Search for SUSY with Delayed Photons at the Compact Muon Solenoid

Thesis by

Gillian Baron Kopp

In Partial Fulfillment of the Requirements for the
degree of
Bachelor of Science

CALIFORNIA INSTITUTE OF TECHNOLOGY
Pasadena, California

2018
Defended May 22, 2018

ACKNOWLEDGEMENTS

I would like to thank Maria Spiropulu for this wonderful research opportunity, and for welcoming me into the group three years ago. Researching in the Caltech High Energy Physics group has been an amazing experience, and has fueled my interest and curiosity for particle physics research. Maria's enthusiasm and brilliance in the field has been continually inspiring to me. Thank you, Maria, for being such an outstanding role model and mentor.

I would also like to thank Cristián Peña, Dustin Anderson, and Zhicai Zhang for their insightful explanations and for answering my many questions throughout this research. Their support has been invaluable. Thanks also to Artur Apresyan and Si Xie for teaching me so much during my first beam test at Fermilab. I would really like to thank Adi Bornheim and Jay Lawhorn for the physics discussions last summer at CERN, greatly expanding my appreciation for the LHC experiments. Thanks to Irene Dutta and Olmo Cerri for being the best office mates this year, and for the fun discussions.

Many thanks to my professors at Caltech who have furthered my passion for physics through their teaching and fantastic lab courses. Professor Frautschi's discussions about early high energy physics piqued my interest in particle physics research when I came to Caltech, and Professor Patterson's engaging particle physics course has made me ever more curious to learn more in this field. Eric Black and Frank Rice have been extraordinary lab professors, and have greatly broadened my appreciation for experimental physics. I am so thankful I have had the opportunity to learn from all these brilliant and inspiring professors. Caltech has been an incredible place to learn — made so by the professors, students, and research here.

Finally, thank you to my friends, parents, and sister who have continually supported and encouraged me.

ABSTRACT

The Compact Muon Solenoid (CMS) experiment records data from proton-proton (pp) and heavy ion (Pb-Pb and Pb-p) collisions at the Large Hadron Collider (LHC) to search for physics beyond the Standard Model, test theories of supersymmetry (SUSY), and measure properties of known particles with higher precision. One of the main objectives of the LHC was to discover the Higgs Boson, the final particle needed to complete the Standard Model. The discovery of the Higgs Boson was announced in 2012, and is a very important achievement and a triumph of particle physics experiment and theory. In 2025, the LHC will be upgraded to the High Luminosity LHC (HL-LHC), where the luminosity will be increased by a factor of 10. This will increase the number of pile-up collisions to 140-200 events per proton-proton bunch crossing, compared to the current 40 events per crossing (where each bunch crossing occurs every 25 ns).

In order to fully exploit the sensitivity of the CMS experiment, the current detectors must be upgraded to mitigate the effects of the large number of pileup interactions expected in collisions at the HL-LHC. New capabilities, such as precision timing measurements in calorimetric devices and minimum ionizing detectors, have been shown to effectively mitigate the effects due to pileup, and are expected to benefit the overall physics reach of the experiment. We present results obtained using a dedicated silicon timing layer identical to that proposed for the High Granularity Calorimeter proposed by CMS. This timing layer (pico-sil detector) was tested with high energy electromagnetic showers produced by electrons at the Fermilab Test Beam Facility. An outstanding time resolution of less than 16 ps was measured for a beam energy of 32 GeV.

A simulation of a benchmark long lived neutralino SUSY search is presented, and it is shown that the generator particle flight times can be faithfully reconstructed using the detector-level information. Identification algorithms for the SUSY model have been significantly improved with the use of a Boosted Decision Tree, and it is demonstrated that this algorithm has many benefits as compared to cut based IDs. The BDT is used to improve the separation of signal and background events in the SUSY model, and thus may similarly benefit many of the physics measurements at the LHC. With use of the BDT for the long lived neutralino SUSY model, the background rejection is increased significantly, with constant signal acceptance of 53.6%. This is an improvement in the significance of the signal selection by 2.38

σ. Further improvement is seen with the inclusion of detector timing information in the BDT. In addition, the detector timing information is shown to be important to the overall classification, contributing ≈25% of the information used in signal event identification. We thus demonstrate that with the BDT, the SUSY neutralino search can be performed with increased signal identification significance, and the searches sensitivity is expected to improve with the time resolution attained by the CMS calorimeter.

PUBLISHED CONTENT AND CONTRIBUTIONS

[1] Gillian Kopp et al. “Precision Timing Studies for the High Granularity Calorimeter Using a Dedicated Timing Layer”. In: *The Compact Muon Solenoid Experiment Detector Note* CMS DN-2016/028 (Nov. 2016).

[2] Gillian Kopp et al. “Search for SUSY with Delayed Photons at CMS”. In: *The Compact Muon Solenoid Experiment Experiment Analysis Note* CMS AN-2017/325 (Dec. 2017).

[3] Artur Apresyan. “Investigation of fast timing capabilities of silicon sensors for the CMS high granularity calorimeter at HL-LHC”. In: *2016 IEEE (NSS/MIC/RTSD), Strasbourg*. doi: 10.1109/NSSMIC.2016.8069873 (2016), pp. 1–8. URL: `http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=8069873&isnumber=8069358`.

[1] Gillian Kopp wrote this CMS Detector Note based on the June 2016 Fermilab testbeam performed with the Caltech CMS group, and analyzed the testbeam data from the prototype high granularity calorimeter. This work was performed with the support of a Caltech Summer Undergraduate Research Fellowship (SURF), with the fellowship supported by the Caltech Associates. This work is summarized in Chapter 6.

[2] Gillian Kopp wrote this CMS Analysis Note based on the simulation work for the neutralino dark matter model performed at CERN in Summer 2017 working with the Caltech CMS group. This work was performed with the support of a Caltech Summer Undergraduate Research Fellowship (SURF). This work is relevant for Chapter 2.

[3] This was written by a mentor, Artur Apresyan. Section IV is based on the 2016 testbeam analysis work, and Gillian Kopp made figures 12, 13, and 15. This work was presented at IEEE 2016, in Strasbourg, France. This work is relevant for Chapter 6.

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS

LIST OF TABLES

Chapter 1

INTRODUCTION

1.1 Motivation

The Large Hadron Collider (LHC) in Geneva, Switzerland, records data from proton-proton collisions to search for physics beyond the Standard Model, test theories of supersymmetry (SUSY) [1], and measure properties of known particles with higher precision. In proton-proton collisions, the primary hard interaction is between quarks or gluons (partons), and we investigate the production of rare new physics processes such as supersymmetric particles and dark matter candidates. In 2025, The LHC will be upgraded to the high-luminosity LHC (HL-LHC), with a target luminosity of $7.2 \cdot 10^{34}$ cm^{-2}s^{-1}, which is 10 times higher than the current LHC luminosity. Luminosity is defined as:

$$\mathcal{L} = \frac{1}{\sigma}\frac{dN}{dt} \tag{1.1}$$

where N is the number of events and σ is the interaction cross section. The number of events, N, can be expressed as

$$N = \sigma \int \mathcal{L}(t)dt \tag{1.2}$$

with $\int \mathcal{L}(t)dt$ as the integrated luminosity and $\mathcal{L}(t)$ as the instantaneous luminosity. Then, the luminosity in terms of the LHC beam parameters is

$$\mathcal{L} = \frac{N_b^2 n_b f \gamma}{4\pi\epsilon_n\beta} \qquad [2] \tag{1.3}$$

with N_b as the particles per bunch, n_b as the bunches in a beam, f as bunch frequency, γ as the relativistic correction factor, ϵ_n as the transverse emittance, and β as the transverse beam size [2]. Thus, an increase in luminosity is achieved by increasing the rate of particle collisions. The target luminosity of $7.2 \cdot 10^{34}$ cm^{-2}s^{-1} corresponds to 140-200 pileup collisions per bunch crossing [3], which is a significant increase in the number of collisions (currently 30-40 collisions per beam crossing). Additionally, this also means there is an increase in the amount of data gathered. Currently, the cross section for producing supersymmetric or other exotic particles in collision interactions is theorized to be very small compared with that for producing Standard Model particles, but with more data, rare and exotic processes can be searched for with higher precision.

In order to continue the current LHC performance given the higher pileup present in the HL-LHC, a precision timing detector is needed. A hermetic precision MIP (minimum ionizing particles) Timing Detector (MTD) has been proposed by CMS for the Phase 2 upgrade [3]. This detector will provide 4D tracking, have a time resolution of 30 ps, and provide particle reconstruction abilities that would otherwise not be available in the HL-LHC [4] [5]. An important added capability the ability to search for long lived particles using the improved timing. The MTD expands the parameter space in which long-lived particles can be searched for, as shown in Figure 1.1. The benefit of the MTD upgrade is in extending the available phase space to long-lived particles with higher masses. This is due to the ability of the detector to measure decay time of long-lived particles. A prototype high granularity calorimeter (HGC) has been demonstrated to have a time resolution of 16 ps, showing that this timing detector is feasible. The HGC will cover the ECAL endcap region ($\eta > 1.6$), and an additional precision timing detector for the barrel region with different technology is proposed. The analysis of the performance of the timing detector is discussed in Chapter 6, and its impact on the SUSY model detection is discussed in Chapter 4.

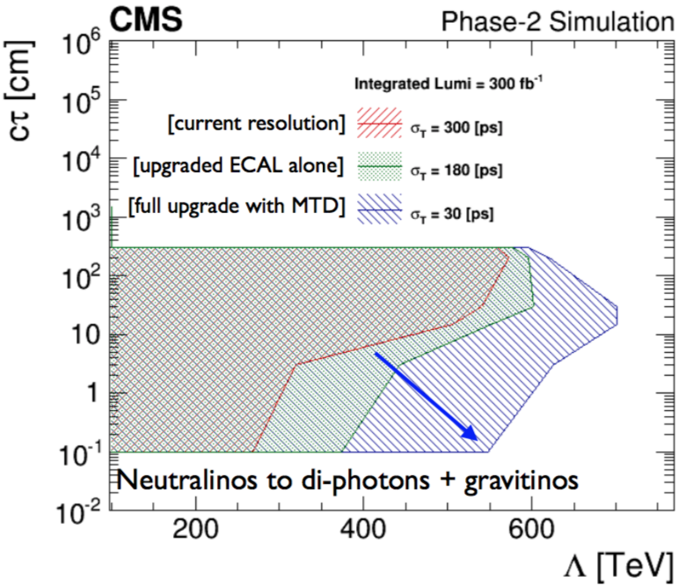


Figure 1.1: Phase II CMS Simulation demonstrating the SUSY mass scale (Λ) and lifetime ($c\tau$) parameters that can be searched for with the upgraded MTD. Figure from Lindsey Gray's talk presented at LHCC, 1 December 2017 [4].

In this thesis work, we investigate the signature of a particular long-lived dark matter SUSY model. This SUSY model decay may be observed in the HL-LHC, and improving the algorithms to identify dark matter events will increase the search sensitivity. Additionally, the precision timing information from the MTD is expected improve the SUSY search by probing more masses and lifetimes with better time resolution.

This precision timing information is crucial to the dark matter search because it will provide information about the decay of each particle. The focus is on improving the dark matter event identification algorithms with machine learning techniques, and investigating the importance of detector time information to these algorithms to understand the improvements offered by the MTD.

1.2 LHC Physics at the Compact Muon Solenoid

The Compact Muon Solenoid (CMS) experiment is a general purpose detector at the LHC, and measures the energy, charge, and momentum of particles resulting from proton-proton collisions up to 13 TeV. The detector is cylindrical, with forward calorimeters at the endcaps of the cylinder to completely enclose the interaction point (4π coverage). The other components (from inside to outside) are the silicon trackers, electromagnetic calorimeter (ECAL), hadron calorimeter (HCAL), superconducting solenoid (with a magnetic field of 3.8 Tesla), and the muon chambers [6] [7]. A diagram of the CMS detector is shown in Figure 1.2.

The CMS inner silicon tracker has a fine granularity and primarily records charged particle trajectories and secondary verticies with a high momentum resolution [8]. Transverse momentum (p_T) resolution is given by

$$\frac{\sigma_{p_T}}{p_T} = \frac{1}{BL^2} \tag{1.4}$$

where B is the magnetic field (3.8 T for CMS) and L is the length.

The inner tracker has a momentum resolution uncertainty of 0.7% at 1 GeV/c^2 (5% at 1000 GeV/c^2) [9]. The ECAL is made of 75,000 lead tungstate (PbWO4) crystals and measures energies of electrons and photons from the scintillation of the crystals with an excellent energy resolution [6]. The energy resolution uncertainty for photons in the ECAL barrel is between 0.7 – 1.8%, and 1.5 – 3% in the endcap region [10]. The HCAL absorbs hadronic showers from charged and neutral particles, and the energies and momentum of the particles are measured. The superconducting magnet is outside of the HCAL, and the magnetic field of 3.8 T curves the tracks of charged particles. The muon chambers are the final detector, and are formed of drift tubes and cathode strip chambers interlaced with iron plates for the magnetic system. The thickness of the iron absorbers prevents particles besides muons and neutrinos from reaching the muon chambers [6].

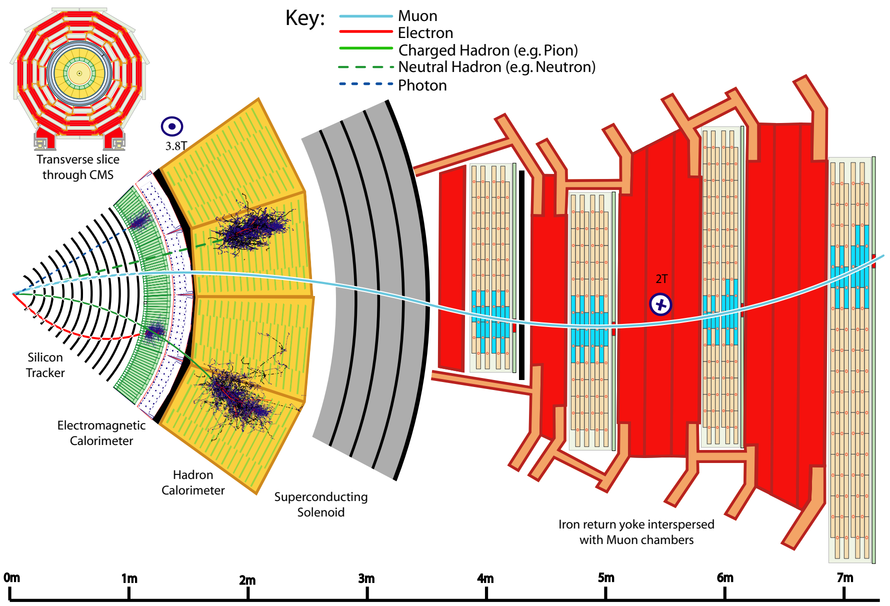


Figure 1.2: Diagram of the transverse section of the CMS detector, with representations of particles traveling through the various trackers and detectors. Diagram from *Particle-flow Reconstruction and Global Event Description with the CMS Detector* [8].

Detecting delayed photons from the neutralino dark matter model (Figure 1.4) benefits from a detector with excellent time resolution and lateral segmentation in order to collect enough high quality data [11]. The Compact Muon Solenoid (CMS) detector is hermetic and has complete azimuthal coverage, so therefore can detect and track outgoing particles in all directions [12]. Additionally, the CMS ECAL has a time resolution of about 150 ps (over time and with higher pileup this has degraded to the current 300 ps). The CMS detector provides large quantities of high quality data meeting these criteria, and therefore is a good detector perform dark matter searches. This analysis focuses on the ECAL information, which is most relevant for the measurement of photons.

1.3 High Luminosity Upgrade

In the upcoming high luminosity LHC upgrade (HL-LHC), a significant challenge is mitigating the effects of pileup. Each bunch crossing (crossing of the circulating proton beams) in the LHC is spaced by 25 ns, and the bunch crossing creates many collisions (140-200 in the HL-LHC). However, only a few of these collisions are “hard” with high transverse momentum indicating that new high mass particles may be involved in the collision [13]. The data gathered from an interaction will be from multiple proton-proton crossings, and the detectors need to be able to differentiate which crossing each decay profile originated from. Therefore, due to the increased number of interactions per crossing, the reconstruction algorithms currently in use will not be as effective, and thus the capabilities of the detectors must be improved.

The MTD will greatly help reduce the effects of pileup by providing precise timing information and tracking. This is very helpful for both the mitigation of pileup for all physics analyses, but is also very relevant for identifying the delayed particles resulting from certain dark matter SUSY models.

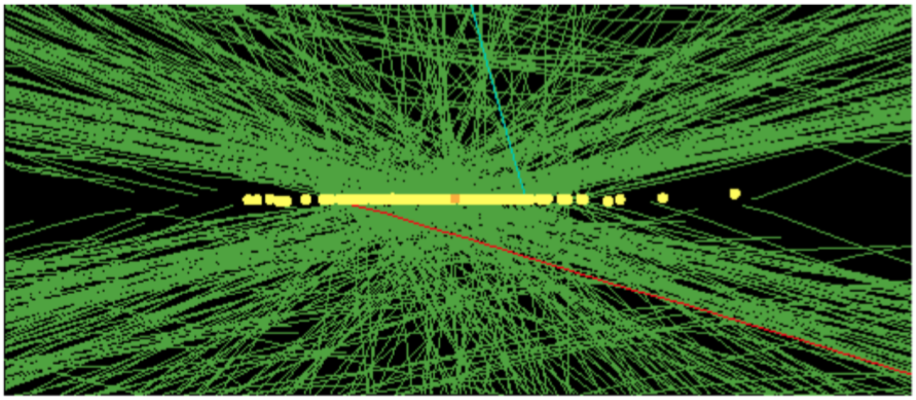


Figure 1.3: Diagram of a high pileup event from 2012 with 78 reconstructed vertices. Figure from *Technical Proposal for the Phase-II Upgrade of the CMS Detector* [13].

1.4 Dark Matter Searches

There is extensive experimental evidence supporting the existence of dark matter, much resulting from gravitational effects in the universe. Galaxy rotation is an example of this; the large velocity of the stars around the center of the galaxy cannot be explained only by the observed visible matter [14]. Additionally, gravitational lensing and galaxy formation give evidence for dark matter. The Lambda-CDM (cold dark matter) model in conjunction with experimental measurements suggests that dark matter comprises about 27% of the energy density of the universe, while ordinary matter is about 5% (and dark energy is 68%). However, assuming that dark matter is a particle with weak scale coupling, it has not yet been produced or directly observed as an elementary particle.

There are three main current searches strategies for dark matter — direct, indirect, and collider searches. These searches have probed dark matter particle masses above a few GeV, but none have found evidence for particle dark matter yet. Direct detection searches have placed bounds on the elastic scattering cross section of dark matter with nucleons, however, these experiments explore a small region of the particle physics models parameter space [15]. Indirect dark matter searches use astronomical observations of SM particles to detect dark matter annihilation or decay, and have constrained the annihilation cross section for weakly interacting dark matter particles [16].

One of the leading theories providing potential candidates for dark matter is supersymmetry (SUSY) [1], which proposes the existence of supersymmetric partners to the known particles in the Standard Model (which includes quarks, leptons, gauge bosons, and the Higgs boson). The Standard Model (SM) is the theory describing the fundamental particles and their interactions through the electroweak, strong, and weak forces. However, the SM has a number of shortcomings, in that it does not provide an explanation for gravitation, dark matter, or dark energy. SUSY is as an extension of the SM and can explain many of these observations.

A SUSY theory of dark matter is considered here, as the proton-proton collisions and interactions at the LHC provide a method to search for SUSY dark matter particles. Dark matter particles cannot be observed directly due to their weak coupling to Standard Model particles [14], so we search for dark matter in events with another signature, specifically delayed photons. If there is another SUSY particle that decays into the SUSY dark matter particle and photons, then the delayed photons provide a distinct signature with which to search for dark matter events.

Dark Matter SUSY Model

This thesis research analyzes dark matter production involving long-lived neutralinos, which the experiments at the LHC so far have not extensively considered. My analysis focuses on the proton-proton collision producing a pair of neutralinos ($\tilde{\chi}_1^0$) that decay to two photons and two gravitinos, shown in Figure 1.4. In this model, the neutralinos are a long-lived SUSY particle, and decay into the gravitino (another SUSY particle). The gravitinos are weakly interacting massive particle (WIMP) dark matter candidates [11]. The mass and lifetime of the neutralino intermediary is unknown, but it has a long lifetime, and is a heavy and therefore slow moving particle ($v \approx .7c$) [17]. The mass of the gravitino is assumed to be 1 GeV [18].

The neutralinos and gravitinos cannot be directly observed as they have a weak coupling to the SM particles, so the signature of this event is two delayed photons. These photons are considered “delayed” as they are not generated directly from the primary vertex, but rather originate from the neutralino decay vertex. Thus, the angle of incidence on the detector of the photons produced from a neutralino is distinct from photons produced at the primary vertex. If the photons had originated from the primary vertex, then they would be “prompt”, instead of delayed since the time to reach the detector would not be increased by the travel time of the neutralino. Selecting delayed photons is used to investigate dark matter events.

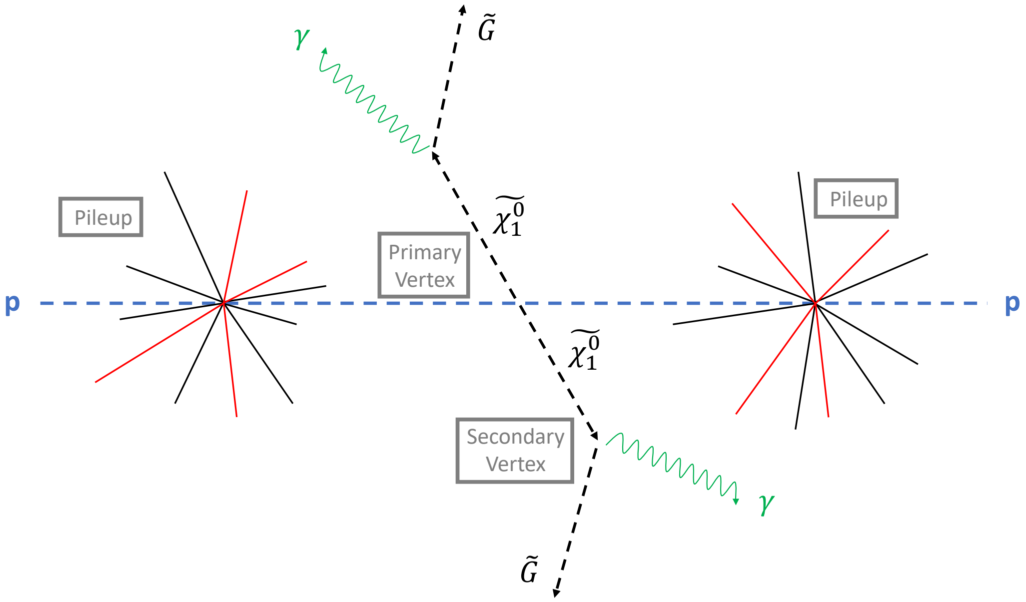


Figure 1.4: Diagram of the proton-proton collision model, where a pair of $\tilde{\chi}_1^0$ (neutralinos) are produced. Each neutralino decays to a gravitino and photon at the secondary vertex. The two final state gravitinos are the dark matter particles, and the delayed photons will be detected by the ECAL. Two pile up verticies are also illustrated.

