

A Search for Long Lived Dark Photons with the
ATLAS Detector
REU Program at Nevis Labs, Columbia University

JENNA RODERICK

August 7, 2024



Contents

1	Introduction	1
1.1	The Standard Model	1
1.2	Dark Matter	1
1.2.1	Rotation Curves	1
1.2.2	Gravitational Lensing	2
2	Instrumentation	3
2.1	The LHC	3
2.2	The ATLAS Detector	3
2.2.1	Geometry	4
2.3	Boosted Decision Trees and TMVA	4
3	Search for Long Lived Dark Photons in the Two Photon Signature	6
3.1	Motivation	6
3.2	Samples Used in BDT Training	6
3.3	BDT Input Variables	8
3.4	Photon Pair Selection	9
3.5	Training over different signals	10
3.5.1	Signal Grid	10
3.6	BDT Training Results	11
4	Conclusion	15
5	Acknowledgements	16
6	Appendix	17

Abstract

A search is conducted for long lived dark photon production in proton-proton collisions at the LHC by examining ATLAS detector data. Previous investigations into this signal have only considered prompt decays of the dark photon. This analysis focuses on the case where the dark photon is considered to be long lived, in which case electrons could be reconstructed as photons by the detector. In this analysis, a Boosted Decision Tree (BDT) is trained to separate the two photon signature of the signal from all other events that occur during proton collisions. Due to the unknown nature of Z_d and S, the BDT is trained on many different signals with different Z_d mass, S mass, and Z_d lifetime. This analysis will probe whether one BDT can be applied to all of the signals, or if the kinematics of certain signals varies enough that a different BDT should be trained over certain signals.

1 Introduction

1.1 The Standard Model

The Standard Model (SM) is a model which describes all known particles and their interactions. In the SM, particles with half-integer spin are fermions. All quarks are fermions, and they interact with the strong force to form hadrons. Leptons are fermions that do not interact via the strong force. The three first generation leptons are electrons, muons, and taus. Each of these leptons also has a corresponding flavor neutrino partner. Neutrinos are leptons that are neutrally charged and have very small mass. Each lepton has a corresponding anti-particle. For first generation leptons, antiparticles have the same mass but opposite charge. The positron is the anti-particle of the electron, with the same mass as the electron but opposite charge. For neutrinos, antiparticles have the opposite sign lepton number.

Gauge bosons are mediators of the fundamental forces. There are four fundamental forces through which SM particles interact. One is the strong force which is mediated by the gluon. Photons are bosons that mediate electromagnetic force. The weak force is mediated by both W and Z bosons. The last, and most well known, fundamental force is gravity. It is theorized that particles interact with the Higgs field to gain mass. The Higgs Boson is an excitation of the Higgs field, so in this way, the Higgs Boson is related to the mechanism through which particles gain mass.

1.2 Dark Matter

There are many mysteries that cannot be explained by just SM particles. For example, cosmological observations have been discovered that gravitational forces are stronger than just what would be produced by standard visible matter. This means that there is more matter in the universe than just what we can see. It is estimated that over 80% of matter in the universe is dark matter. Dark matter is named for the fact that it doesn't interact electromagnetically, making it "invisible". However, dark matter does interact with gravity and has mass. If dark matter interacts with SM particles at all, it does so very weakly, which is why it is so difficult to detect. However, there are various pieces of evidence that prove dark matter is out there.

1.2.1 Rotation Curves

One piece of evidence of the existence of dark matter is found through measurement of the orbital velocities at which objects rotate inside galaxies. Classical mechanics predicts the rotational velocity v_r of an object inside of a galaxy to be

$$v_r = \sqrt{\frac{GM(r)}{r}}$$

Where G is the gravitational constant, r is the radial distance from the center of the galaxy, and $M(r)$ is the enclosed mass as a function of the radial distance from the center. $M(r)$ is expected to increase for radii inside the radius of visible mass of the galaxy, as most of the visible mass in a galaxy is located around the center. For values of r past where the visible mass is, $M(r)$ is expected to remain constant. This means that v_r is expected to decrease proportional to $\sqrt{\frac{1}{r}}$. However, what has been observed is that the rotational velocity of objects in the galaxy remains constant at distances outside the

galaxy, meaning mass increases proportional to r at distances beyond the visible mass of the galaxy[1]. A graph of the findings is shown in Figure 1. This means that there must be additional invisible matter surrounding the galaxy.

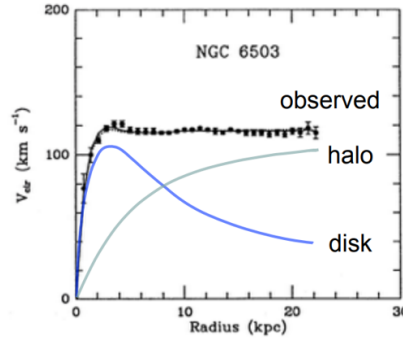


Figure 1: A graph of rotational velocity over distance[2]. The expected rotational velocity is shown in the blue line. The black line shows the actual relationship between rotational velocity and radius observed. The mass correction of an additional halo of invisible matter is shown in the green line.

1.2.2 Gravitational Lensing

Light travelling through the universe to an observer is bent by gravity from massive objects. The magnitude and location at which light is bent as it travels through a galaxy is an indication of the distribution of matter the galaxy. This calculation of mass distribution in a galaxy is called gravitational lensing. Gravitational lensing of the bullet cluster provides another piece of evidence for the existence of dark matter.

The bullet cluster consists of two galaxies which collided billions of years ago. The visible matter of the bullet cluster is located at the center, as the gases in the galaxies interacted and slowed. This is shown by the pink region in the picture of the bullet cluster in figure 2. However, gravitational lensing calculations indicate that most of the mass should be located in halos around the outside of the cluster[3], shown by the blue regions in figure 2. The shape and location of these halos indicates that the dark matter halos passed right through each other and the visible mass during the collision. This is proof that most matter in a galaxy must be a form of matter which doesn't interact electromagnetically or with visible matter, but does interact with gravity.



