A - E were used in the generation and the dark pion lifetimes were generated with cτ in the set of (1, 2, 5, 10, 25, 50, and 100 mm). These signal samples were reconstructed with an overlay of simulated pileup samples pulled from a probability distribution that mimics the pileup distributions for each of 2015+2016, 2017, and 2018 (referred to as mc16a,

mc16d, and mc16e campaigns). The identical EXOT3 derivation filter was applied to both these samples and to the data.

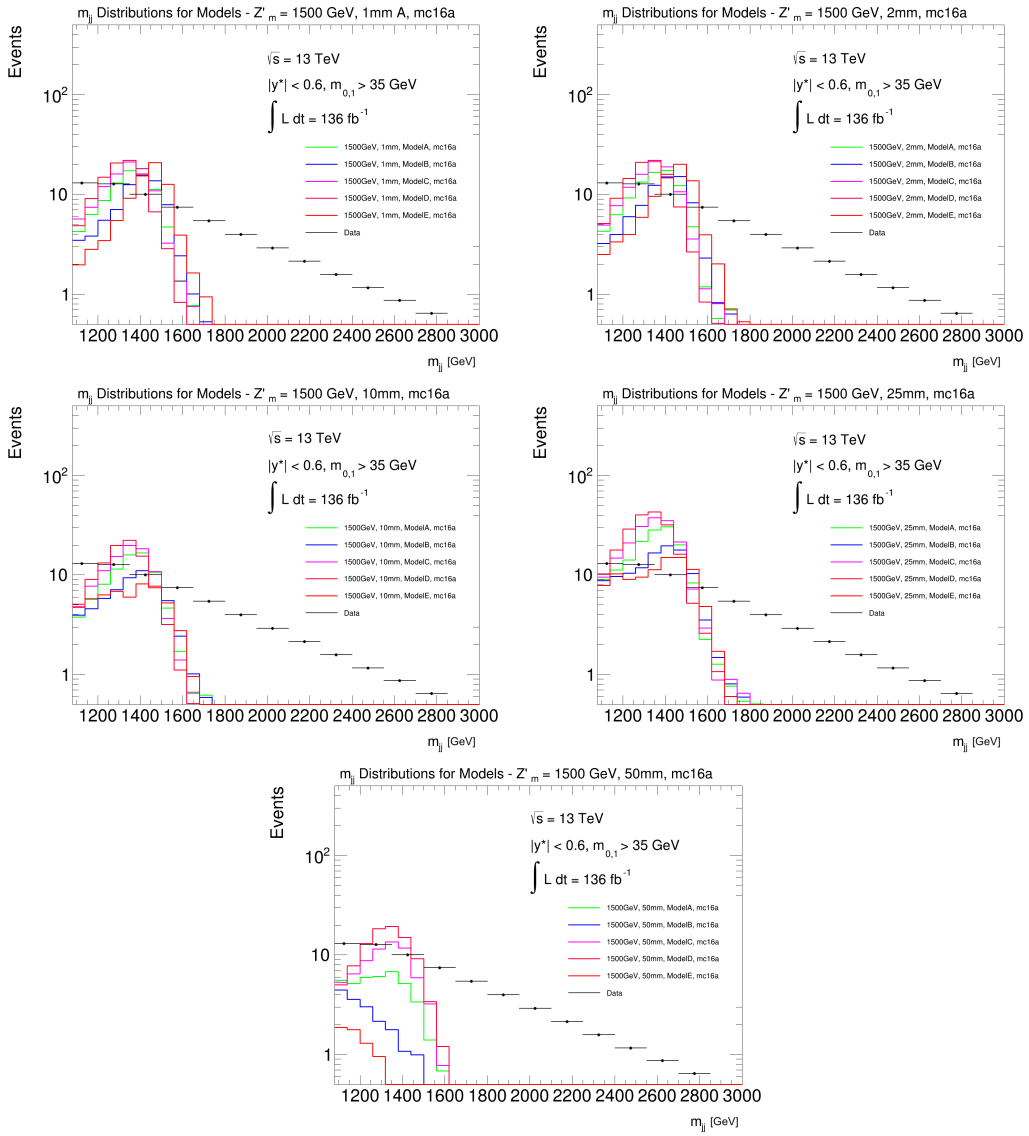


FIGURE 4.16. Invariant mass distributions of the signal samples sorted into 5 different lifetime sets. Models A-E are shown at each defined lifetime. The overall number of signal events are defined by the signal cross section which is a function of the Z' quark coupling g_q, which is set to 0.001 for these signal samples. The data are scaled to fit the plot's scale.

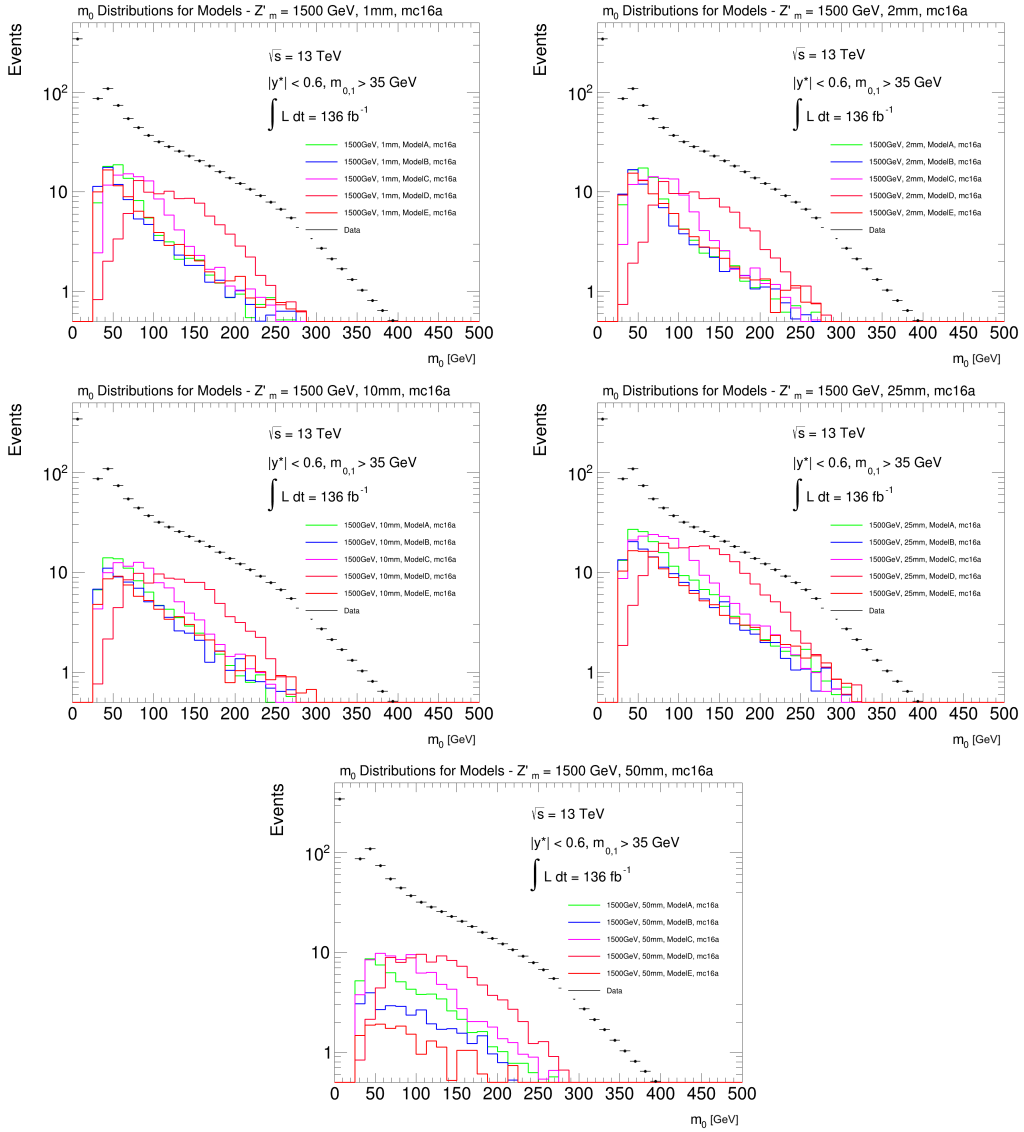


FIGURE 4.17. Leading jet mass distributions for 5 different signal lifetimes. Models A-E are shown at each defined lifetime for production during Run 2. The data distributions are scaled to show on the same plot.

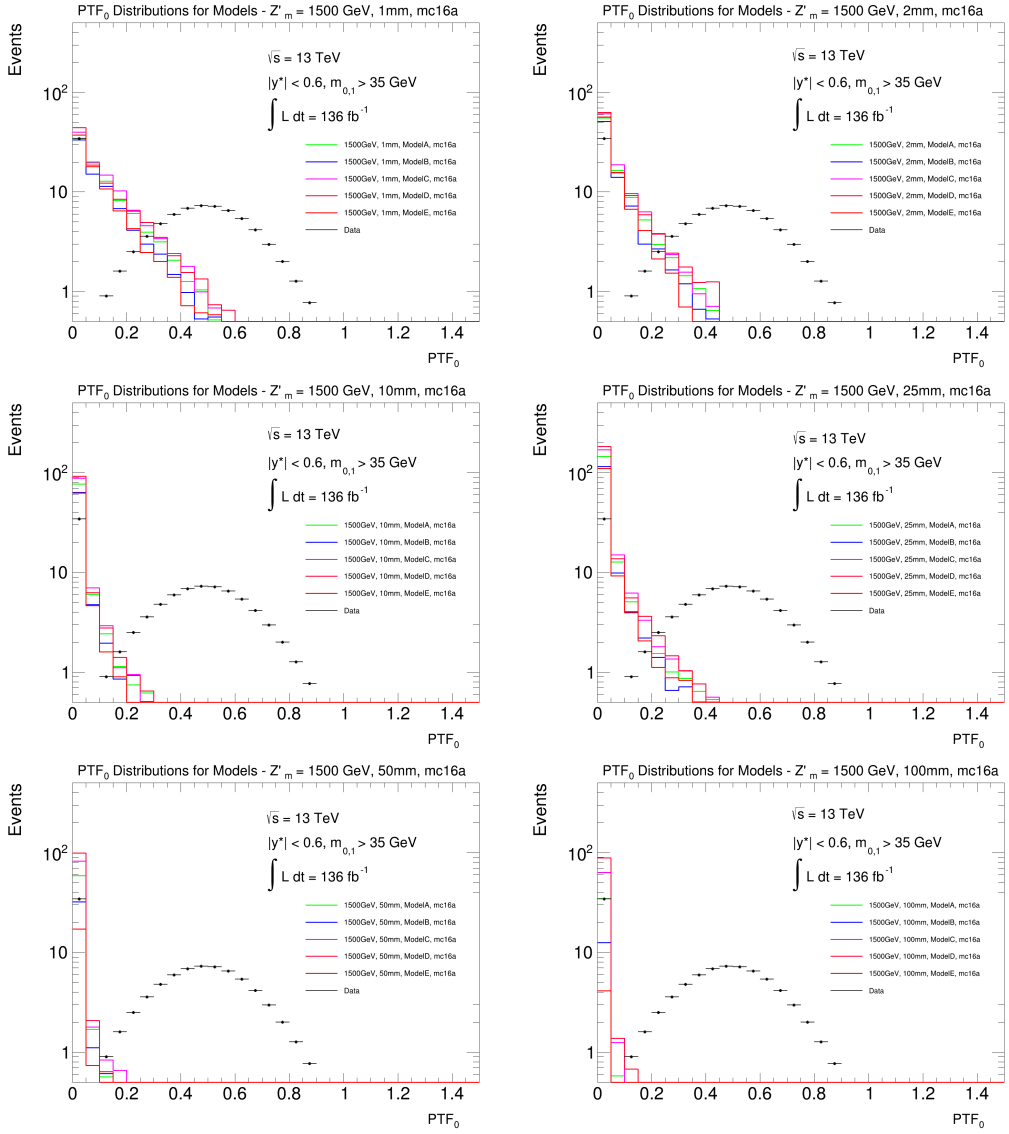


FIGURE 4.18. Leading PTF distributions for 6 different signal lifetimes. Models A-E are shown at each defined lifetime, and the data are scaled to fit the plot.

The training revealed that the neural networks tend to discriminate the signal regions from the sidebands primarily from the jet mass information. The signal efficiencies tended to peak at regions of high leading and subleading jet mass, and the networks did not readily learn to use the PTF variable appropriately with just the derivation and the analysis cuts, such as the $|y^*|$ cut. The NN

selection efficiencies also tended to peak in regions of high PTF, indicating that spurious events with unusually high PTF values tended to bias the NN learning. To reduce both of these effects, a PTF < 0.3 cut and the jet mass cuts were placed on the leading and subleading jets in the events. The signal sample PTF distributions tend to reach a maximum at 0.3, as shown in Figure 4.18. Therefore, a tighter PTF cut was not applied in order to prevent dependence of the limits on the sample lifetime, in addition to ensuring that there were reasonable statistics available to fit the m_{jj} spectra in the higher mass bins.

4.9.1. Signal Injection Studies

To test the learning capabilities of the trained networks, signal events were injected into the data spectrum at varying strengths and the efficiencies of the neural networks selections were evaluated. The overall cross-sections of the signal spectra are quark-coupling dependent, but to avoid regenerating the signal samples at each strength, the spectra were weighted by their respective strengths before injection. The samples were generated at a quark coupling of 0.001, but events were injected with strengths in the range (1x, 10x, 100x, 1000x). The efficiencies listed in Table 4.3 reflect the signal efficiencies without factoring in the effects of the derivation and analysis cut inefficiencies from the reconstruction and analysis.

Inj. Mass	Inj. $c\tau$	Inj. Model	Strength	Selection with 10% NN Cut
2000 GeV	1 mm	A	0x	9.1 %
2000 GeV	1 mm	A	20x	15.1 %
2000 GeV	1 mm	A	200x	36.8 %
2000 GeV	1 mm	A	2000x	42.1 %
2000 GeV	1 mm	A	20000x	47.6 %
3000 GeV	10 mm	B	0x	23.9 %
3000 GeV	10 mm	B	20x	22.8 %
3000 GeV	10 mm	B	200x	26.3 %
3000 GeV	10 mm	B	2000x	37.1 %
3000 GeV	10 mm	B	20000x	39.1 %

TABLE 4.3. Global signal selection efficiencies for a range of signal injection strengths.

The 3 TeV signal samples starts at a selection efficiency of 23.9% when no signal is injected, indicating that there is a correlation between the events naturally in the data that are selected by the NNs and the injected signal. The NNs' predominance in selecting events with high jet masses is likely the cause of this autoencoder-like effect. The 2 TeV signal starts at an approximately random selection (10%) and increases to 47.6 % at the highest injected signal strength, which is likely the fraction of the peak that is not rejected due to the anti-tagging effect. If there is no enhancement in signal efficiency due to the response of the NNs, the signal significance, $S/\sqrt{B}$, is expected to decrease by $10/\sqrt{10} = 3.2$ if signal is randomly selected. Therefore, at the point where a signal efficiency more than 32% at a NN score of 10%, the baseline signal efficiency is improved relative to a random selection. Similarly, at the NN cut of 1%, $100/\sqrt{100} = 10$, indicating that the signal significance is boosted once a signal efficiency of 10% is achieved.

The injections were performed on search spectrum data at varying strengths, and the global signal selection efficiencies are listed in Figures 4.34 and 4.33. The NN selection efficiencies and resulting signal sensitivities for 1500 GeV ($c\tau$

= 1 mm) injected samples for individual bins are shown in Figures 4.23 - 4.31. They illustrate the significant model dependencies in the NN performances. The anti-tagging quality of the CWoLa method is also demonstrated, where both the significance and efficiencies of the NN selection are worse than random selection in the tails far from the location of the signal peaks.

The background fits get skewed by the presence of localized signal above a certain injection threshold, and the discovery potential is limited by the maximum injected signal that can be statistically detected. This is demonstrated in Figures 4.19 and 4.20, although these plots illustrate injections done with approximately 29 fb^{-1} of data and a limited set of trigger decisions considered. The discovery threshold lied between an injection strength of 2000x and 20000x at a 10% NN score cut and between 10x and 100x at a 1% NN score cut.

In addition, the anti-tagging effect of these NNs resulted in an overall reduction in the signal widths. This manifested in the analysis as a reduction in the Gaussian widths for the determined limits. They resulted in a reduction of the signal width from 12 - 18% to 5 - 10%, as can be seen in Tables 4.4 and 4.5 and Figures 4.23 - 4.32. In these figures, the dashed lines indicate the "break-even" point - the point where the signal sensitivity is enhanced relative to a random selection.

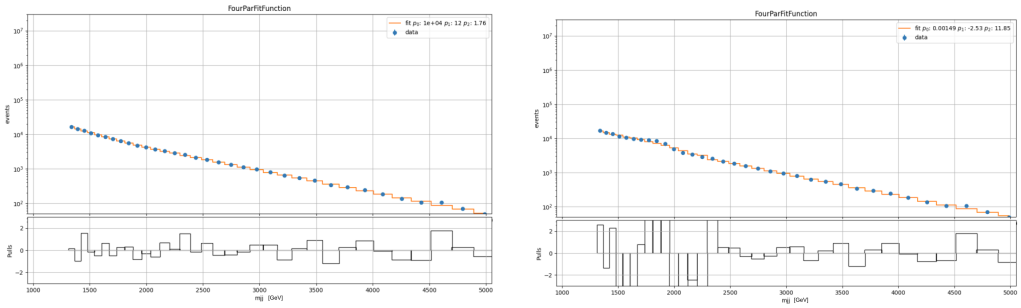


FIGURE 4.19. Signal injections (2000 GeV Z', $c\tau = 1\,\text{mm}$, Model A) done at 2000x and 20000x strength relative to the cross-sections defined from $g_d = 0.001$. The BumpHunter p-values for these injections were 0.77 and 0.0 for these spectra after application of the 10% NN score cut.