In this model, the $\tilde{\chi}_1^0$ decays into a photon and gravitino (the lightest supersymmetric partner, LSP for this model). The observed particles are the two time delayed photons, along with missing transverse energy (MET) due to the escaping gravitinos. This exotic signature of the delayed photon will expand the search sensitivity to a broader range of dark matter models not included in the previous direct, indirect, and collider searches.

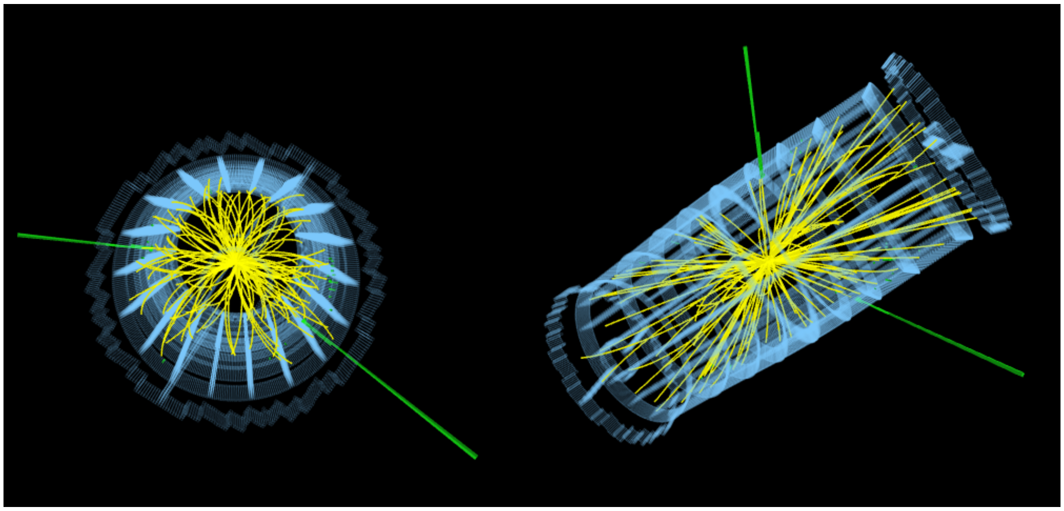


Figure 1.5: Event simulation demonstrating an event with two high transverse momentum photons, as is expected from the $\tilde{\chi}_1^0 \rightarrow \tilde{G} + \gamma$ model. The two green tracks are the photons in the ECAL, and the yellow tracks are showers in the inner tracker. Event simulation created with CERN CMS Open Data.

1.5 Searches for Physics Beyond the Standard Model

In an event where neutralinos are produced and decay to photons, the CMS detector would be able to measure the photon's energy, time of arrival, and momentum. However, the photons will be "time delayed", since they result from the neutralino ($\tilde{\chi}_1^0$) decay, not the original proton-proton collision vertex. This causes the total time of flight to be larger than expected from the collision vertex. If the mass or the lifetime of the neutralino is changed, then the time the photons are observed will be affected, along with other measured parameters (energy and momentum).

This particular motivating model of neutralinos decaying to photons has not been extensively considered at the LHC. Additionally, the ability to identify a delayed particle signature is vital for many beyond the Standard Model searches. Many SUSY theories predict massive long-lived particles decaying to non-prompt jets, with a very similar signature to the neutralino model considered in this research [19]. This signature is predicted from models with weak R parity violations, hidden valley models, and Z' with long-lived neutrinos [19]. Thus, the signature of two delayed photons is relatively model independent, and therefore it is crucial to focus on identifying the resulting delayed particles.

In this thesis, I investigate the SUSY dark matter model with simulation data in order to improve the photon and event identification algorithms, and investigate the impact of ECAL timing information on the photon identification.

Chapter 2 discusses the event simulation for the dark matter model, Chapter 3 discusses the specifics of the SUSY model, Chapter 4 presents the Boosted Decision Tree machine learning algorithm applied to the event identification problem, Chapter 5 compares the BDT performance to the current data analysis, and Chapter 6 discusses improvements in detector precision timing and its impact on similar searches.

Chapter 2

EVENT SIMULATION FOR DARK MATTER EVENTS

The neutralino event modeling is done with CMS Full Simulation using Geant4 [20] and CMSSW [21] [22]. Monte Carlo techniques are used to determine the shower propagation after each small time step, and Geant4 models the CMS detector response [23].

2.1 Pythia and Geant4

The Full Simulation is used to model the $p + p \longrightarrow \tilde{\chi}_1^0 + \tilde{\chi}_1^0$ and $\tilde{\chi}_1^0 \longrightarrow \gamma + \tilde{G}$ events (Figure 1.4). The neutralino ($\tilde{\chi}_1^0$) is the supersymmetery partner for the Z / Higgs boson, is uncharged, and is the lowest mass neutralino. The final decay product, $\tilde{G}$, the gravitino, is the lightest supersymmetric particle (LSP) for this model and therefore the stable dark matter candidate.

The neutralino ($\tilde{\chi}_1^0$) mass and lifetime and the gravitino ($\tilde{G}$) mass are set in Pythia [24] [25] and SLHA [26] files, and $q\bar{q} \longrightarrow \tilde{\chi}_1^0\tilde{\chi}_1^0$ is set to "on" for the simulation [27] (Appendix B). Then the collision and subsequent decay pathway are modeled, and Pythia models the parton showering from QCD processes. The Monte Carlo event simulation outputs a ROOT file [28], where the momentum, position, energy, particle type, ϕ, and η for each event are stored (Figure 2.1). This information is considered the "true" or generator level information, and will later be compared to the detector reconstruction to validate the simulation.

After the Geant4 simulation of particle interactions in the CMS detector, further event processing must occur [23]. The following processing steps are completed after the event simulation: GEN (generator events from Monte Carlo), SIM (reconstructing the energy deposit of simulated hits in the detector), DIGI (converting simulated hits into detector digitization), RECO (reconstructing particles and detector hits), and MiniAOD (used for analysis). After the data processing, the final ROOT file has both "true" (generator level) and reconstructed information. These are compared to validate the simulation and ensure events and timing information are processed and stored correctly.

2.2 Simulation Validation

Two of the parameters used for the validation are η and ϕ, which describe the spatial position of the particles in the ECAL detector. ϕ ranges from 0 to 2π, azimuthally around the beamline direction. η ranges from 0 to 5, with 0 perpendicular to the beamline (Figure 2.1) [29]. η is pseudorapidity, and can be expressed in terms of the total momentum and longitudinal momentum as

$$\eta = \tanh^{-1}\left(\frac{p_L}{|p|}\right) \tag{2.1}$$

η is a Lorentz invariant quantity, meaning that between each η interval, the same amount of energy is dissipated. Along with r (the azimuthal distance from the beamline), η and ϕ give full information about the particle position. This position information is used to validate the simulation and reconstruction, by comparing the reconstruction information to the generator "truth" information.

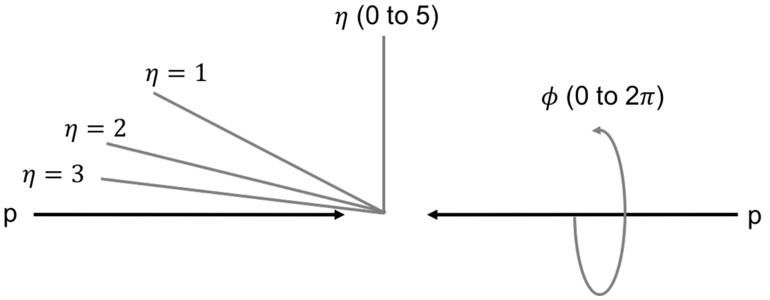


Figure 2.1: Diagram of η and ϕ, the variables used to label the position of the observed particle in the CMS detector. η ranges from 0 to 5, while ϕ is azimuthal and ranges from 0 to 2π.

Electromagnetic Calorimeter Detector Signals

Reconstruction algorithms are used to determine the particle type based on the energy deposit in the calorimeter. For photons, the reconstruction information comes from the CMS electromagnetic calorimeter (ECAL).

The ECAL is tiled by many crystals, and when a photon interacts in the lead tungstate crystal it scintillates [30]. A shower of electrons, positrons, and photons (EM shower) is formed, which then produces the scintillation light. The energy deposited in the lead tungstate excites the electrons in the crystal lattice, and when they decay to the ground state, scintillation light is emitted and then observed by the Avalanche PhotoDiodes (APDs). Energy from the EM shower then spreads to other ECAL crystals. This could also occur if the photon interacts before the crystal, and in this case, the radiated energy will also create a transverse spread in the ECAL crystals. One crystal, called the "seed" has most of the shower energy, and the timing information is found from the seed crystal. Each photon creates an electrical pulse in the ECAL, and the peak of this pulse is used as the timestamp. For photons, information about the ECAL crystal the photon hits and its time of arrival is stored.

For photons directed at the ECAL crystal, 94% of the energy of the photon energy is contained in the 9 crystals (3×3 grid) closest to the seed crystal, and 97% in the 25 closest crystals [11]. However, for delayed photons, there can be an angle between the incoming photon and the front face of the ECAL crystal, which will alter the shower shape and the energy spread. Thus, for delayed photons, the shower shape and ECAL timing information is useful for selection of events.

Understanding the change in the ECAL timestamp from prompt to delayed photons is crucial. This involves understanding how it is reconstructed, its time resolution, and how the timestamp changes with delayed photons. The best way to understand the ECAL timestamp is to compare the reconstructed timestamp with the true time from the Monte Carlo sample. For an individual crystal channel, a digital signal is taken every 25 ns, and a single pulse has 10 data points. The timing reconstruction is done by finding the maximum amplitude the weighted signal reached [11]. This is given by Equation 2.2 [11].

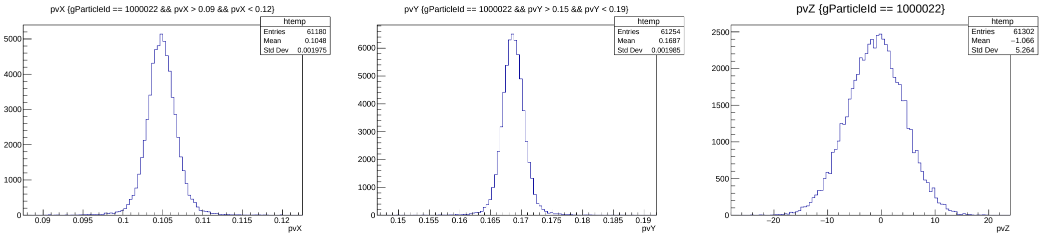
$$A = \sum_{i=0}^{N} w_i \times S_i \tag{2.2}$$

A is the weighted signal, N is the total number of crystals combined, w is the weight for each crystal, and S is the individual signal [11]. The time the weighted signal is maximized is used as the photon timestamp.

Generator Information

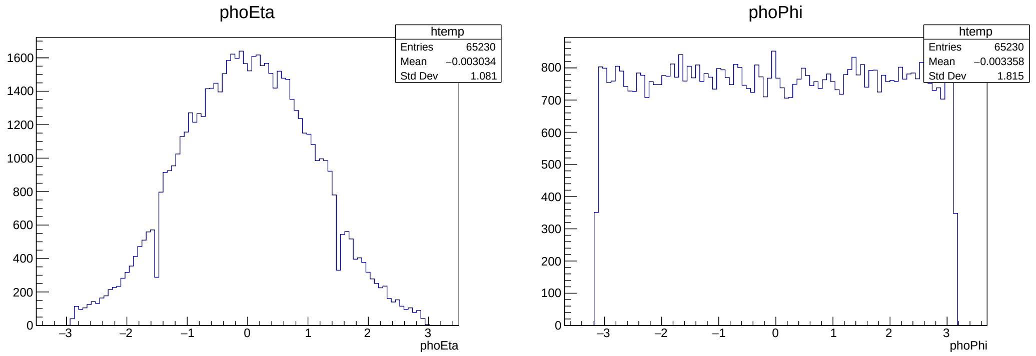
The generator level information for the primary vertex position (spatial x, y, z coordinates where the neutralinos are created), and the decay positions (x, y, z coordinates of where the two neutralinos decay and the two photons are created) is extracted from the simulation files. The primary vertex position is plotted in Figure 2.2. The spread in the primary x and y vertices is about 0.01 cm, and 5 cm in the primary z vertex, which corresponds to the beam spot size.

The decay vertex of the neutralino is found by tracing back the decay pathway and identifying the particles based on "ParticleMotherIndex" and "ParticleID" (particle



(a) Primary X vertex. (b) Primary Y vertex. (c) Primary Z vertex.

Figure 2.2: Primary X, Y, and Z vertices of the collision (and neutralino creation vertex), with the x axis units in cm. The vertices agree with the size of the beam — a 0.01 cm spread in X and Y, and a 20 cm spread in the Z, corresponding to the beam spot size (longitudinal direction, along the beam axis).



(a) η distribution. (b) ϕ distribution.

Figure 2.3: η and ϕ positions of the neutralino creation vertex. The distribution is as expected — equal around ϕ and η is focused around 0, with the declines at ±1.5 due to the intersection of the barrel and the endcap of the ECAL detector.

ID is 1000022 for $\tilde{\chi}_1^0$) [31]. The (x,y,z) of each neutralino's decay vertex are found in this way. Thus, the total lifetime of the neutralino can be calculated, and is used to validate the reconstruction. Based on the lifetime set (500 mm/c = 1.67 ns), the neutralino lifetime distribution should follow an exponential decay, with

$$\text{slope}^{-1} = \text{lifetime} \tag{2.3}$$

The time of flight (TOF) is calculated based on the distance between the generator primary and secondary vertices:

$$\text{distance} = \sqrt{\Delta_x^2 + \Delta_y^2 + \Delta_z^2} \tag{2.4}$$

$$\Delta_x = x_{secondary} - x_{primary} \tag{2.5}$$

where $x_{secondary}$ is the neutralino decay vertex, and $x_{primary}$ is the neutralino creation vertex. The velocity is given by $p = \gamma m v$, giving

$$v_{neutralino} = \frac{cp}{\sqrt{c^2m^2 + p^2}} \tag{2.6}$$

where p is the total momentum, and is related to the transverse momentum, p_T, by

$$p = p_T \cdot \cosh(\eta) \tag{2.7}$$

Therefore, the TOF of $\tilde{\chi}_1^0$ is $t = \frac{d}{v}$ in lab frame:

$$TOF_{\tilde{\chi}_1^0} = \frac{\sqrt{\Delta_x^2 + \Delta_y^2 + \Delta_z^2}}{pc}\sqrt{m^2c^2 + p^2} \tag{2.8}$$

For each of the two neutralinos produced in a collision event (a selection is made such that only events with two photons and two neutralinos are considered), the TOF is calculated. However, the direct result from the simulation is in the lab frame (t_{LF}), and the lifetime is set in the rest frame (t_{RF}) of the $\tilde{\chi}_1^0$ particle. To compare the results of the simulation with the lifetime set, the γ conversion factor is used:

$$t_{\text{RF}} = \frac{1}{\gamma}t_{\text{LF}} = t_{neutralino}\frac{m}{\sqrt{c^2m^2 + p^2}} \tag{2.9}$$

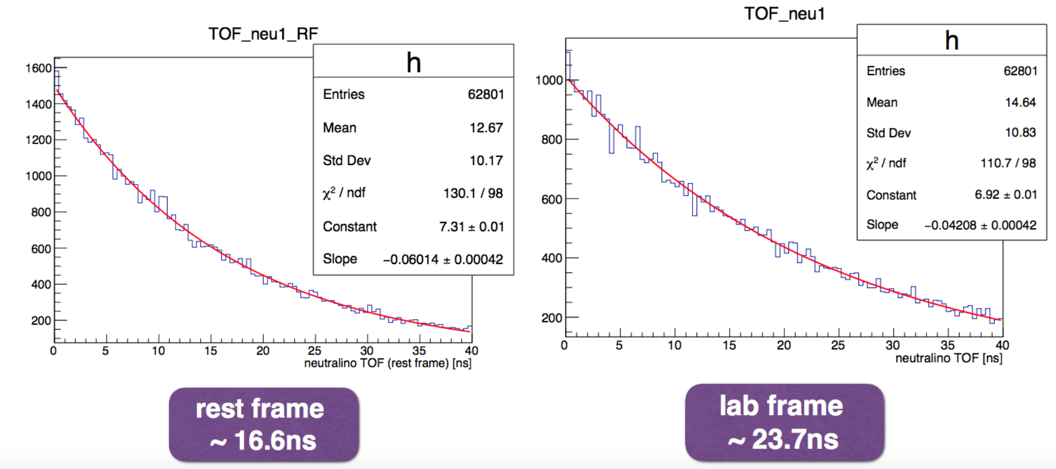
The TOF histograms exhibits the expected exponential decay as the calculated lifetime from the generator vertices is is 16.6 ns, matching the 16.67 ns set in the configuration file (or 23.7 ns in the lab frame). The result is shown in Figure 2.4.

The agreement between the calculated and set neutralino lifetime indicates that the neutralino lifetime is stored and processed correctly in the Full Simulation and Monte Carlo at generator level.

Reconstruction and Generator “truth” Total TOF Validation

In order to confirm that the event reconstruction is valid for this decay pathway, the total event timing information must be further validated. This is done by comparing the “truth” information (from the generator level) and the reconstruction information (based on the reconstruction algorithms from the ECAL detector) for the total time from proton-proton collision to photon detection in the ECAL.

From the generator level information, the primary vertex and secondary vertices are known. Therefore, the total generator TOF can be found by adding the neutralino creation time (this is distributed around 0 ns, and stored as part of the generator



(a) Neutralino TOF in the rest frame.

(b) Neutralino TOF in the lab frame.

Figure 2.4: TOF for the neutralinos, with an exponential fit for the rest frame and lab frame adjusted TOF. The lifetime of the neutralino was set as 16.67 ns in the configuration file (rest frame). This validates that the neutralino lifetime is stored correctly in CMSSW [21] [22].

level information), the neutralino TOF (calculation discussed in previous section), and the photon TOF. The photon that hits the crystal is matched with the correct neutralino by matching the photon momenta between generator and reconstruction level information. The neutralino TOF calculation is described above, with Equation 2.9, and the photon TOF calculation is similar. The photon TOF is given by

$$TOF_\gamma = \frac{\sqrt{\Delta_x^2 + \Delta_y^2 + \Delta_z^2}}{c} \tag{2.10}$$

$$\Delta_x = x_{secondary} - x_{crystal} \tag{2.11}$$

where $x_{secondary}$ is the photon creation vertex (equivalent to the neutralino decay vertex from Equation 2.5) and $x_{crystal}$ is the position of the ECAL crystal where the photon is observed.

A distance correction is applied, so TOF is corrected for the time it would take a photon to travel from the origin to the crystal where the photon is observed. Effectively, this shifts the graph to be centered 0, as a photon that was created at the origin would have a corrected time of 0. However, if the photon results from a neutralino traveling slower than c, then the TOF correction will result in an overall positive TOF, with larger shifts for larger neutralino lifetimes. Therefore, the corrected generator TOF calculation is

$$\mathrm{TOF} = T_{vertex} + TOF_{neutralino} + TOF_{photon} - \frac{\sqrt{x_{crystal}^2 + y_{crystal}^2 + z_{crystal}^2}}{c} \tag{2.12}$$

This time calculation corresponds to the time measurement the ECAL crystals output in real data in the actual ECAL detector. The time a photon is detected is reported as the difference of the length of time from the collision and the TOF of a photon traveling directly from the collision vertex to the ECAL crystal. Therefore, this time measurement examined in the simulation mimics the time given by the ECAL.

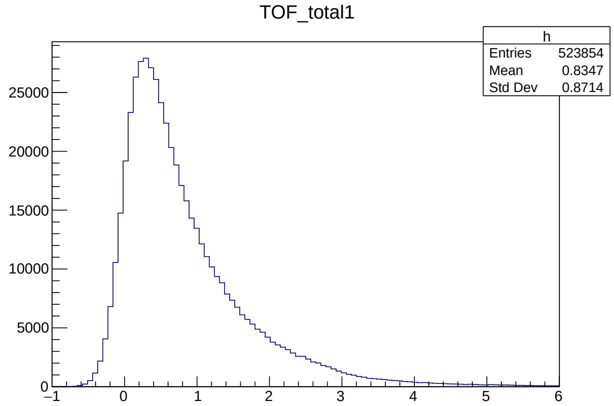


Figure 2.5: Total generator TOF for the gluino → neutralino → photon + gravitino decay pathway, in ns. This is the summation of the neutralino creation vertex time, TOF of the neutralino (in the lab frame), and TOF of the photon, and then the correction for origin to crystal time of flight is applied (Equation 2.12).

In the total TOF histogram in Figure 2.5, there are events with both positive and negative TOF. Positive events occur due to the neutralino traveling slower than the speed of light, and therefore the TOF correction is less than the sum of the particle's TOF. The negative events occur if the neutralino primary vertex is closer to the crystal where the photon is observed than $(0, 0, 0)$, since the origin is used for the TOF distance correction. If the neutralino is created away from the assumed $(0, 0, 0)$ origin, then the sum of the particle's TOF may be shorter than the assumed TOF correction between $(0, 0, 0)$ and the crystal. A plot of neutralino decay positions (equivalent to photon origin positions) is shown in Figure 2.6.

The generator arrival time and reconstruction arrival time is plotted against each other in Figure 2.7, and the expected linear dependence is seen.