Figure 2: An image of the bullet cluster [3]. Shown in pink is the visible mass of the galaxy. In blue is the estimate of where most mass should be according to the effect of gravitational lensing.

2 Instrumentation

2.1 The LHC

The Large Hadron Collider (LHC), located at the European Organization for Nuclear Research (CERN), is the largest particle collider in the world. At roughly 27 km in circumference, it spans parts of both Switzerland and France. Inside, two beams of protons are sent in different directions. The LHC contains thousands of magnets which produce a strong magnetic field of 8.3 T that directs and accelerates the proton beams through the collider. Protons reach nearly the speed of light before they are collided in highly energetic (13.6 TeV) collisions. Proton-proton collisions occur at four points at the LHC, and a large detector is positioned at each point to take measurements [4].

2.2 The ATLAS Detector

One of the four large detectors at the LHC is the ATLAS detector. The ATLAS detector is a multipurpose detector meaning it is useful for a wide range of different experiments. Its ability to take high precision measurements make ATLAS perfect for searching for new physics beyond the standard model, such as dark matter.

ATLAS consists of an inner detector, an outer calorimeter, a muon spectrometer, and a magnet system. The inner detector consists of a pixel detector and two trackers which provide detailed trajectories of charged particles that pass through the detector. The Semiconductor Tracker (SCT) is made of extremely small strips of silicon which can measure particle tracks with precision less than half the width of a human hair. Wrapped around the SCT is the Transition Radiation Tracker (TRT), which consists of 4mm in diameter straws filled with gas. When charged particles pass through the TRT, they ionize the gas, which creates an electrical signal that reveals the particles' location[5]. The ATLAS detector also features a strong system of magnets. The Central Solenoid magnet wraps around the inner detector and produces a 2 T magnetic field. There are also three toroidal magnets, one at each end of the detector and one in the middle, that each produce a 4 T magnetic field. The purpose of these magnetic fields is to bend charged particles [6]. When the track is measured by the trackers, the magnitude and direction of curvature in the magnetic field provides information about the charge and momentum of the particle that passed through the detector.

After passing through the tracker, particles hit the calorimeter. The calorimeter's job is to stop the particles and measure their momentum. There are two calorimeters in the ATLAS detector: the Liquid Argon (LAr) calorimeter and the Tile Hadronic Calorimeter. The LAr calorimeter wraps around the inner detector and consists of layers of different kinds of metal with layers of liquid argon between each layer of metal. When electrons and photons hit a metal layer, they produce a shower of particles which ionize the liquid argon and produce a current which can be used to measure energy. The outermost calorimeter is the Tile Hadronic Calorimeter, which consists of a steel layers which stop hadrons such as protons and neutrons and produce a shower of particles. Sandwiched between the steel layers are layers of plastic scintillating tiles, which measure the energy of the resulting particle showers[7].

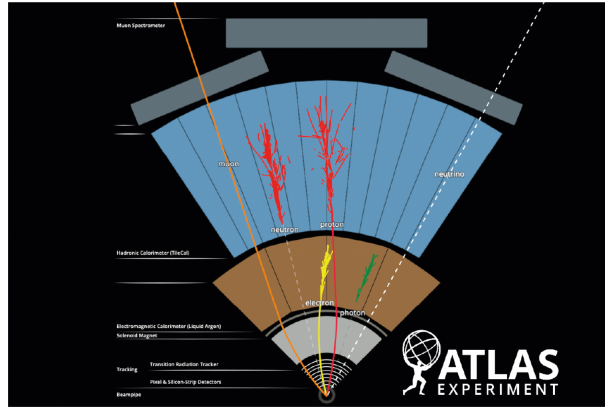


Figure 3: A schematic drawing of the ATLAS detector[8]

2.2.1 Geometry

An x-y-z coordinate plane is defined in the ATLAS detector. The z-axis is set along the beam. The x and y axes are perpendicular to the beamline, and the x-y plane is called the transverse plane. ϕ is the azimuthal angle in the transverse plane. θ is the angle from the beamline. However, θ is not an invariant quantity, so pseudorapidity is used instead. Pseudorapidity, η is calculated using $\eta = -\ln[\tan(\frac{\theta}{2})]$, and is the invariant angle from the beamline. An η value of zero corresponds to $\theta = 90^\circ$, and as η approaches infinity, it approaches the beamline.

2.3 Boosted Decision Trees and TMVA

Boosted Decision Trees (BDTs) are a machine learning technique used to perform classification on a dataset. In particle physics, BDTs are used to split a dataset into signal and background events. The signal is the process that is being examined, while the background is all other events. The process through which the BDT separates signal and background events is as follows. Say the list of variables is described by a vector, $\vec{v}$, of length j . First, the BDT orders all of the events, signal and background, by their values for variable $v_i, i \in j$. Then, it identifies the value, n , of v_i which causes greatest separation between signal and background events. This process is repeated for all $i \in j$, until it finds the v_i which has a cut that results in the greatest separation in signal and background events. The BDT makes a cut on the criteria for this value, in this example $v_i < n$. After

the cut, the full set of events (the first node) is split into two new nodes, one containing events for which the variable meets the criteria and the other containing events for which it does not. After this, the decision determines whether another cut on a variable improves signal and background separation in either of the two new subsets. If it doesn't, then it declares the node as finished and moves on to the next node. It repeats this until no nodes are left. After this, the decision tree defines each resulting leaf as either signal-like or background-like. It does this by calculating $S/(S+B)$, where S is the number of signal events and B is the number of background events. This gives the ratio of signal events to total events in the new sample. A ratio close to 1 is more signal-like, where a ratio closer to 0 is background-like. These values are given as outputs in the tree[9]. An example of a decision tree is shown in Fig [4].