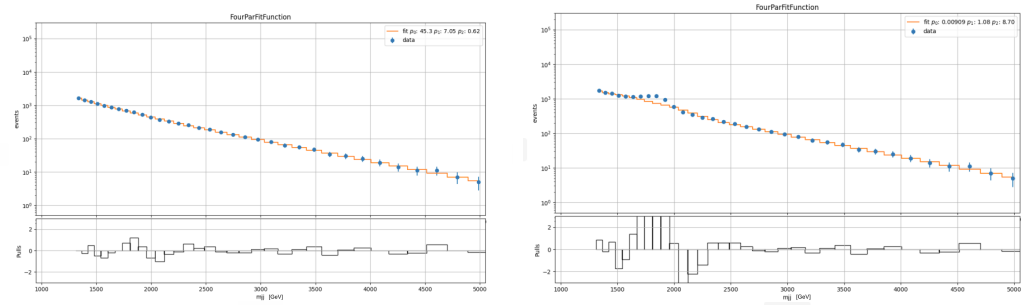


FIGURE 4.20. Signal injections (2000 GeV Z', $c\tau = 1\,\text{mm}$, Model A) done at 10x and 100x strength relative to the cross-sections defined from $g_d = 0.001$ after application of the 1% NN score cut.

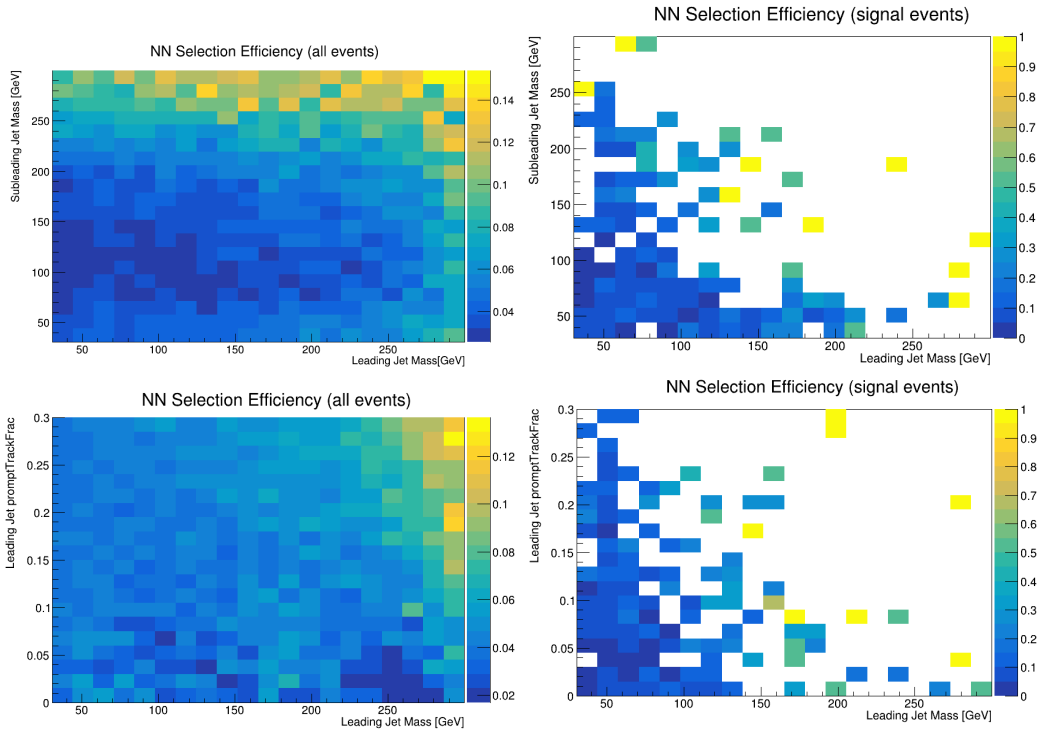


FIGURE 4.21. Selection efficiencies as a function of leading and subleading jet mass and leading jet promptTrackFrac distributions for data and an injected 1500 GeV Z' signal. The trained neural networks tended to learn information about high jet mass events and preferred to use these mass variables as a discriminant.

Z'_m	Model	Width	Z'_m	Model	Width
1500	A	18%	3000	A	13%
	B	18%		B	13%
	C	18%		C	13%
	D	17%		D	14%
	E	15%		E	12%
2000	A	19%	4000	A	11%
	B	18%		B	10%
	C	18%		C	11%
	D	19%		D	10%
	E	13%		E	10%

TABLE 4.4. Unfiltered signal widths from Gaussian fits to $c\tau = 1$ mm samples. Models B and E have mass hierarchies closest to QCD and therefore generally have smaller widths.

Z'_m	Model	Width	Z'_m	Model	Width
1500	A	5.6%	2000	A	6.9%
	B	7.8%		B	6.9%
	C	5.6%		C	7.8%
	D	n/a		D	8.0%
	E	7.0%		E	5.6%

TABLE 4.5. Signal widths from Gaussian fits to $c\tau = 1\,\text{mm}$ samples after application of the trained NNs.

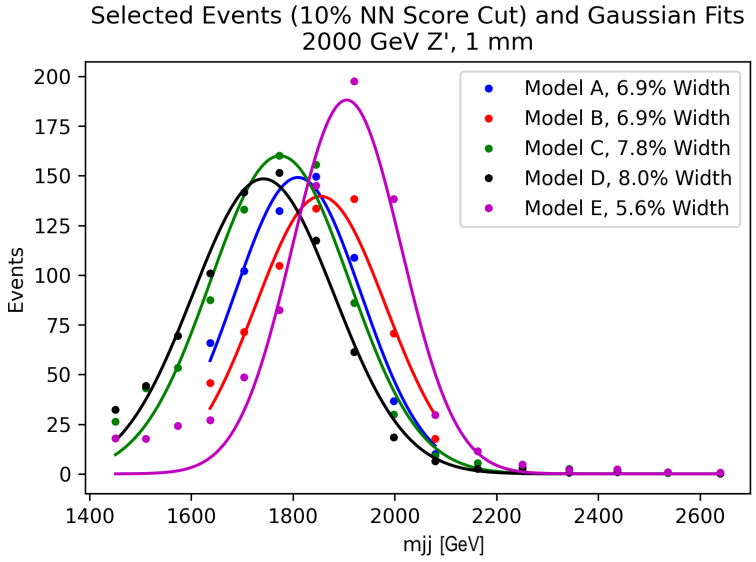


FIGURE 4.22. Selected signal events and Gaussian fits from a 2000 GeV Z' injected with signal strength 50 after application of a 10% NN score cut.

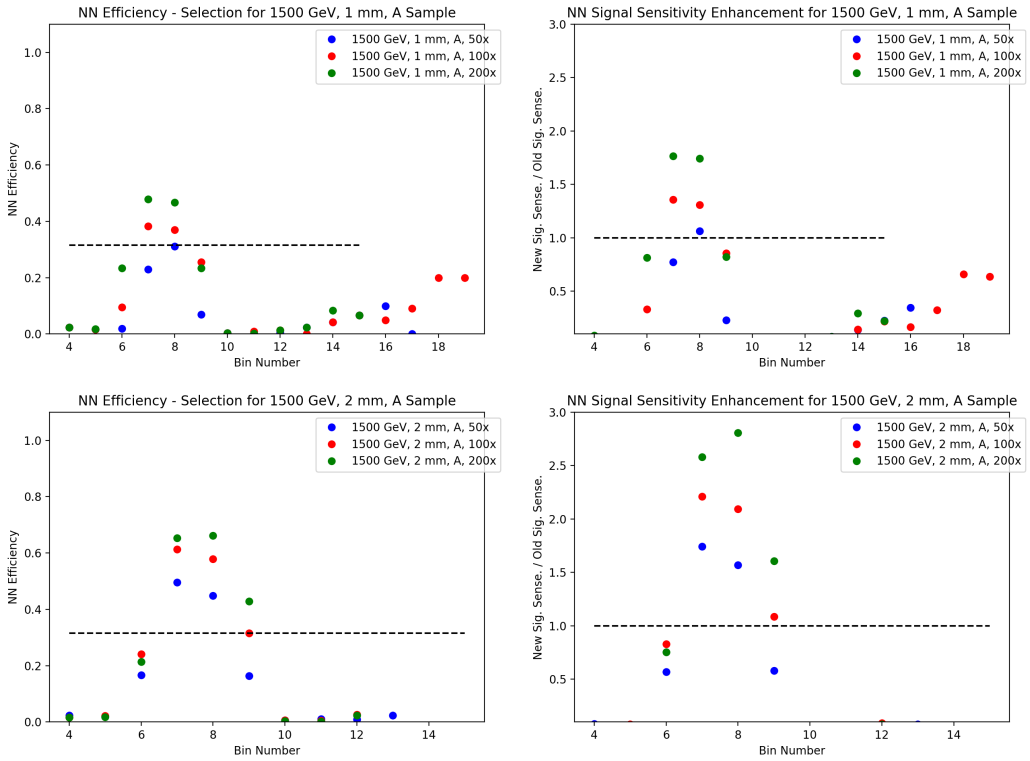


FIGURE 4.23. Signal injections (1500 GeV Z', Model A) done at 50x (2σ fluctuations in peak bin), 100x, and 200x strength relative to the cross-sections defined from $g_d = 0.001$. The top and bottom plots correspond to cτ=1 mm and cτ=2 mm samples respectively.

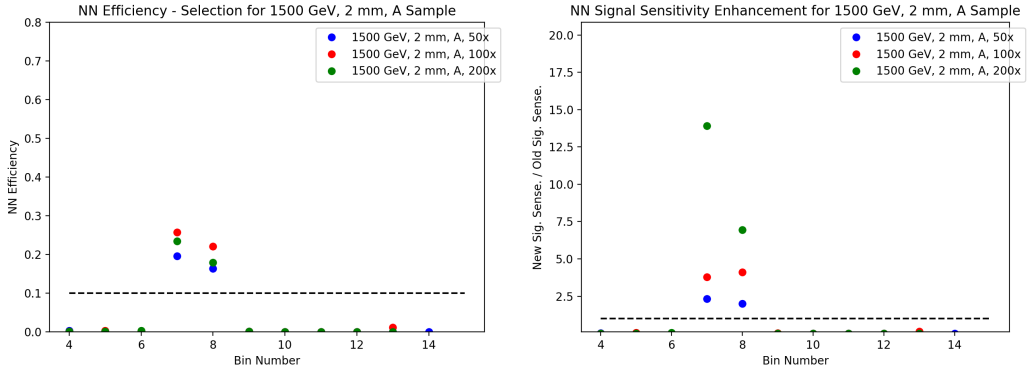


FIGURE 4.24. NN Signal efficiencies and Signal Significances for 1% NN score cut criteria on (1500 GeV Z', Model A, c$\tau = 2$ mm). Large enhancements in the signal sensitivity are seen relative to the 10% NN score cuts.

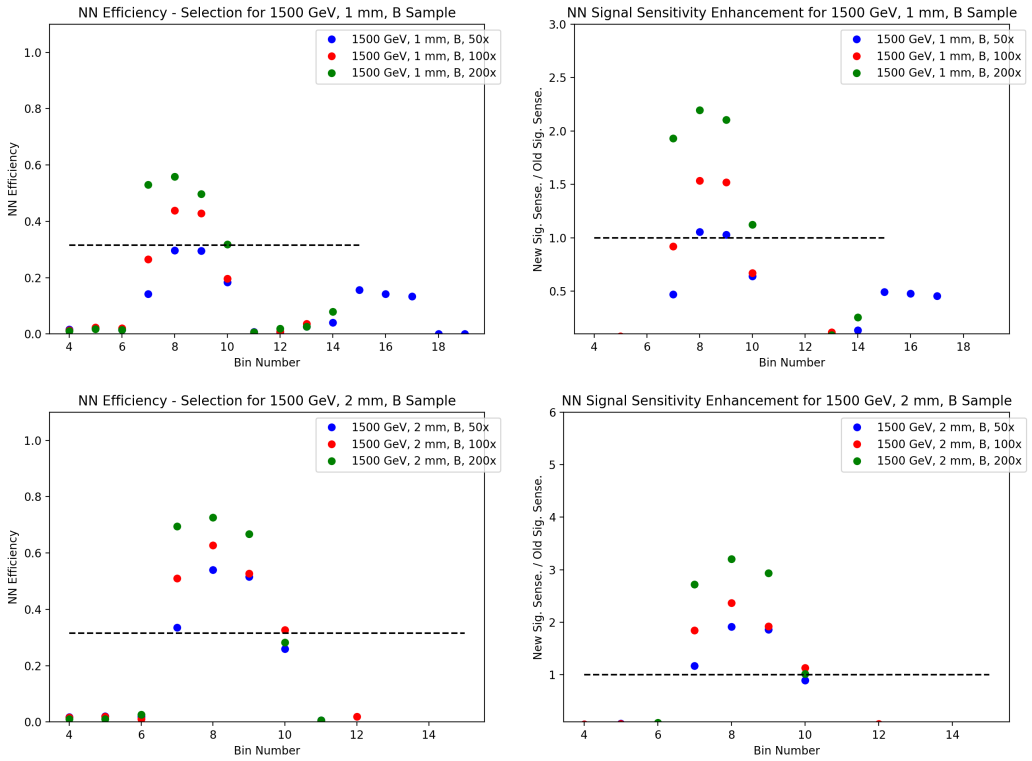


FIGURE 4.25. Signal injections (1500 GeV Z', Model B) done at 50x (2σ fluctuations in peak bin), 100x, and 200x strength relative to the cross sections defined from $g_d = 0.001$. The top and bottom plots correspond to $c\tau = 1$ mm and $c\tau = 2$ mm samples respectively.

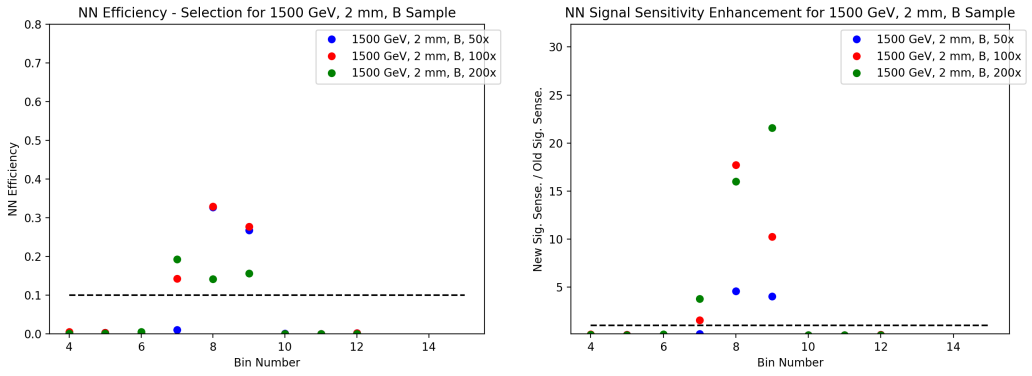


FIGURE 4.26. NN Signal efficiencies and Signal Significances for 1% NN score cut criteria on (1500 GeV Z', Model B, $c\tau = 2$ mm). Large enhancements in the signal sensitivity are seen relative to the 10% NN score cuts.

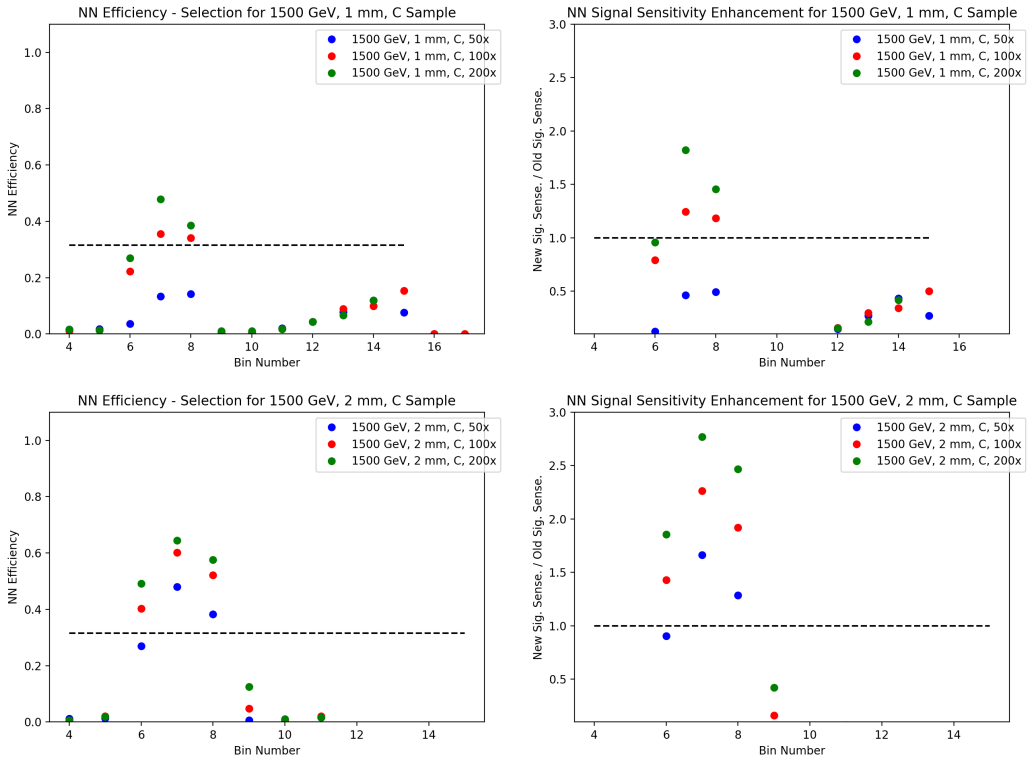


FIGURE 4.27. Signal injections (1500 GeV Z', Model C) done at 50x (2σ fluctuations in peak bin), 100x, and 200x strength relative to the cross sections defined from $g_d = 0.001$. The top and bottom plots correspond to $c\tau = 1\,\text{mm}$ and $c\tau = 2\,\text{mm}$ samples respectively.