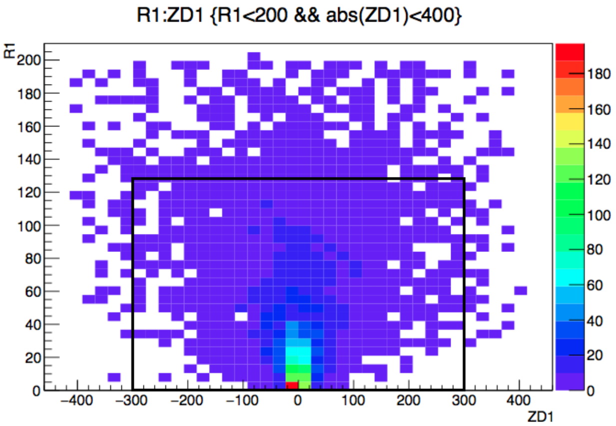


Figure 2.6: Neutralino decay position, with the ECAL barrel indicated in the black box, from a sample with neutralino lifetime of τ = 500 mm/c = 1.67 ns.

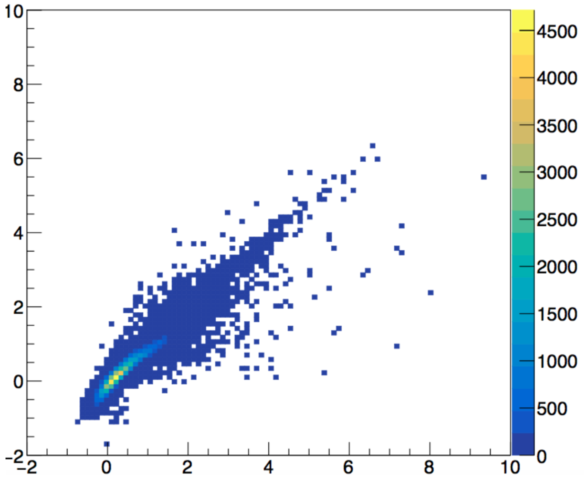
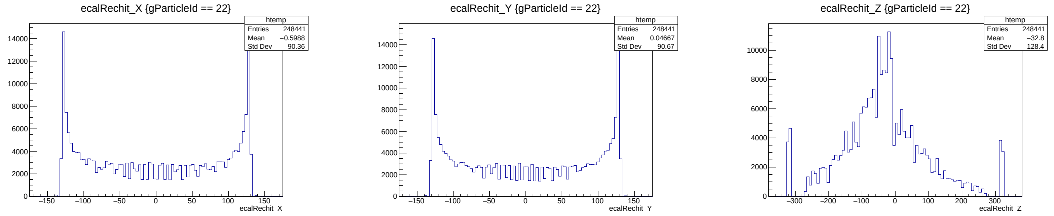


Figure 2.7: Reconstructed and generator corrected time for the photons. The expected linear dependence is seen, indicating that the reconstruction matches the generator information. Additionally, the time is shifted toward positive times as a result of the neutralino lifetime of $c\tau$ = 500 mm for this sample.



(a) Reconstructed X vertex of the photon on the ECAL detector.

(b) Reconstructed Y vertex of the photon on the ECAL detector.

(c) Reconstructed Z vertex of the photon on the ECAL detector.

Figure 2.8: The position the reconstructed photon hits the ECAL detector. The barrel is 129 cm in radius, and from -300 to 300 cm in length.

2.3 Time of Flight (TOF) Validation

To compare the generator and reconstruction photon time information, generator and reconstruction photons are matched. Photon matching is done based on a requirement on small ΔR and Δp_t, given by

$$\Delta R = \sqrt{\Delta\phi^2 + \Delta\eta^2} \tag{2.13}$$

$$\Delta p_T = p_{T_1} - p_{T_2} \tag{2.14}$$

The ΔR photon matching was cross-checked with the photon energy, momentum, and position. ΔR is calculated based on the difference in the reconstruction photon seed hit position (known to be accurate) and the calculated x, y, z position the generator photon hits the detector. The position the generator photon hits the ECAL is found from solving for the intersection of the photon's path of flight and the inner radius of the ECAL cylinder (Figure 2.9).

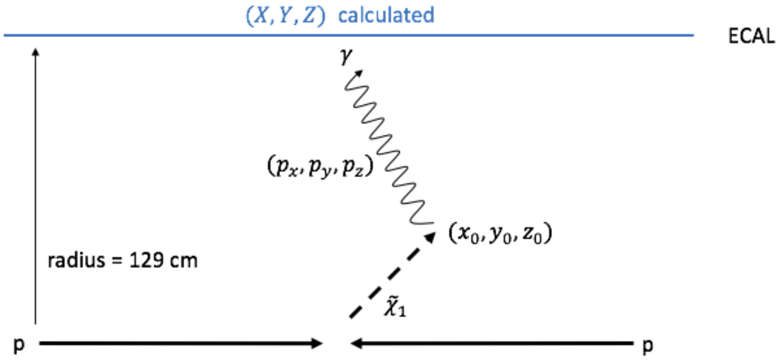


Figure 2.9: Figure demonstrating how the x, y, z position of the generator photon is calculated. The information used is the inner radius of the ECAL barrel (assuming the photon is measured at the first contact with the barrel), the photon p_x, p_y, and p_z values, and the neutralino decay x_0, y_0, z_0 position.

This generator position information relies on the assumption that the photon is measured as soon as it reaches the ECAL inner radius, instead of further inside a crystal. This assumption is accurate enough to use for matching of reconstructed and generator photons.

Event Selection

Events are selected where the photon decays within the ECAL barrel, and cuts are made on ΔR and Δp_T to match reconstructed and generator photons. These selection cuts are:

$$\text{ECAL radius} = \sqrt{\Delta x^2 + \Delta y^2} \leq 1.29 \text{ m} \tag{2.15}$$

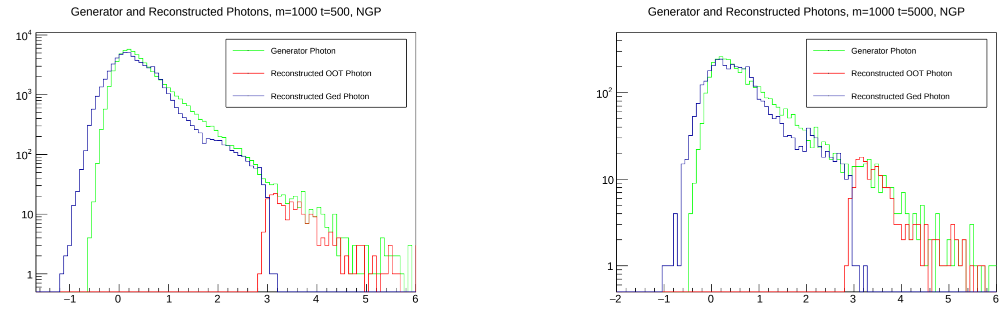
$$|z| \leq 3.0 \text{ m} \tag{2.16}$$

$$\Delta R < 0.02 \tag{2.17}$$

$$\Delta p_T < 5 \text{ GeV} \tag{2.18}$$

The generator and reconstructed photon time information is expected to match. This is confirmed by overlaying the TOF plots for a matched generator and reconstructed photons. The reconstructed photons are tagged for delayed (“out of time” or OOT) and standard photons. In the following plots, it can be seen that the standard + delayed reconstructed photons sum to the generator photon distribution, indicating that the reconstructed information matches the generator information.

One sample used is a simple $\tilde{\chi}_1^0 \longrightarrow \tilde{G} + \gamma$, using a neutralino mass of 1000 GeV and $c\tau$ = 500 mm (1.67 ns) and $c\tau$ = 5000 mm (16.7 ns). Results are plotted in Figure 2.10.



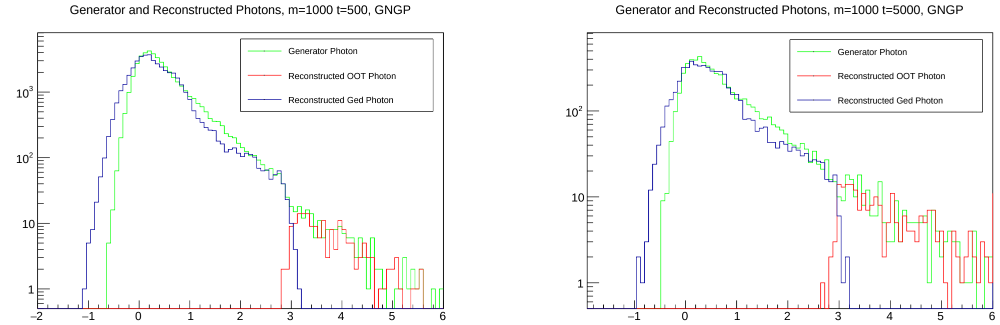
(a) The neutralino mass is 1000 GeV and the lifetime is $c\tau$ = 500 mm (1.67 ns).

(b) The neutralino mass is 1000 GeV and the lifetime is $c\tau$ = 5000 mm (16.7 ns).

Figure 2.10: Reconstructed TOF information compared to the generator time information. This is a sample with a simple decay of $\tilde{\chi}_1^0 \longrightarrow \tilde{G} + \gamma$, and the reconstruction matches the generator information.

A similar sample of $\tilde{g} \longrightarrow \tilde{\chi}_1^0 \longrightarrow \tilde{G} + \gamma$ was also used, with the same parameters for the neutralino. A gluino is the supersymmetric partner of a gluon, and is hypothesized to be pair produced in the LHC. Thus, this model is another possibility of neutralino production after a proton-proton collision. The results of the generator and reconstruction timing information are shown in Figure 2.11.

These plots demonstrate that the generator and reconstruction timestamps match well, and the reconstruction has been validated.



(a) The neutralino mass is 1000 GeV and the lifetime is $c\tau = 500$ mm (1.67 ns).

(b) The neutralino mass is 1000 GeV and the lifetime is $c\tau = 5000$ mm (16.7 ns).

Figure 2.11: Reconstructed TOF information compared to the generator time information. This is a sample with a $\tilde{g} \longrightarrow \tilde{\chi}_1^0 \longrightarrow \tilde{G} + \gamma$, and the reconstructed distribution of times shows a high level of agreement with the generator distribution.

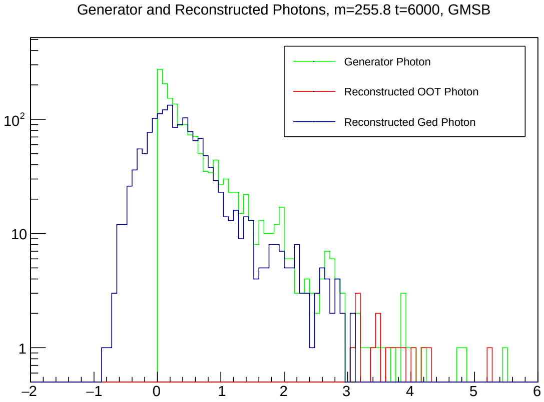


Figure 2.12: Reconstructed time information (flagged for standard and out of time photons) compared to the generator time information. The reconstructed distribution of times shows a high level of agreement with the generator distribution. This is a GMSB sample, where after the initial proton-proton collision, a neutralino can be created from any of 30 mother particles, and then subsequently decays to a photon (time delayed) and gravitino (discussed further in Chapter 3). For the GMSB model, the generator time does not have any events with a negative time. This is due to the simulation not saving the generator vertex time (the time when the collision occurs), so it is assumed to be 0 and at the origin. Therefore, when the time of flight correction is applied, there are only events with positive times.

Understanding the neutralino masses and lifetimes that ECAL is capable of detecting is important — in particular, the ECAL is limited due to its physical size. Long lifetimes may not be detected, as the neutralino can travel outside of the ECAL barrel before decaying, or the photon event will be treated as noise if it is too close in time to the next photon bunch crossing. This leads to a lower efficiency with longer lifetime samples.

The Monte Carlo simulations have been validated and shown to generate events correctly. With this method, an analysis for time delayed photons indicating dark matter particles can be done, and we are now well posed to carry out the delayed photon analysis and perform a comparison to a 2017 LHC data analysis.

Chapter 3

LONG-LIVED NEUTRALINO SUSY DARK MATTER MODEL

The Standard Model (SM) of particle physics describes all the fundamental particles we have observed, and it describes three of the four fundamental forces — weak, strong, and electromagnetic force. However, there is a large amount of physics not explained by the Standard Model, such as gravity and dark matter. In order to expand the current model to explain these missing elements, the theory of sypersymmetry (SUSY) has been used [1].

In SUSY, all SM particles are given a partner particle, and there are a number of other possible particles — some of which are viable particle dark matter candidates. To explain the cosmological evidence for dark matter, the dark matter particle would need to have a super-weak coupling to normal matter and have a mass around or less than 100 GeV [2]. The proposed particle is called a *weakly interacting massive particle* (WIMP) dark matter candidate.

3.1 SUSY and Symmetry Breaking

The minimal extension of the SM to include SUSY is the Minimal Supersymmetric Standard Model (MSSM). The MSSM gives the minimum particle content for SUSY and the SM, and adds parameters to account for the additional particles and their interactions [18]. However, MSSM is agnostic to how the symmetry is broken, and therefore there are a number of possibilities for supersymmetry breaking. Each leads to different phenomenologies and particle cross sections.

Three of the symmetry breaking options are gauge mediated supersymmetry breaking (GMSB), gravity mediated supersymmetry breaking, and anomaly mediate supersymmetry breaking. The model relevant for this work is GMSB.

Gauge Mediated Supersymmetry Breaking

Gauge mediated supersymmetry breaking (GMSB) models involve a number of “sectors”. The *observable sector* has all the SM particles and their SUSY partner particles [32]. The *secluded sector* mediates the symmetry breaking, and the *messenger sector* affects the model phenomenology [32]. In GMSB models, the gravitino is the lightest supersymmetric particle (LSP), and the neutralino is the next-to-lightest supersymmetric particle (NLSP) [32]. Assuming R parity conservation, the neutralino is expected to decay into the gravitino, which is a spin 3/2 particle [33].

The gravitino is formed from a spin 1/2 fermion (created from the spontaneous SUSY breaking) coupling to gravity [32]. The gravitino has a SUSY breaking mass as well, and is assumed to be 1 GeV. Additionally, the gravitino is the supersymmetric partner particle of the graviton, proposed to mediate the gravitational force. Since the gravitino is the LSP, it cannot decay to anything else (since it would have to decay to a lighter SUSY particle and a SM particle) [2]. Therefore, the gravitino ($\tilde{G}$) is a stable particle and a valid dark matter candidate.

In SUSY searches, the focus is on detecting the lighter particles (NLSP or LSP). In collider physics, the initial particles are produced by strong production. This produces quarks, squarks, and gluinos. These particles then decay to particles with a lower production cross section, such as third generation squarks and electroweak particles [18]. The neutralino is an electroweak particle, as it is a mix of SUSY gauge bosons (higgsino, bino, and 2 winos).

Previous Dark Matter Collider Searches

In many SUSY models besides GMSB, the neutralino is the LSP, and is a stable non-baryonic dark matter candidate [33]. Most previous collider searches have searched for the neutralino as the LSP. Therefore, the models where the neutralino is long-lived and decays to the LSP (the gravitino) have been relatively under-explored. The unique signature of delayed particles from the GMSB model would have been missed by analyses focusing on the neutralino as the LSP, and therefore this is a very interesting region of phase space to investigate further.

Previous GMSB Searches

The signature of a GMSB model in a collider search heavily depends on the lifetime of the intermediate neutralino particle. For this search, we consider the particularly interesting range of lifetimes where the neutralino decays inside the ECAL, and therefore the decay products are candidates for detection.

At the Large Electron Positron (LEP) collider, searches were performed for diphotons and missing energy events produced by neutralinos [34]. This excluded $\tilde{\chi}_1^0$ masses below 73 GeV (given a lifetime such that the neutralino decays inside the detector) [32]. However, this analysis was done at $\sqrt{s} = 172$ GeV [32], whereas the current LHC is running at $\sqrt{s} = 13$ TeV. Neutralino masses greater than 73 GeV are considered for this analysis, and the increased collision energy of the current LHC also expands the parameter space available for this search.

3.2 GMSB Signal

The GMSB signal event for the $\tilde{\chi}_1^0 \rightarrow \tilde{G} + \gamma$ decay is from a model where the neutralino results from decay of any of 30 individual particles. These intermediary particles are 12 squarks, 9 sleptons, 3 neutralinos, 2 charginos, a gluino, and 3 Higgs particles:

- Left and right squarks: $\tilde{u}, \tilde{d}, \tilde{s}, \tilde{c}, \tilde{b}, \tilde{t}$
- Left and right electron, muon, tau: $\tilde{e}^-, \tilde{\mu}^-, \tilde{\tau}^-$
- Left neutrinos: $\tilde{\nu}_{eL}, \tilde{\nu}_{\mu L}, \tilde{\nu}_{\tau L}$
- Neutralinos: $\tilde{\chi}_2^0, \tilde{\chi}_3^0, \tilde{\chi}_4^0$
- Charginos: χ_1^+, χ_2^+
- Gluino: $\tilde{g}$
- Higgs: $H^0/H_2^0, A^0/H_3^0, H^+$

However, the signature of the event is always two delayed photons, regardless of the neutralino origin — the delayed photons are a *model independent* signature. Thus, the search for SUSY neutralino events focuses on identification of events with displaced and delayed photons. Since the photons are time delayed and approach the ECAL with a unique angle, they provide a distinct event signature.

3.3 Neutralino Model Backgrounds

As in any search, the backgrounds must be considered. “Backgrounds” refers to the numerous SM events and pileup that occur from the interactions in a proton-proton beam crossing. Many of the pileup interactions are “soft” interactions that are already understood and do not contain new physics [13] — however, they still hadronize and create jets that are observed by the detector. This can make identifying signal events difficult, as the contributions to the overall detector response are primarily from background events.

For the search for long-lived neutralinos decaying to gravitinos and photons, the main background sources are QCD (quantum chromodynamics) and γ + jets events (also a QCD process). The QCD events can have jets that are reconstructed similar to photons, so these are *fake* photons from mis-measurement. The γ + jets produces *prompt* photons, while the dark matter model produces *delayed* photons. The two types of background (QCD and γ + jets) and the signal (neutralino) events are simulated with Monte Carlo methods. Using the simulated events, the detector information about the energy deposit, photon momentum, and shower shape are plotted for these three models to understand how the background and signal events can be distinguished.

The γ + jets and QCD simulation events were produced in HT bins (scalar sum of transverse momentum) to ensure that there are enough statistics for the analysis. Higher momentum bins are less likely to be produced, but producing the events in bins and then weighting the samples provides sufficient statistics for the analysis.

3.4 Photon Variables

A related CMS Run 1 analysis [11] focused the photon identification on the photon incident angle and elliptical energy spread in the ECAL. Since the signal photons are not from the primary collision vertex, they are expected to have a broader energy spread and a larger incidence angle as compared to standard photons. However, there are a number of variables that the Run 1 analysis did not consider, so these additional pieces of information from the detector are considered here. These variables are:

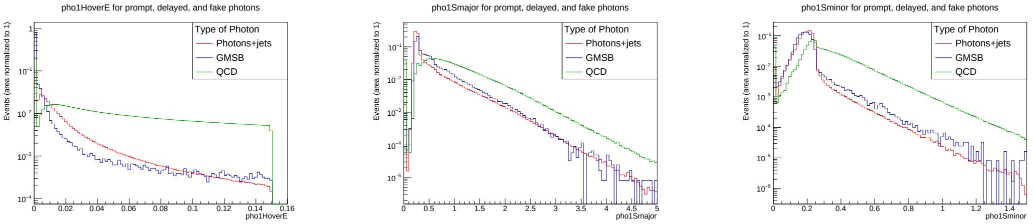
- S_{major}: the semi-major axis of the elliptical ECAL energy deposit of a photon
- S_{minor}: the semi-minor axis of the elliptical ECAL energy deposit of a photon
- $\sigma_{i\eta i\eta}$: measures the width of the shower in the η direction
- HoverE: the ratio of the hadronic energy to the electromagnetic energy deposited in the calorimeters per photon
- p_T^{ch}: the sum of the transverse momentum of the charged hadrons, given by particle flow
- E_T^n:: the sum of the transverse energy of the neutral hadrons, given by particle flow
- E_T^γ: the sum of the transverse energy of the photons from an event, given by particle flow

- R9: the ratio of energy deposited in a 3 × 3 crystal square around a the photon seed crystal in the ECAL to the energy deposited in the larger ECAL supercluster

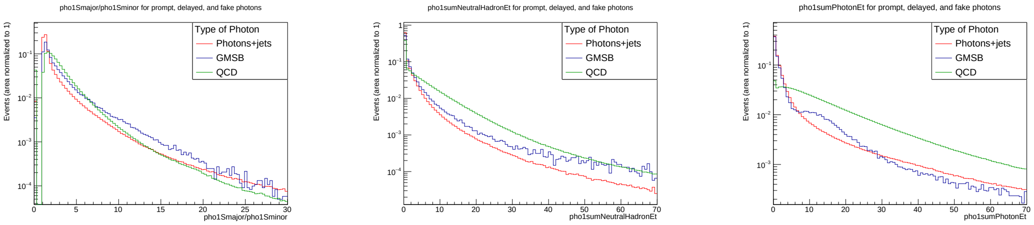
- $\gamma_{isolation}$ (pho1ecalPFClusterIso): the cluster isolation of the shower in the ECAL, given by particle flow

- p_T^{cone} (pho1trkSumPtHollowConeDR03): the sum of the transverse momentum of particles in a given cone within the tracker (before ECAL)

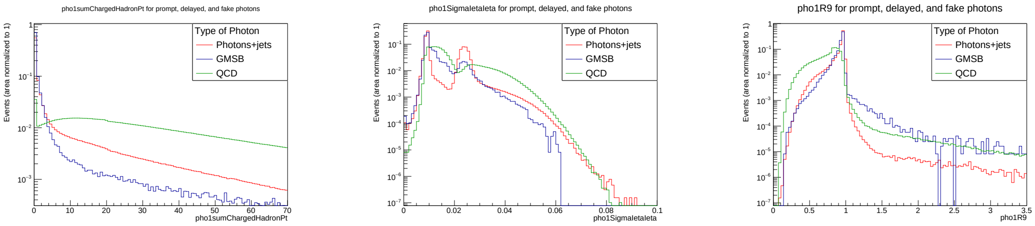
and their distributions for GMSB, γ + jets, and QCD events are plotted in Figure 3.1.