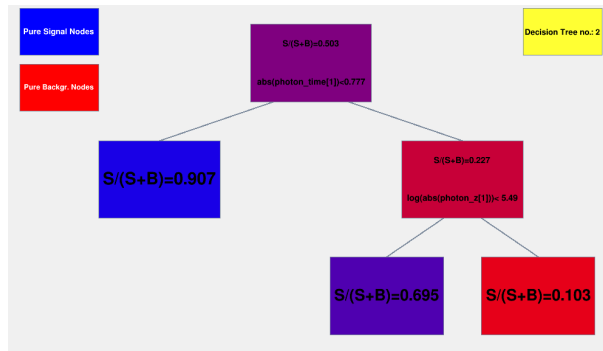


Figure 4: An example of a boosted decision tree.

Boosting refers to the use of information about previous trees to inform the training of new trees. There are many different algorithms for boosting, but one of the most effective, and relevant to this analysis, is Adaptive Boosting, or AdaBoost. If a tree has poorer performance, meaning it is a weak learner, the BDT will try to learn from it. It does this by looking at the rate at which events are misclassified in all of the trees. The misclassification rate of an event in each of the trees is used to give a weight to each event. More weight is given to events in with a higher misclassification rate. The tree is then retrained with the events that have the weights applied. Since events that tended to be misclassified in prior training have higher weights, the new tree will be more focused on performing cuts that correctly classify them [9].

Before training, the BDT splits the full set of events up into testing and training events. It performs the aforementioned training over the training events. Then, it tests the results by making the same cuts on the testing events. This checks that the cuts made aren't specific to the training data, in which case the BDT would be overtrained.

Toolkit for Multivariate Data Analysis (TMVA) is a ROOT data analysis package used to train BDTs for separating signal and background in particle physics. It contains the BDT training algorithm as well as a user interface that is used to create classifier output plots. TMVA also stores information about the training and creates classifier outputs that are used to evaluate BDT performance. Following training, TMVA also prints out a list of the variables by ranked by their importance in separating signal and background.

3 Search for Long Lived Dark Photons in the Two Photon Signature

3.1 Motivation

The purpose of this search is to investigate a potential dark matter particle called a dark photon. The dark photon is introduced in the Hidden Abelian Higgs Model (HAHM), which predicts a new type of symmetry called dark gauge symmetry. This calls for the existence of a dark gauge boson called a dark photon (Z_d), as well as a new dark scalar (S) [10]. The mass and lifetime of Z_d is unknown, as is the mass of S. This analysis examines the possibility that this dark photon is produced during proton-proton collisions at the LHC. We look at the process where S decays to two Z_d , which then each decay to an electron-positron pair. This process is described by the Feynman diagram in figure 5.

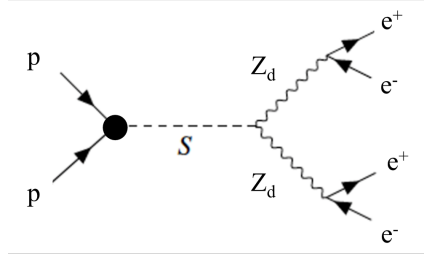


Figure 5: A Feynman diagram of the signal process used in this analysis: a scalar of unknown mass decaying to two dark photons, which each decay to a pair of leptons

Prior searches for this signal process have been conducted. However, in these searches, Z_d was considered to be prompt, meaning it decays almost instantly. This analysis searches for the case where Z_d is considered to be long lived [11]. In this case, Z_d decay could occur near the edge of the TRT in the ATLAS detector. Electrons produced at the edge of the TRT won't have much, if any, tracking information. As can be seen in Fig. 3, the only way the ATLAS detector distinguishes electrons and photons is that electrons have a track, since they are charged and interact with the tracker, while photons do not. So, for this analysis, many of the electrons we are looking for will be reconstructed as photons. The purpose of this project is to train a BDT to separate the signature of a long-lived dark photon decaying to two reconstructed photons from the many other types of processes that occur during proton-proton collisions in the LHC.

3.2 Samples Used in BDT Training

Each signal sample consists of MonteCarlo (MC) v1.0 FactoryTools ntuples. Each signal contains around 50,000 events. These MC events are generated using the HAHM. Reconstruction of the events is based on 2022 data-taking conditions at the ATLAS detector. The control region (CR) background is around 3 million events of 2022 ATLAS data. Preselection is applied to the samples to ensure we only look at events which are relevant to this analysis. Preselection criteria is:

- ≥ 2 photons: Since we look for the 2 photon final state, events should contain at least 2 photons

- < 2 electrons
- $|\text{photon time}| < 12.5 \text{ ns}$: Because there is a 25 ns delay in pp collisions at the LHC, this ensures we aren't looking at photons from a prior collision

Signal events are required to have photon time > 0 , while CR background events are required to have photon time < 0 . The purpose of these cuts is to blind the data by making sure the background contains very few signal events. This cut blinds the data because of the way time is calibrated. Time is calibrated by subtracting the time it takes a particle traveling at the speed of light (c) from the time that the calorimeter measures that the object hit the calorimeter. Since most particles resulting from pp collisions have extremely high energies and travel at a speed very very close to c , the distribution of photon time for the background (CR) is expected to be zero. Due to errors in calibration, the actual timing CR distribution is a Gaussian centered at zero. The signal photons, however, are produced by a long-lived particle, so they are expected to have a time above zero. Thus, photon time distribution for signal is not symmetrical at zero, but has a higher percentage of events with positive times than negative times. This is illustrated by the normalized histogram of electron time in figure 6. This figure uses the 2 electron final state as the signal, but the principle is the same for photons.

By only looking at photons with time < 0 in the background, we ensure that the background is expected to have very few signal events. Thus, when we train the BDT, we are actually separating signal and background, and not signal and a mixture of signal and background.