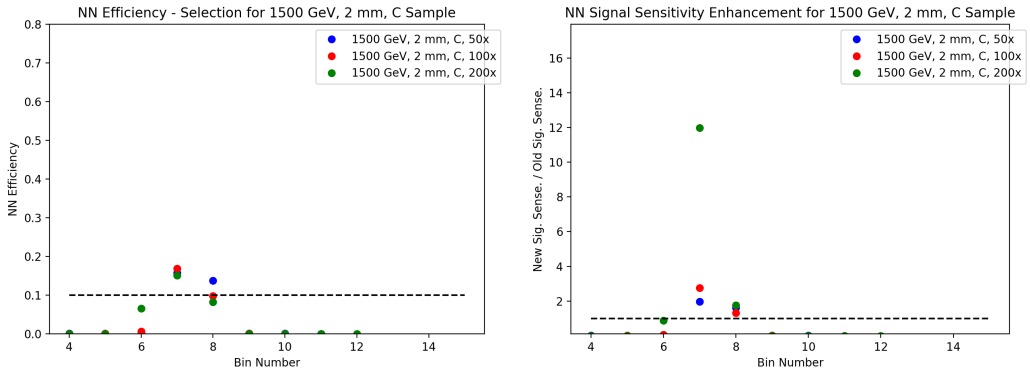


FIGURE 4.28. NN Signal efficiencies and Signal Significances for 1% NN score cut criteria on (1500 GeV Z', Model C, $c\tau = 2\,\text{mm}$). Large enhancements in the signal sensitivity are seen relative to the 10% NN score cuts.

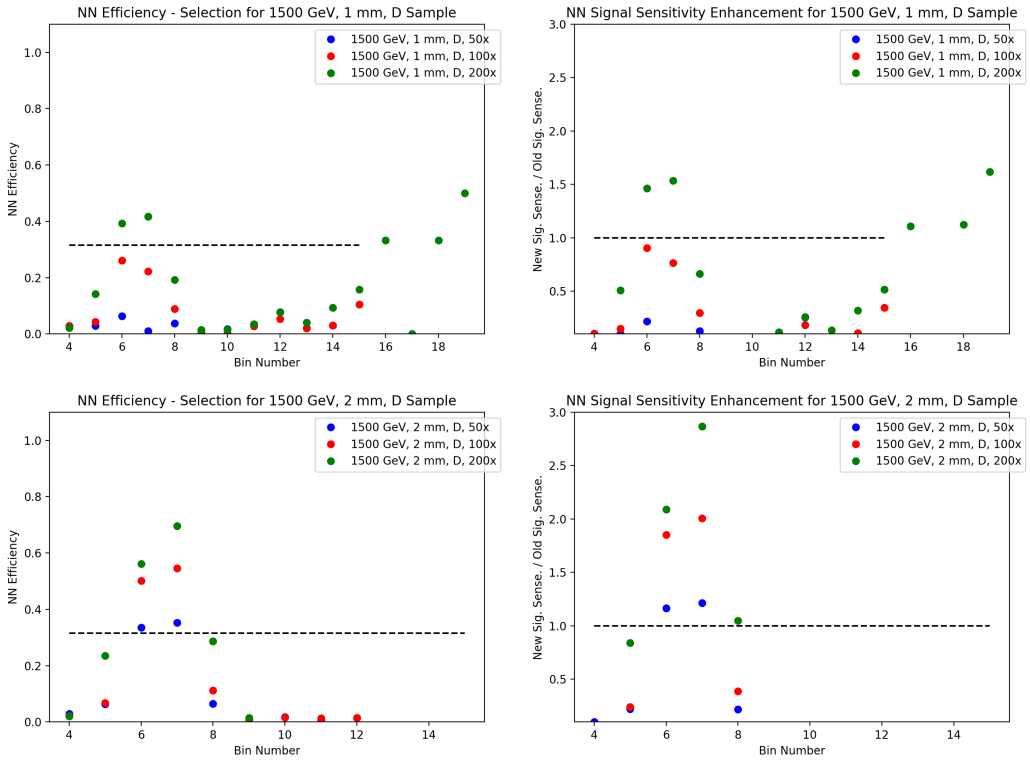


FIGURE 4.29. Signal injections (1500 GeV Z', Model D) done at 50x (corresponding to 2σ fluctuations in peak bin), 100x, and 200x strength relative to the cross sections defined from $g_d = 0.001$. The top and bottom plots correspond to $c\tau = 1$ mm and $c\tau = 2$ mm samples respectively.

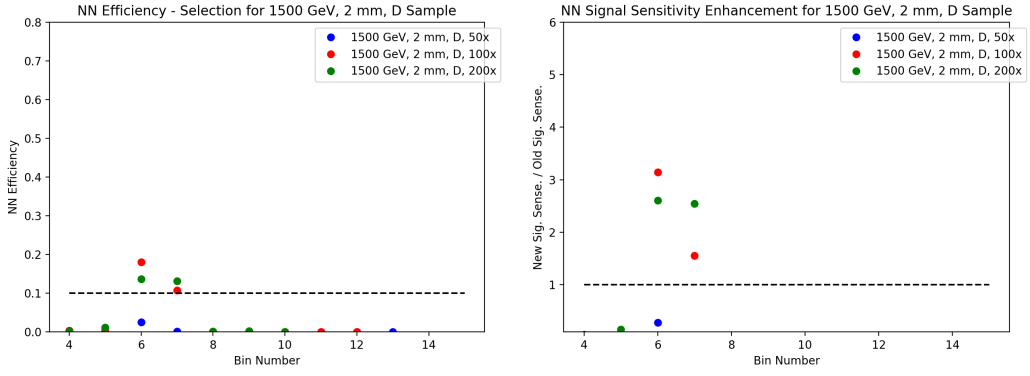


FIGURE 4.30. NN Signal efficiencies and Signal Significances for 1% NN score cut criteria on (1500 GeV Z', Model D, $c\tau = 2$ mm). Large enhancements in the signal sensitivity are seen relative to the 10% NN score cuts.

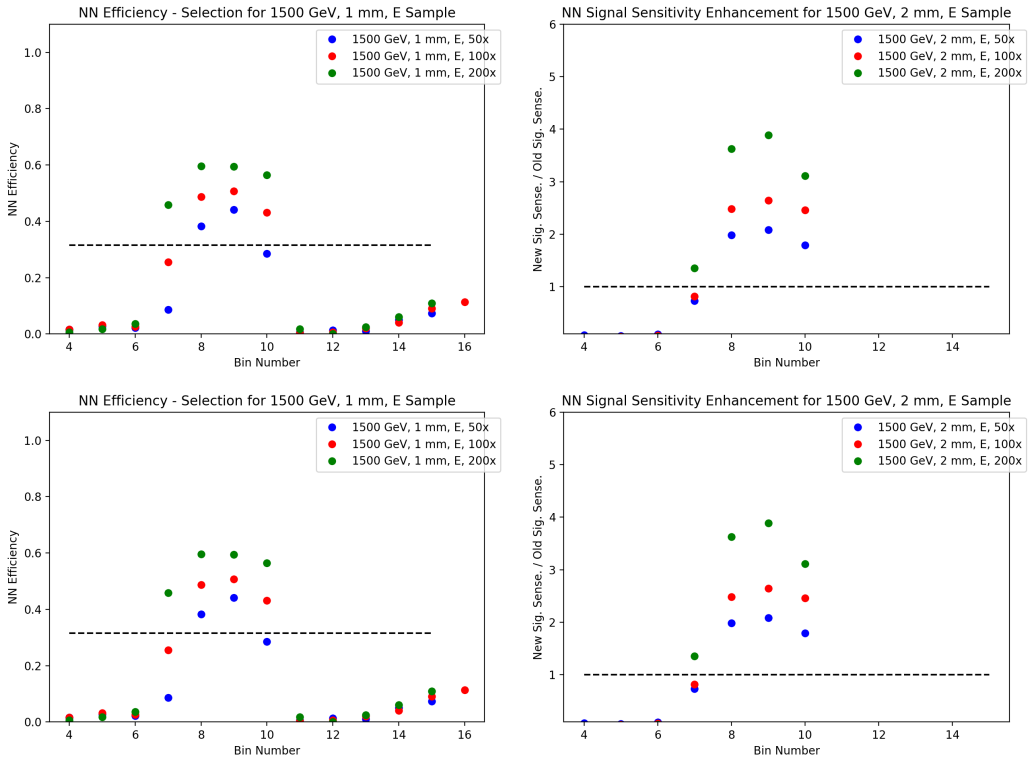


FIGURE 4.31. Signal injections (1500 GeV Z', Model E) done at 50x (2σ fluctuations in peak bin), 100x, and 200x strength relative to the cross sections defined by $g_d = 0.001$. The top and bottom plots correspond to $c\tau = 1$ mm and $c\tau = 2$ mm samples respectively.

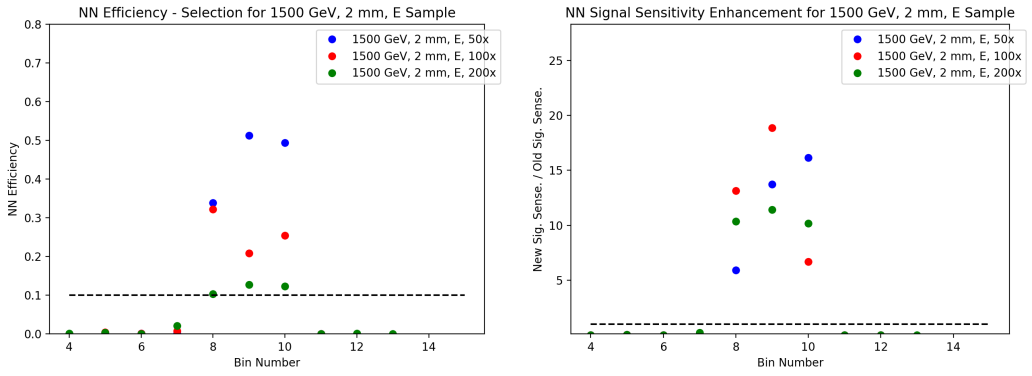


FIGURE 4.32. NN Signal efficiencies and Signal Significances for 1% NN score cut criteria on (1500 GeV Z', Model E, $c\tau = 2$ mm). Large enhancements in the signal sensitivity are seen relative to the 10% NN score cuts.

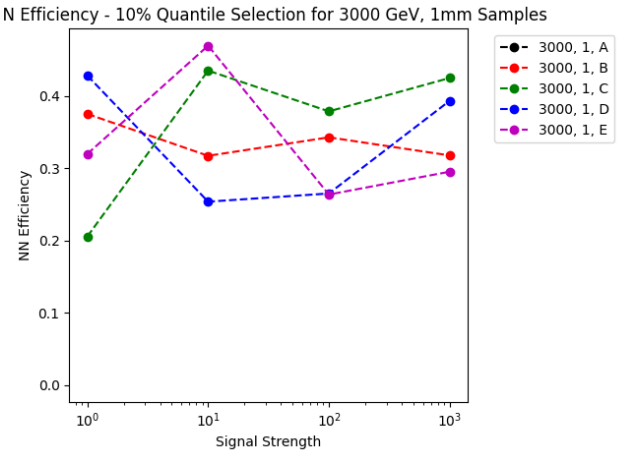


FIGURE 4.33. Global signal selection efficiencies for 3000 GeV ($c\tau = 1$ mm) samples by NNs in overall spectrum as a function of injected signal strength (efficiencies after 10% NN score quantile cut.) The overall signal efficiency is not necessarily the target metric of interest in bump hunt searches, but do indicate the general trend of the NN performance.

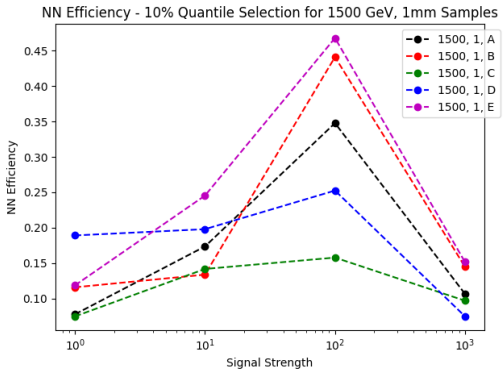


FIGURE 4.34. Global signal selection efficiencies after a 1% NN score quantile cut for 1500 GeV ($c\tau = 1$ mm) samples by NNs in overall spectra as a function of injected signal strength.

4.10. Signal Search Phase

The goal of the Search Phase of this analysis was to determine the probability that the observed data was consistent with statistical fluctuations from the smoothed background model. So given a smooth and analytic function that

is fit to a given window, variations from the background estimate at the probe (central) bin are expected to be due to statistical fluctuations from a Poissonian distribution in the presence of no signal.

The BumpHunter program [65] is a toolkit designed to search for these bumps in a given spectrum. The program handles two important aspects of a dijet analysis that are difficult to account for, combined up or down fluctuations and the "look elsewhere" effect. Poissonian fluctuations over the interpolated background estimate are expected, however these fluctuations must be correlated under the assumption of the presence of a signal, which will fluctuate the spectrum upwards within the signal width. For example, six bin-sequential upward fluctuations in the spectrum are far more signal-like than the same-sized fluctuations alternating in sign. Therefore, any probability computation must use the sign of the fluctuation as a measure of the signal-likeness. BumpHunter accounts for this by combining bins within a combination width in sets of pseudoexperiments; this accounts for correlations in these fluctuations.

In addition, in any statistical measure involving a large number of counting experiments (i.e. a bump-hunt search with a large number of bins), a certain fraction of the pseudoexperiments (bins) will fluctuate at a statistically significant amount. For example, in a counting experiment where the threshold for discovery is a 5% probable deviation from the background estimate, a discovery will be made, on average, in 1 out of 20 bins. BumpHunter accounts for this "look elsewhere" effect by performing an additional number of pseudoexperiments on the spectrum itself. It estimates the fraction of times that a given statistical fluctuation would produce a discovery at any other bin location in the spectrum and utilizes this fraction in the p-value calculation. Pairs of bins are alternated for

each psuedoexperiment; this accounts for situations where a random fluctuation might align in one local region with other fluctuations and induce a bump.

BumpHunter computes probability values by looking at the set of pseudoexperiments that are performed. Each bin is fluctuated by a Poisson distribution relative to the background estimate and the test statistic is computed. The p-value is the probability that, given the background estimate, the BumpHunter test statistic will be greater than or equal to the test statistic computed from the actual data. The statistical basis of this method is based on a theorem that shows that the Type-I error (probability to wrongly reject the null hypothesis / background estimate) is α if a well-defined test statistic has a discovery threshold of α on its p-value [65]. In the case of this analysis, the p-value criteria for discovery is set to 0.01 based on long-standing arguments from the BumpHunter developers. The interpretation of this number is that an alternate hypothesis other than the null hypothesis would have a 99% chance of fitting the model better, and since signal plus background models are inclusive in that set of hypothetical models, the presence of signal in the evaluated data can not be ruled out by this criteria.

4.11. Validation Region Studies

To validate the methodology applied in the analysis, NNs were trained and evaluated on a Validation region spectrum. Any "signal" discovery in this Validation region would be attributable to an issue within the applicability or the implementation of the CWoLa method within the analysis, or potentially due to another issue such as spurious bumps induced by miscalibration. The signal-sparse and background-enriched region that was selected for the Validation region was

defined where both jets in the events satisfy $0.3 \leq \text{PTF}_{0,1} < 0.5$. An alternate Validation region was tested that used an inversion of the $|y^*|$ cut, $|y^*| > 0.6$, however this region was severely affected by trigger inefficiencies up to 1550 GeV. The selected Validation region contains events more kinematically similar to the Search region events, with a selection that resulting in similar numbers of events to the Search region, in order to validate this method for a similar luminosity.

Following the plan for the Search Phase in the Search region, a data set was formed using the cuts specified in the Validation region column in Table 5.1. Events were sorted into the same bins as the Search region, and NNs were trained to separate the events from the signal bin from those in the sidebands. The Search Phase was performed with BumpHunter searching for significant bumps in the resulting spectra.

The results from BumpHunter indicate that the application of these NNs does not seem to induce spurious bumps in the spectrum. The BumpHunter p-values for the 10% and 1% spectra were 0.91 and 0.69 respectively, and the results are shown in Figures 4.35 and 4.36. The conclusion drawn from these Validation region studies is that detector-level issues or the CWoLa method did not causing spurious bumps at the luminosities that were considered. In addition, the methodology of extrapolating the NN score quantile from the sidebands to the signal region does not seem to produce any spurious signal as long as sufficient statistics are available for the training and evaluation. In the upper regions of the spectra, the training statistics are O(100) applied to O(10) events in each signal bin, forcing the potential upper mass limits of this analysis to below 4 TeV. Appropriate spectra fits could not be done above 4 TeV due to the statistical

fluctuations caused by the statistics-poor extrapolation of the NN score quantiles to the signal regions.