(a) HoverE (b) S$_{major}$ (c) S$_{minor}$



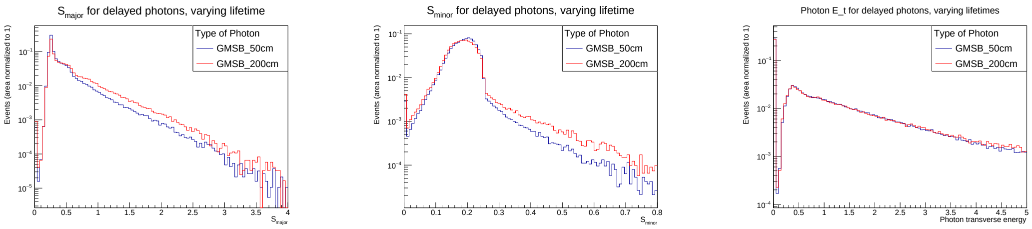
(d) S$_{major}$ / S$_{minor}$ (e) E_T^n (f) E_T^γ



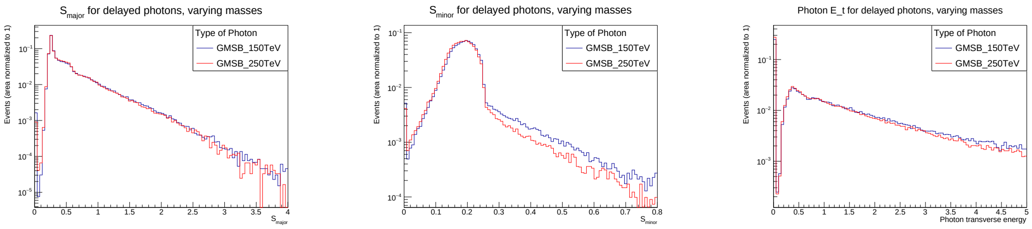
(g) p_T^{ch} (h) $\sigma_{i\eta i\eta}$. (i) R9

Figure 3.1: Variables for QCD, γ + jets, from the GMSB neutralino dark matter decay model, with a neutralino mass of 358 GeV (Λ=250 TeV) and a lifetime of $c\tau$ = 200 cm. All plots have a log y scale, and the areas under the curves are normalized to 1.

In addition, the change in these distributions with different neutralino masses and lifetimes is reasonably small — with changing mass and lifetime, approximately the same distribution is seen for the GMSB signal events. This is shown in Figure 3.2.



(a) S_{major} with two lifetimes (b) S_{minor} with two lifetimes (c) E_T^{γ} with two lifetimes



(d) S_{major} with two masses (e) S_{minor} with two masses (f) E_T^{γ} with two masses

Figure 3.2: Three of the photon ID variables shown with different neutralino lifetimes and masses. The two neutralino lifetimes are $c\tau$ = 50 cm and $c\tau$ = 200 cm (1.7 and 6.7 ns), and the two neutralino masses plotted are 212 and 358 GeV (150 and 250 SUSY Λ scale). These are representative of the other variables, where little change in the ID variable distribution is seen with a change in the lifetime of mass.

3.5 Photon Identification

Particularly, we focus on features that can discriminate between the signal photons (resulting from neutralino decay) and the background photons (from QCD and γ + jets). However, from the plots of potential photon ID variables on signal and background samples in Figure 3.1, there is not a clear discrimination between the two. Therefore, with the current detector reconstructions and the results from these plots, there are not clear optimal cuts to discriminate between neutralino signals and backgrounds. In order to better optimize the selection cuts for discriminating between the background and signal events, a Boosted Decision Tree is used.

Chapter 4

PHOTON IDENTIFICATION USING BOOSTED DECISION TREES

A Boosted Decision Tree (BDT) is a machine learning algorithm for classification problems, using an ensemble of decision trees [35]. A BDT is investigated here for delayed photon event identification.

4.1 Boosted Decision Trees

A decision tree is a flowchart of the event classification and gives final predictions based on the outcomes of each decision on the input variables. A variable is an event level feature, and every event has numerical values for each variable. Specific variables are selected as inputs to the BDT, and these are used in the classification process. Decision trees are often represented as a set of “if-then” rules indicating how an individual event will be classified [36]. Each node or branch point in the tree (indicated by circles in Figure 4.1) is determined from selections on an input variable to the tree. The decisions are based on if a given variable of an input event is larger or smaller than the decision value at the node. Decision trees are generally constructed top down, starting with a node for the attribute that is found to be most effective at distinguishing the training data [36]. Similarly, an event is classified starting at the initial node of the tree, and then proceeds by following the branches corresponding to the decision made at each node [36]. The final outcome of the tree is the probability that an event is part of the signal.

The BDT relies on an ensemble of diverse and accurate classifiers (multiple decision trees). This allows the machine learning method to be highly accurate since the individual decisions of each classifier are weighted in the final result [37]. The ensemble classification method has been shown to have significant improvement in prediction accuracy over regular decision tree models [38].

The BDT is an example of a boosted classification algorithm: it generates a sequence of decision trees, each dependent on the previous ones. The BDT is a type of Adaboost algorithm, where weights on the training data set are adjusted to improve the classification of events initially misclassified in the first decision tree [37].

BDTs are a type of supervised learning, since the input is labeled data sets [39]. The

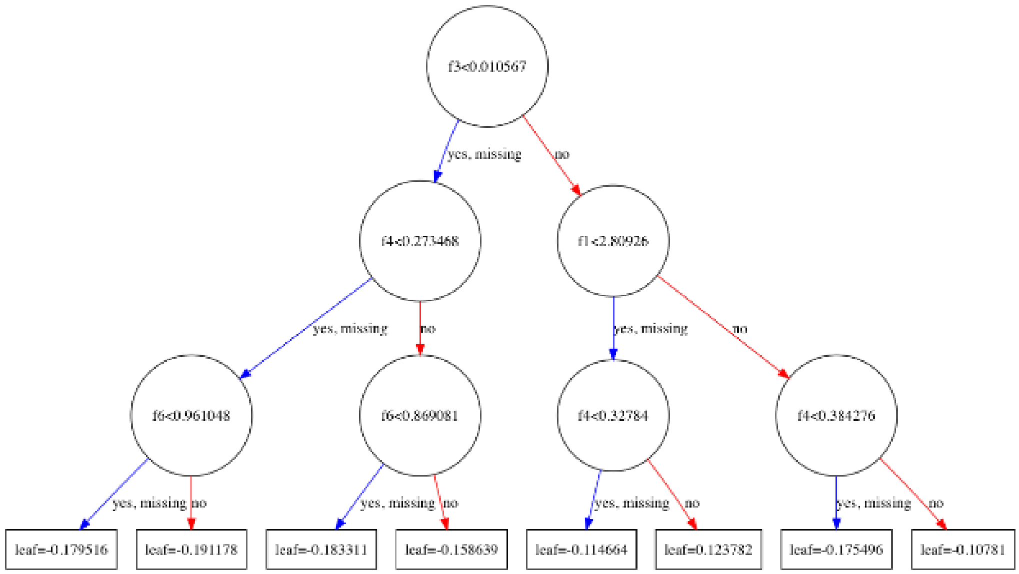


Figure 4.1: Example of a decision tree. Each node is labeled with the decision (cut on a variable) being made.

purpose of a BDT is to create a model that predicts an outcome given input data, and this prediction is based on parameters learned from the input variables. This is done by optimizing an objective function that estimates the ability of the model to correctly predict the outcome of the training data. The objective function is defined as:

$$\text{obj}(\theta) = L(\theta) + \Omega(\theta) \qquad [39] \tag{4.1}$$

$L(\theta)$ represents the loss function that is minimized, similar to minimizing an entropy [39]. The loss function is often the mean squared error between the correct output and the predicted output, $\sum_i (y_i - \hat{y}_i)^2$ [39]. $\Omega(\theta)$ is a regularization term, with penalties for large numbers of nodes or highly oscillatory weights (the real number output on each leaf). Regularization is used to limit over-training, which is when the model fits too precisely to the input training data, and is therefore a bad estimator of the outcome of other input data. Preventing over-training is done by re-sampling the data and limiting the depth and number of nodes — which makes the performance of the classifier worse on the training data, but better on test data.

The BDT creates multiple trees and each tree re-samples the data and variables. The results of each tree are added in decreasing importance, weighted by $\alpha < 1$:

$$f(x) = f_0(x) + \alpha f_1(x) + \alpha^2 f_2(x) + \cdots \tag{4.2}$$

This is called an ensemble model, and for the final classification result the predictions of the individual trees are summed (weighting controlling their relative contributions) [39]. This means the objective function being maximized becomes a sum as well:

$$\text{obj}(\theta) = \sum_{i}^{n} L(y_i, \hat{y}_i) + \sum_{k=1}^{K} \Omega(f_k) \qquad [39] \qquad (4.3)$$

where $\hat{y}_i$ is the prediction given input x_i, and y_i is the true outcome (labeled from the input training data). n is the number of input events, so the loss function is summed over all events. K is the number of trees, and f_k is a function of the space of all classification trees. The regularization term is summed over all decision trees [39].

Each tree outputs a real number on every leaf, which is converted to a probability of the input event being signal. This is done by using a logistic function to map the real line onto the interval $0 \rightarrow 1$, which can now be interpreted as a probability. The logistic function, or sigmoid curve, is:

$$f(x) = \frac{L}{1 + e^{-k(x-x_0)}} \qquad (4.4)$$

where x_0 is the midpoint, L is the maximum value, and k is the steepness [40].

Essentially, the Boosted Decision Tree combines weak variables (the variables plotted in Figure 3.1) into an aggregate strong selection created from a combination of input variables.

Receiver Operating Characteristic Curves

From the performance of the testing data on the BDT algorithm, a ROC (Receiver Operating Characteristic) curve is made, and used to evaluate the discrimination abilities of the algorithm between signal and background samples [41]. On a ROC curve, the true positive rate (signal acceptance) is plotted against the false positive rate (background acceptance). In order to quantify BDT performance, the area under the ROC curve is used as the figure of merit to evaluate the model performance [41]. An integral of 1 corresponds to perfect signal and background discrimination, while an integral of 0.5 is random event selection. For this analysis, the area under the ROC curve is not explicitly calculated, but is used to qualitatively compare the performance of BDT models.

True positives are when a positive event (signal) is correctly classified as positive; a false positive is when the negative event (background) is classified as positive. Similarly, true negatives are when a negative event (background) is correctly classified as negative; a false negative is when a positive event is classified as negative [41]. Then the true and false positive rates are (from [41]):

$$\text{true positive rate} = \frac{\text{correctly classified positive events}}{\text{total positives}} \tag{4.5}$$

$$\text{false positive rate} = \frac{\text{incorrectly classified negative events}}{\text{total negatives}} \tag{4.6}$$

The ROC curve is the true positive vs. false positive rate, so depicts how much incorrect classification must be accepted in order to achieve a higher correct classification. This means that $x = y$ corresponds to random classification, and points approaching (0, 1) have a higher true positive rate.

Each point on the ROC curve is found from signal and background distributions over a discriminator. The discriminator ranges from 0 to 1, and is an aggregate variable. As the cut on the discriminator value decreases from 1, the signal and background acceptance on the ROC curve increases (higher x and y values). When the discriminator value is at 0, all of the signal and background have been accepted (top right corner of the ROC curve). If the discriminator distributions of signal and background have little overlap, then a discriminator cut separates the signal and background distributions well and the ROC curve is closer to the top left corner, as shown in Figure 4.6.

4.2 XGBoost

For this classification problem, the XGBoost package was used [42]. XGBoost, or Extreme Gradient Boosting, is a machine learning algorithm that minimizes a loss function by updating the prediction using the negative gradient of the loss (called steepest-descent). The algorithm for gradient boosting is described in detail in Friedman's paper *Greedy Function Approximation: A Gradient Boosting Machine* [42].

For this research, 60% of the Monte Carlo data set is used to train the BDT, and the remaining 40% is used to test the algorithm the BDT found to be optimal. The train-test split is done to avoid over-training, where the algorithm trains too specifically on a data set and focuses on statistical fluctuations that are too specific for the broader data of interest. Additionally, the depth of the trees is set at 6, the learning rate α is 0.3, and the number of trees is set to 100 (these are the default values for XGBoost). The learning rate is a factor applied for each new tree added in the model — meaning that each additional tree is less important in the overall classification (Equation 4.2).

4.3 BDT for Photon Identification

The data sets are from QCD, γ + jets, and the neutralino Monte Carlo simulations, and each event is labeled according to which process it resulted from.

The BDT algorithms for photon selection are optimized to distinguish GMSB signal photons from backgrounds (real photons from γ + jets, or jet activity from QCD recorded as photons). A two way BDT is used, meaning that the signal sample can only be compared against one of the background samples at a time. The comparisons performed are signal vs. QCD, signal vs. γ + jets, and QCD vs. γ + jets. In the case of QCD vs. γ + jets, the QCD is considered background and γ + jets is considered signal. The signal sample (GMSB) is a neutralino to gravitino and photon decay, with the mass of the neutralino set by the SUSY mass scale $\Lambda = 250$ TeV and the lifetime as $c\tau = 200$ cm.

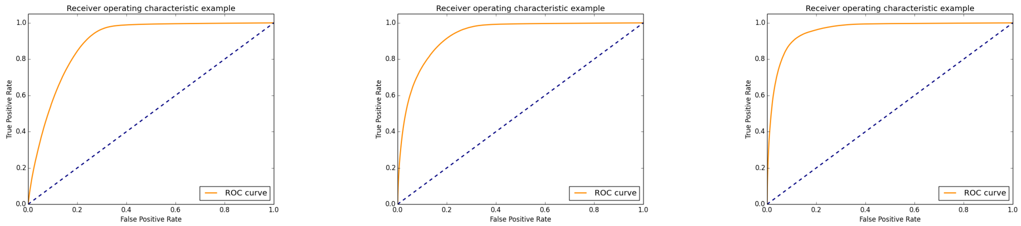
4.4 BDT Classification Results

The variables considered in the BDT are S_{major}, S_{minor}, $\sigma_{I\eta I\eta}$, the hadronic to electromagnetic energy ratio (HoverE), transverse momentum of charged hadrons (p_T^{ch}), transverse energy of neutral hadrons (E_T^n), transverse energy of photons (E_T^{pho}), and the energy deposit spread ratio R9 (variables plotted in Figure 3.1).

Some variables are calculated from Particle Flow, which is the data reconstruction providing a complete global event description, using information from the ECAL, HCAL, and trackers [8]. Particle Flow uses basic information from all detector layers of CMS to identify particles and reconstruct particle properties. This allows for a holistic event reconstruction, and the combination of information from each detector allows for accurate energy measurements.

As variables are added to the BDT, the discrimination abilities improve, as demonstrated for QCD vs. γ + jets in Figure 4.2. The increase in area under the ROC curve indicates that the BDT is an effective method of combining weak selections (the observables from the CMS detectors) into a strong aggregate selection. The relative contributions of each variable to the model are plotted in Figure 4.3.

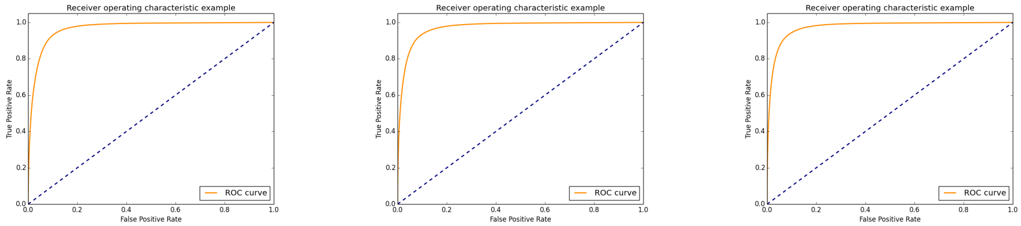
The background (QCD and γ + jets) samples are produced in bins according to the transverse momentum in order to have adequate statistics in each bin. This is done since higher transverse momentum bins have fewer events, but these regions of the backgrounds are important, and this method insures there are enough events in all bins. For the BDT, the background files for all bins are combined such that the overall background can be separated from the signal.



(a) Variables: p_T^{ch}, E_T^n, E_T^{pho} (all calculated from particle flow)

(b) Variables: added HoverE

(c) Variables: added $\sigma_{I\eta I\eta}$



(d) Variables: added S_{major}

(e) Variables: added S_{minor}

(f) Variables: added R9

Figure 4.2: ROC curves for QCD vs. γ + jets discrimination, successively adding more variables. As more variables are added, the discrimination abilities of the BDT are improved, as seen by the increase in area under the ROC curve. Neutralino mass is set as $\Lambda = 250$ TeV and the lifetime as $c\tau = 200$ cm.

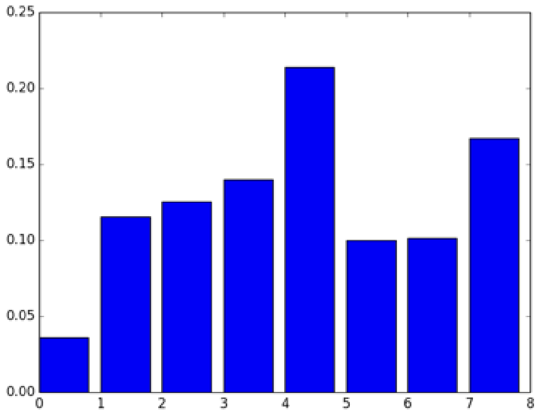


Figure 4.3: Relative importances of each variable for the QCD vs photons + jets BDT. This is for the variables p_E^n, p_T^{ch}, p_T^{pho}, HoverE, $\sigma_{I\eta I\eta}$, S_{major}, S_{minor}, and R9. This corresponds to the ROC curve shown in Figure 4.2f.

4.5 EGamma Cut Comparison (CMS 2016) and Data Analysis Comparison

The performance of the BDT is compared to the EGamma cut based photon ID developed in 2016 by CMS, and also to the current delayed photons data analysis (cut based photon ID). This is done in order to understand the improvement of the BDT over the cut based selection methods. Cut based methods select delayed photon events based on cuts on variables.

EGamma Comparison

First, the BDT is compared against the cuts developed for 2016 [1] by CMS [43]. This was done for ECAL barrel photons, and the efficiency (number of QCD or γ + jets events passing the cuts / total events) of the selection cuts is plotted on the BDT in Figure 4.4.

A table of the CMS developed cut based ID (referred to as EGamma cuts) is listed in Table 4.1.

Background rejection	**Loose (83%)**	**Medium (87%)**	**Tight (89%)**
H over E	0.0597	0.0396	0.0269
$\sigma_{i\eta i\eta}$	0.01031	0.01022	0.00994
ρ corrected PF charged hadron isolation	1.295	0.441	0.202
ρ corrected PF neutral hadron isolation	$10.910 + 0.0148p_t$ $+0.000017p_t^2$	$2.725 + 0.0148p_t$ $+0.000017p_t^2$	$0.264 + 0.0148p_t$ $+0.000017p_t^2$
ρ corrected PF photon isolation	$3.630 + 0.0047p_t$	$2.571 + 0.0047p_t$	$2.362 + 0.0047p_t$
ρ corrected neutral hadron (E_t^n)	6	6	6
Photon S_{minor}	>0.3	>0.3	>0.3
Photon S_{minor}	<0.15	<0.15	<0.15

Table 4.1: Table of cuts for the CMS 2016 EGamma developed photon ID. PF refers to Particle Flow, the reconstruction giving a global event description. Only the tight cuts are used. These cuts are from [43].

Current Data Analysis Comparison

Similar cuts have been developed for the current delayed photon cut based photon ID data analysis. These cuts are also used to compare to the BDT performance to discriminate between the delayed photons and the QCD or γ + jets background. The current delayed photon search cuts are listed in Table 4.2, and a performance comparison is in Figure 4.4.

---

[1] https://twiki.cern.ch/twiki/bin/view/CMS/CutBasedPhotonIdentificationRun2#Selection_implementation_details

Variable	**Cut**
ρ corrected photon isolation	0.0028 p_t + 8
ρ corrected PF charged hadron isolation	0.0056 p_t + 6
ρ corrected PF neutral hadron isolation	0.264 + 0.0148 p_t + 0.000017 p_t^2
ρ corrected neutral hadron (E_t^n)	6
Photon S_{minor}	>0.3
Photon S_{minor}	<0.15

Table 4.2: Table of cuts for the current cut based photon ID data analysis. PF refers to Particle Flow, the reconstruction giving a global event description.

ρ Corrections

The ρ corrected variables in the EGamma and cut based photon ID have an energy correction based on the energy density in the ECAL, which is correlated with pileup. This correction is calculated with:

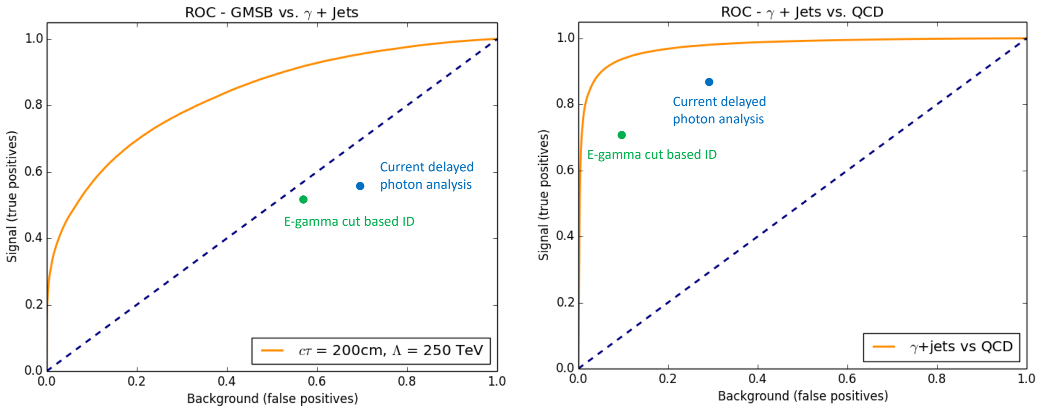
$$\rho \text{ corrected} = \text{Max}\left(\frac{\text{variable} - \rho \cdot \text{EA}}{\text{photon } p_t}, 0\right) \tag{4.7}$$

where EA is the effective area, and depends on the pseudorapidity η. The effective areas are listed in Table 4.3, with values from [43]. ρ is given by the variable "fixedGridRhoFastjetAll" which is a measure of the median energy density in the event, and is an energy/area with area in $\eta - \phi$ space. The photon p_t is the measured transverse momentum of the photon in the ECAL.