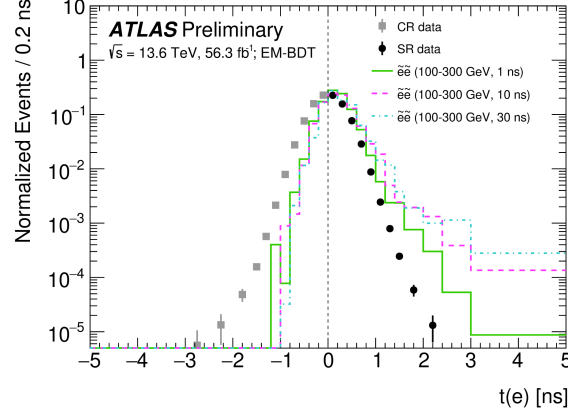


Figure 6: Normalized distribution of electron time. Background data, shown by the black and gray dots, is a Gaussian distribution symmetric around 0. The signals are represented by the colored lines [12].

3.3 BDT Input Variables

The following are variables are used to train the BDT:

- $|t|$: The absolute value of the calibrated time that the photon hits the calorimeter.
- p_t : The transverse momentum of the photon measured by the calorimeter
- $\ln|z|$: A pointing variable, the "track" of the photon given by the calorimeter
- f_3 : The ratio of the photon's energy deposited in the third layer of the calorimeter to the total energy of the photon.
- f_1 : Ratio of photon's energy deposited in the first layer of the calorimeter to the total energy of the photon.
- η : The pseudorapidity of the photon - the angle of the photon relative to the beamline
- ID: The ID of the photon, indicates likelihood of being a truth object
- Conversion type: Indicates whether or not a photon has been converted

The value of each of these variable is inputted into the BDT for each of the two photons. Additional variables, which are either functions of the above variables or individual variables, are implemented into the BDT for training as well.

- $\ln|z_1 - z_2|$: Natural logarithm of the differences in the pointing variables of the two photons
- $\ln|z_1 + z_2|$: Natural logarithm of the summation of the pointing variables of the two photons
- $\Delta\phi$: Difference in azimuthal angle of the two photons, $|\phi_1 - \phi_2|$

- $\Delta\eta$: Difference in angle from the beamline of the two photons, $|\eta_1 - \eta_2|$
- ΔR : The angular separation between the two photons, calculated as $\sqrt{\Delta\phi^2 + \Delta\eta^2}$
- MET: Missing transverse energy
- n: Number of photons

3.4 Photon Pair Selection

As we look at variables for each of the two photons, how we choose the two photons to perform the analysis is important. While most signal events result in only two photons being reconstructed, some have three or all four. We wish to train the BDT on photons which came from the same parent particle (vertex). This will give optimal separation between signal and background distributions of angular variables such as ΔR . There are two different methods of photon pairing investigated in this analysis.

The first way to pair up photons is by their transverse momentum p_t . The photon with the highest p_t is called the leading photon, and the second highest is the subleading photon. Another option for photon pairing is by the leading photon and the photon which minimizes the ΔR calculation with the leading photon (the minDR photon). This pair provides the photon pair with the leading photon which has the lowest angular separation. This method of photon matching is well motivated because a photon that is produced from the same vertex as the leading photon will be close to it in the detector. Photons produced close together will have small angular separation. A photon produced by a different vertex will be farther away from the leading photon and will have a greater angular separation.

To determine which method of photon pairing is more effective at correctly pairing photons from the same vertex, we use a process called truth-matching. During truth-matching, the reconstructed photons are checked against truth-level information about the vertex from which they originate. If the two photons each came from the same vertex (which could be either of the two possible vertices), they are truth-matched to the same vertex. A comparison of how many signal events are truth-matched to the same vertex for each of the methods of photon pairing is given below. Figure 7 shows a histogram of ΔR distributions for signal events after preselection. The black line shows the distribution for all signal events, the red line shows the distribution for just pairs truth-matched to the same vertex.

Proof of the effectiveness of truth-matching can be seen in the shape of the ΔR histograms. The ΔR distribution has peaks at two different values. The first peak at a lower angular separation corresponds to the case where photons are closer together and thus more likely to have come from the same vertex. The second peak at a higher ΔR corresponds to the case where the photons are far apart and likely to have come from a different vertex. We can see in the plot of both distributions that after truth-matching is performed on the events, the second peak is flattened. This indicates that truth-matching removes some of the events with higher ΔR where the pairs did not come from the same parent particle.

The number of events before and after the truth-matching can be determined by taking the integral of each histogram. Using these values, an efficiency of photons truth

matched to the same vertex can be calculated. The fraction of leading and subleading photons which are truth-matched to the same vertex is 0.734, while the fraction of leading and minDR photons which are truth-matched to the same vertex is 0.777. Therefore, matching up photons by minimizing the ΔR calculation results in more correctly vertexed photon pairs. This indicates that if we wish to improve the likelihood of selecting photons which came from the same vertex, we should look at the leading and minDR photons. For the remainder of this analysis, variable values for leading and minDR photons will be used to train the BDT.

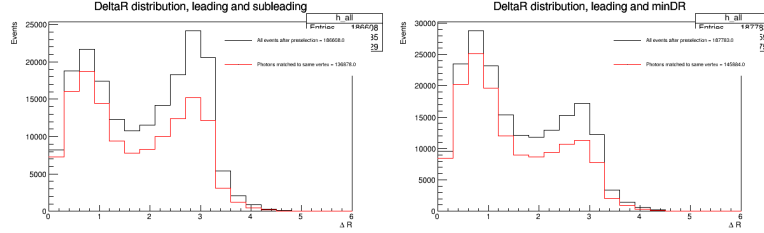


Figure 7: The ΔR distribution for all signal events compared to just events truth-matched to the same vertex. Left: Leading and subleading photon pairs. Right: Leading and minDR photon pairs.