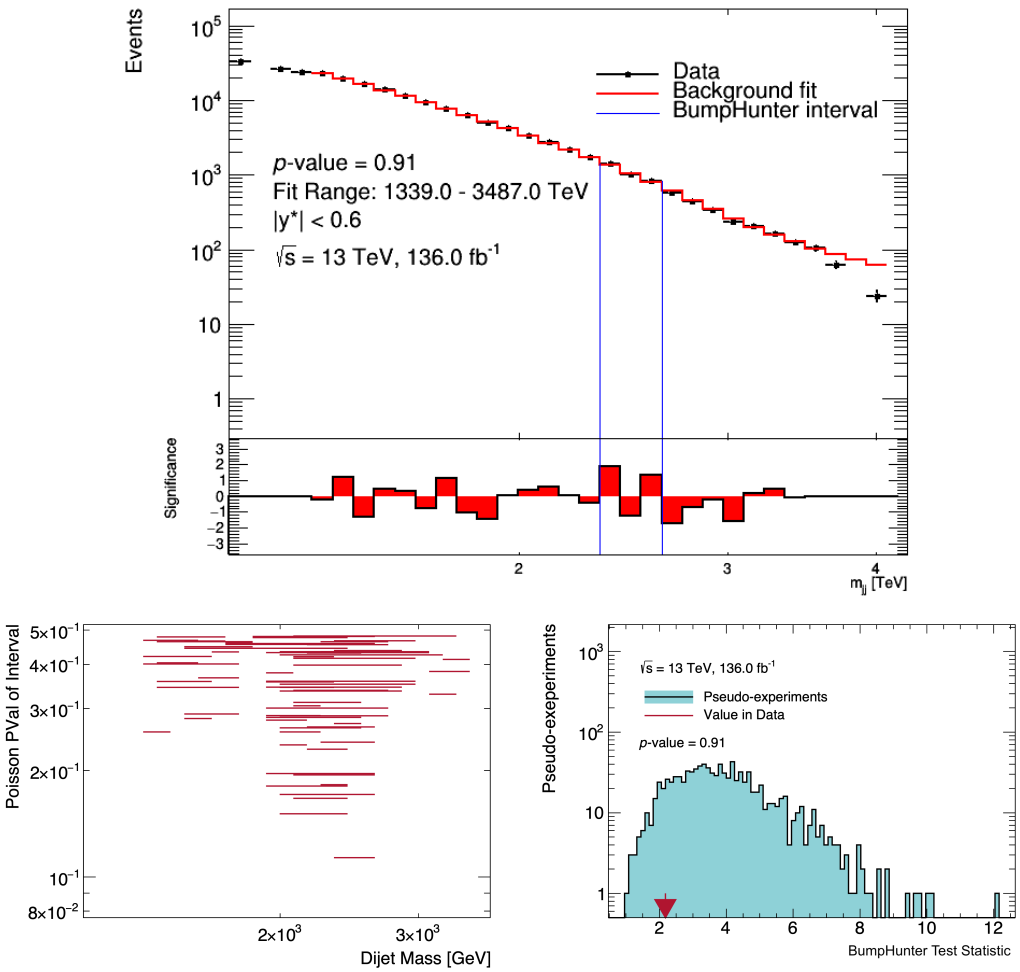


FIGURE 4.35. BumpHunter outputs from application to the 10% NN score cut Validation region spectrum. The upper plot indicates the formed spectrum and the maximal BumpHunter bump region, its tomography plot is shown on the bottom right, and the distribution of the BumpHunter test statistic is shown on the bottom right. Bumphunter evaluates the p-value for a range of intervals across the spectrum and uses the range with the minimal p-value as the maximum discovery region; these intervals are shown in the tomography plot.

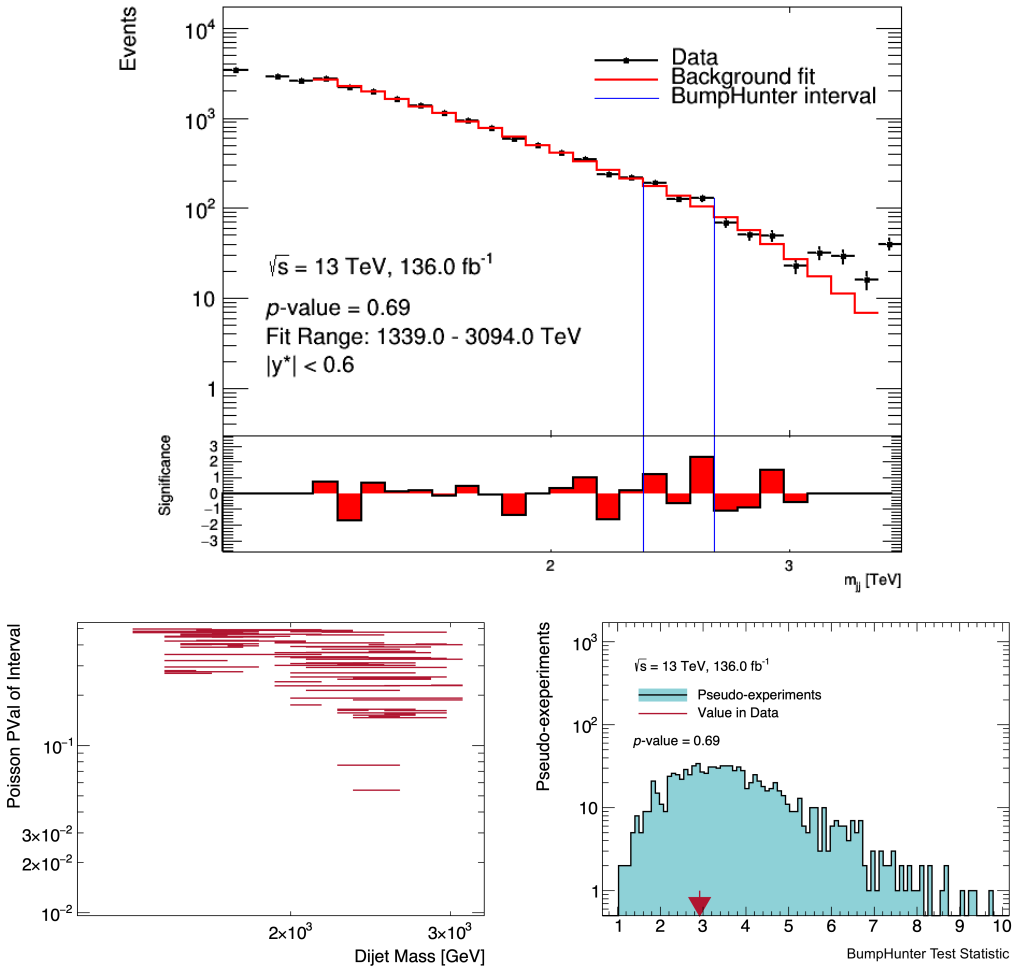


FIGURE 4.36. BumpHunter outputs from application to the 1% NN score cut Validation region spectrum. The upper plot indicates the formed spectrum and the maximal BumpHunter bump region, its tomography plot is shown on the bottom left, and the distribution of the BumpHunter test statistic is shown on the bottom right.

4.12. Limit-Setting Phase

Given the lack of indication of excesses or bumps in the spectrum, the limit-setting phase follows. The general question asked in this phase is "Given that no signal was observed, to what cross-sections can we place limits on the signal model?" This is a complex procedure involving the many facets and variations

possible in the analysis. For example, small changes in the JES can affect the signal shapes observed, modifying the overall sensitivity to a bump of that sort, and thereby affecting the final limits obtained.

The Bayesian Analysis Toolkit (BAT) is a high-energy physics framework designed to evaluate model limits [72]. The BAT evaluates injected signals of defined or approximated signal shapes and evaluates the maximum injection strength before discovery is possible. The posterior distributions of the signal cross-sections are estimated using a Monte-Carlo sampling method; this method assesses the signal strength given an injection with plus and minus 1σ variations from the systematics. The final limits obtained from the toolkit are uncorrected cross section that do not incorporate model and analysis considerations such as the signal branching ratio BR, total selection efficiency ϵ, and acceptance to produce the fiducial cross-section σ_{fid}. The relationship between these variables is

$$\sigma_{\mathrm{BAT}} = BR \cdot \epsilon \cdot \sigma_{\mathrm{fid}}. \tag{4.11}$$

A determination of the final cross-section limits must therefore incorporate the branching ratio and signal efficiency. For the assumption of a leptophobic Z' with decays purely to dark quark pairs, the Branching Ratio (BR) is 1.0. The signal efficiency is more complicated, as a series of truth-level, derivation-level, and analysis-level inefficiencies are aspects of this analysis that must be incorporated. The truth-level inefficiencies arise from truth-level filters that are placed on generated events to ensure that the produced events are likely to be resolved in the detector. These truth-level cuts are listed in Section 4.4, but are not easily reinterpretable by theorists, and so the final cross-sections can be more difficult to interpret. So, a fiducial cross-section σ_{fid} was defined that incorporated the initial

truth-level filters, and theorists can then define final-state efficiencies and fiducial cross-sections to compare to these cross-section limits.

The injected signals were Gaussian signals that varied in their injected mass, width, and amplitude. The largest number of injected events pulled from these signal shapes becomes the observed limit for that particular signal. The systematic parameters in the signal model, such as the JES and JER contributions, all contribute to the posterior probability of each tested signal point. A Markov-Chain Monte-Carlo (MCMC) algorithm [73] is used to define the central value of the observed limit at each mass point, and the uncertainties on that observed limit can be obtained by extraction of the CLs of each measured point. These observed and expected limits, however, are not the fiducial cross-section limits. The overall efficiencies of the signal selection and NN selection can be used to correct for the analysis procedure to extract fiducial limits on the emerging jets production process.

At the truth level, the fiducial cross-section is the product of the truth filter efficiencies, the production cross-section, and the Branching Ratio (BR) to the channel of interest.

$$\sigma_{\text{fid}} = \epsilon_{\text{truth}} \cdot \sigma_{\text{prod}} \cdot \text{BR}. \tag{4.12}$$

At the analysis level, the BAT cross-section limits σ_{BAT} are a product of the analysis efficiencies, the NN selection efficiencies, the derivation efficiencies, and the fiducial cross-sections. ALso, the leptophobic assumption of the model ensures that the BR can be set to 1.0 in the limit inference.

$$\sigma_{\text{BAT}} = \epsilon_{\text{analysis}} \cdot \epsilon_{\text{NN}} \cdot \epsilon_{\text{deriv}} \cdot \sigma_{\text{fid}}. \tag{4.13}$$

So the measured limits can be corrected to match the fiducial production by the correction,

$$\sigma_{\text{fid}} = \frac{\sigma_{\text{BAT}}}{\epsilon_{\text{deriv}} \cdot \epsilon_{\text{analysis}} \cdot \epsilon_{\text{NN}}}. \tag{4.14}$$

The BAT also produces expected limits that are used to give context to the observed limits. The expected limits are calculated by Poisson fluctuating the smooth background estimate in a number of pseudoexperiments. The ensemble of the Poisson-fluctuated spectra that can be generated without triggering the p-value discovery threshold becomes the expected limits, and 1σ and 2σ error bands can be determined from this procedure. Plots of the various efficiencies used in this analysis to convert the BAT limits to a fiducial cross-section are shown in Figures 4.37 - 4.38.

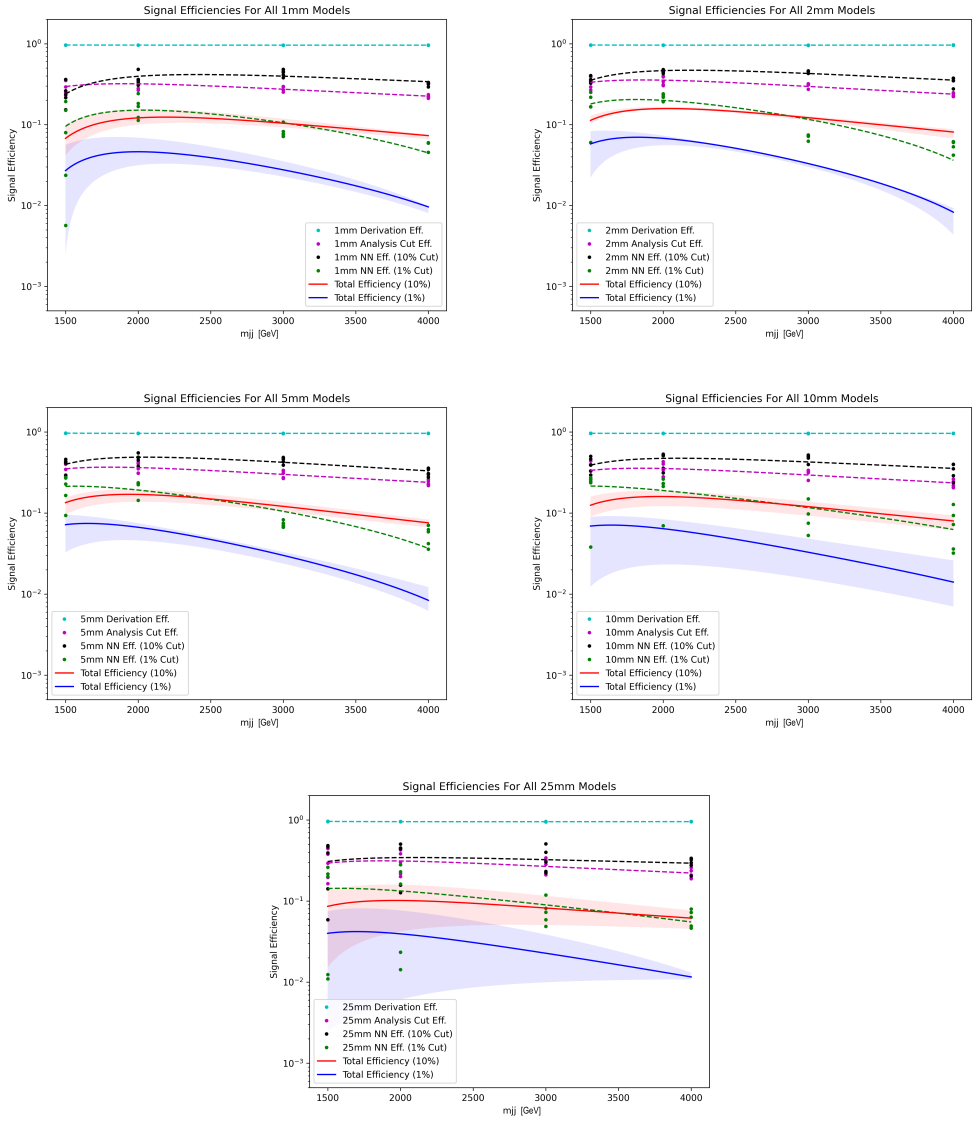


FIGURE 4.37. Signal efficiencies from application of the derivation-level cuts, the analysis cuts, and the NN selection for different lifetime samples. The envelope of the five model variations defines the error bands of the total efficiency and is an approximation of the hadronization systematic.

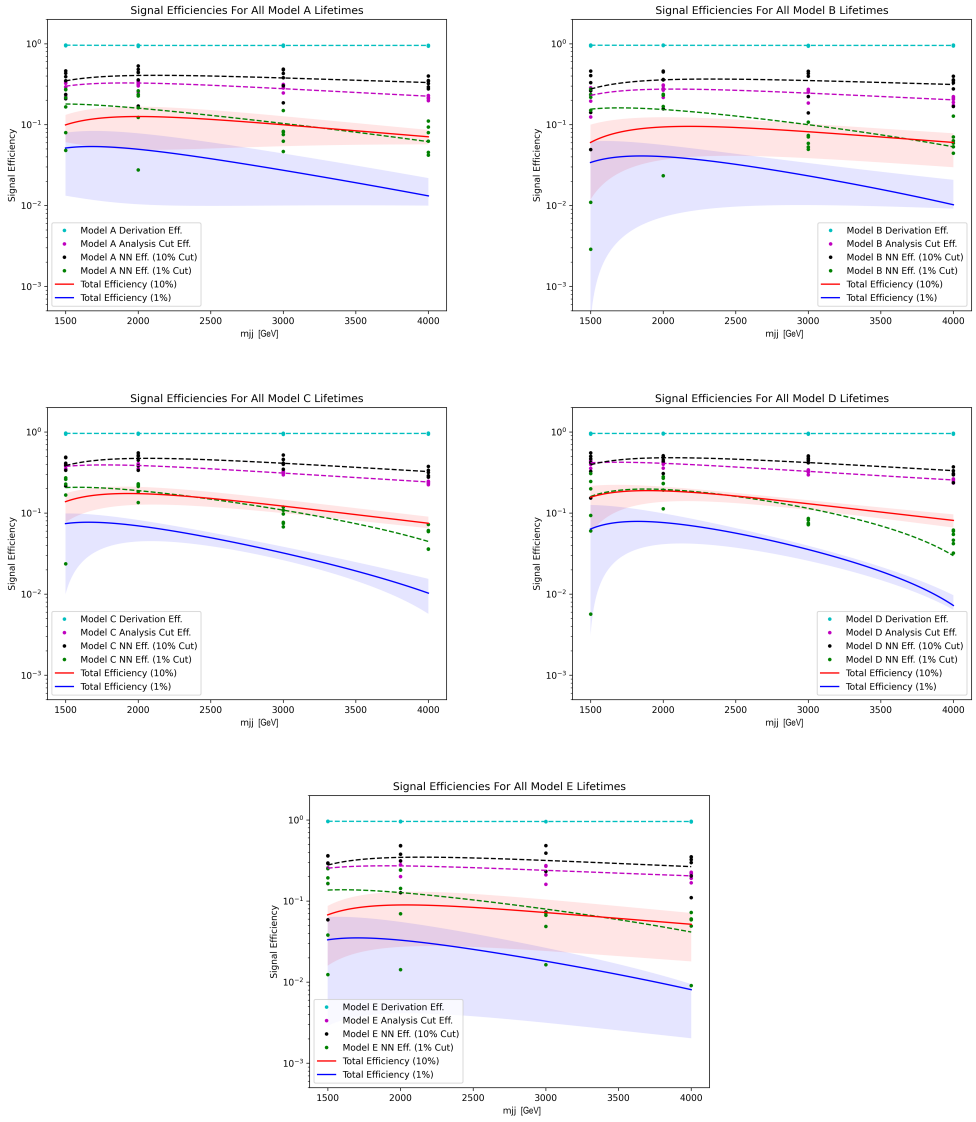


FIGURE 4.38. Signal efficiencies from application of the derivation-level cuts, the analysis cuts, and the NN selection for Model A - E samples. The envelope of the five model variations defines the error bands of the total efficiency. This can also be interpreted as an approximation of the effects of lifetime and width variance of the models on the limits, although this is beyond the scope of the analysis presented here.

4.13. Systematics

A number of systematics in this analysis affect the overall resulting limits and p-values. Uncertainties in measured quantities are affected by variances due to statistical fluctuations, but other effects ranging from the calibration procedure needed to reconstruct jets, the fitting procedure needed for a background estimate, and the hadronization model used during the signal generation can have large effects and must be accounted for in determining the final limits.