Pseudorapidity range	**EA charged hadrons**	**EA neutral hadrons**	**EA photons**
$\lvert\eta\rvert$<1.0	0.0360	0.0597	0.1210
1.0<$\lvert\eta\rvert$<1.479	0.0377	0.0807	0.1107
1.479<$\lvert\eta\rvert$<2.0	0.0306	0.0629	0.0699
2.0 <$\lvert\eta\rvert$ < 2.2	0.0283	0.0197	0.1056
2.2 <$\lvert\eta\rvert$ < 2.3	0.0254	0.0184	0.1457
2.3 <$\lvert\eta\rvert$ < 2.4	0.0217	0.0284	0.1719
$\lvert\eta\rvert$ > 2.4	0.0167	0.0591	0.1998

Table 4.3: Table of effective area (EA) for hadrons and photons, based on the η value of the particle. These effective area values are from [43].

The cuts in Tables 4.2 and 4.1 are used to compare the cut based and BDT photon identification, based on the signal and background rejections of each. Results are shown in Figure 4.4, using the “tight” EGamma cuts. The BDT separates signal and background with a higher efficiency than either the EGamma cuts or the cut based photon ID, as the cut based points are inside the ROC curves. The cut based IDs perform considerably worse on GMSB vs. γ + jets since they are optimized to distinguish between QCD and γ + jets events. The improved photon identification by the BDT indicates that the BDT is better able to distinguish between the signal and backgrounds than the cut based selections.



(a) ROC curve for GMSB vs. γ + jets.

(b) ROC curve for γ + jets vs. QCD.

Figure 4.4: These ROC curves show point comparisons to the current analysis and EGamma cuts. There is no ECAL timing information (just the seven initial variables) so a direct comparison can be made. In both cases, the BDT performs better than the cut based ID, as the cut based results fall within the ROC curve. The cuts without S_{minor} are used here.

4.6 ECAL Timing Impact

The importance of ECAL timing information in the BDT is investigated by comparing the ROC curve from a BDT with the ECAL timing included and one without timing information. Both BDTs have the same seven initial variables:

- $\gamma_{isolation}$ (pho1ecalPFClusterIso): the cluster isolation of the shower in the ECAL, given by particle flow
- p_T^{cone} (pho1trkSumPtHollowConeDR03): the sum of the transverse momentum of particles in a given cone within the tracker (before ECAL)
- p_T^{ch}: the sum of the transverse energy of the charged hadrons, given by particle flow

- $\sigma_{I\eta I\eta}$: measures the width of the shower in the η direction
- S_{major}: the semi-major axis of the elliptical ECAL energy deposit of a photon
- S_{minor}: the semi-minor axis of the elliptical ECAL energy deposit of a photon
- R9: the ratio of energy deposited in a 3×3 crystal square around a the photon seed crystal in the ECAL to the energy deposited in the larger ECAL supercluster

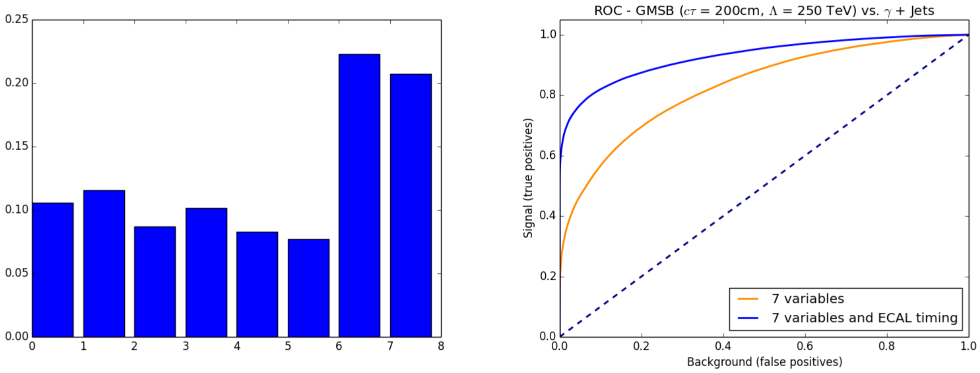
Significant improvement is seen when the ECAL timing information is added to the BDT, as shown in Figure 4.5.

The performance of the GMSB vs. QCD discrimination is significantly better than the GMSB vs. γ + jets in Figure 4.5. This is likely due to the shower shape being a very helpful distinguishing feature in comparing GSMB and QCD, which depends on the cluster isolation and R9 (shower shape distribution). The importance of these variables is seen in the relative importance graph in Figure 4.5c. For GMSB vs. QCD separation, the variables rated as being important are: pho1ecalPFClusterIso (1), pho1trkSumPtHollowConeDR03 (2), R9 (7), and ECAL timing (8). The cluster isolation, cone momentum sum, and energy ratio give information about the shower distribution and energy, which is a more important distinguishing feature for QCD than for γ + jets. However, timing information makes a larger relative contribution to the performance of the GMSB vs. γ + jets BDT, due to this BDT primarily distinguishing between prompt and delayed photons.

4.7 BDT Model Capabilities

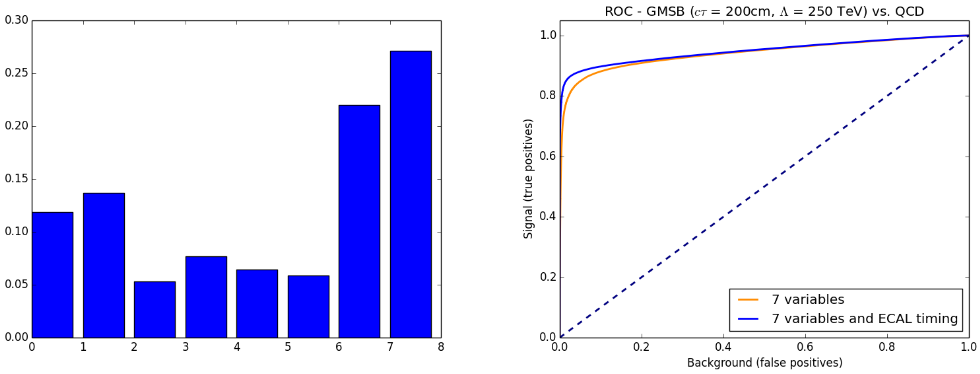
I next investigated the ability of a model that was trained on separating GMSB and QCD to distinguish between GMSB and γ + jets. This is useful, because if the initial model performs well on also discriminating γ + jets from the signal, then a single model can be used to separate the GMSB signal from both backgrounds. Another approach to this would be to train a model on GMSB vs. QCD, and then train a separate model on the remaining GMSB vs. γ + jets, though this is a more complicated approach (partly due to splitting the simulation data into multiple training and testing sets for the two models to avoid biasing the BDT). See Appendix D for a more in depth discussion of the reason for this requirement.

For each event, the BDT outputs a discriminator variable between 0 and 1, with higher numbers indicating that the event is more signal-like. The events in the various



(a) Relative importance plot for GMSB vs γ + jets.

(b) ROC curve for GMSB vs γ + jets.



(c) Relative importance plot for GMSB vs QCD.

(d) ROC curve for GMSB vs QCD.

Figure 4.5: ROC curves and relative importance plots for GMSB vs. QCD and GMSB vs. γ + jets. The order of variables in the relative importance plot is: pho1ecalPFClusterIso, pho1trkSumPtHollowConeDR03, p_E^n, $\sigma_{I\eta I\eta}$, S_{major}, S_{minor}, R9, and ECAL photon timing. Both ROC curves are done for a neutralino model with lifetime of $c\tau = 200$ cm and SUSY scale of $\Lambda = 250$ TeV. These plots show the large improvement by adding ECAL timing information — this is rated as one of the most important variables and significant improvement in the BDT is seen with the addition of ECAL timing. ECAL timing was not considered in Figures 4.2, 4.4, or 4.3.

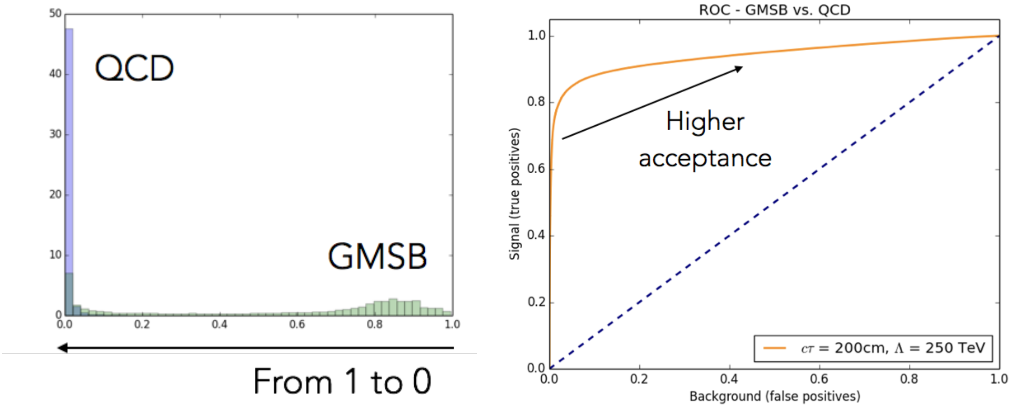
signal and background dataset have different distributions over the discriminator, as can be seen in Figure 4.6a with the different colors indicating a different event type. When a value of the discriminator is selected, a number of signal and background events will have higher discriminator values. As the cut on the discriminator decreases, more signal and background events are included (higher acceptance for both). Ideally, the distributions of signal and background over the discriminator would have very little overlap such that a discriminator value could be chosen to separate signal and background nearly perfectly.

The percentage of signal and background events passing the discriminator cut is plotted as the ROC curve (background vs. signal). As the cut on the discriminator is moved from 1 to 0, the acceptance of signal and background increases (going upwards and right on the ROC curve). This is shown in Figure 4.6. The initial large slope of the ROC curve is due to the GMSB event distribution cluster near discriminator values of 1.

In order to test the performance of a model (in this case, GMSB vs. QCD) on the separation of another type of background (in this case, γ + jets), the efficiency of the new background at various discriminator values is needed. This is obtained by loading the previously trained BDT model and running the new background through the model. Then the efficiencies at all discriminator values are obtained for the signal and two backgrounds, which is the information needed to create ROC curves for all signal and background comparisons, shown in Figure 4.8.

When the GMSB vs. QCD model is used to separate GMSB vs. γ + jets, the model initially performs well in the region where the GSMB discriminator distribution looks significantly different than γ + jets. The ROC curve and discriminator distribution are shown in Figure 4.7. Since part of the distribution of γ + jets looks similar to that of GMSB (near discriminator values of 1), and part looks similar to QCD (near discriminator values of 0), the model does not perform as well as it did distinguishing GMSB from QCD in Figure 4.6.

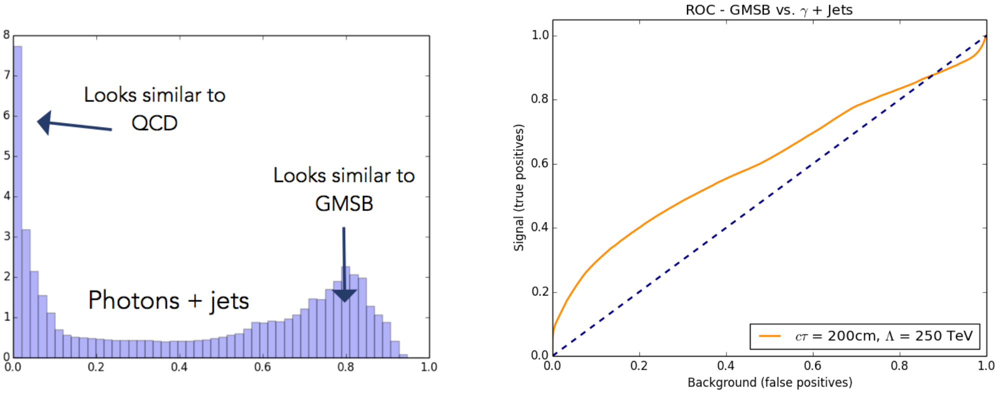
From the three distributions (GMSB, QCD, and γ + jets) over the discriminator variable, three ROC curves are made (GMSB vs. QCD, GMBS vs. γ + jets, and γ + jets vs. QCD). The performance of these ROC curves are compared to the current cut based photon ID and the EGamma cuts (Tables 4.2 and 4.1) by plotting the points from the cut based analyses on the ROC curve plot. These comparisons are done including the S_{minor} variable cut, as this decreases the dominant QCD background. In both cases, the ROC curve performs better than the cut based ID, as the points are



(a) Discriminator distribution for GMSB and QCD.

(b) ROC curve for GMSB vs. QCD.

Figure 4.6: The results for a model trained on GMSB vs. QCD. As the discriminator cut moves from 1 to 0, initially more GMSB is accepted (leading to the rapid increase in signal acceptance on the left of the ROC curve), and then closer to a discriminator value of 0, the background acceptance rapidly increases. This demonstrates the correspondence between the ROC curve (efficiencies of various discriminator cuts) and the discriminator distribution.



(a) Discriminator distribution for γ + jets.

(b) ROC curve for GMSB vs. γ + jets.

Figure 4.7: Results for a model trained on distinguishing GMSB vs. QCD. When this model attempts to separate GMSB and γ + jets, it initially performs well (bottom left of ROC curve), but as the discriminator value moves toward 0, the acceptance of GMSB and γ + jets becomes more similar, and therefore the ROC curve tends toward the 50-50 acceptance line. For the GMSB vs. γ + jets plot, the signal separation becomes much worse toward the top right. This is the region where low momentum jets are reconstructed as photons. However, high momentum photons are more likely to be from GMSB, and therefore the performance in the lower left region is more relevant to the analysis.

contained within the ROC curve. This indicates that the BDT has increased event discrimination abilities as compared to either cut based selection. Table 4.4 gives the cut efficiencies.

Note that in these BDT models, timing information is not used. This is in order to make a fair comparison to the cut based analysis, which only implements the timing information at the end.

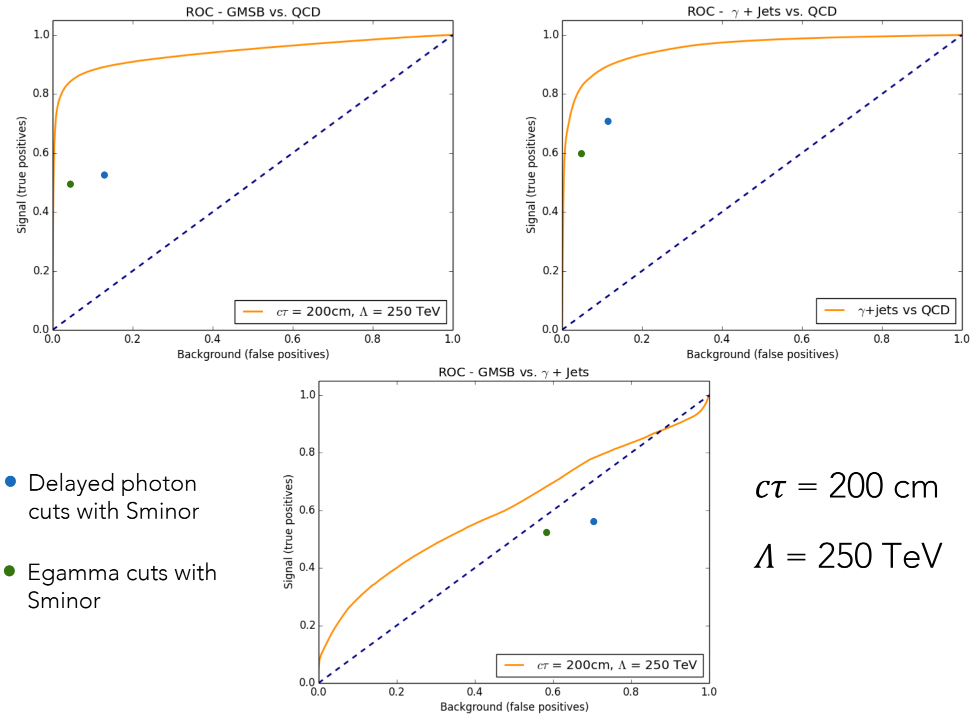


Figure 4.8: Results for a model trained on distinguishing GMSB vs. QCD. In all cases, the BDT results from this model perform better at separating the events as compared to the cut based IDs. This BDT is trained without ECAL timing information in order to directly compare to the cut based IDs, which implement timing information at the end, not during event selection. Cut based photon ID with the S_{minor} requirement are used here.

I also investigated the ability of a model that was trained on separating GMSB and γ + jets to distinguish between GMSB and QCD. However, the ability of the model to separate of GMSB vs. QCD was low. In this case, the cut based ID performs better, and thus, the model initially trained on GMSB vs. QCD is a better model to pursue for photon identification improvements.

	GMSB	**Photons and Jets**	**QCD**
Cut based photon ID	72.7%	85.6%	29.5%
EGamma cuts	68.9%	72.4%	11.5%
Cut based photon ID with $\mathbf{S_{minor}}$	53.6%	68.6%	13.4%
EGamma cuts with $\mathbf{S_{minor}}$	51.3%	58.2%	5.78%

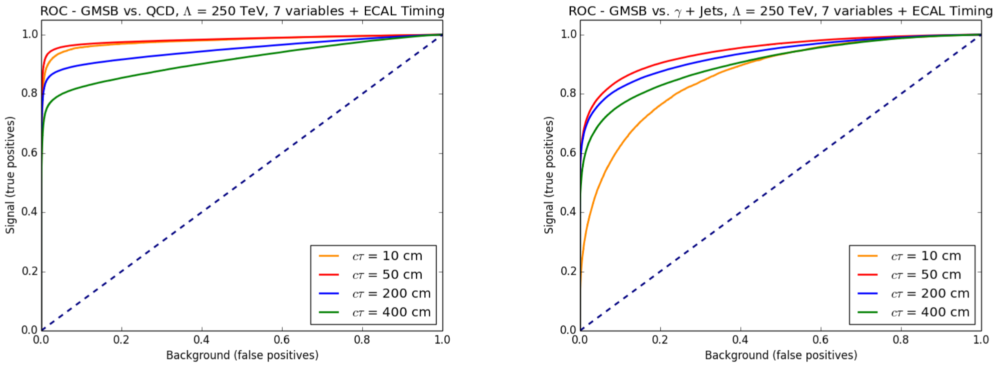
Table 4.4: Table for the cut results (percentage of events passing) for the three types of events and the four types of cuts. This corresponds to the information plotted in Figure 4.8. Only the cuts with S_{minor} are plotted, since these cuts have much better QCD background rejection.

4.8 Neutralino Lifetime and Mass Comparisons

The analysis above has been performed for a GMSB model with a SUSY scale Λ = 250 TeV, and a neutralino lifetime of $c\tau$ = 200 cm. The Λ scale corresponds to neutralino mass, given in Table 4.5. However, very similar results are expected for models trained on various masses and lifetimes, since the ROC curves follow an expected pattern — for increasing lifetime, it is easier to separate GMSB vs. background (until greater than $c\tau$ = 200 cm where more neutralinos begin decaying outside of the ECAL barrel); and for increasing neutralino mass, it is easier to separate GMSB vs. background. ROC curves with various neutralino masses and lifetimes are in Figure 4.9. Though the analysis is performed for a specific lifetime and mass model to demonstrate the results, it is designed to be relatively model independent, and is expected to perform similarly for a large range of lifetimes and masses. This is supported by the observed abilities of BDTs to distinguish between signal and background over a large range of masses and lifetimes.

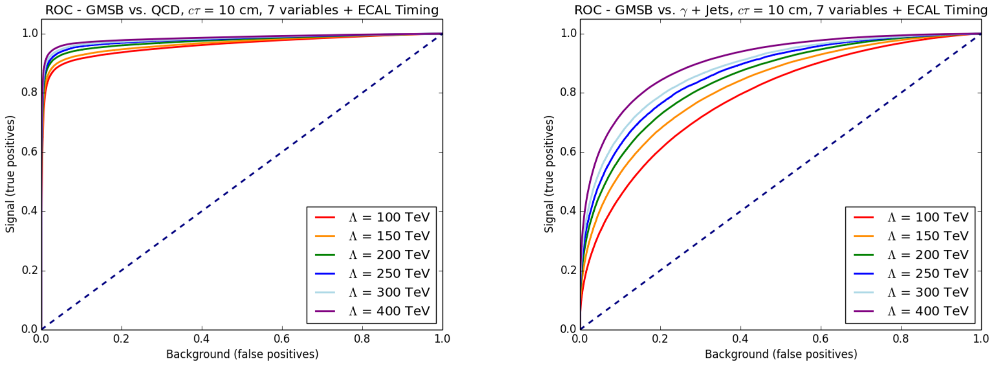
Λ **(TeV)**	**100**	**150**	**250**	**400**
$m_{\tilde{\chi}_1^0}$ **(GeV)**	139.4	212.1	357.5	576.4
$m_{\tilde{g}}$ **(GeV)**	837.8	1206.7	1915.9	2934.5

Table 4.5: Correspondence between SUSY theory cutoff mass and neutralino and gluino masses.



(a) ROC for GMSB vs. QCD, various lifetimes.

(b) ROC curve for GMSB vs. γ + jets, various lifetimes.



(c) ROC for GMSB vs. QCD, various masses.

(d) ROC curve for GMSB vs. γ + jets, various masses.

Figure 4.9: ROC curves for various neutralino lifetime and mass models, showing that the separation of signal and background follows expectations.

Chapter 5

PERFORMANCE COMPARISON TO CURRENT ANALYSIS

In order to compare the BDT performance to the current cut based analysis, the photon ID of the current analysis will be replaced with a cut on the discriminator from the BDT model. This work is currently in progress, and reported here is a quantification of the improvement in event discrimination offered by the BDT model. The improvement is tested with and without ECAL timing information.

The BDT trained on GMSB vs. QCD creates a model that other simulation data files can be run through such that they have a discriminator value associated with every event. A cut on the discriminator can be used for the photon identification. A single discriminator cut gives the acceptances of the GMSB signal and both QCD and γ + jets backgrounds.

5.1 Data Analysis Comparison

When simulation files are used in the BDT, they are skimmed such that only events with leading photon $p_T > 0$, photon η position in the ECAL barrel, and events matching a Monte Carlo truth are considered. This is done with the cuts:

$$\text{pho1Pt} > 0 \tag{5.1}$$

$$|\text{pho1Eta}| < 1.479 \tag{5.2}$$

$$\text{pho1isPromptPhoton} == 0 \tag{5.3}$$

These same requirements will be put on the events used for data analysis comparison.