3.5 Training over different signals

3.5.1 Signal Grid

The mass of S , mass of Z_d , and lifetime of Z_d are unknown. This analysis looks at 14 different combinations of S mass and Z_d mass, shown on the signal mass grid in figure 8. We examine four possible Z_d lifetimes as well: 0.1, 0.5, 2, and 10 ns. This leads to 56 total signal grid points. We want to determine if one BDT can effectively distinguish each of these signals from background data, or if a different BDT should be trained for certain signals. One way signals can be divided is by mass splitting, (ΔM). ΔM refers to the mass difference between S and Z_d . The reason why signals are split into high and low ΔM is because high and low ΔM signals are expected to have different kinematics. In high ΔM signals, each Z_d will get a lot of energy from the parent S , where in low ΔM

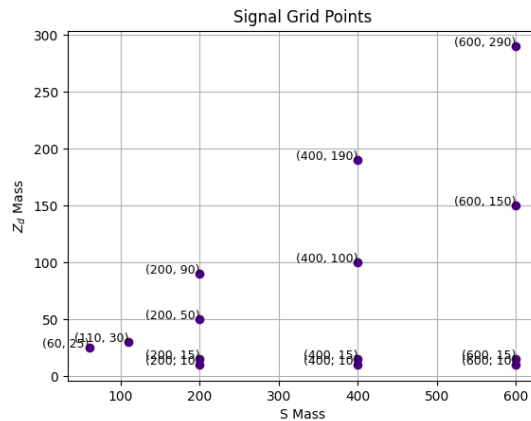


Figure 8: A grid of the signal mass points used in this analysis.

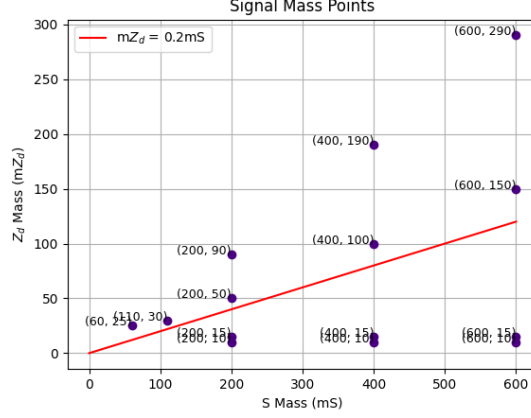


Figure 9: Signal grid with line at $mZ_d = 0.2mS$ dividing up high and low ΔM signals.

Events	All	Low ΔM	High ΔM
Signal (train/test)	93891/93891	89381/89381	6955/6955
Background (train/test)	40264/40264	40264/40264	40264/40264

Table 1: The amount of events that are used for testing and training the BDT over each signal file. Note that the signal events in the low ΔM training and high ΔM training do not add up to the signal events used in the merged file containing all the signals. This might indicate a signal is missing from the all signals file.

signals, each Z_d gets less energy. For this analysis, high ΔM signals have $\frac{Z_d}{S} < 0.2$. The line dividing high and low ΔM signals is shown in figure 9. From this figure, it can be seen that for this analysis, high ΔM events are all those with Z_d mass of 10 or 15 GeV. These signals have low mass Z_d all have a high difference in masses between Z_d and S. The low ΔM signals are all other signals because they have minimal difference in S and Z_d mass. To determine if a separate BDT should be trained over low and high ΔM signals, the same BDT is trained over the low ΔM signal points, the high ΔM signal points, as well as a merged signal file containing each of the 56 signals.

3.6 BDT Training Results

A table of the number of events used in testing and training for each grouping of signals is provided in table 1. High ΔM signals contain much fewer events than low ΔM signals. This is because the dark photons have relatively tiny mass compared to S. Conservation of energy mandates that each Z_d receive large amounts of transverse energy since very little of the S energy goes towards Z_d mass. Each Z_d produced in these events is boosted (nearly relativistic). Most of the dark photons fly right through the detector without decaying, so very few photons are reconstructed in the signal. Because of this, few high ΔM events pass preselection.

Differences in important variables used in separating signal and background events among the different signals was observed. A list of full variable rankings by importance for each signal is shown in Table 2.

Overall, top variables are consistent for both the all and low ΔM signals. This is because high ΔM signals make up a very small percentage of the total signal events after

Rank	Variable, All	Variable, Low DM	Variable, High DM
1	$ t[\text{minDR}] $	$ t[\text{minDR}] $	ΔR
2	$ t[0] $	$ t[0] $	$f_3[\text{minDR}]$
3	$p_t[0]$	$p_t[0]$	$f_3[0]$
4	$\ln z[\text{minDR}] $	$\ln z[\text{minDR}] $	MET
5	$\ln z[0] $	$\ln z[0] $	$p_t[0]$
6	$p_t[\text{minDR}]$	$f_3[\text{minDR}]$	ID[0]
7	$f_3[\text{minDR}]$	$f_3[0]$	$f_1[0]$
8	n	$p_t[\text{minDR}]$	$ \eta[0] $
9	$f_3[0]$	n	$\ln z_1 - z_2 $
10	$ \eta[\text{minDR}] $	MET	$\ln z[\text{minDR}] $
11	MET	$ \eta[\text{minDR}] $	$p_t[\text{minDR}]$
12	ID[minDR]	ΔR	$\ln z[0] $
13	$\ln z_1 + z_2 $	$ \eta[0] $	$\Delta\phi$
14	$\Delta\phi$	$\ln z_1 + z_2 $	$\Delta\eta$
15	$ \eta[0] $	conversionType[minDR]	$f_1[\text{minDR}]$
16	conversionType[minDR]	$\Delta\phi$	ID[minDR]
17	ΔR	conversionType[0]	conversionType[0]
18	conversionType[0]	ID[minDR]	$ \eta[\text{minDR}] $
19	$f_1[\text{minDR}]$	$f_1[\text{minDR}]$	$ t[0] $
20	ID[0]	ID[0]	conversionType[minDR]
21	$\ln z_1 - z_2 $	$f_1[0]$	$\ln z_1 + z_2 $
22	$f_1[0]$	$\ln z_1 - z_2 $	$ t[\text{minDR}] $
23	$\Delta\eta$	$\Delta\eta$	n

Table 2: Variables ranked by importance in separating signal and background for BDT training over each of the three signal files.

preselection in the all signal file. After preselection, high ΔM signal events make up only around 7% of the total events for all the signals.