In any calibration procedure, uncertainties on the calibration method and the applied cuts are varied in order to determine their systematic effects on the final measurable quantities. There are a number JES systematics considered in analyses that affect the invariant mass that is reconstructed from the jets in the event. The standard JES variations for jet systematics in these analyses use variations of plus and minus 1σ, and so the BAT toolkit will vary the JES as applied to the signal shapes, shifting the spectra up and down from their nominal peak position [11] by a percentage that's defined by the JES uncertainty. For this analysis, the JES uncertainty was conservatively set to 2%.

The overall calibration affects the p_T resolution and the JES in a manner that will shift and smear the Gaussian injected signals. Figure 3.14 shows the uncertainty bands on the JER measurements, which hit a maximum of 2% close to the low edge of the p_T distribution. So convservatively, the JER systematic can be approximated by a shift in the JER analytical function by 2%.

The fit type systematic approximates the variation in the background estimate due to the choice of fit function. The standard procedure for dijet searches is to use a different class of fit function as the systematic variation of

the background estimate. To be consistent with previous dijet analyses, the UA1 function was chosen as the functional form for this alternate fit function.

$$f(x) = p_0 \cdot x^{-p_1} \cdot \ln\left(p_2 \ / \ x\right) \cdot \ln\left(p_3 \ / \ x^2\right) \tag{4.15}$$

The fit error systematic defines the uncertainty on the background fit. If a local fitter is used, marginally different background estimates can be formed depending on the fitting algorithm and the starting point of the fit parameters. The BAT accounts for this effect by performing fit pseudoexperiments to determine the effect of varying starting parameters by the fitting algorithm on the overall limits.

The luminosity uncertainty indicates the total uncertainty of the integrated luminosity observed by the detector [74]. The luminosity calibration procedure started with a baseline luminosity that was measured in low-luminosity detector conditions using van der Meer scans [75]. These beam scans are intended to measure the beam cross section and maximum visible luminosity using forward detectors such as LUCID [76]. The overall uncertainty of the Run 2 luminosity measurement is 1.7% after corrections to the baseline luminosity measurement are made to account for pileup conditions.

The hadronization model affects the overall width of the Gaussian signal shapes, as well as the signal strength. So the spread of observed limits for a variety of signal widths becomes a measure of the hadronization uncertainty on the BAT limits. The fiducial cross-sections are estimated by dividing by the analysis' total signal efficiency, and the envelope of the signal efficiencies approximates the effects of the hadronization systematic on the final fiducial limits. Therefore, the final

fiducial limits shown in Figures 4.41 - 4.42 have the hadronization systematic incorporated into the bands of the plotted limits for each dark model.

4.14. Results

The Search Phase evaluated the spectrum to determine if there were any statistically significant bumps using BumpHunter. The spectra were extracted with the application of the NNs with 10% and 1% cuts on the NN scores, and BumpHunter was applied to the resulting spectra with ranges $1300 < m_{jj} < 4088\,\text{GeV}$ for the 10% cut spectrum and $1300 < m_{jj} < 3628\,\text{GeV}$ for the 1% cut spectrum. The limited range is due to the statistics available above each corresponding cut after implementation of the NNs and analysis cuts, and the spectra and BumpHunter statistical plots are available in Figures 4.39 and 4.40.

The standard evaluation for p-values that is used in ATLAS dijet searches uses a threshold of 0.01 before evidence of a discovery can be determined. In the event that a below-threshold p-value is found, a series of cross-checks of the data reconstruction, calibration, and the fitting procedure would be done to check for any source that can explain a spurious bump in the spectrum. In the event that an experiment- or methodology-based explanation cannot be found for the excess causing this p-value, this is considered evidence of an unexplained spurious signal and more exhaustive checks would be required to understand why these simple analytical functions cannot fit the spectra. However, the p-values of these fits were 0.90 and 0.92 indicating that there is good confidence in no observed excesses in the spectra.

These computed p-values were obtained from a SWiFt-like fit to the data with a window width of 13 bins with the 3-parameter fit function, with 10000

pseudoexperiments performed by BumpHunter. The fit systematic uncertainties were evaluated with 1000 BumpHunter pseudoexperiments.

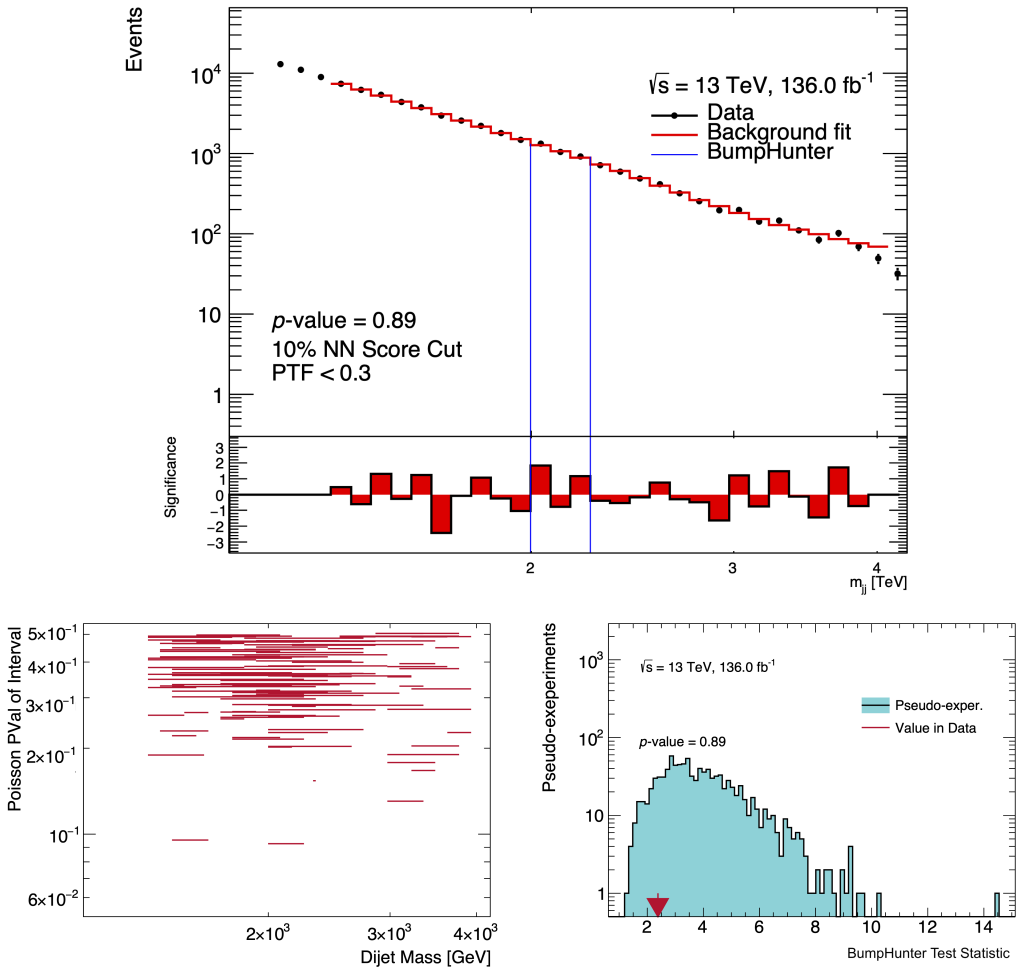


FIGURE 4.39. The three parameter analytical function fit to the spectrum after the 10% NN score cut. The BumpHunter p-value score for this spectrum was 0.90 for the nominal background, but the final value after inclusion of the fit systematics was 0.917 ± 0.009. The BumpHunter tomography plot is shown in the bottom plot, indicating the local p-value intervals in the algorithm. The bump of maximum local significance was between 1998 GeV and 2251 GeV with local significance 2.37932. The global fit parameters were [7.84855×10^{-11}, -21.1505, -12.4406] with the SWiFt fit χ^2/ndf = 0.711501.

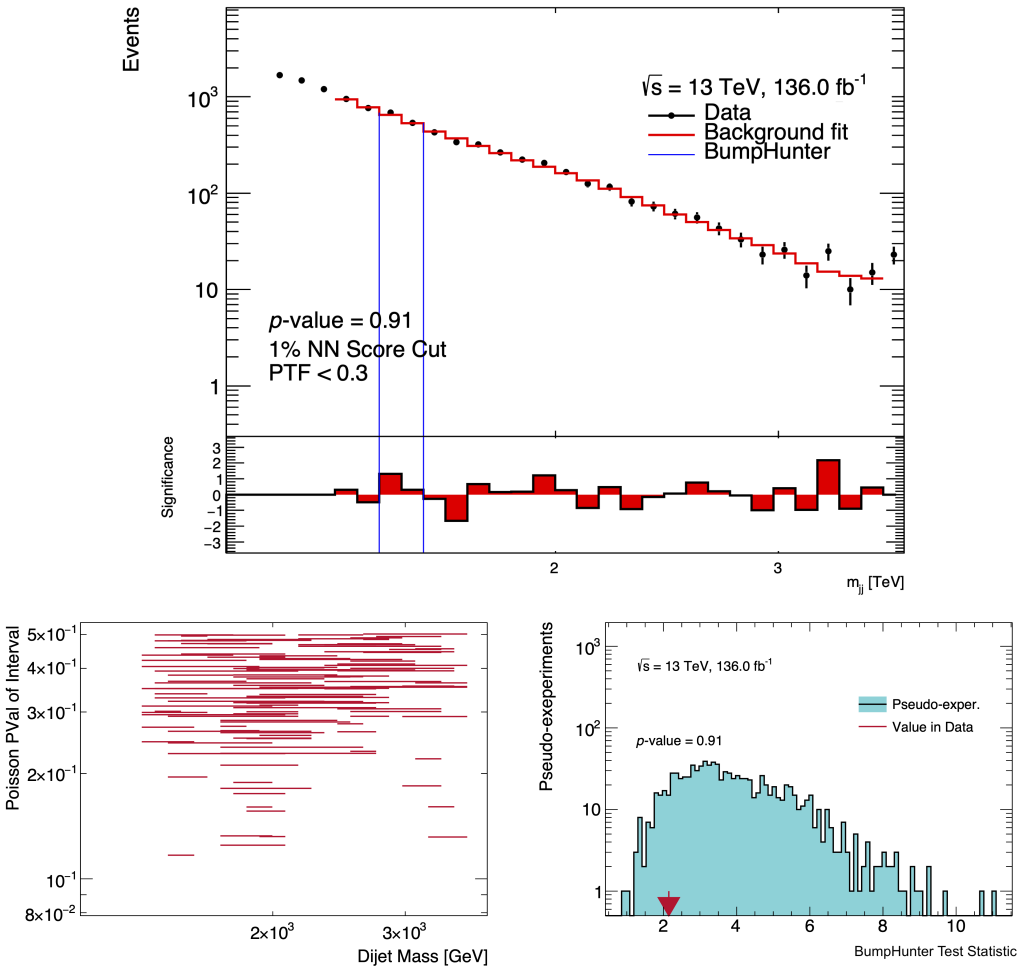


FIGURE 4.40. The three parameter analytical function fit to the spectrum after the 1% NN score cut. The BumpHunter p-value score for the nominal spectrum was 0.918, and after inclusion of systematics, the evaluated p-value was 0.887 ± 0.01. The BumpHunter tomography plot is shown in the bottom plot, indicating the local p-value intervals in the algorithm. The bump of maximum local significance was between 1451 GeV and 1573 GeV with local significance 1.41454. The global fit parameters were [4.3384×10^{-21}, -51.8578, -21.5426] with the SWiFt fit χ^2/ndf=0.483154.

The limits obtained from this analysis are presented for the spectrum with a 10% NN score cut in Figure 4.41 and for the spectrum with a 1% NN score cut in Figure 4.42. The observed limits and expected limits were scaled by the signal efficiency to get the fiducial cross-sections as shown. The signal fiducial

cross-sections were excluded from approximately 20 fb at 1300 GeV down to 2 fb at 3000 GeV by the 10% NN score cut spectrum, and the signal fiducial cross-sections were excluded from 10 fb at 1400 GeV down to 2 fb at 3000 GeV by the 1% NN score cut spectrum.

This analysis was designed to provide model-independent limits on these processes at a fixed width and at fixed lifetimes, to make it possible to evaluate the limits for a variety of models. The black line on the fiducial cross-section limit plots in Figures 4.41 and 4.42 indicate the Z' production cross-sections as a function of the truth Z' mass with a g_q coupling of 0.015. This indicates that Z' production with masses between 1.3 TeV and 3.1 TeV can be excluded for dark sector couplings down to $g_q = 0.015$. The dark sector coupling g_d was set to 1 - g_q in this estimate to ensure that g_d is greater than g_q. The value does not have a significant effect on the branching ratio or on the model curve shown in Figures 4.41 and 4.42 until the coupling to dark hadrons approaches that of the SM. The Z' width was set to the fixed 20 GeV width implemented within the nominal Monte-Carlo production.

With the Z' width set to 20 GeV, the dark quark couplings are 0.2, still at least an order of magnitude above the SM coupling. This gives essentially the same solid curve as in Figures 4.41 and 4.42. Reasonable lifetimes can be found by allowing a wider width of the Z' to allow dark sector couplings on the order of $g_d = 1.0$. In that case, for a 2 TeV Z' with dark sector confinement scales at around 40 GeV, a strong coupling from the SM to the dark sector with strength O(10) produces O(10 mm) lifetimes, within the detector acceptance region. Note that if the Z' is the only interaction contributing to the dark pion lifetime, then the resulting lifetimes could be longer than 10 mm. This effect is fixed by the manual

setting of the Z' width to 20 GeV and the setting of the dark pion decay distance to 1 mm.

The O(10 mm) lifetimes in some of the parameter space considered here require an additional coupling of the dark sector to the standard model, for example from a heavy scalar X_d particle which can induce dark pion decays via a Yukawa coupling [1]. Also, in this work, the effects on the analysis presented here of longer lived dark hadrons with $c\tau > 100$ cm have not been studied. In this case, detailed detector Monte-Carlo studies are needed for the range of possible dark hadron lifetimes and are beyond the scope of this work. In cases where the dark hadrons are so long-lived that they escape the detector, the ATLAS mono-jet and mono-photon searches apply, but would become relevant for Z' particles with masses less than 2 TeV and SM couplings less than 0.25 [77].

O(1.0) couplings in both sectors at the 500 GeV mass scale readily produce dark pion lifetimes at around O(10 mm), so the detailed Run 3 opportunities shown in Chapter V present a good pathway to exploring a larger range of the dark sector parameter space.

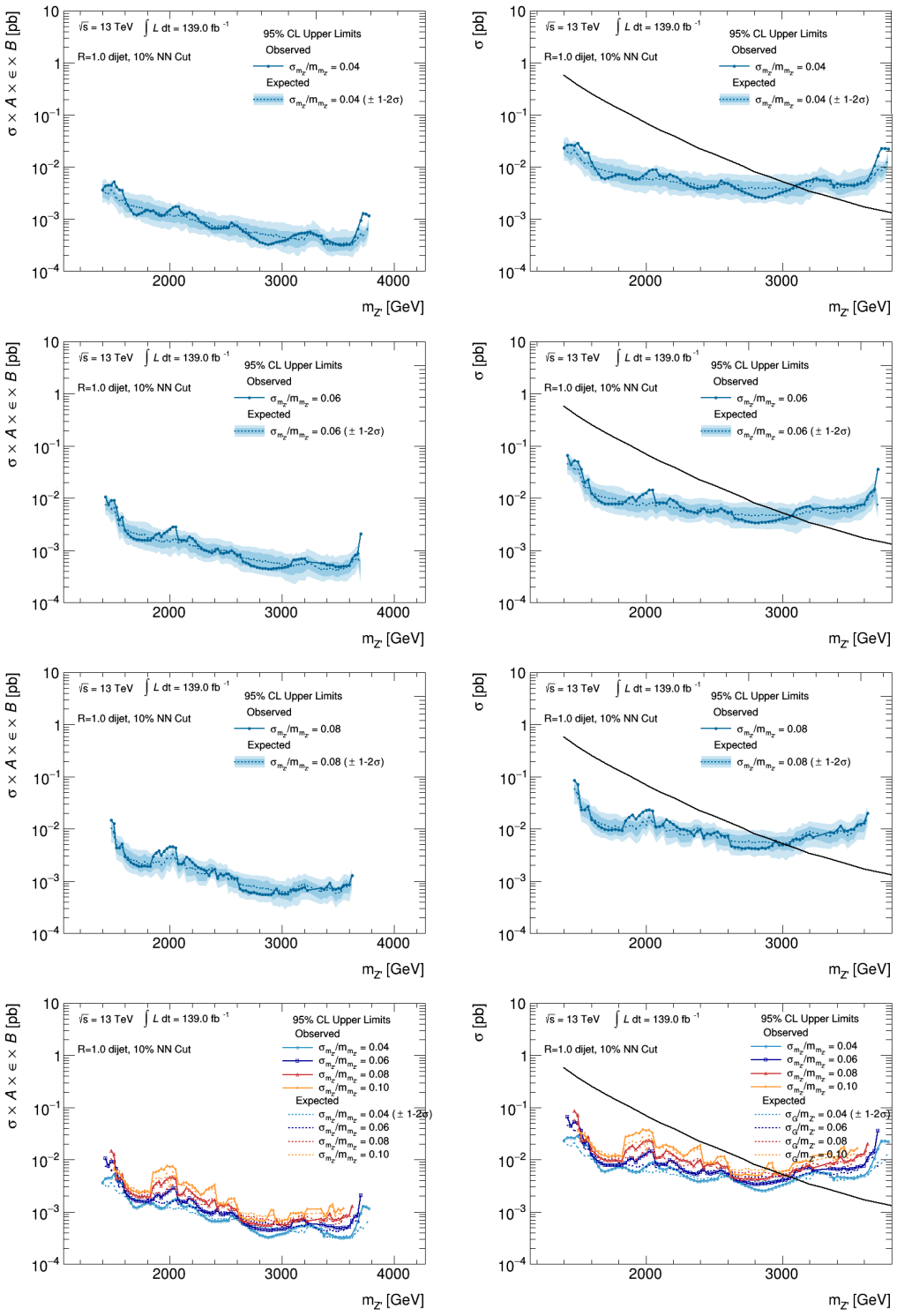


FIGURE 4.41. (left) Gaussian BAT limits on Z' production from the 10% NN score cut spectrum. (right) Fiducial cross-section limit estimates after signal efficiency corrections. 1 and 2 sigma bands on the expected limits are shown. The black line indicates the fiducial cross-section for Z' production with a g_q coupling of 0.015.