In order to compare the BDT performance to the cut based ID, a discriminator value cut is used in place of the cut based ID (cut based IDs listed in Tables 4.2 and 4.1). From the values in Table 4.4, it is seen that the current cut based photon ID with S_{minor} have an acceptance of 13.4% on QCD, 68.6% on γ + jets, and 53.6% on GMSB. There are three approaches to compare this performance to the BDT:

1. the GSMB signal acceptance is fixed at 53.6% and the reduction in QCD and γ + jets backgrounds are found
2. the background γ + jets acceptance is fixed at 68.6% and the affects on GMSB and QCD acceptance are found

3. the background QCD acceptance is fixed at 13.4% and the affects on GMSB and γ + jets acceptance are found

For the GMSB vs. QCD BDT model to reach the same acceptance on QCD, a discriminator of 0.0093 is needed. Similarly, for the BDT model to reach the same GMSB signal acceptance, a discriminator value of 0.668 is needed. For the BDT to reach the same γ acceptance (68.6%) as the cut based approach, a discriminator value of 0.105 is needed. These values are shown in Table 5.1, for the BDT models with and without ECAL timing information.

Using ECAL timing information in the BDT would mean the analysis strategy is changed, while a similar analysis strategy could be used by implementing the BDT without timing. However, to determine if this is a approach worth exploring, the improvements offered by the BDT with ECAL timing are analyzed as well.

	GMSB	**Photons and Jets**	**QCD**
Cut based photon ID with S_{minor}	53.6%	68.6%	13.4%
Constant GMSB Signal:			
No timing, disc = 0.668	53.6%	37.4%	0.26%
With timing, disc = 0.9578	53.6%	0.736%	0.00545%
Constant γ + Jets Background:			
No timing, disc = 0.105	77.1%	68.6%	1.4%
With timing, disc = 0.05115	84.9%	68.6%	1.5%
Constant QCD Background:			
No timing, disc = 0.0093	89.3%	90.6%	13.4%
With timing, disc = 0.00635	90.5%	90.6%	13.4%

Table 5.1: Table of acceptances from current cut based photon ID and BDT with a given discriminator for GMSB, photons + jets, and QCD. One discriminator value is chosen to have a constant QCD value, one for a constant GMSB value, and one for a constant γ + jets. This is done for BDTs with and without ECAL timing information.

From Table 5.1 it is seen that by maintaining the same GMSB acceptance of 53.6%, the use of the BDT both with and without timing decreases the acceptance of both of the backgrounds. All comparison calculations will be done with the constant GMSB signal value of 53.6%, analyzing how significant the reduction in background is. The constant γ + jets background and QCD background results are presented in Table 5.1 but are not used for the significance calculations.

5.2 Significance of BDT Improvement

To quantify the improvement with the BDT model, the signal to background selection significance is calculated. This is done by comparing the signal GMSB with the background QCD and γ + jets acceptances for cut based and BDTs (with and without timing).

Under the assumption of a Poisson distribution, the fluctuation in an event bin is equal to the square root of the number of events.

$$\text{fluctuation}_N = \sqrt{N} \tag{5.4}$$

To approximate the significance of a signal, the quantity $\sigma = \frac{S}{\sqrt{B}}$ is calculated, where S is the difference from the background prediction to the observation with signal (fluctuation), and B is the background (number of events). $\frac{S}{\sqrt{B}}$ is a significance approximation in units of standard deviations from expected value. Then the significance of improvement between two points is

$$\frac{\sigma_2}{\sigma_1} = \frac{S_2}{\sqrt{B_2}} \frac{\sqrt{B_1}}{S_1} \tag{5.5}$$

However, since the comparison is done with constant GMSB signal (53.6% from Table 5.1), then $S_1 = S_2$, and this becomes $\frac{\sigma_2}{\sigma_1} = \frac{\sqrt{B_1}}{\sqrt{B_2}}$. The background B is given by the number of background events, N_B, which is reduced by the use of the BDT. Thus, this becomes

$$\frac{\sigma_2}{\sigma_1} = \sqrt{\frac{N_{B_1}}{N_{B_2}}} \tag{5.6}$$

The background N_{B_1} is given by QCD and γ + jets, so

$$N_{B_1} = N_{QCD} + N_{\gamma+jets} \tag{5.7}$$

The background N_{B_2} is also the sum of QCD and γ + jets, but these background have been reduced by the improvements from the BDT demonstrated in Table 5.1. Thus,

$$N_{B_2} = N_{QCD} \frac{QCD_2}{QCD_1} + N_{\gamma+jets} \frac{\gamma_2}{\gamma_1} \tag{5.8}$$

where QCD_2 is the percentage of QCD passing the BDT selection, and QCD_1 is the percentage of QCD passing the cut based photon ID. This is the same for γ_1 and γ_2 with the γ + jets background. Therefore,

$$\frac{\sigma_2}{\sigma_1} = \sqrt{\frac{N_{B_1}}{N_{B_2}}} = \sqrt{\frac{N_{QCD} + N_{\gamma+jets}}{N_{QCD} \frac{QCD_2}{QCD_1} + N_{\gamma+jets} \frac{\gamma_2}{\gamma_1}}} \tag{5.9}$$

For the backgrounds in the GMSB model, it is assumed that 70% of the background is from QCD, and 30% is from γ + jets [44]. Therefore,

$$N_{QCD} = 0.7N_{B_1} \tag{5.10}$$

$$N_{\gamma+jets} = 0.3N_{B_1} \tag{5.11}$$

Using this relationship and Equation 5.7 in Equation 5.9:

$$\frac{\sigma_2}{\sigma_1} = \sqrt{\frac{N_{B_1}}{0.7N_{B_1}\frac{QCD_2}{QCD_1} + 0.3N_{B_1}\frac{\gamma_2}{\gamma_1}}} = \sqrt{\frac{1}{0.7\frac{QCD_2}{QCD_1} + 0.3\frac{\gamma_2}{\gamma_1}}} \tag{5.12}$$

Using this, the significance of improvement between the cut based analysis and the BDT performance can be calculated by comparing the amounts of QCD and γ + jets included by each method. This approach considers the effects of each background on the overall improvement.

The results of the BDT models trained on GMSB vs. QCD with and without ECAL timing information are shown in Figure 5.1, with the performance comparisons to the cut based IDs indicated as points. For the BDT comparison to the current cut based photon ID analysis, values from Table 5.1 are used, and only the comparison with constant GMSB signal is done.

Improvement without ECAL Timing

When no ECAL timing is used, the BDT performance can be directly compared to the cut based ID performance, as the cut based does not use ECAL timing. This significance comparison is done using the data points from the ROC curve compared to the signal and background acceptance from the cut based photon ID.

First the BDT without ECAL timing information is considered. With a constant GMSB signal value of 53.6%, the BDT accepts 37.4% of the γ + jets and 0.26% of the QCD. This is compared to the 68.6% and 13.4%, respectively, accepted with the cut based approach. Therefore, the improvement is

$$\frac{\sigma_2}{\sigma_1} = \sqrt{\frac{1}{0.7 \cdot \frac{0.26}{13.4} + 0.3 \cdot \frac{37.4}{68.6}}} = 2.38 \qquad \sigma_2 = 2.38\sigma_1 \tag{5.13}$$

The significance improvement with BDT model is by a factor of 2.38 for fixed GMSB signal, considering both backgrounds. This provides a significant improvement over the cut based ID. Therefore, for the data analysis comparison, the current photon ID will be replaced with a discriminator cut of

$$\text{disc} > 0.668 \tag{5.14}$$

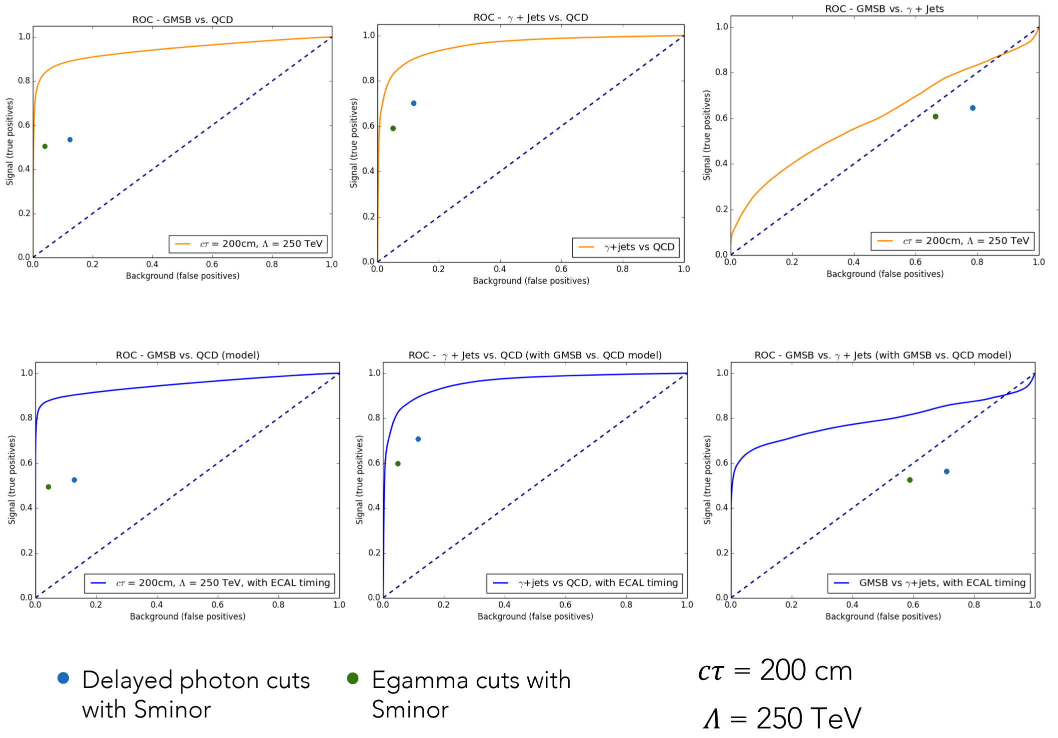


Figure 5.1: ROC curve results for the BDT model with and without ECAL timing — GMSB vs. QCD, γ + jets vs. QCD, and GMSB vs. γ + jets from left to right. Performance comparisons are plotted for the EGamma 2016 and current cut based photon ID selections (with S_{minor}). The calculations presented in the following sections are done using the performance of these BDTs.

If the discriminator cut on events used in the full analysis leads to a larger sensitivity region, then this will demonstrate that in the search for delayed photons, the BDT algorithm is a more precise way to identify SUSY neutralino events.

Improvement with ECAL Timing

The improvement offered by the BDT model with ECAL timing included is quantified using $\frac{S}{\sqrt{B}}$, comparing the performance to the cut based ID. Using timing in the photon ID as compared to later in the analysis would require changing the analysis strategy. The improvements in significance from the BDT with timing is investigated, as it may be worth exploring the possibility to use timing in the BDT instead of later in the analysis if this offers a large improvement.

With a constant GMSB signal of 53.6%, the cut based photon ID can be compared to the BDT with timing information. When this is compared, the QCD background acceptance is decreased from 13.4% to 0.00545%, and the γ + jets background acceptance is decreased from 68.6% to 0.736% resulting in:

$$\frac{\sigma_2}{\sigma_1} = \sqrt{\frac{1}{0.7 \cdot \frac{0.00545}{13.4} + 0.3 \cdot \frac{0.736}{68.6}}} = 16.9 \qquad \sigma_2 = 16.9\sigma_1 \tag{5.15}$$

Use of the ECAL timing information increases the significance of the signal vs. background acceptance by $\sigma = 16.9$ as compared to the cut based ID. This demonstrates that the BDT with ECAL timing information has a large potential to improve the event selection. Having said this, the current analysis uses timing after the photon identification, so the improvement calculated here by including ECAL timing is not a direct one-to-one comparison. $\sigma_2 = 16.9\sigma_1$ demonstrates that ECAL timing in the BDT gives significant improvement, but a more careful study on the inclusion of timing in the final analysis is needed. The final comparison in the full analysis would need to be between the BDT without timing (with timing information included after the photon ID selection) and the BDT with ECAL timing. To do this and establish the final sensitivity, the full analysis will need to be run on both to determine which method is better. The result of $\sigma_2 = 16.9\sigma_1$ demonstrates the potential improvement through a different analysis approach — using machine learning and the ECAL timing information in the photon identification.

Additionally, as can be seen from the plots in Figure 5.1, there is a large change in the GMSB vs. γ + jets ROC curve when ECAL timing is added. Without ECAL timing, the curve is slightly above the $x = y$ line, until the background acceptance reaches $\approx$ 90%. With timing information, about 60% of the GMSB signal is accepted before large amounts of γ + jets background is accepted. Though the ROC curve does still cross the $x = y$ line around a γ + jets acceptance of 90%, it is the lower background regions that are more relevant for the delayed photon search. Low momentum jets are often mis-reconstructed as photons, and these are the events with poor signal to background discrimination (top right of the ROC curve). The separation of high momentum jets and delayed photons is important for detecting the GMSB signal, and this falls primarily in the low γ + jets acceptance region. This is the region improved by the addition of ECAL timing information.

Training Data Note

For all events used as training data (60% of the total Monte Carlo events), the discriminator was automatically set to -1. This is to ensure that these events are not used in the comparison, as using these events when the BDT was trained on them would make the result biased.

5.3 Analysis Results

As can be seen from the results in Table 5.1 and the significance calculations, the BDT algorithm has a higher efficiency in selecting delayed (GMSB) photons from the QCD and γ + jets backgrounds. Compared at the same GMSB signal acceptance, the decreased acceptance values for QCD and γ + jets with the BDT indicates that this model distinguishes between signal and both background more precisely.

The discriminator value chosen is motivated by the performance of the current analysis. The current analysis cuts accept 53.6% of the GMSB, while also accepting 13.4% of the QCD and 68.6% of the γ + jets. When the BDT model (without ECAL timing) accepts 53.6% of GMSB, the γ + jet and QCD acceptances both decrease from the cut based ID values. In this situation, 37.4% of γ + jet and 0.26% of QCD is accepted — both an improvement over the cut based photon ID. This gives a selection efficiency increase of 2.38 times.

Adding ECAL timing to the BDT model has the potential to increase the selection efficiency by 16.9 times (with a constant GSMB acceptance).

5.4 Future Applications and Improvements

The discriminators for GSMB, QCD, and γ + jets are plotted for comparison and shown in Figure 5.2. This illustrates the similar distributions seen for GMSB and γ + jets, while QCD has a distinct distribution (and is therefore easier to separate).

A direct comparison to the current cut based delayed photon data analysis is being completed. This is done with the discriminator cut from Table 5.1 (disc $>$ 0.668) replacing the current photon ID cut in the analysis in order to determine how the final exclusion plot is affected.

The exclusion plot in Figure 5.3 plots SUSY Λ scale vs. neutralino lifetime (cm) for the neutralino to di-photons + gravitino decay pathway. The shaded red area indicates the region of possible neutralino mass and lifetime excluded by the current analysis algorithms, and a comparison will be done for the BDT algorithm performance.

A number of the masses and lifetimes tested are near the right edge of the current ECAL performance (red region on Figure 5.3). Figure 4.9 demonstrates that a BDT is able to accurately distinguish between signal and background events with a large range of masses and lifetimes, and therefore with the BDT model, the sensitivity is expected to be greatly extended from the current regions in Figure 5.3. Additionally, given the significance comparison presented in Equation 5.13, the BDT offers large

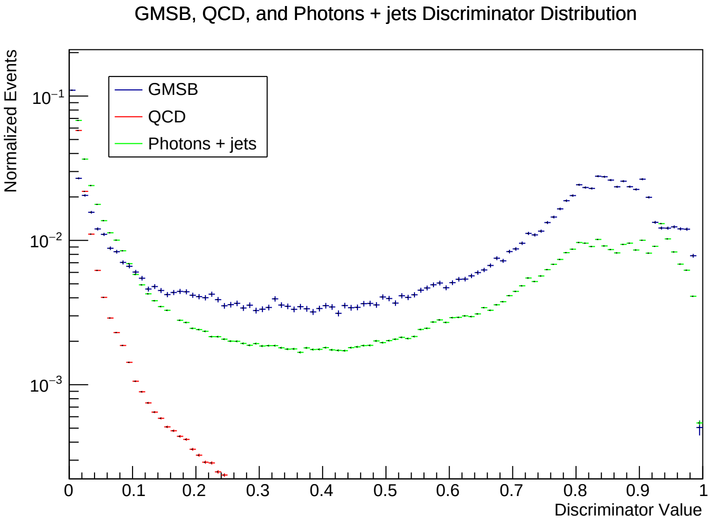


Figure 5.2: Plot of discriminator values for GMSB, QCD, and γ + jets for a model trained on GMSB (Λ = 250 TeV and $c\tau$ = 200 cm) vs QCD. QCD and GMSB are well separated by a discriminator cut, however, this same cut will accept significantly more γ + jets as this has a similar distribution to GMSB. All areas are normalized by the integral of the distribution.

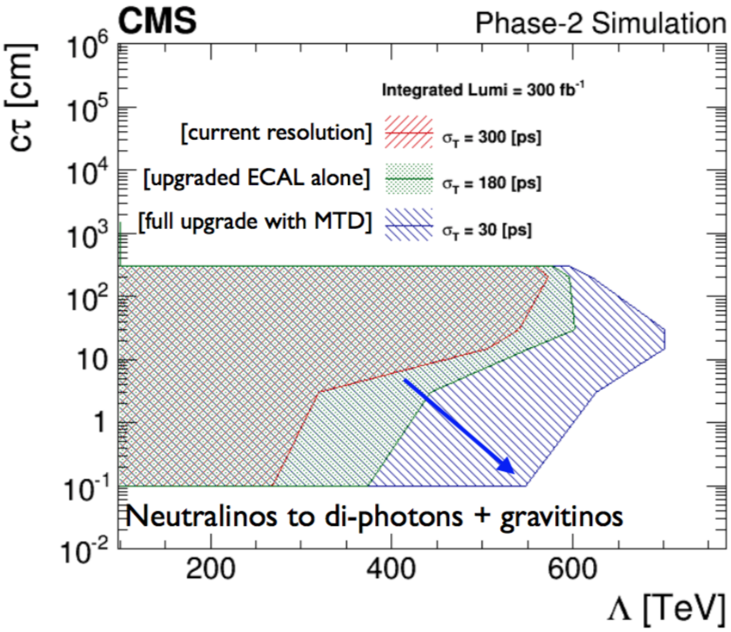


Figure 5.3: Phase II CMS Simulation demonstrating the SUSY mass scale (Λ) and lifetime (τ) parameters that can be searched for with the upgraded MTD. Figure from Lindsey Gray's talk presented at LHCC, 1 December 2017 [4]. The red area is the current ECAL performance.

improvements over the cut based photon ID approaches.

ECAL timing was shown to be a very important variable (consistently ranking as the variable with the highest relative importance in the BDT), and thus significant improvement in event selection is also expected in the full MTD upgrade region with the increased time resolution. This is further supported by the event selection significance improvement by a factor of 16.9 achieved by including timing information in the BDT.

Chapter 6

PRECISION TIMING STUDIES WITH THE HIGH GRANULARITY CALORIMETER

The work presented in this chapter was completed at Fermilab in June-August 2016, and is relevant to my thesis work, though not a part of the research completed for my senior thesis.

At the higher collision rates in the high luminosity LHC (HL-LHC), pileup collisions will greatly increase, making event reconstruction more difficult. This can be mitigated by increasing the spatial resolution of the collision vertex — either by improving the detector capabilities (detectors with finer granularity have a better spatial resolution) or by using precision timing detectors, employing time-of-flight (TOF) techniques. This research focuses on precision timing detectors that measure the particle time stamp at the interaction point. From this time stamp, the collision vertex the particle originated from can be reconstructed [45]. This research analyzes the timing capabilities of calorimeters and to demonstrates their ability to achieve a time resolution on the order of 30 ps, which allows for 1 cm spacial resolution [45].

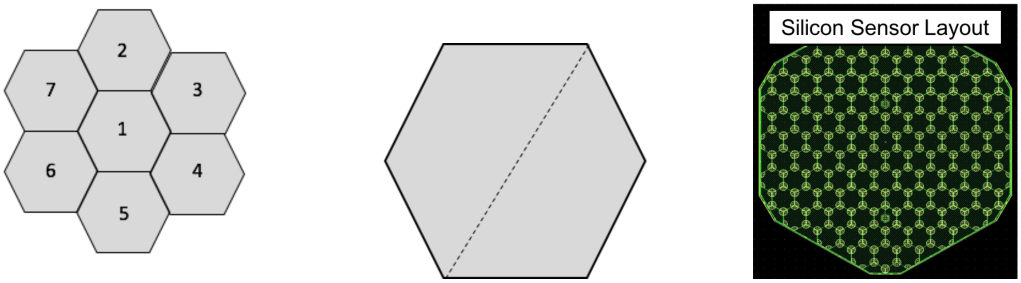
Precision timing is extremely relevant to the SUSY dark matter model, since identification of events with long-lived neutralinos rely on detecting the neutralino decay vertex and resulting photons. Improved detector timing capabilities are expected to improve the search capabilities, as has been demonstrated by the proposed MTD.

The HL-LHC and Phase II upgrades of the CMS detectors are scheduled for 2023, and prior to the upgrade, the detectors must be replaced and improved [46]. The High Granularity Calorimeter (HGC) is the chosen design for the CMS Phase II Detector Upgrade and will have a high granularity due to the hexagonal silicon sensors [46]. The HGC will replace the current CMS electromagnetic calorimeter (ECAL) endcap, which currently has a time resolution of 150 ps. This research investigates the time resolution capabilities of a prototype HGC.

6.1 Test Beam

For this investigation, detectors were tested in the MTest beamline at Fermi National Accelerator Lab (Fermilab) in June 2016. Proton and electron beams from 4-32 GeV were used. To test the capabilities and limitations of a HGC detector, we used a detector with hexagonal silicon "pixels" (Figure 6.1). This dedicated silicon timing layer is identical to that proposed for the HGC, and we measured its intrinsic time resolution and limitations. This detector, referred to as pico-sil HGC (picosecond timing silicon detector) was built by engineers at Fermilab. The pico-sil detector was set up with a DRS4 evaluation board for signal digitization during the test beam.