Photon times are ranked first and second in low ΔM signals, but are ranked near the bottom for high ΔM signals. This is because in low ΔM signals, photon time is expected to be much higher than background events because the dark photon from which they decay is considered to be long lived. However, this same pattern is not seen for high ΔM signals which is due to the nature of these events. Since each Z_d produced in these events is boosted, the photons that are reconstructed must have a very small time of photon production in order to be detected. This is why the signal distribution of photon time for the high ΔM signals looks similar to the background, shown in Figure 10.

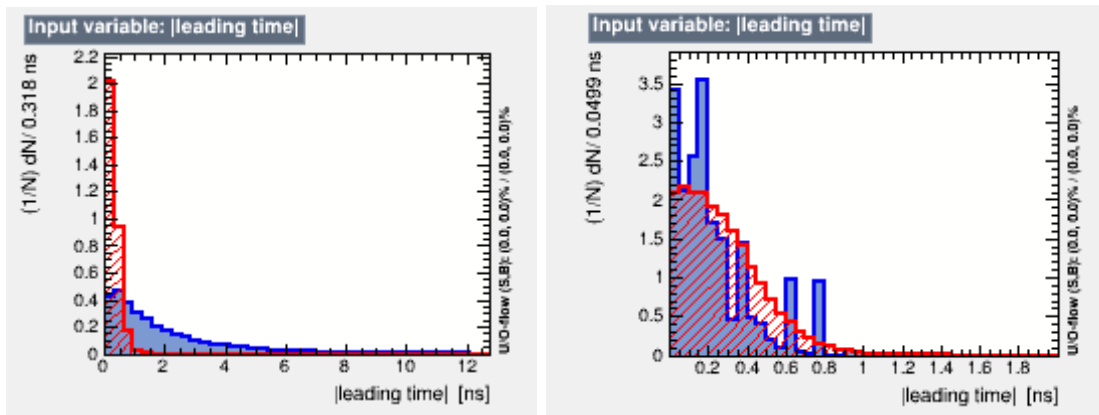


Figure 10: From TMVA, a comparison of the photon time variable distribution in signal (blue) vs background (red). Left: Low ΔM signals. Right: High ΔM signals

ΔR ranks most important in training the BDT over the high ΔM signals, but 17th and 12th in importance for the all and low ΔM signals respectively. The distributions of ΔR for signal and background are shown in Figure 11. The ΔR distribution for the high ΔM signal is sharply peaked at a very low ΔR value. This is because the dark photons are boosted so their daughter electrons are collimated (extremely close together). This makes a cut on ΔR much more effective at separating signal and background for the high ΔM signal region than for the low ΔM signals.

MET also ranks highly in importance for high ΔM signals. This is because many of the nearly relativistic dark photons escape the detector, contributing to larger amounts of missing energy. This can be seen in the MET distribution for signal and background given by TMVA in Figure 12.

There are many different outputs from TMVA that are used to analyze the performance of the BDT. One is the Receiving Operator Characteristic (ROC) curve, shown for each of the three signals in figure 13. ROC curves display background rejection over signal efficiency. Background rejection is the fraction of background events that are removed and signal efficiency is the fraction of signal that is kept. The desired shape of the ROC is

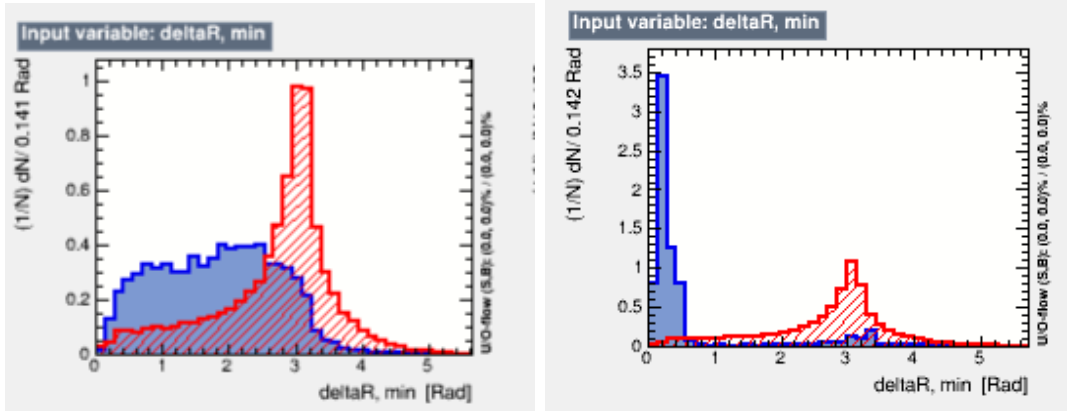


Figure 11: A comparison of the ΔR variable distribution in signal (blue) vs background (red), given by TMVA. Left: Low ΔM signals. Right: High ΔM signals