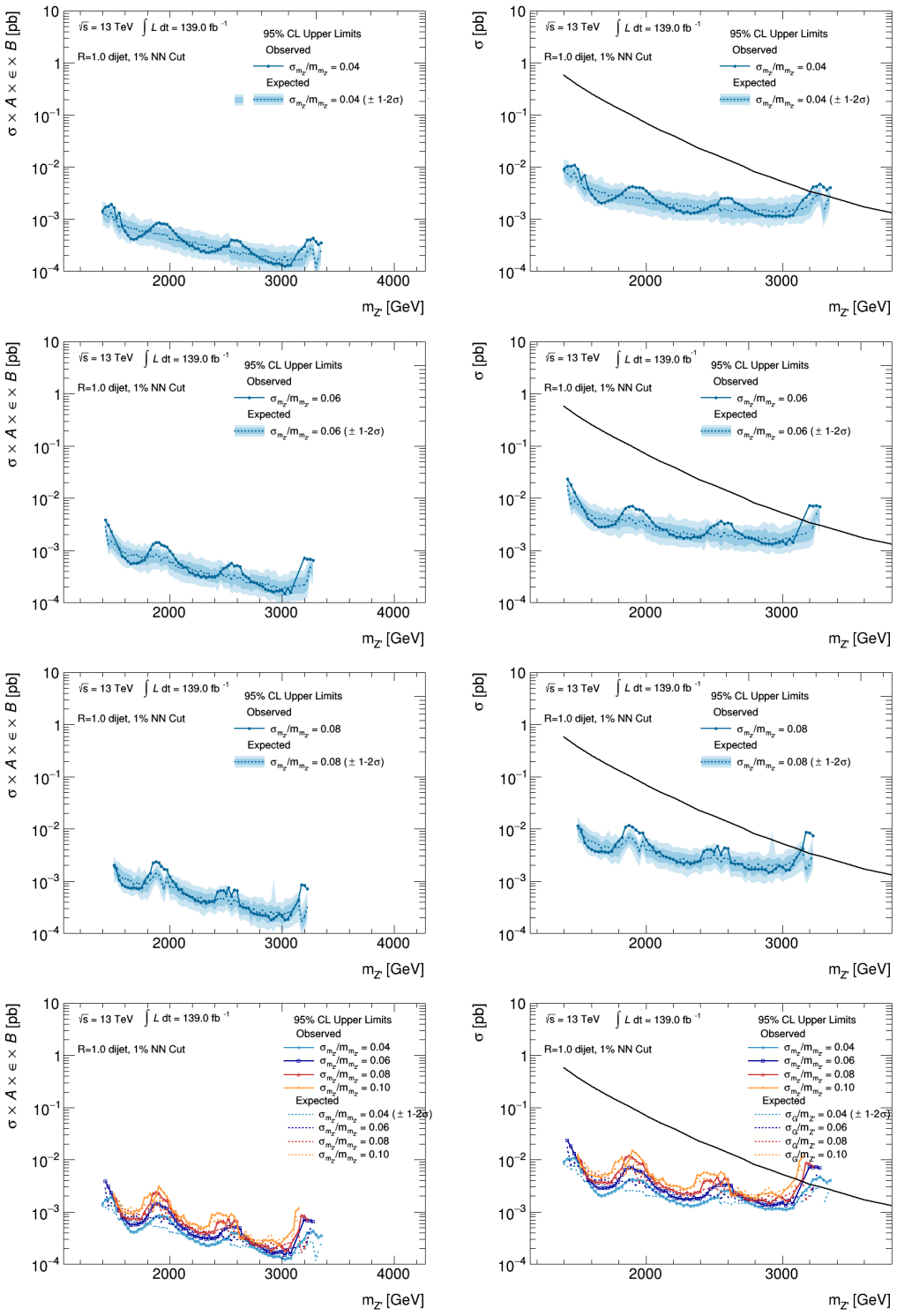


FIGURE 4.42. (left) Gaussian BAT limits on Z' production from the 1% NN score cut spectrum. (right) Fiducial cross-section limit estimates after signal efficiency corrections. 1 and 2 sigma bands on the expected limits are shown. The black line indicates the fiducial cross-section for Z' production with a g_q coupling of 0.015.

The limits obtained by this analysis are very comparable to the limits obtained by the previous ATLAS CWoLa analysis. Although a different signal case was considered, similar data with similar cuts were evaluated, although the original analysis was less inclined to provide additional variables in the models such as the tracking variables used here. This analysis obtained very competitive limits using both the 10% and the 1% NN score spectra; the 95% CL exclusion limits obtained by the original analysis with a 10% cut were on the order of 1 - 20 fb, with significant dependence on the chosen models for the 3 TeV and 5 TeV signals considered. The limits obtained here are all competitive at those mass scales, where they range from 2 - 5 fb at 3 and 4 TeV.

Significant statistical effects were seen in the 1% spectrum by the original CWoLa analysis, but competitive limits exist in the 1 - 5 fb range at 3 and 5 TeV. This analysis achieved emerging jets cross section limits on the order of 2 - 3 fb at that scale, which indicates the strength of this analysis method, especially applied with track-based discriminating variables such as PTF.

CHAPTER V

FUTURE STUDIES

5.1. Run 3 Emerging Jets Triggers

A specialized set of triggers were developed for the purpose of triggering on emerging jets signatures during ATLAS Run 3 operations. These target the dijet and dijet-plus-photon emerging jet signatures that are potentially detectable by triggering on low-PTF events, allowing the potential mass and cross-section limits to be enhanced in future analyses.

Type	Dijet
Chain	HLT_j200_0eta180_emergingPTF0p08dR1p2_a10sd _cssk_pf_jes_ftf_preselj200_L1J100
$\|\eta\|$ Range	< 1.8
PTF Cut	< 0.08
PFlow Threshold	200 GeV
EMTopo Threshold	200 GeV
L1 Seed	L1J100
Rate	8.9 Hz
Type	Dijet+photon
Chain	HLT_g45_tight_icaloloose_2j55_0eta200_emergingPTF0p1dR0p4 _pf_ftf_L1EM22VHI
η Range	< 2.0
PTF Cut	< 0.1
PFlow Threshold	2×55 GeV
EMTopo Threshold	n/a
L1 Seed	L1EM22VHI
Rate	11.4 Hz

TABLE 5.1. HLT-level cuts applied for each emerging jet trigger topology.

Substantial improvements to ATLAS computing has allowed tracking to be run at the HLT level at much more competitive jet p_T scales, allowing for greatly

enhanced trigger efficiencies for signatures that have a substantial track-based component to their signatures [78]. For jet chains at ATLAS, tracking can only be run on events that have a leading jet with p_T greater than 200 GeV, which is is due to the CPU resources available for the software trigger, at the time of writing. This threshold is placed on the EMTopo jets (the standard jets used that utilize ECAL-scaled energy deposits) before tracking is run. The online PTF variable can therefore by computed on the jet Region of Interest after Fast Track Finding (FTF) is run in the trigger, and an HLT-level selection on events with low PTF values can be made at much lower mass scales than were ever possible. Due to the greater precision of the Particle-Flow (PFlow) jet algorithm in reconstructing jet energy, the PTF selection trigger algorithm will be applied to the PFlow jets available above 200 GeV, once the fast tracking has been applied to the event. This chain will be seeded on the unprescaled L1_J100 Level-1 jet trigger, which has an L1 rate of 3.6 kHz at $\mu = 50$.

The efficiencies of these chains with and without tracking resource restrictions are shown in Figure 5.1. Above the 600 GeV mass point, this trigger algorithm has relatively competitive signal efficiencies for several generated emerging jet models and for a baseline BSM Higgs decay. The expected trigger rate for this chain is 8.9 Hz at $\mu = 50$ with a PTF < 0.08 cut.

The unfortunate situation where an EMTopo jet is a preselection before an algorithm running on a PFlow jet will likely introduce some trigger inefficiencies at the lower-mass region, pushing potential mass limits higher than desired. In addition, these EMTopo and PFlow thresholds are close to the turn on of the L1 trigger system, which lies between 200 and 220 GeV, depending on the η selection.

For the dijet+γ signature, a trigger chain was developed that performs the PTF calculation on the two leading jets in the event. In the event that the photon passes the jet p_T cut, the photon may be reconstructed as one of the leading or subleading jets in the event. By applying the PTF algorithm to the two leading objects in the event, this ensures that the trigger is not just selecting events with a leading unconverted photon that mimics a non-prompt jet. The trigger is seeded on the unprescaled EM22VHI L1 trigger that has an expected average L1 rate of 2 kHz in Run 3 conditions. The HLT chain itself is expected to have a rate of 11.4 Hz.

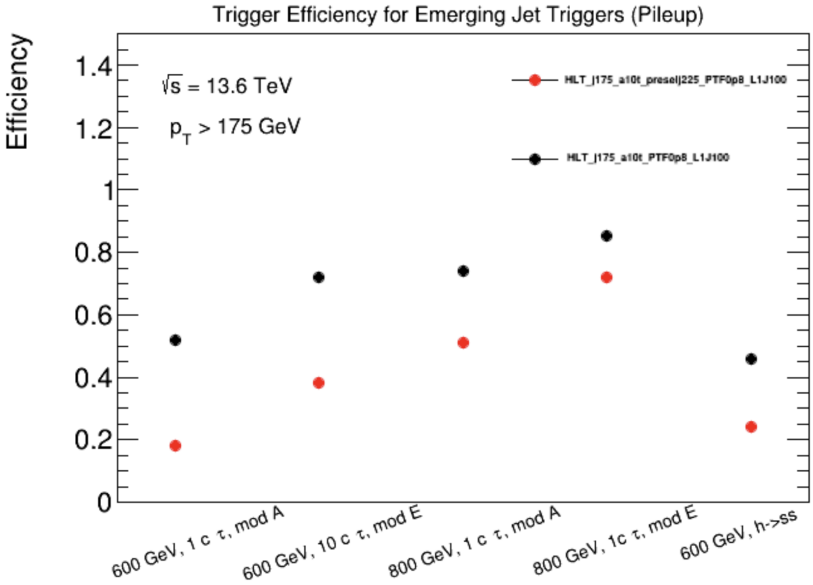


FIGURE 5.1. Emerging jet trigger signal efficiencies with and without jet preselection applied for a j175 seeded trigger with a 225 GeV preselection. Efficiencies are computed for baseline low-mass Z' models and one baseline H→ss sample (the decay of a 600 GeV BSM Higgs into two 150 GeV scalars with a $c\tau = 3.3$ m.).

These triggers are intended to probe two general mass ranges. The single-jet trigger was designed to target invariant mass ranges above 500 - 600 GeV, and the dijet+photon trigger was designed to target mass ranges down to 200 - 250

GeV. The actual efficiency points will need to be determined after data-taking has commenced, as sufficient luminosity will need to be taken by low-p_T performance triggers before these can be calculated.

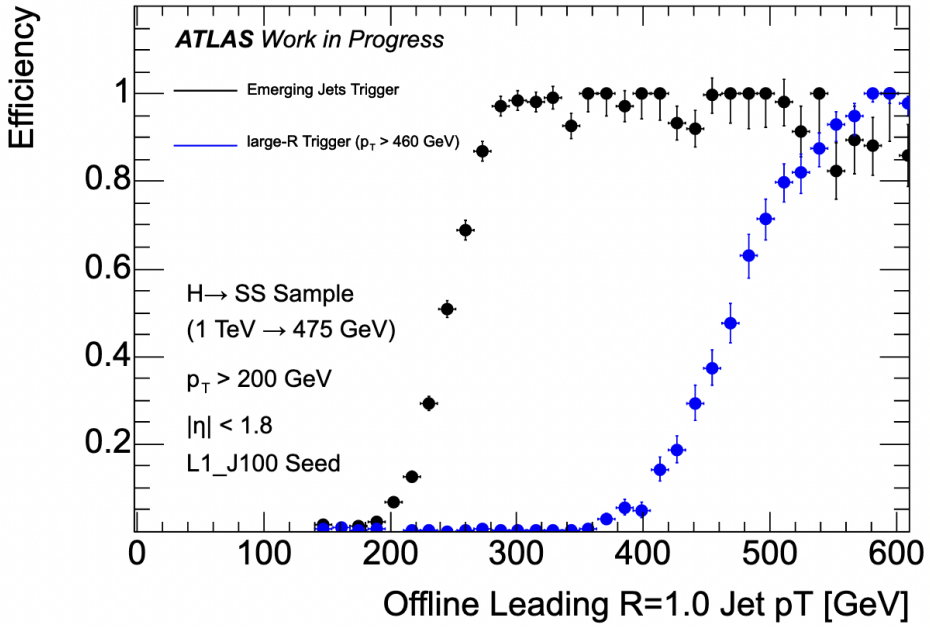


FIGURE 5.2. Trigger efficiency plots for the 200 GeV seeded emerging jets trigger chain determined via the trigger simulated in BSM Higgs to LLP scalar (1 TeV $\rightarrow$ 475 GeV) signal events. These indicate that the trigger will be fully efficient in p_T at approximately 250 GeV, as compared to over 500 GeV for a standard unprescaled large-R jet trigger. The trigger curve is scaled to the fraction that pass the PTF cut to allow a plateau at 1.0.

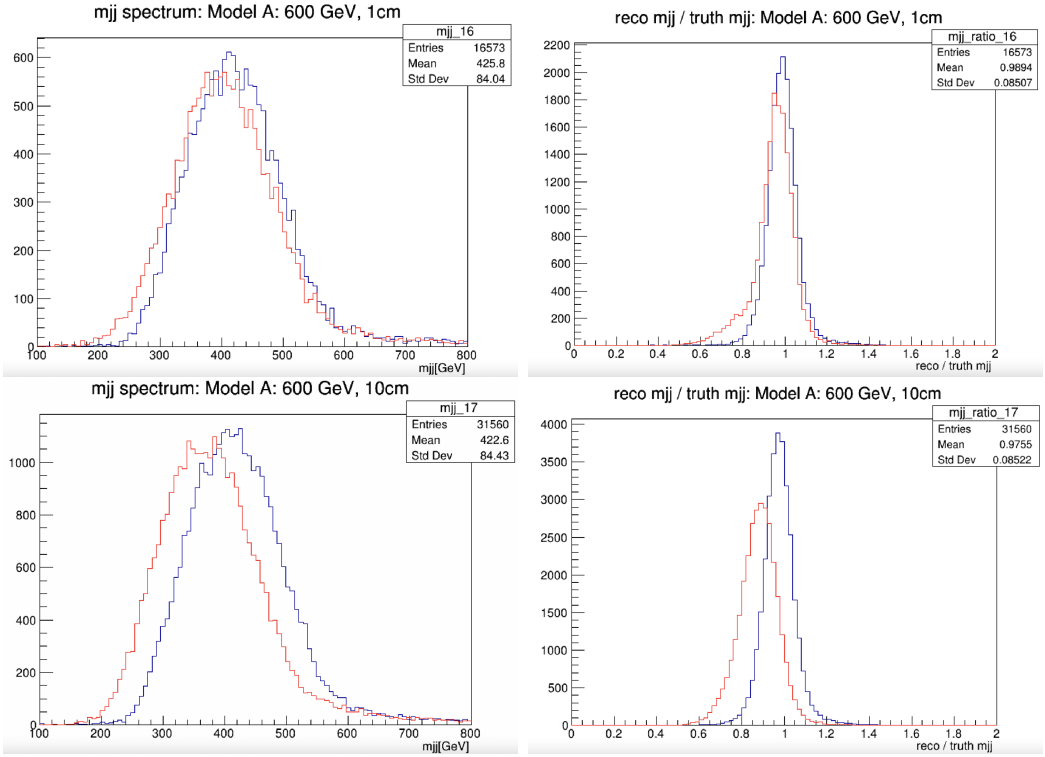


FIGURE 5.3. Invariant mass distributions and invariant mass resolution for 600 GeV Z' samples with lifetimes 1 mm (top) and 10 mm (bottom). The PFlow algorithm has a definite lifetime dependence that affects the number of reconstructed tracks associated to the jet, making the non-prompt jet appear like a photon to the algorithm at sufficient lifetimes.

An additional consideration for the single-jet trigger is from the implementation of new Level-1 hardware trigger systems at ATLAS. The global Feature EXtractor (gFEX) is a hardware trigger that is designed to provide an event-wise large-R jet trigger system to the detector. Part of the Phase-I hardware upgrade, It is composed of three FPGA systems that input coarse 0.2×0.2 calorimeter regions and compute $R = 1.0$ Level-1 jets in each FPGA's corresponding detector region. The overall event is reconstructed from these three regions in an additional FPGA and global event decisions are made at the hardware level. Dark jet signatures such as emerging jets are expected to be

very efficiently selected at Level-1 due to this system, particularly due to the flat pileup/noise energy subtraction applied in the reconstruction, which enhances sensitivity to soft and wide energy deposits. The development and evaluation of these emerging jet triggers with the use gFEX was not possible due to the lack of a gFEX trigger simulation in Athena at the time of the development. However, this system is likely a future promising candidate as a hardware seed for the developed emerging jet triggers.

CHAPTER VI

CONCLUSION

This analysis placed competitive cross-section limits on emerging jets production from exotic dark sector Z's. These limits were evaluated down to 10 fb at masses of 1300 GeV and to 2 fb at 3000 GeV. Production of Z' particles with masses up to 3.1 TeV were excluded for g_q dark sector couplings down to 0.015. A novel machine-learning technique using unsupervised learning was utilized to enhance potential signal sensitivities by over 10× relative to cut-based analyses, and this method is shown to be a powerful tool for the probe of these dark jet signal types.