The HGC time resolution is tested by observing the propagation of an EM shower from an electron beam in the detectors and analyzing how the signal is amplified by the multi-channel plates and silicon sensor. The detectors used are shown in Figure 6.2, and this analysis focuses on the transverse analysis of the pico-sil detector.



(a) Diagram of the pixels in the pico-sil detector. The pixels are in a hexagonal arrangement, and the beam is focused on the center pixel (labeled #1).

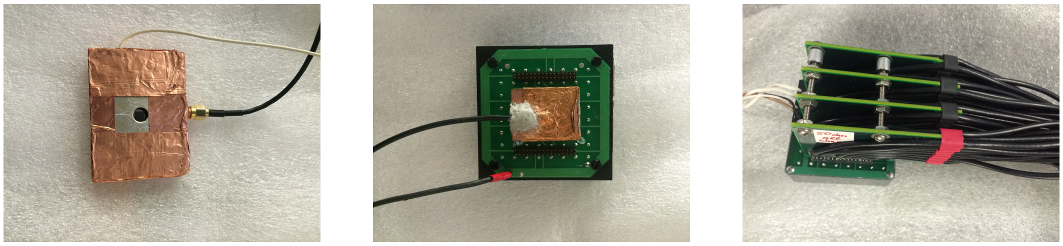
(b) A single pixel in the pico-sil detector. The diagonal (marked with a dashed line) of each hexagon is 1cm.

(c) The silicon sensor electronics used for the HGC. The silicon sensor was bonded to an electronic readout board for the experiment.

Figure 6.1: Diagrams of the pico-sil detector and the hexagonal arrangement of the pixel electronics.

Five detectors were used in the test beam runs: pico-sil HGC, silicon pad, Photonis (single channel), Photonis MCP-PMT (64 channel microchannel plate photo multiplier), and a Photek detector (reference timer), with a trigger in front (Figures 6.2 and 6.3). The trigger is used to determine when events are observed, essentially as a logic check for the system. All detectors were aligned with the center of the test beam, and the entire detector set up was placed on a cart for alignment with the beam (Figure 6.3). A DRS4 was used for detector electronic readout since it is optimized for intrinsic time resolution studies.

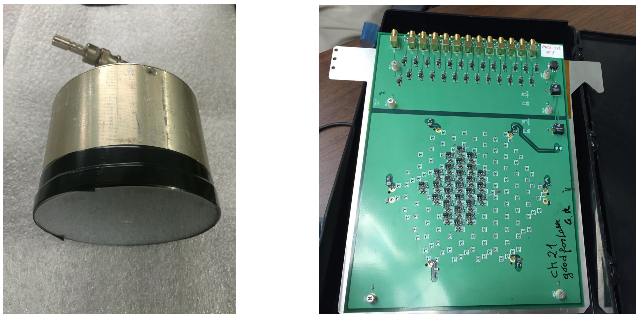
Tungsten absorbers were placed in front of the detectors (Figure 6.3c), and the distance between the detectors and the absorbers was varied throughout the test beam run. All of this analysis is from runs using a $6X_0$ tungsten absorber placed in front of the pico-sil detector. $6X_0$ means 6 radiation lengths of absorber, which



(a) Silicon pad detector.

(b) Photonis single channel detector.

(c) Photonis 64 channel detector, MCP-PMT (microchannel plate photo multiplier), side view.



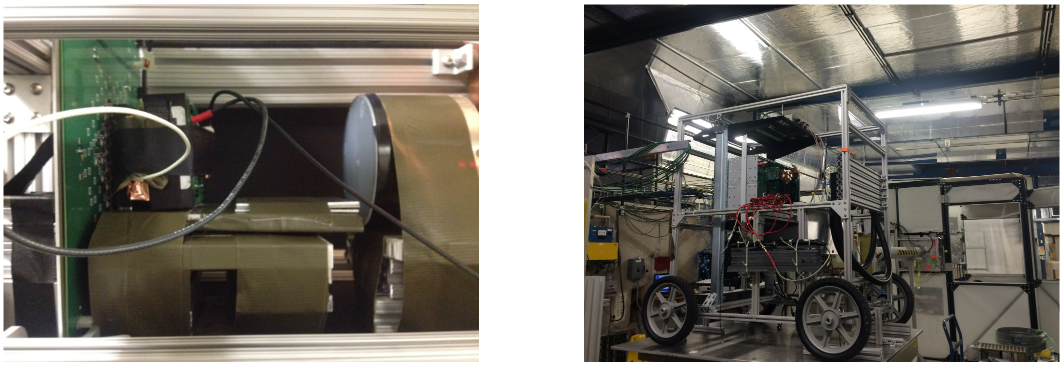
(d) Photek detector, with the black covering to block out background light visible.

(e) Pico-sil HGC detector, showing the electronic components of the detector. Photo by Javier Duarte.

Figure 6.2: Pictures of the five detectors used in the test beam. All detectors had black coverings on the front to eliminate any background light seen.

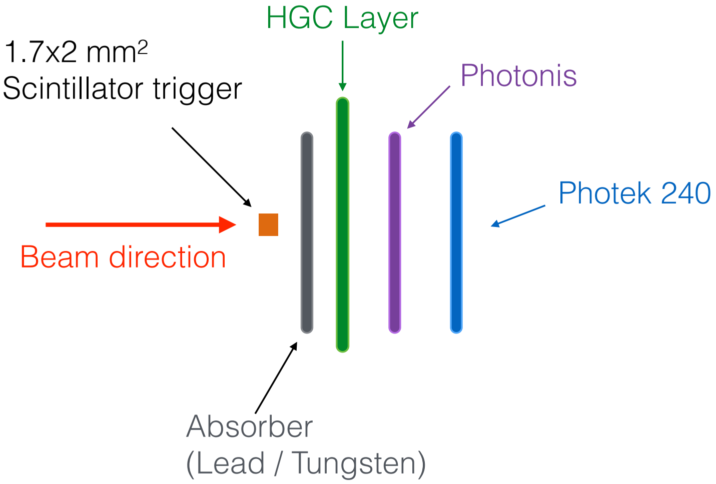
is the thickness of material a high energy electron must travel through to retain only $\frac{1}{e}$ of its energy on average, where the energy is lost by Bremsstrahlung. When the electron beam hits the absorber, a shower of electrons and photons is created through Bremsstrahlung and pair production, and is then observed by the detectors. Bremsstrahlung occurs when an electron approaches a nucleus and the electron loses kinetic energy, electromagnetic radiation is produced, and the electron is converted to a photon and a lower energy electron. The Bremsstrahlung cross section is proportional to Z^2 (Z is the atomic number of the nucleus) [47]. Pair production is when a photon is converted to an electron-positron pair, and the cross section is also proportional to Z^2 [47].

When the distance between the pico-sil HGC detector and the absorber is varied, the pico-sil observes the shower at a wider or more narrow point (Figure 6.3c) — which can be quantified by observing the transverse spread of the shower. The pico-sil detector is made of multiple pixels, arranged in a hexagonal tiling (Figure



(a) Set up of the detectors for the beam line test. The pico-sil detector was placed at the very front, just behind the scintillator trigger.

(b) The cart that was used to align the detectors with the beamline.



(c) Diagram of the beamline and the detector alignment. This depicts the main set up used for the runs in this analysis. In some other runs, an absorber was also placed behind the pico-sil HGC. Diagram by Cristián Peña.

Figure 6.3: Photo of the detector alignment, cart set up, and a diagram of the detectors, trigger, and absorber in the beamline test.

6.1). Each pixel is recorded independently, so the transverse spread of the shower is recorded. The integral (integral of the current vs. time plot for an event, giving units of electric charge, and proportional to the total energy observed by the pixel) of each pixels was calculated and plotted for various positions of the tungsten absorber and electron beam energies (Figure 6.4). These plots show the total energy observed by each pixel.

When the tungsten absorber is further away from the pico-sil detector, less total energy is seen by the detector (Figure 6.4). This is seen by the decrease in the "ring2 sum" as the distance between the detector and absorber is increased. However, the amount of energy seen by the pixels further from the center (in the first or second ring) does not increase significantly, but remains constant or decreases. This is not expected — further away pixels were expected to see an increased amount of energy as the detector is moved further from the absorber, since the pixels are observing a wider part of the shower. However, it is seen that the energy is concentrated in the center pixel for all absorber distances.

When the energy contained is plotted for 32 and 16 GeV electrons with the tungsten at 1 mm, the beam is focused on the center pixel, and the first ring has about half of the energy for the 16 GeV beam as compared to the 32 GeV beam (Figure 6.5). These plots give the typical beam spread for the various beam energies and absorber distances.

The spread in the shower can be seen by plotting the ratio of the charge observed by the central pixel to the charge observed by all pixels. As the $6X_0$ tungsten absorber is moved further away from the pico-sil, the ring pixels contain more of the total charge per event on average. The ratio between the charge contained in the central pixel and all pixels is plotted in Figure 6.6. R_7 is defined as

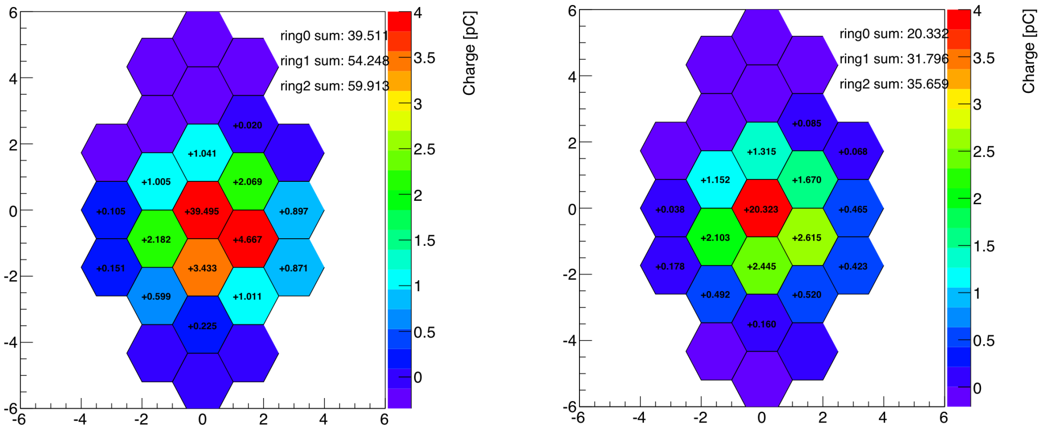
$$\mathrm{R}_7 = \frac{charge_1}{\sum_{i=1}^{7} charge_i} \tag{6.1}$$

where $charge_i$ is the charge in the i^{th} pixel, from Figure 6.1.

6.2 Data Analysis and Results

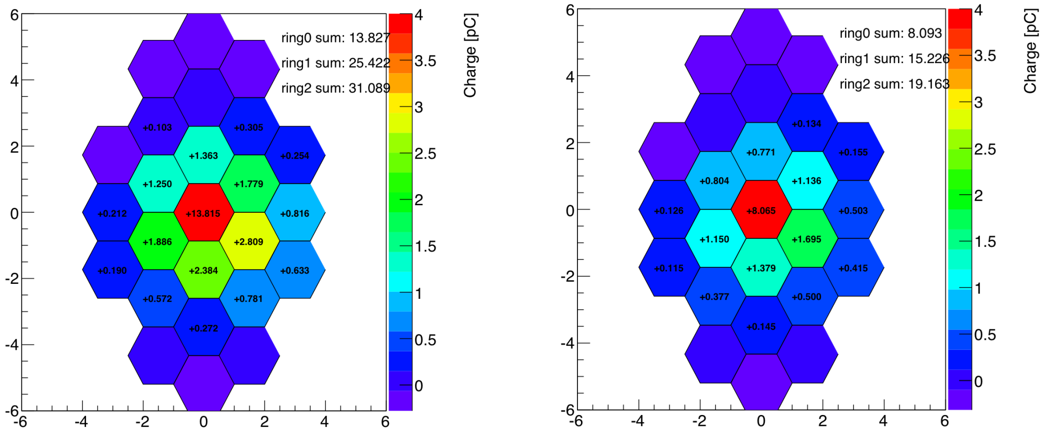
This research focuses on analyzing the effects of combining information from the pico-sil transverse pixels on the time resolution.

The Photek detector is used as the reference timer for all analysis, as it has an outstanding time resolution of 10 ps (measured in previous test beam analyses). The



(a) $6X_0$ tungsten absorber placed 1 mm from the pico-sil detector, with 32 GeV electron beam. The charge in the central pixel is 39 pC, average charge in the first ring is 2.4 pC, and the average charge in the second ring is 0.49 pC.

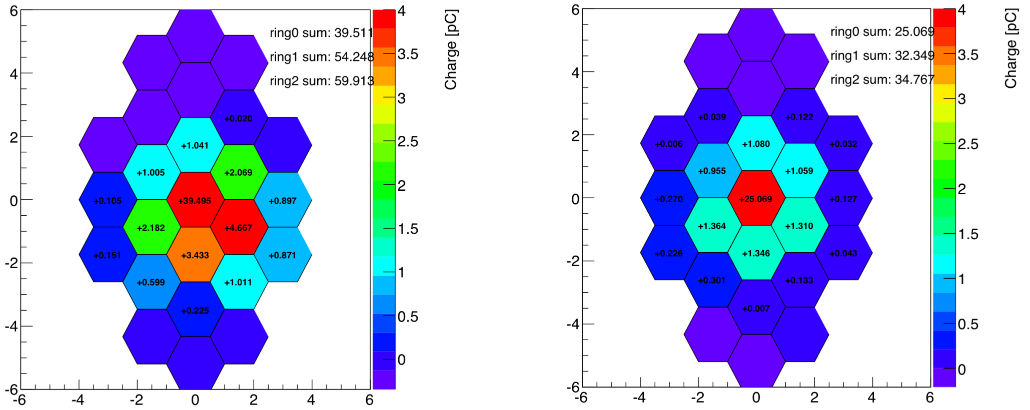
(b) $6X_0$ tungsten absorber placed 1 cm from the pico-sil detector, with 32 GeV electron beam. The charge in the central pixel is 20 pC, average charge in the first ring is 1.9 pC, and the average charge in the second ring is 0.27 pC.



(c) $6X_0$ tungsten absorber placed 32 mm from the pico-sil detector, with 32 GeV electron beam. The charge in the central pixel is 13 pC, average charge in the first ring is 1.2 pC, and the average charge in the second ring is 0.41 pC.

(d) $6X_0$ tungsten absorber placed 75 mm from the pico-sil detector, with 32 GeV electron beam. The charge in the central pixel is 8 pC, average charge in the first ring is 1.2 pC, and the average charge in the second ring is 0.27 pC.

Figure 6.4: Plots of the total charge for each pixel on the pico-sil detector, with the tungsten absorber at varying distances from the detector. As the tungsten is moved further away from the pico-sil, less energy is observed.



(a) $6X_0$ tungsten absorber placed 1 mm from the pico-sil detector, with 32 GeV electron beam. The charge in the central pixel is 39 pC, average charge in the first ring is 2.4 pC, and the average charge in the second ring is 0.49 pC.

(b) $6X_0$ tungsten absorber placed 1 mm from the pico-sil detector, with 16 GeV electron beam. The charge in the central pixel is 25 pC, average charge in the first ring is 1.2 pC, and the average charge in the second ring is 0.12 pC.

Figure 6.5: Plots of the total charge for each pixel on the pico-sil detector, with the tungsten absorber 1 mm from the detector, and with an electron beam of 32 and 16 GeV. With the 16 GeV beam, about half the energy is observed as compared with the 32 GeV beam.

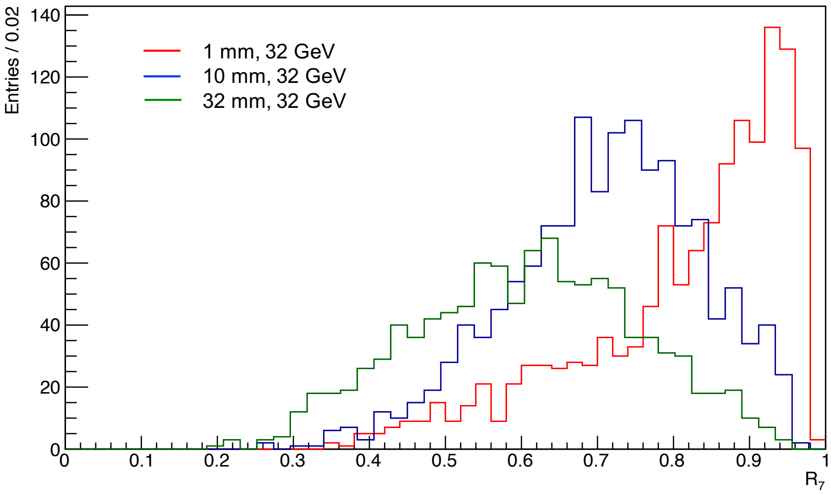
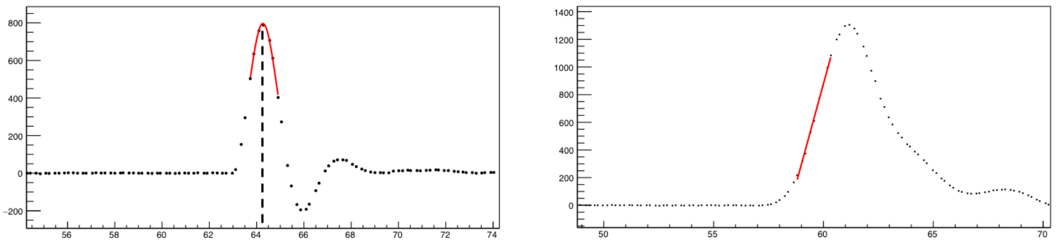


Figure 6.6: Histogram of the number of events vs. R_7 (Equation 6.1), the ratio of the charge contained in the central pixel to all pixels, for various separation distances between the pico-sil HGC and the $6X_0$ tungsten absorber. 32 GeV 75 mm data is not plotted, but overlaps the 32 mm R_7 plot significantly.

time difference of the detection of an event between the pico-sil and the Photek is due to the distance between the two. Therefore, the time difference distribution is expected to be a Gaussian. The absolute difference in time (the mean of the Gaussian) does not matter, but rather the standard deviation of the time difference (σ) gives the time resolution as this is the spread in the time difference between the two event measurements.

In an event, a "pulse" is recorded by each detector, and by the individual pixels in the pico-sil. The Photek (reference timer) has a Gaussian response to an event, and thus its timestamp is determined as the mean of the Gaussian fit to the pulse. The pico-sil pixels do not have a Gaussian response (the rise time is shorter than the fall time), and thus a linear fit is used on the rising edge from 10% to 90% of the maximum amplitude values. The timestamp of the pixel is set as the time when the linear fit reaches 45% of the maximum amplitude. Examples of the timestamp fits are shown in Figure 6.7.

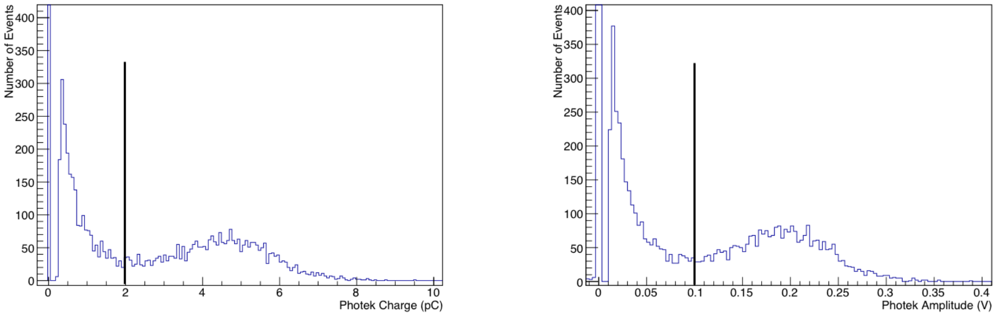


(a) Fit of the pulse observed in the Photek to a Gaussian, with the mean of the Gaussian indicated.

(b) Linear fit of the rising edge of the pico-sil pixel response.

Figure 6.7: Pulses observed in the pico-sil HGC detector and the Photek with the fits to determine the timestamp shown. The fits for the Photek pulse and the pico-sil pixel pulse are a Gaussian and linear fit, respectively.

Many events are recorded, however, not all events are from the electron beam. Thus, electron events must be separated from the background events (pions and muons primarily). The background events have a much lower charge than the electron events, and thus charge and amplitude cuts on the central pixel and the Photek are used to select electron events. Events are selected based on a coincidence observation both in the Photek and in the center pixel of the pico-sil, since this is where the electron beam is focused. Histograms of events vs. event charge are shown in Figure 6.8, with the cuts indicated. Electron events are the Gaussian to the right of the cut, while the background events are centered around 0 pC. Charge and amplitude cuts for the Photek and the central pixel vary based on the beam energy and absorber distance from the pico-sil, and the cuts are shown in Table 6.1.



(a) Histogram of number of events vs. Photek charge, with the cut indicated by the black line.

(b) Histogram of number of events vs. Photek amplitude, with the cut indicated by the black line.

Figure 6.8: Histograms and cuts indicated for the charge and amplitude cuts on the Photek to select electron events. These plots are from the 32 GeV $6X_0$ tungsten 1mm away from the pico-sil runs.

When information from multiple pixels is combined determine the time resolution, each pixel must pass both charge and amplitude cuts (charge > 1 pC and amplitude > 0.01 V). If a pixel does not pass these cuts, it is added with a weighting of 0, such that the event can still be used even if not every pixel passes the cuts. This means that when all pixels are combined, in most events, actually 3 or 4 pixels are combined, since it is rare for all 7 pixels to pass the cuts within the same event. Plots of the number of pixels passing the cuts per event are shown in Figure 6.12.

Run	**Beam Energy (GeV)**	**Separation Distance (mm)**	**Photek amp cut (V)**	**Photek charge cut (pC)**	**Center amp cut (V)**	**Center charge cut (pC)**
104-110, 115, 116	32	1	0.1	2	0.15	11
129-138	32	10	0.1	2	0.1	8
117-122	32	32	0.09	2	0.05	3
123-128	32	75	0.09	2	0.03	2
144-155	16	1	0.03	0.8	0.01	2.5
167-171	8	1	0.015	0.4	0.01	2.5
46-52	8	32	0.04	1	0.01	1

Table 6.1: Table of charge and amplitude cuts for the center and Photek to select electron events.