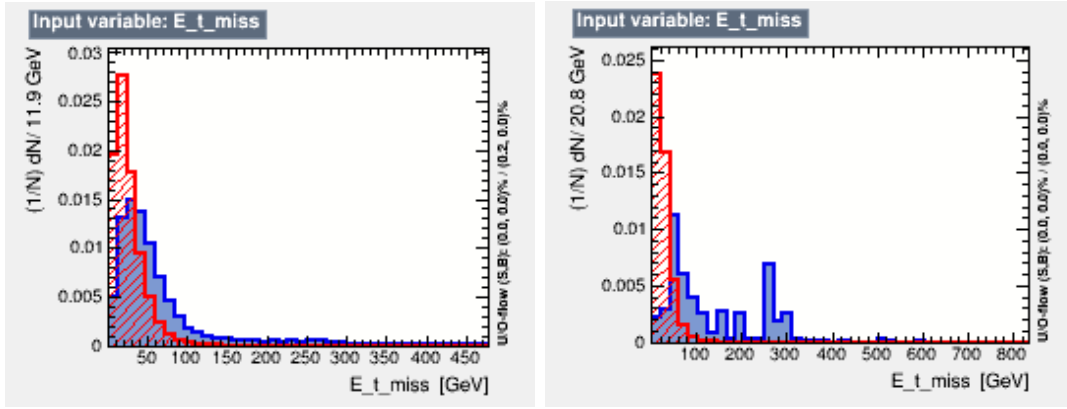


Figure 12: From TMVA, a comparison of the missing transverse energy distribution in signal (blue) vs background (red). Left: Low ΔM signals. Right: High ΔM signals

a horizontal line at background rejection = 1, since ideally we want to keep all the signal when we remove all the background. This would correspond to a perfect separation between signal and background. ROC curves for all signals and low ΔM signals are similar and indicate good performance of the BDT. The reason BDT performance is similar for training over all signals put together and just the low ΔM signals is because the high ΔM signals make up a very low percentage of the total signal events. Not many high ΔM signal events pass the preselection cuts due to the dark photons being boosted, so it is less likely that two photons will be reconstructed. The ROC curve for the high ΔM has a significant dip on the top right corner. This indicates worse performance for the BDT trained on the high ΔM signals than the other two groups of signals. The area under the ROC curve (AUC) can also be computed as a quantitative check of quality of the ROC curve. An AUC of 1 corresponds to perfect separation. AUC is printed by TMVA following the training. The AUCs for all, low ΔM , and high ΔM signals are 0.999, 0.999, and 0.938 respectively. While AUC is the same for all signals as for just low ΔM signals, it is much lower in the high ΔM signals.

Not much distinction can be made in quality of BDT performance in the all and low ΔM signals based on ROC curves and AUCs. A more precise measurement of BDT performance is examined to see if there is a benefit separating the low and high ΔM signals.

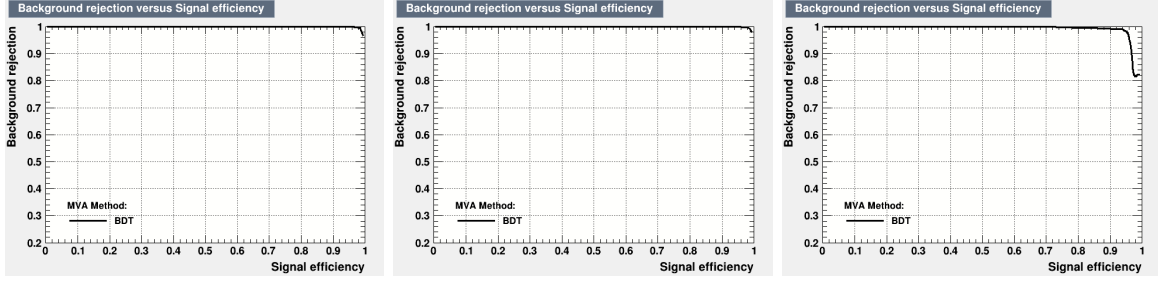


Figure 13: ROC curves for all signals (left), low ΔM signals (center), and high ΔM signals (right)

For this the BDT score distribution is evaluation. BDT score is a classifier given by the BDT to each event it trains over that indicates how likely the BDT thinks an event is to be signal or background. A BDT score of 1 indicates that the BDT is absolutely certain that the event is signal, and a BDT score of -1 indicates the BDT is absolutely certain the event is background. The BDT score for each event is averaged among all the trees trained. In an ideal BDT score distribution, all of the signal events would have a BDT score above 0, and all the background events would have a BDT score below 0. The BDT score distributions for each signal file trained over are shown in figure 14. There are similar BDT score distributions for the all and low ΔM signals and both show good separation between signal and background. Large error bars and a spiky shape of the BDT score distribution for training over the high ΔM signal file indicates poor statistics. To quantitatively compare BDT score distributions, a cut is made on the BDT score which corresponds to a background efficiency of 0.01. This cut removes the events on the BDT score distribution which have a BDT score below the requested value. The signal efficiency after the BDT score cut is the fraction of signal events that remain when we remove 99% of the background. Signal efficiency at 0.01 background efficiency is printed by TMVA after BDT training. The signal efficiencies for all signals, low ΔM signals, and high ΔM signals are 0.986, 0.988, and 0.938 respectively. From these values, we observe a slightly lower quality BDT performance for all signals put together than for the low ΔM signals. Furthermore, we see a much lower signal efficiency at 0.01 background efficiency for the high ΔM signals. It's important to note again the poor statistics of the high ΔM signals. Poor statistics could be a cause of the low signal efficiency after the BDT score cut, as can be seen in the outlier around BDT score = -0.35 in the high ΔM BDT score distribution in figure 14. This outlier corresponds to only one events, and could be the main reason why the signal efficiency is so low. So, it cannot be definitively claimed that BDT performance is worse in high ΔM signals. BDT training should be redone with improved high ΔM statistics to confirm the result.