Despite the lack of discovery of emerging signals in the dijet topology at invariant masses greater than 1.3 TeV, exciting physics opportunities for discovery are available in much lower mass and cross-section regimes. The computing and detector upgrades at ATLAS will open a realm of track-jet exotic particle triggers for discovery. Dark jets, LLPs, and boosted soft exotic decays will all be enhanced in discovery potential through the use of specialized jet-track triggers. The discovery potential and limits of emerging jets, in particular, will be greatly enhanced in Run 3 with the utilization of the specialized emerging jets triggers in the dijet and dijet+photon channels. Mass sensitivity down to 550 GeV are expected for the single-jet emerging jets trigger, and the dijet+photon triggered data are expected to allow good sensitivity for Z' bosons down to 250 GeV.

These new trigger chains have also been proven to be very efficient for the selection of LLP signatures that include BSM Higgs decays to LLP heavy scalars and long-lived R-Hadrons. A future analysis will require a dedicated analysis of

the scope of the data with respect to the correlated trigger inefficiencies due to the overlap of the L1, EMTopo HLT jet, and PFlow HLT jet thresholds, but the extremely competitive mass limits allow for a significant probe of a range of model spaces.

APPENDIX

APPENDIX1

A.1. Appendix A

Scaling factors were applied to the data for NN training and evaluation. These were applied to reweight the signal to the appropriate injection cross-section, and they were also applied to weight events in the sidebands to be identical to those in the signal region bins.

Before training, the m_0, m_1, PTF_0, PTF_1, $\Delta\phi$, y^*, and number-of-jets distributions were scaled by the respective factors: 150 GeV, 150 GeV, 0.15, 0.15, 1.5, 0.3, 5. The corresponding event features were subtracted by these means and divided by these means to make an approximate scaling between -1 and 1.

The chosen binning was based on resolution bins found by previous dijet searches such as the high-mass dijet search and the ATLAS trigger-level analysis: 1049, 1093, 1139, 1186, 1235, 1286, 1339, 1394, 1451, 1511, 1573, 1637, 1704, 1773, 1845, 1920, 1998, 2079, 2163, 2251, 2342, 2437, 2536, 2639, 2746, 2857, 2973, 3094, 3220, 3351, 3487, 3629, 3776, 3929, 4088, 4254, 4427, 4607, 4794, 4989, 5191 GeV

A.1.1. Trigger Studies

To optimize the PTF definition, its selection on R-Hadron samples was studied. R-Hadrons are the supersymmetric equivalent to glue-balls; they are neutral balls of gluinos constrained by the model of choice. Two R-Hadron samples were considered, one with lifetime 10 ns and one with 0.1 ns. These R-Hadron samples were configured with a small mass difference between the

R-Hadron pairs produced and the Neutralinos in the event. Neutralinos are neutral supersymmetric (SUSY) particles that are predicted to be the lightest supersymmetric particle (LSP) in SUSY theory, therefore they are the canonical dark matter candidate. The small mass splitting between these samples produces the long-lived properties of these R-Hadrons. The resulting low-p_T byproducts of this decay will be displaced and well-selected by the PTF algorithm.

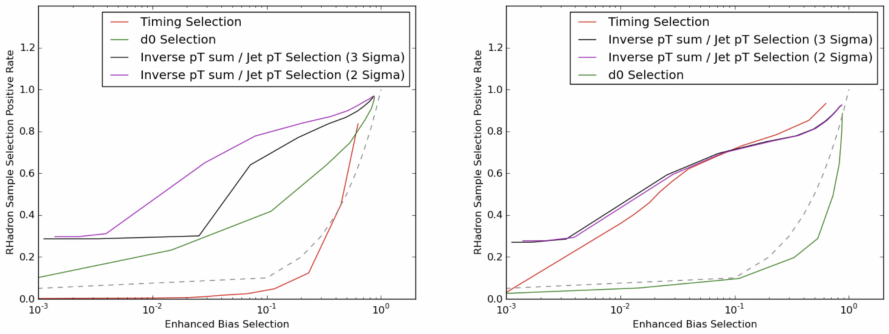


FIGURE A.1. Various jet and jet-track selections on R-Hadron samples with 0.1 ns (left) and 10 ns (right) lifetimes. The timing selection refers to a selection on the offline jet timing, the d_0 selection refers to a selection with a cut on the highest d_0 track associated to the jet, and the inverse p_T cut is the PTF algorithm applied with varying cofactors in front of the resolution function.

The emerging jets trigger is therefore expected to have good sensitivity to these signal types.

A.1.2. GRLs and Configurations

GRLs used in data selection:

`data15_13TeV.periodAllYear_HEAD_`

`DQDefects-00-02-02_PHYS_StandardGRL_All_Good_25ns`

`data16_13TeV.periodAllYear_DetStatus-v89-pro21-01_DQDefects-`

`00-02-04_PHYS_StandardGRL_All_Good_25ns`

data17_13TeV.periodAllYear_HEAD_

Unknown_PHYS_StandardGRL_All_Good_25ns_ Triggerno17e33prim

data18_13TeV.periodAllYear_HEAD_Unknown_

PHYS_StandardGRL_All_Good_25ns_ Triggerno17e33prim

A.1.3. Pythia Model Generation

The software tags used for the generation are the following: Event Generation: e8381 Event Simulation: s3126 Event Reconstruction: r9364 (mc16a) r10201 (mc16d) r10724 (mc16e) Derivation Production: p4432

The overlaid pileup samples were: mc16_13TeV.361239.Pythia8EvtGen_ A3NNPDF23LO_minbias_inelastic_high.simul.HITS.e4981_s3087_s3111

mc16_13TeV.361238.Pythia8EvtGen_A3NNPDF23LO

_minbias_inelastic_low.simul.HITS.e4981_s3087_s3111

The following is the configuration used within Athena for the event generation.

```
##############################################################
# Emerging Jets Event Generation
# Pythia 8: Zd --> Qd Qd_bar --> 2EJ
# contact: Aaron Kilgallon (aaron.joseph.kilgallon@cern.ch)
#============================================================


evgenConfig.description = "emerging jets from pair-produced dark quarks"
evgenConfig.keywords = ["exotic", "hiddenValley", "4jet"]
evgenConfig.process = "p p --> Zd --> Qd Qd_bar --> 2EJ"
evgenConfig.contact = ["aaron.joseph.kilgallon@cern.ch"]
```

```
include("Pythia8_i/Pythia8_A14_NNPDF23LO_EvtGen_Common.py")



# set sample / model parameters automatically based on jo name
print("ARGS: ", runArgs.jobConfig[0])
print("JO ARGS: ", jofile.rstrip('.py').split('_'))

m_Xd = float(jofile.rstrip('.py').split('_')[3])
ctau_pi_d = float(jofile.rstrip('.py').split('_')[4])

print("SCALAR MEDIATOR MASS: %f " % m_Xd)
print("DARK PION LIFETIME: %f " % ctau_pi_d)

mod = jofile.rstrip('.py').split('_')[2]
print("MODEL: %s " % mod)
if mod == "ModelA":
    m_pi_d = 5.0
elif mod == "ModelB":
    m_pi_d = 2.0
elif mod == "ModelC":
    m_pi_d = 10.0
elif mod == "ModelD":
    m_pi_d = 20.0
elif mod == "ModelE":
    m_pi_d = 0.8
print("DARK PION MASS: %f " % m_pi_d)
print("DARK RHO MASS: %f " % (m_pi_d*4))
```

```
print("LAMBDA / DARK QUARK MASS: %f " % (m_pi_d*2))
print("PT MIN FSR: %f " % (m_pi_d*2*1.1))

# show 5 events for testing
genSeq.Pythia8.Commands += ["Next:numberShowEvent = 1"]

## OVERRIDE STANDARD ATLAS TAU0 LIMIT ##
genSeq.Pythia8.Commands += ["ParticleDecays:limitTau0 = off"]

# settings for dark sector
genSeq.Pythia8.Commands += ["HiddenValley:spinFV = 0",
                            "HiddenValley:Ngauge = 3",
                            # n dark QCD colors
                            "HiddenValley:alphaFSR = 0.7"]
                            # dark coupling

# Model settings
genSeq.Pythia8.Commands += ["4900101:m0 = " + str(m_pi_d*2),
                            # qd mass
                            "4900111:m0 = " + str(m_pi_d),
                            # pi_d mass
                            "4900113:m0 = " + str(m_pi_d*4),
                            # rho_d mass
                            "4900211:m0 = " + str(m_pi_d),
                            # pi_d off-diag
                            "4900213:m0 = " + str(m_pi_d*4),
                            # rhod offdiag
```

```
                            "HiddenValley:Lambda = " + str(m_pi_d*2),
                            "HiddenValley:pTminFSR = "
                            + str(m_pi_d*2*1.1)]
                            # pT cutoff for dark shower

# dark pion lifetime
genSeq.Pythia8.Commands += ["4900111:tau0 = " + str(ctau_pi_d)]
# pi_d lifetime -- variable
# off-diagonal dark pion lifetime
genSeq.Pythia8.Commands += ["4900211:tau0 = " + str(ctau_pi_d)]

# non-model dependent settings
genSeq.Pythia8.Commands += ["PartonLevel:MPI = on",
                            "PartonLevel:ISR = on"]

# emerging jet event processes
genSeq.Pythia8.Commands += ["HiddenValley:ffbar2Zv = on"]
genSeq.Pythia8.Commands += ["4900023:m0 = " + str(m_Xd)]
genSeq.Pythia8.Commands += ["4900023:tau0 = 1E-20",
                            "4900023:mWidth=10",
                            "4900023:isResonance = on",
                            "4900023:mayDecay = on",
                            "4900023:0:bRatio = 1",
                            "4900023:0:meMode = 102"]

genSeq.Pythia8.Commands += ["4900023:onMode = off",
                            "4900023:offIfAny 1 2 3 4 5 6 -1 -2 -3 -4 -5
```

```
                            -6 7 8 9 10 11 12 13 14 15 16 -7 -8 -9 -10
                             -11 -12 -13 -14 -15 -16",
                             "4900023:onIfAny 4900101 -4900101",
                             "4900023:oneChannel 1 0.999 102 4900101
                             -4900101",
                             "4900023:addChannel 1 0.001 102 1 -1"]

# dark meson decays
genSeq.Pythia8.Commands += ["4900111:0:all on 1.0 102 1 -1",
                             \newline # dark pion to down quarks
                             "4900113:0:all on 0.999 102 4900111 4900111",
                             # dark vector to dark pions 99.9%
                             "4900113:addchannel on 0.001 102 1 -1"]
                             # dark vector to down quarks 0.1%
# dark meson off-diagonal decays
genSeq.Pythia8.Commands += ["4900211:oneChannel on 1.0 91 1 -1",
                             # dark pion to down quarks
                             "4900213:oneChannel on 0.999 102 4900211
                             4900211",
                             # dark vector to dark pions 99.9%
                             "4900213:addchannel on 0.001 102 1 -1"]
                             # dark vector to down quarks 0.1%



# dark QCD coupling (alphaHV) running
genSeq.Pythia8.Commands += ["HiddenValley:alphaOrder = 1",
                             "HiddenValley:nFlav = 7"]
```

```
# HV parton shower settings
genSeq.Pythia8.Commands += ["HiddenValley:FSR = on",
                            "HiddenValley:fragment = on"]


genSeq.Pythia8.Commands += ["Main:timesAllowErrors = 500"]


genSeq.Pythia8.Commands += ["ProcessLevel:all = on",
"ProcessLevel:resonanceDecays = on",
"PartonLevel:all = on",
"PartonLevel:ISR = on",
"HadronLevel:all = on",
"PhaseSpace:useBreitWigners = on"]


# workarounds for TestHepMC
testSeq.TestHepMC.MaxVtxDisp=5000000.
testSeq.TestHepMC.MaxTransVtxDisp = 5000000.


## JET FILTERING ##
include ("GeneratorFilters/FindJets.py")
CreateJets(prefiltSeq, 0.4)

if not hasattr( filtSeq, "TruthJetFilter" ):
```

```
    from GeneratorFilters.GeneratorFiltersConf import TruthJetFilter
    filtSeq += TruthJetFilter()
    pass

filtSeq.TruthJetFilter.TruthJetContainer = "AntiKt4TruthJets"
filtSeq.TruthJetFilter.Njet = 2
filtSeq.TruthJetFilter.NjetMinPt = 125*GeV
filtSeq.TruthJetFilter.NjetMaxEta = 2.4
```

A.2. Appendix B

The following are the full set of signal sample efficiencies.

Z'_m [GeV]	$c\tau$ [mm]	Model	DAOD	Analysis	Eff. 10% NN	Eff. 1% NN
1500.0	1.0	A	96.4	29.2	23.6	8.0
1500.0	1.0	B	96.1	23.1	26.2	15.0
1500.0	1.0	C	96.2	35.1	21.6	2.4
1500.0	1.0	D	96.3	35.6	15.3	0.6
1500.0	1.0	E	96.4	25.1	36.1	19.3
1500.0	2.0	A	96.3	32.5	34.9	16.6
1500.0	2.0	B	96.1	26.9	40.4	21.9
1500.0	2.0	C	96.4	38.9	33.6	16.7
1500.0	2.0	D	96.3	39.7	32.6	6.0
1500.0	2.0	E	96.2	29.2	36.0	25.1
1500.0	5.0	A	96.3	34.4	43.2	27.0
1500.0	5.0	B	96.3	28.6	45.9	27.5
1500.0	5.0	C	96.5	40.0	41.2	22.8
1500.0	5.0	D	96.3	42.9	42.0	9.4
1500.0	5.0	E	96.3	29.5	29.1	16.4
1500.0	10.0	A	96.1	32.8	46.2	27.4
1500.0	10.0	B	95.9	26.4	33.1	23.6
1500.0	10.0	C	96.1	39.7	39.0	26.2
1500.0	10.0	D	96.1	43.8	50.1	24.5
1500.0	10.0	E	95.6	25.7	29.4	3.8
1500.0	25.0	A	95.8	29.2	39.3	21.6
1500.0	25.0	B	95.2	19.6	14.1	1.1
1500.0	25.0	C	95.9	37.8	48.2	26.0
1500.0	25.0	D	95.3	44.2	46.6	20.0
1500.0	25.0	E	95.9	16.4	5.9	1.2
1500.0	50.0	A	94.3	22.5	20.9	4.8
1500.0	50.0	B	93.9	12.5	4.9	0.3
1500.0	50.0	C	93.9	34.5	48.9	27.2
1500.0	50.0	D	93.6	43.1	55.1	30.7
1500.0	50.0	E	95.4	7.3	2.9	0.0
1500.0	100.0	A	91.9	13.3	4.5	0.8
1500.0	100.0	B	92.8	4.9	8.3	1.7
1500.0	100.0	C	90.9	25.6	19.3	5.7
1500.0	100.0	D	88.9	38.2	43.6	13.7
1500.0	100.0	E	94.7	1.5	16.7	0.0

TABLE A.1. Signal efficiencies for 1.5 TeV samples.

Z'_m [GeV]	$c\tau$ [mm]	Model	DAOD	Analysis	Eff. 10% NN	Eff. 1% NN
2000.0	1.0	A	95.9	31.3	35.9	12.3
2000.0	1.0	B	96.2	26.7	36.3	16.8
2000.0	1.0	C	96.3	35.3	33.7	18.3
2000.0	1.0	D	96.0	35.7	30.7	11.3
2000.0	1.0	E	96.0	28.0	48.3	24.3
2000.0	2.0	A	96.0	34.0	44.5	22.7
2000.0	2.0	B	96.3	30.4	45.9	23.8
2000.0	2.0	C	96.1	38.9	44.7	21.9
2000.0	2.0	D	95.9	39.8	42.9	19.2
2000.0	2.0	E	95.8	31.2	48.1	24.1
2000.0	5.0	A	95.5	35.2	48.1	23.7
2000.0	5.0	B	96.1	31.0	44.8	23.2
2000.0	5.0	C	95.9	40.6	55.3	22.4
2000.0	5.0	D	95.8	42.2	48.9	23.1
2000.0	5.0	E	95.9	31.0	37.8	14.3
2000.0	10.0	A	95.8	33.9	53.3	26.2
2000.0	10.0	B	95.7	28.1	35.8	23.0
2000.0	10.0	C	96.2	40.2	51.7	21.3
2000.0	10.0	D	95.9	42.9	51.1	23.2
2000.0	10.0	E	95.5	27.9	31.3	7.0
2000.0	25.0	A	95.0	30.1	44.1	16.2
2000.0	25.0	B	94.9	21.8	15.7	2.3
2000.0	25.0	C	95.4	38.3	50.3	23.0
2000.0	25.0	D	95.3	42.8	45.2	28.0
2000.0	25.0	E	95.1	20.1	12.7	1.4
2000.0	50.0	A	92.4	25.3	17.0	2.8
2000.0	50.0	B	94.5	16.3	2.3	0.3
2000.0	50.0	C	94.1	34.9	38.2	13.5
2000.0	50.0	D	94.4	41.9	49.4	26.9
2000.0	50.0	E	95.0	12.6	2.9	0.4
2000.0	100.0	A	91.7	17.3	11.3	2.5
2000.0	100.0	B	92.8	8.4	6.5	0.8
2000.0	100.0	C	90.1	28.4	17.7	3.0
2000.0	100.0	D	89.5	38.3	36.3	21.4
2000.0	100.0	E	93.3	3.9	12.7	1.8

TABLE A.2. Signal efficiencies for 2 TeV samples.