4 Conclusion

In this analysis, a starter BDT was trained to separate the 2 photon signature of the $S \rightarrow Z_d Z_d \rightarrow 2e^+ 2e^-$ signal process from other events that occur at the LHC. The BDT was applied to two categories of signals which were expected to have different kinematics: high ΔM and low ΔM . The varying kinematics led to different variables being effective

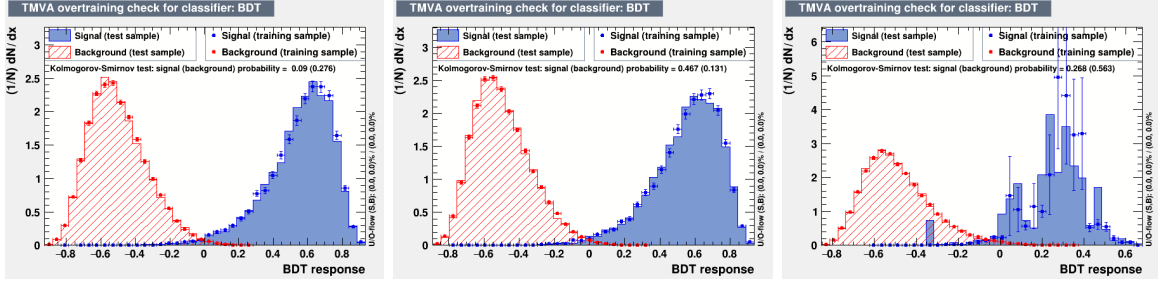


Figure 14: BDT score distributions for all signals (left), low ΔM signals (center), and high ΔM signals (right)

at separating each of these groups of signals from background events. Outputs of the BDT indicated that when the BDT was applied to the high ΔM signals, performance was slightly worse compared to the high ΔM signals, indicating we may want to train a different BDT for high and low ΔM signals. However, this result may have been influenced by the low statistics for the high ΔM signals. This training should be redone with more events in the high ΔM signals so the claim can fully be made.

5 Acknowledgements

I would like to thank Prof. John Parsons, Dr. Jonathan Long, and PhD students Eleanor Woodward and Maria Bressan for their guidance throughout the project. I would also like to thank everyone in the Columbia ATLAS group for assistance with BDT training and helpful feedback in weekly meetings. Thank you to Prof. Georgia Karagiorgi, Prof. Reshmi Mukherjee, and Amy Garwood for their hard work running the Nevis Labs REU program. Special thanks to the National Science Foundation (NSF) for funding this research under grant no. PHY-2349438.

References

- [1] V. C. Rubin and J. Ford, W. Kent, “Rotation of the Andromeda Nebula from a Spectroscopic Survey of Emission Regions,” , vol. 159, p. 379, Feb. 1970.
- [2] K. G. Begeman, A. H. Broeils, and R. H. Sanders, “Extended rotation curves of spiral galaxies: dark haloes and modified dynamics,” *Monthly Notices of the Royal Astronomical Society*, vol. 249, no. 3, pp. 523–537, 04 1991. [Online]. Available: <https://doi.org/10.1093/mnras/249.3.523>
- [3] NASA. (2022) 1e 0657-56: Nasa finds direct proof of dark matter. [Online]. Available: <https://chandra.harvard.edu/photo/2006/1e0657/index.html>
- [4] CERN. The large hadron collider. [Online]. Available: <https://home.web.cern.ch/science/accelerators/large-hadron-collider>
- [5] ——. The inner detector. [Online]. Available: <https://atlas.cern/Discover/Detector/Inner-Detector>

- [6] ——. Magnet system. [Online]. Available: <https://atlas.cern/Discover/Detector/Magnet-System>
- [7] ——. Calorimeter. [Online]. Available: <https://atlas.cern/Discover/Detector/Calorimeter>
- [8] ——. Atlas schematics. [Online]. Available: <https://atlas.cern/Resources/Schematics>
- [9] Y. Coadou, *Boosted Decision Trees*. WORLD SCIENTIFIC, Feb. 2022, p. 9–58. [Online]. Available: http://dx.doi.org/10.1142/9789811234033_002
- [10] D. Curtin, R. Essig, S. Gori, and J. Shelton, “Illuminating dark photons with high-energy colliders,” *Journal of High Energy Physics*, vol. 2015, no. 2, Feb. 2015. [Online]. Available: [http://dx.doi.org/10.1007/JHEP02\(2015\)157](http://dx.doi.org/10.1007/JHEP02(2015)157)
- [11] D. Boye, K. A. Assamagan, S. Connell, S. Snyder, C. Weber, L. L. Leeuw, F. Fassi, K. Hamano, M. Lefebvre, J. Chiu, W. Buttinger, T. N. L. Truong, D. Casadei, V. I. Martinez Outschoorn, k. baker, X. Mapekula, R. Coelho Lopes De Sa, Z. Soumami, and A. Aboulhorma, “Search for a new scalar decays to beyond-the-Standard-Model light bosons in four-lepton final states with 140 fb^{-1} at $\sqrt{s} = 13 \text{ TeV}$,” CERN, Geneva, Tech. Rep., 2021. [Online]. Available: <https://cds.cern.ch/record/2773055>
- [12] “Search for displaced leptons in 13 TeV and 13.6 TeV pp collisions with the ATLAS detector,” CERN, Geneva, Tech. Rep., 2024, all figures including auxiliary figures are available at <https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/CONFNOTES/ATLAS-CONF-2024-011>. [Online]. Available: <https://cds.cern.ch/record/2905292>

6 Appendix

List of variables that were tried and later removed from BDT training.

- z1: Z coordinate of photon in first layer of the calorimeter used in pointing calculation
- z2: Z coordinate of photon in second layer of the calorimeter used in pointing calculation
- photon_E
- photon_dz
- photon_r1
- photon_r2
- photon_maxEcell_t
- photon_clusterE
- photon_maxEcell_x
- photon_maxEcell_y

- photon_maxEcell_z
- photon_isConversion

BDT Hyperparameter values:

- NTrees=800
- MinNodeSize=7
- MaxDepth=3
- BoostType=AdaBoost
- AdaBoostBeta=0.1
- UseBaggedBoost
- BaggedSampleFraction=0.6
- SeparationType=GiniIndex
- nCuts=15
- Transformations = I;D;P;G,D

Link to code used to run BDT:

https://gitlab.cern.ch/ewoodwar/dark-photons-bdt/-/tree/2gBDT?ref_type=heads