Z'_m [GeV]	$c\tau$ [mm]	Model	DAOD	Analysis	Eff. 10% NN	Eff. 1% NN
3000.0	1.0	A	96.4	29.2	38.0	7.6
3000.0	1.0	B	95.4	25.2	45.7	10.7
3000.0	1.0	C	95.7	29.7	41.1	7.7
3000.0	1.0	D	95.7	29.5	44.4	8.2
3000.0	1.0	E	95.6	26.8	48.4	7.2
3000.0	2.0	A	95.9	31.0	43.0	6.3
3000.0	2.0	B	95.3	27.3	45.8	7.4
3000.0	2.0	C	95.7	32.1	45.6	7.5
3000.0	2.0	D	95.8	31.8	46.1	7.2
3000.0	5.0	A	95.8	31.0	48.6	8.3
3000.0	5.0	B	95.2	26.7	42.9	7.0
3000.0	5.0	C	96.0	33.1	46.0	6.8
3000.0	5.0	D	95.9	33.6	47.3	7.5
3000.0	5.0	E	95.8	27.3	38.9	6.7
3000.0	10.0	A	95.7	31.4	47.5	14.9
3000.0	10.0	B	95.1	25.2	39.8	5.3
3000.0	10.0	C	96.1	32.7	51.8	9.8
3000.0	10.0	D	95.8	33.7	49.9	7.5
3000.0	25.0	A	94.1	28.3	30.4	8.1
3000.0	25.0	B	94.3	22.4	22.2	5.9
3000.0	25.0	C	95.5	31.7	39.9	11.9
3000.0	25.0	D	94.7	34.1	50.6	7.3
3000.0	25.0	E	95.4	21.0	23.2	4.9
3000.0	50.0	A	93.1	24.7	18.7	4.7
3000.0	50.0	B	93.3	18.5	14.0	5.0
3000.0	50.0	C	93.1	31.0	34.6	10.8
3000.0	50.0	D	93.6	33.3	41.7	8.6
3000.0	50.0	E	95.0	16.1	7.3	1.6
3000.0	100.0	A	90.8	20.7	9.7	1.5
3000.0	100.0	B	91.5	12.9	6.6	1.2
3000.0	100.0	C	89.8	27.8	17.3	1.6
3000.0	100.0	D	88.9	31.7	29.5	5.4
3000.0	100.0	E	94.0	8.0	11.0	1.3

TABLE A.3. Signal efficiencies for 3 TeV samples.

Z'_m [GeV]	$c\tau$ [mm]	Model	DAOD	Analysis	Eff. 10% NN	Eff. 1% NN
4000.0	1.0	A	95.8	21.5	29.2	4.6
4000.0	1.0	B	95.9	21.1	32.8	5.9
4000.0	1.0	C	96.0	22.3	31.7	5.9
4000.0	1.0	D	95.8	23.5	33.2	6.0
4000.0	1.0	E	95.9	22.0	32.4	6.0
4000.0	2.0	A	95.6	22.9	27.6	4.2
4000.0	2.0	B	96.0	22.2	35.0	5.3
4000.0	2.0	C	96.6	24.3	37.6	6.1
4000.0	2.0	D	96.2	24.7	37.1	6.2
4000.0	2.0	E	96.0	22.6	34.9	6.0
4000.0	5.0	A	95.7	22.9	34.4	6.2
4000.0	5.0	B	96.0	21.9	35.7	7.1
4000.0	5.0	C	96.2	24.7	27.6	3.6
4000.0	5.0	D	95.9	26.1	30.7	4.2
4000.0	5.0	E	95.8	22.6	29.8	5.9
4000.0	10.0	A	95.7	21.8	40.0	9.3
4000.0	10.0	B	95.7	20.5	39.6	12.7
4000.0	10.0	C	96.3	24.3	28.8	3.6
4000.0	10.0	D	96.0	26.2	23.7	3.2
4000.0	10.0	E	95.4	21.7	35.1	7.2
4000.0	25.0	A	94.7	20.9	32.3	8.0
4000.0	25.0	B	95.0	18.9	27.7	6.3
4000.0	25.0	C	95.5	23.6	33.7	7.2
4000.0	25.0	D	95.4	26.0	29.6	4.6
4000.0	25.0	E	95.3	19.1	20.4	4.9
4000.0	50.0	A	93.4	19.8	35.0	11.1
4000.0	50.0	B	93.6	17.3	16.8	4.4
4000.0	50.0	C	93.8	23.2	27.9	6.0
4000.0	50.0	D	93.8	26.4	29.7	5.5
4000.0	50.0	E	94.4	16.7	11.0	0.9
4000.0	100.0	A	91.1	17.6	11.8	2.0
4000.0	100.0	B	92.5	13.2	14.5	1.5
4000.0	100.0	C	90.7	21.8	26.9	5.1
4000.0	100.0	D	88.4	26.0	25.3	6.9
4000.0	100.0	E	94.1	10.4	8.8	0.5

TABLE A.4. Signal efficiencies for 4 TeV samples.

REFERENCES CITED

[1] P. Schwaller, D. Stolarski, and A. Weiler, "Emerging jets," Journal of High Energy Physics, vol. 2015, May 2015.

[2] H. Murayama and K. Riesselmann, "Doe explains...the standard model of particle physics."

[3] M. Thomson, Modern Particle Physics. Modern Particle Physics, Cambridge University Press, 2013.

[4] E. Mobs, "The CERN accelerator complex. Complexe des accelerateurs du CERN," Jul 2016. General Photo.

[5] The ATLAS Collaboration, "Atlas experiment public results," Jul 2015.

[6] The CMS Collaboration, "The atlas experiment at the cern large hadron collider," Journal of Instrumentation, Aug 2008.

[7] A. Salzburger, "The ATLAS Track Extrapolation Package," tech. rep., CERN, Geneva, Jun 2007. All figures including auxiliary figures are available at https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PUBNOTES/ATL-SOFT-PUB-2007-005.

[8] The ATLAS Collaboration, "Operation of the atlas trigger system in run 2," Journal of Instrumentation, vol. 15, no. 10, 2020.

[9] D. Krohn, J. Thaler, and L.-T. Wang, "Jet trimming," Journal of High Energy Physics, vol. 2010, feb 2010.

[10] M. Cacciari, G. P. Salam, and G. Soyez, "The anti-kt jet clustering algorithm," Journal of High Energy Physics, vol. 2008, p. 063–063, Apr 2008.

[11] The ATLAS Collaboration, "Jet energy scale measurements and their systematic uncertainties in proton-proton collisions at 13 tev with the atlas detector," Physical Review D, vol. 96, Oct 2017.

[12] The ATLAS Collaboration, "In-situ calibration of large-radius jet energy and mass in 13 TeV proton–proton collisions with the ATLAS detector," The European Physical Journal C, vol. 79, feb 2019.

[13] The ATLAS Collaboration, "Atlas stand-alone event displays," EventDisplayStandAlone; AtlasPublicResults;.

[14] B. W. Allen, "Search for New Phenomena in Dijet Events using 13 TeV Proton-Proton Collision Data Collected With the ATLAS Detector," Jun 2018. Presented 30 May 2018.

[15] Weinberg, S., "The making of the standard model," Eur. Phys. J. C, vol. 34, no. 1, pp. 5–13, 2004.

[16] "Measurement of the positive muon anomalous magnetic moment to 0.46 ppm," Phys. Rev. Lett., vol. 126, p. 141801, Apr 2021.

[17] "High-precision measurement of the w boson mass with the cdf ii detector," Science, vol. 376, no. 6589, pp. 170–176, 2022.

[18] The Planck Collaboration, "Planck 2018 results," Astronomy and Astrophysics, vol. 641, p. A6, Sep 2020.

[19] K. Garrett and G. Duda, "Dark matter: A primer," Advances in Astronomy, vol. 2011, pp. 1–22, 2011.

[20] Y. Bai and P. Schwaller, "Scale of dark QCD," Physical Review D, vol. 89, Mar 2014.

[21] G. Frattari, "Dark Matter searches with the ATLAS Detector," tech. rep., CERN, Geneva, Jul 2021.

[22] CMS Collaboration, "Search for resonant production of strongly-coupled dark matter in proton-proton collisions at 13 TeV," tech. rep., CERN, Geneva, 2021.

[23] T. Clifton, P. G. Ferreira, A. Padilla, and C. Skordis, "Modified gravity and cosmology," Physics Reports, vol. 513, no. 1, pp. 1–189, 2012. Modified Gravity and Cosmology.

[24] D. H. Weinberg, M. J. Mortonson, D. J. Eisenstein, C. Hirata, A. G. Riess, and E. Rozo, "Observational probes of cosmic acceleration," Physics Reports, vol. 530, pp. 87–255, sep 2013.

[25] M. Rocha, A. H. Peter, J. S. Bullock, M. Kaplinghat, S. Garrison-Kimmel, J. Oñorbe, and L. A. Moustakas, "Cosmological simulations with self-interacting dark matter – i. constant-density cores and substructure," Monthly Notices of the Royal Astronomical Society, vol. 430, no. 1, p. 81–104, 2013.

[26] W. J. G. de Blok, "The core-cusp problem," Advances in Astronomy, vol. 2010, pp. 1–14, 2010.

[27] M. Mateo, "DWARF GALAXIES OF THE LOCAL GROUP," Annual Review of Astronomy and Astrophysics, vol. 36, pp. 435–506, sep 1998.

[28] T. R. Slatyer and C.-L. Wu, "Early-universe constraints on dark matter-baryon scattering and their implications for a global 21 cm signal," Physical Review D, vol. 98, jul 2018.

[29] J. Goldstone, A. Salam, and S. Weinberg, "Broken symmetries," Phys. Rev., vol. 127, pp. 965–970, Aug 1962.

[30] S. L. Glashow, "Partial-symmetries of weak interactions," Nuclear Physics, vol. 22, no. 4, pp. 579–588, 1961.

[31] H. Fritzsch, M. Gell-Mann, and H. Leutwyler, "Advantages of the Color Octet Gluon Picture," Phys. Lett. B, vol. 47, pp. 365–368, 1973.

[32] F. Englert and R. Brout, "Broken symmetry and the mass of gauge vector mesons," Phys. Rev. Lett., vol. 13, pp. 321–323, Aug 1964.

[33] P. W. Higgs, "Broken symmetries and the masses of gauge bosons," Phys. Rev. Lett., vol. 13, pp. 508–509, Oct 1964.

[34] The ATLAS Collaboration, "Observation of a new particle in the search for the standard model higgs boson with the ATLAS detector at the LHC," Physics Letters B, vol. 716, pp. 1–29, sep 2012.

[35] The CMS Collaboration, "Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC," Physics Letters B, vol. 716, pp. 30–61, sep 2012.

[36] T. T. Takahashi, H. Matsufuru, Y. Nemoto, and H. Suganuma, "Three-quark potential in su(3) lattice qcd," Phys. Rev. Lett., vol. 86, pp. 18–21, Jan 2001.

[37] D. Linthorne and D. Stolarski, "Triggering on emerging jets," Physical Review D, vol. 104, no. 3, 2021.

[38] L. Carloni and T. Sjöstrand, "Visible effects of invisible hidden valley radiation," Journal of High Energy Physics, vol. 2010, sep 2010.

[39] S. Knapen, J. Shelton, and D. Xu, "Perturbative benchmark models for a dark shower search program," Physical Review D, vol. 103, jun 2021.

[40] C. Bierlich, S. Chakraborty, N. Desai, L. Gellersen, I. Helenius, P. Ilten, L. Lönnblad, S. Mrenna, S. Prestel, C. T. Preuss, T. Sjöstrand, P. Skands, M. Utheim, and R. Verheyen, "A comprehensive guide to the physics and usage of pythia 8.3," 2022.

[41] B. Andersson, G. Gustafson, G. Ingelman, and T. Sjöstrand, "Parton fragmentation and string dynamics," Physics Reports, vol. 97, no. 2, pp. 31–145, 1983.

[42] B. Andersson and P. Andersson, The Lund Model. Cambridge Monographs on Particle Physics, Nuclear Physics and Cosmology, Cambridge University Press, 1998.

[43] CERN, "Linac4 Technical Design Report," Dec 2006. revised version submitted on 2006-12-14 09:00:40.

[44] J. Wenninger, "Operation and Configuration of the LHC in Run 2," Mar 2019.

[45] The CMS Collaboration, "The cms experiment at the cern lhc," Journal of Instrumentation, Aug 2008.

[46] The ALICE Collaboration, "The alice experiment at the cern lhc," Journal of Instrumentation, Aug 2008.

[47] The LHCb Collaboration, "The lhcb experiment at the cern lhc," Journal of Instrumentation, Aug 2008.

[48] Y. Takubo, "ATLAS IBL operational experience," tech. rep., CERN, Geneva, Nov 2016.

[49] M. Backhaus, "The upgraded pixel detector of the atlas experiment for run 2 at the large hadron collider," Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment, vol. 831, pp. 65–70, 2016. Proceedings of the 10th International "Hiroshima" Symposium on the Development and Application of Semiconductor Tracking Detectors.

[50] The ATLAS Collaboration, "Operation and performance of the ATLAS semiconductor tracker in LHC Run 2," JINST, vol. 17, p. P01013. 55 p, Sep 2021. All figures including auxiliary figures are available at http://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PAPERS/SCTD-2019-01.

[51] A. Vogel, "ATLAS Transition Radiation Tracker (TRT): Straw Tube Gaseous Detectors at High Rates," tech. rep., CERN, Geneva, Apr 2013.

[52] ATLAS LAr Collaboration and M. Spalla, "ATLAS LAr Calorimeter Performance in LHC Run 2," tech. rep., CERN, Geneva, Jan 2020.

[53] ATLAS Collaboration,
ATLAS liquid-argon calorimeter: Technical Design Report. Technical design report. ATLAS, Geneva: CERN, 1996.

[54] A. M. Henriques Correia, "The ATLAS Tile Calorimeter," tech. rep., CERN, Geneva, Mar 2015.

[55] S. D. Jones, "The ATLAS Electron and Photon Trigger," tech. rep., CERN, Geneva, Oct 2017.

[56] A. Collaboration, "Athena," Apr. 2019.

[57] G. P. Salam, "Towards jetography," The European Physical Journal C, vol. 67, pp. 637–686, may 2010.

[58] The ATLAS Collaboration, "Dijet resonance search with weak supervision using s=13 TeV pp collisions in the ATLAS detector," Physical Review Letters, vol. 125, sep 2020.

[59] The ATLAS Collaboration, "Topological cell clustering in the ATLAS calorimeters and its performance in LHC run 1," The European Physical Journal C, vol. 77, jul 2017.

[60] E. Hansen, "Combination methods for in-situ jet calibration in ATLAS," Oct 2018.

[61] The ATLAS Collaboration, "Jet energy scale and resolution measured in proton–proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector," Eur. Phys. J. C, vol. 81, p. 689. 73 p, Jul 2020. 73 pages in total, author list starting page 57, 31 figures, 2 tables, submitted to Eur. Phys. J. C. All figures including auxiliary figures are available at http://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PAPERS/JETM-2018-05.

[62] J. H. Collins, K. Howe, and B. Nachman, "Extending the search for new resonances with machine learning," Physical Review D, vol. 99, jan 2019.

[63] R. D. Ball, , V. Bertone, S. Carrazza, C. S. Deans, L. D. Debbio, S. Forte, A. Guffanti, N. P. Hartland, J. I. Latorre, J. Rojo, and M. Ubiali, "Parton distributions for the LHC run II," Journal of High Energy Physics, vol. 2015, apr 2015.

[64] The GEANT Collaboration, "Geant4–a simulation toolkit," Jul 2003.

[65] G. Choudalakis, "On hypothesis testing, trials factor, hypertests and the bumphunter," 2011.

[66] M. Abadi, A. Agarwal, P. Barham, E. Brevdo, Z. Chen, C. Citro, G. S. Corrado, A. Davis, J. Dean, M. Devin, S. Ghemawat, I. Goodfellow, A. Harp, G. Irving, M. Isard, Y. Jia, R. Jozefowicz, L. Kaiser, M. Kudlur, J. Levenberg, D. Mané, R. Monga, S. Moore, D. Murray, C. Olah, M. Schuster, J. Shlens, B. Steiner, I. Sutskever, K. Talwar, P. Tucker, V. Vanhoucke, V. Vasudevan, F. Viégas, O. Vinyals, P. Warden, M. Wattenberg, M. Wicke, Y. Yu, and X. Zheng, "TensorFlow: Large-scale machine learning on heterogeneous systems," 2015. Software available from tensorflow.org.

[67] F. Chollet et al., "Keras." https://keras.io, 2015.

[68] R. Goodman, J. Miller, and P. Smyth, "Objective functions for neural network classifier design," in Proceedings. 1991 IEEE International Symposium on Information Theory, pp. 87–87, 1991.

[69] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," 2014.

[70] D. Choi, C. J. Shallue, Z. Nado, J. Lee, C. J. Maddison, and G. E. Dahl, "On empirical comparisons of optimizers for deep learning," CoRR, vol. abs/1910.05446, 2019.

[71] K. Fukushima, "Visual feature extraction by a multilayered network of analog threshold elements," IEEE Transactions on Systems Science and Cybernetics, vol. 5, no. 4, pp. 322–333, 1969.

[72] A. Caldwell, D. Kollár, and K. Kröninger, "BAT – the bayesian analysis toolkit," Computer Physics Communications, vol. 180, pp. 2197–2209, nov 2009.

[73] W. K. Hastings, "Monte carlo sampling methods using markov chains and their applications," Biometrika, vol. 57, no. 1, pp. 97–109, 1970.

[74] The ATLAS Collaboration, "Luminosity determination in pp collisions at $\sqrt{s} = 13$ TeV using the ATLAS detector at the LHC," tech. rep., CERN, Geneva, Jun 2019. All figures including auxiliary figures are available at https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/CONFNOTES/ATLAS-CONF-2019-021.

[75] S. van der Meer, "Calibration of the effective beam height in the ISR," tech. rep., CERN, Geneva, 1968.

[76] C. Sbarra, "The LUCID Detector for LHC Run-2," Universe, vol. 5, p. 11. 7 p, 2019.

[77] The ATLAS Collaboration, "Search for new phenomena in events with an energetic jet and missing transverse momentum in pp collisions at 13 tev with the atlas detector," Physical Review D, vol. 103, jun 2021.

[78] A. Poreba, "Trigger performance monitoring and rate predictions preparation for run 3 at atlas experiment," Computing and Informatics, vol. 40, no. 4, p. 833–849, 2021.