Time Resolution Improvement with Pixel Combination

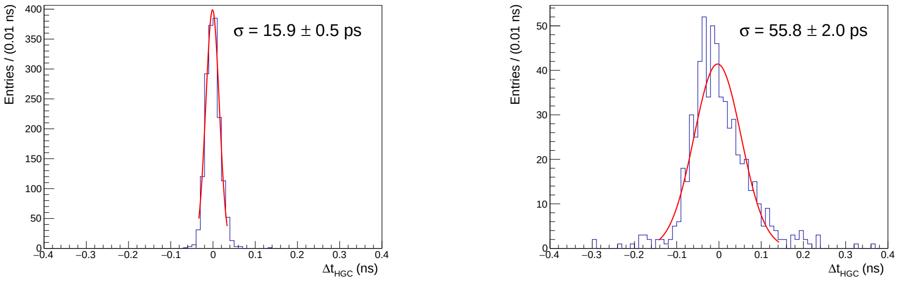
The central pico-sil pixel time resolution is found to be 15.9 ps (Figure 6.9a) with a 1 mm separation between the absorber and detector and a 32 GeV electron beam. This is an outstanding time resolution (corresponding to less than 1 cm spatial resolution), and next the impact of timing information from multiple pixels is investigated.

It is expected that the time resolution will improve as more pixels are combined. When the timing information from each pixel is combined to determine the overall time resolution, they are added in order of how much energy was observed (starting with the pixel with the most energy). Time information from individual pixels is either added with equal weighting, or weighted based on the magnitude of the charge contained.

To measure the time resolution improvement, the time resolution of the center pixel alone is compared to the time resolution of all 7 closest pixels (central plus the first ring of 6 pixels, Figure 6.1). The time resolution is determined by fitting the Δt vs. number of events histogram to a Gaussian, using the range ±2 RMS around the mean of the histogram (Figure 6.9). All fits were done with the log-likelihood method, as this is a better statistical method for bins with low numbers of events (as many of the histograms have approaching the tails).

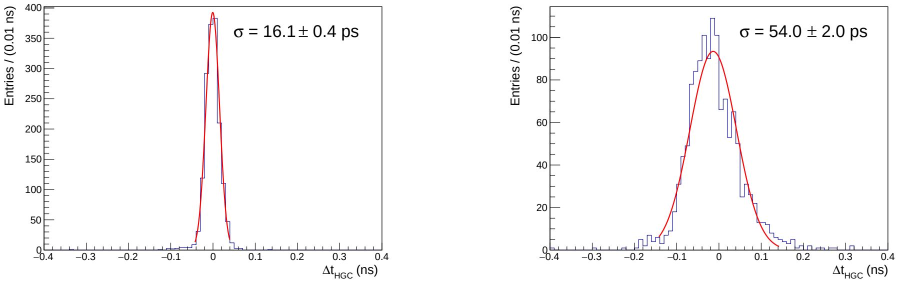
It was found that combining the pixels with equal weighting does not improve the time resolution, but rather makes it considerably larger (Figure 6.11). This is due to the individual time resolutions of the ring pixels (around 60 ps) being considerably larger than that of the central pixel (15.9 ps), and thus a combination of all pixels with an equal pixel weighting gives a larger time resolution than the central pixel alone. Charge weighting the time information (with all pixels combined) did not improve the time resolution significantly (Figure 6.11). This is likely because most of the time information was already contained in the central pixel, so adding additional pixels does not add significant information.

The same analysis was done with the $6X_0$ tungsten absorber at 10 mm, 32 mm, and 75 mm from the pico-sil detector, and results are shown in Figure 6.11. In all cases, combining pixels by equal weighting makes the time resolution larger, while charge weighting keeps the time resolution approximately the same as the center pixel resolution alone. Additionally, there was no significant improvement in the shape of the histograms or their tails seen when the pixels are combined by charge weighting — indicating that combining pixels does not provide an advantage over determining time resolution from the center pixel alone for these arrangements. If



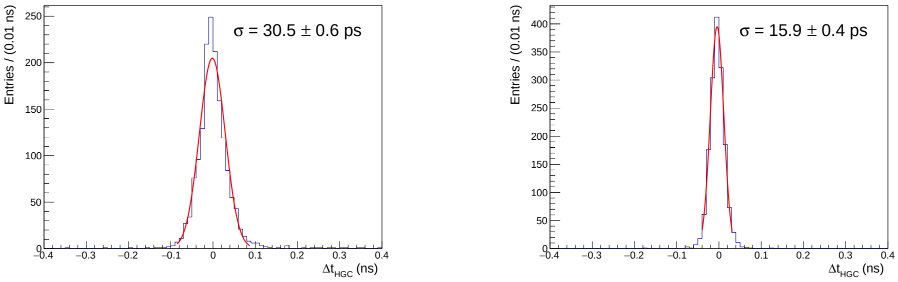
(a) The time resolution fit for the central pixel only.

(b) The time resolution fit for one of the ring pixels (pixel number 4, Figure 6.1a).



(c) The time resolution fit for the pixel with the largest amount of charge. As shown in Figure 6.10, in nearly all the events, the central pixel is the pixel containing the most charge.

(d) The time resolution fit for the pixel with the second largest amount of charge.



(e) The time resolution fit for all 7 pixels, combined with equal weighting for each pixel.

(f) The time resolution fit for all 7 pixels, combined with weighting based on the charge contained in each pixel.

Figure 6.9: Time resolution plots for the 32 GeV electron data, with the tungsten absorber at 1mm from the pico-sil absorber. The central pixel alone (a) has a time resolution of about 16 ps, while the ring pixels have a time resolution of about 50-65 ps (b). In comparing the time resolutions from (a) and (d), the time resolution for the central pixel and the time resolution for the pixel with the second largest amount of charge, it is clear that the center pixel alone is considerably better. When all 7 pixels are combined equally (e), the time resolution is made worse as compared to the central pixel (a). When the pixels are combined with charge weighting (f), the time resolution remains the same as in (a).

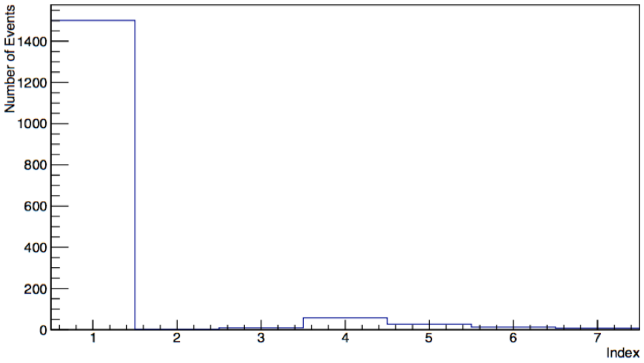


Figure 6.10: Histogram of the index of the pixel with the most charge. In almost all events (≈95%), the central pixel is the pixel with the most charge.

the shower were more spread and the first ring pixels had a larger signal, there may be an improvement in the time resolution seen by combining pixels, since in this case the additional pixels would be adding more time information. A wider shower spread is what is expected in the LHC, and therefore charge weighting may provide a significant advantage here. Further tests with a larger shower spread would be needed to confirm this.

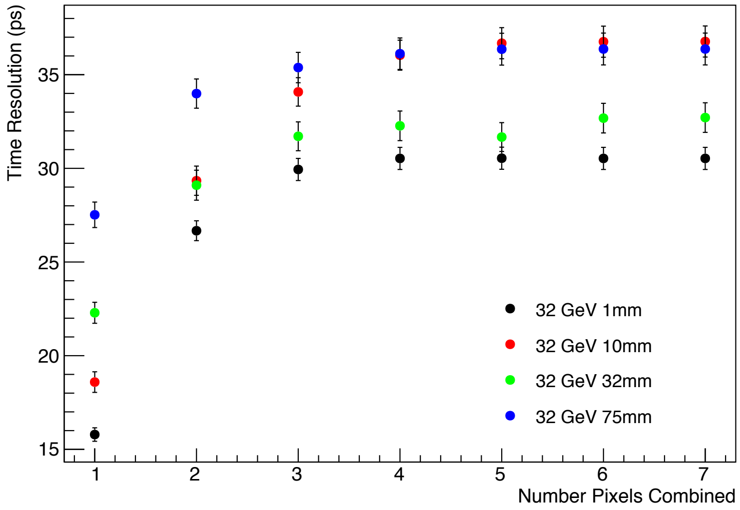
It is expected that the time resolution is related to the number of pixels combined by the relation

$$\frac{\text{initial time resolution}}{\text{final time resolution}} = \sqrt{N} \tag{6.2}$$

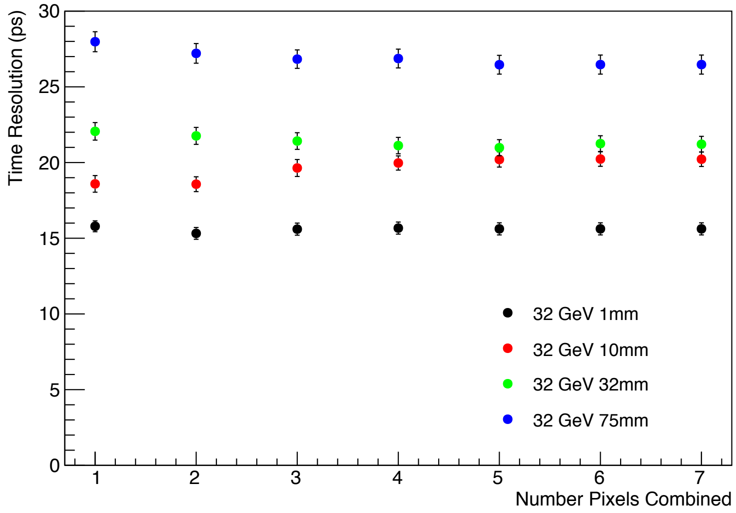
where N is the number of pixels combined (or number of pixels above the charge and amplitude cuts). In an event, less than seven pixels are often active, based on which pixels pass the cuts (pixels that do not pass the cuts are added with a weight of 0 so the event can still be used). The average number of pixels combined per event is a good representation of N, and is determined by the mean of the histograms shown in Figure 6.12.

6.3 Summary and Relevance to Dark Matter Search

Thus, it has been shown that the pico-sil HGC detector has the capability to achieve a time resolution of 16 ps, and with more active pixels from a wider shower shape, time resolution improvement with $\sqrt{N}$ will be seen. The precision timing goals of 30 ps with the high granularity upgrade for CMS are achievable with the high granularity detector, and as demonstrated with the dark matter simulation and analysis, the increased time precision will have large benefits for long-lived particle searches, particularly for precise vertex determination. This will greatly expand the

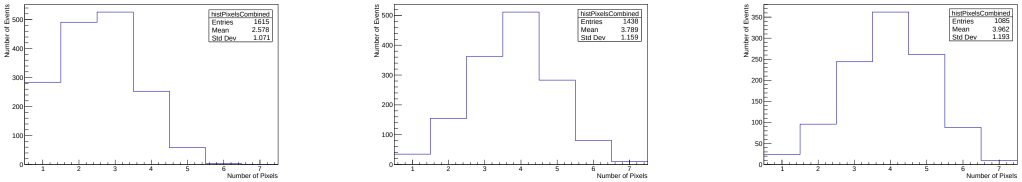


(a) Plot of the time resolutions vs. number of pixels combined (for tungsten at various distances from the pico-sil absorber). The combination is done with equal pixel weighting.



(b) Plot of the time resolutions vs. number of pixels combined (for tungsten at various distances from the pico-sil absorber). The combination is done with charge pixel weighting (based on the charge contained in the pixels).

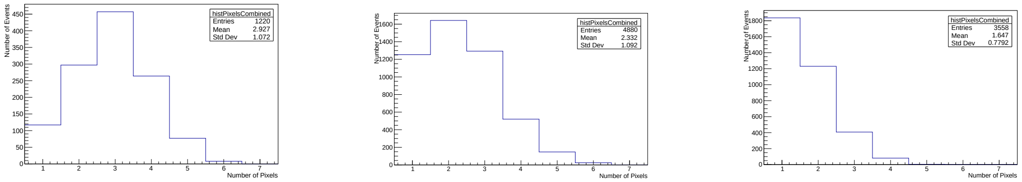
Figure 6.11: The time resolution (in ps) plotted vs. the number of pixels combined. With equal pixel weighting, the time resolution is larger with the combination of all 7 pixels, and with charge weighting, the time resolution is approximately the same as with just the central pixel.



(a) Histogram of largest number of pixels passing the charge and amplitude cuts when the absorber is at 1mm, 32 GeV.

(b) Histogram of largest number of pixels passing the charge and amplitude cuts when the absorber is at 10mm, 32 GeV.

(c) Histogram of largest number of pixels passing the charge and amplitude cuts when the absorber is at 32mm, 32 GeV.



(d) Histogram of largest number of pixels passing the charge and amplitude cuts when the absorber is at 75mm, 32 GeV.

(e) Histogram of largest number of pixels passing the charge and amplitude cuts when the absorber is at 1mm, 16 GeV.

(f) Histogram of largest number of pixels passing the charge and amplitude cuts when the absorber is at 1mm, 8 GeV.

Figure 6.12: The number of pixels passing the charge and amplitude cuts for each absorber distance. Generally, around 3 pixels pass the cuts for a given event. The mean of these histograms is taken as a good representation of the number of pixels combined for each analysis.

phase space available at the LHC and increase its sensitivity to SUSY events. The implementation of the MTD with precision timing and 4D vertex reconstruction capabilities will greatly increase the sensitivity in the GMSB dark matter search.

Chapter 7

CONCLUSION

Searches for physics beyond the Standard Model at the LHC with the CMS experiment have been discussed, with particular focus on the experiment's capability to detect long-lived particles. Gauge mediated supersymmetry breaking models producing neutralinos and gravitinos from proton-proton collisions provide a viable WIMP dark matter candidate, as the gravitino is the LSP and stable. This model is very well motivated by cosmological observations, and also is a model not heavily considered in previous collider searches, as most previous searches use MSSM models where the neutralino is the LSP.

In order to investigate this model, signal ($\tilde{\chi}_1^0 \to \tilde{G} + \gamma$) and background (QCD and photons + jets) events were simulated with CMS Full Simulation, from Pythia and SLHA files. The simulation was validated, and event identification was focused on. To identify the $\tilde{\chi}_1^0 \to \tilde{G} + \gamma$ events from background, the unique signature of two time delayed and displaced photons was used. This relies on the neutralino decaying within the ECAL barrel, but otherwise is relatively model independent. In this analysis, a model where the neutralino could be created from 30 other SUSY particles was used.

To improve the photon identification algorithms, the machine learning classification technique of Boosted Decision Trees (BDTs) was implemented. The BDT takes labeled data, and combines weak classifiers (observables from the CMS detectors) to develop a strong aggregate classifier. The two way BDT model performs better selecting delayed photon events as compared to the previously developed CMS EGamma and current cut based photon identification techniques. Particularly, with the same amount of GMSB signal acceptance, the BDT model reduces the QCD and γ + jets backgrounds considerably, for an overall increase in the significance of the GMSB signal selection by 2.38σ. When timing information is included in the BDT, the significance of the signal to background discrimination is increased to 16.9σ.

Additionally, it was shown that the ECAL timing information has a strong impact on the BDT classification abilities. Adding the timing information from the ECAL allowed the BDT to develop a more accurate classifier, and ECAL time information was one of the most important variable included in the BDT. This further emphasizes the importance of a precision timing detector in the CMS upgrade. The proposed MTD (minimum ionizing particle (MIP) timing detector) will provide 4D tracking and precise time information for vertex reconstruction. In addition to extending the phase space over which the detector will be sensitive to long-lived SUSY particles, the MTD will be particularly useful for models with two delayed photons with these improved identification techniques.

The BDT model performance was shown to be a significant improvement over current cut based photon IDs. Given the ability of the BDT to distinguish between GMSB signal and background at a range of masses and lifetimes, a full comparison of the BDT to the current analysis capabilities is expected to extend the sensitivity of the ECAL to higher masses. The full comparison of the BDT performance to the current cut based photon ID performance is currently being completed by replacing the photon ID cuts with a cut on the discriminator value from the GMSB vs. QCD BDT model.

Results from a precision timing detector test beam analysis are also presented. The detector tested was a silicon high granularity calorimeter prototype for the Phase II CMS endcap upgrade. This detector achieved an outstanding time resolution of 16 ps, and with a wider EM shower spread (as is expected in the jets produced in the LHC) time resolution improvement by combining channels with charge weighting is expected. Thus, it has been demonstrated that a time resolution below 30 ps (1 cm spatial resolution) is achievable. This precision timing capability will be vital for performing a SUSY dark matter search with displaced vertex photons at CMS with a high sensitivity.

Overall, this thesis research has demonstrated that precision timing detectors can achieve a time resolution of 16 ps, and this ability is highly applicable to searches for physics beyond the Standard Model. In particular, a well motivated gauge mediated supersymmetry model with candidate dark matter particles produced in conjunction with displaced and delayed photons is investigated. The event identification (based on photon identification) is greatly improved beyond the current cut based identification algorithms. The improved SUSY dark matter event identification from the BDT combined with the increased tracking and timing abilities of the MTD will allow this search to be performed with high sensitivity in the HL-LHC.

Future continuations of this research include completing the BDT model comparison to the current data analysis results to understand the sensitivity improvement the

BDT with timing information allows for. Additionally, the ECAL timing with the MTD could be simulated to fully understand how the higher precision timing information improves the BDT. Overall, this research demonstrates that BDTs are able to distinguish between signal from a GMSB model and the background (QCD and photons + jets) with a higher accuracy than the current cut based IDs. BDT models will be a very valuable tool in the HL-LHC, and combined with precision timing, will provide a vital way to mitigate the pile-up effects for a variety of physics analyses.

Appendix A

PYTHIA SAMPLES AND SIMULATION

The simulation samples for GMSB, photons + jets, and QCD are created in CMS Full Simulation, starting from Pythia and SLHA files. The mass of the particles (in GeV) are set in the SLHA files, and the lifetimes ($c\tau$ in cm/c) are set in the Pythia files (Appendix B). The allowed processes and decays are also set in the SLHA file. Pythia reads in the SLHA file [27]. To identify particles in the event generators, the Monte Carlo particle numbering scheme is used. In the Monte Carlo numbering system, each particle is given a seven digit number, with information about its spin, flavor, and quantum numbers [31]. Additionally, the Monte Carlo particle numbering system allows for hypothetical and SUSY particles.

Labels for all the particles used in the GMSB model [31]:

35 is the Higgs (H^0/H_2^0)
36 is the Higgs (A^0/H_3^0)
37 is the Higgs (H^+)
1000001 is the left down squark ($\tilde{d}_L$)
1000002 is the left up squark ($\tilde{u}_L$)
1000003 is the left strange squark ($\tilde{s}_L$)
1000004 is the left charm squark ($\tilde{c}_L$)
1000005a is the left bottom squark ($\tilde{b}_1$)
1000006a is the left top squark ($\tilde{t}_1$)
1000011 is the left electron ($\tilde{e_L}^-$)
1000012 is the left electron neutrino ($\tilde{\nu}_{eL}$)
1000013 is the left muon ($\tilde{\mu_L}^-$)
1000014 is the left muon neutrino ($\tilde{\nu}_{\mu L}$)
1000015a is the left tau ($\tilde{\tau}_1{}^-$)
1000016 is the left muon neutrino ($\tilde{\nu}_{\tau L}$)
2000001 is the right down squark ($\tilde{d}_R$)
2000002 is the right up squark ($\tilde{u}_R$)
2000003 is the right strange squark ($\tilde{s}_R$)
2000004 is the right charm squark ($\tilde{c}_R$)
2000005a is the right bottom squark ($\tilde{b}_2$)
2000006a is the right top squark ($\tilde{t}_2$)
2000011 is the right electron ($\tilde{e_R}^-$)
2000013 is the right muon ($\tilde{\mu_R}^-$)
2000015a is the right tau ($\tilde{\tau}_2{}^-$)
1000021 is the gluino ($\tilde{g}$)
1000022 is the lightest neutralino ($\tilde{\chi}_1^0$)
1000023 is a neutralino ($\tilde{\chi}_2^0$)
1000024 is a chargino ($\tilde{\chi}_1^+$)
1000025 is a neutralino ($\tilde{\chi}_3^0$)
1000035 is a neutralino ($\tilde{\chi}_4^0$)
1000037 is a chargino ($\tilde{\chi}_2^+$)
1000039 is the gravitino ($\tilde{G}$)

Appendix B

SLHA FILES

The $\tilde{\chi}_1^0 \rightarrow \tilde{G} + \gamma$ was produced as a GMSB model with SLHA and Pythia files [26] [25]. In the Pythia file, various SUSY processes are turned on:

- "qqbar2chi0chi0" is set to on, which allows "pair production of neutralinos by quark-antiquark annihilation" [48]
- "gg2squarkantisquark" is set to on, allowing "pair production of a scalar quark together with a scalar antiquark by gluon annihilation via s-channel gluon exchange, t- and u-channel squark exchange, and the direct 4-point coupling" [48]
- "qqbar2squarkantisquark" allows "Pair production of a scalar quark together with a scalar antiquark by quark-antiquark annihilation" [48]

In the Pythia file, the neutralino lifetime is set as 200 cm/c.

In the SLHA file, the particle masses set are:

PDG Code	**Particle**	**Mass (GeV)**
24	W Boson	80.40
25	Higgs Boson	125
1000022	$\tilde{\chi}_1^0$ (neutralino)	358
1000039	$\tilde{G}$ (gravitino)	1
	Other SUSY particles	10^4

Table B.1: Values set in the SLHA file for the SUSY particles. All other SUSY particles have a mass set to 10^4 GeV.

The branching ratio of the decay $\tilde{\chi}_1^0 \rightarrow \tilde{G} + \gamma$ is set as 1.

Appendix C

GAUGE MEDIATED SUPERSYMMETRY BREAKING MODEL

Details of the gauge mediated supersymmetry breaking model are given in *Theories with Gauge-Mediated Supersymmetry Breaking* by G.F. Giudice1 and R. Rattazzi [49].

The neutralino mass is approximately linear in SUSY Λ scale, though also depends on the messenger mass and $\tan\beta$.

For the Monte Carlo simulation:

$$\frac{M_{messenger}}{\Lambda} = 2 \tag{C.1}$$

$$\tan\beta = 15 \tag{C.2}$$

The gravitino mass is related to the neutralino lifetime and mass scale Λ, approximately by

$$mass \approx \sqrt{c\tau} \tag{C.3}$$

Appendix D

MULTIPLE BDT SIMULATION REQUIREMENTS

This considers the requirements for handling simulation training data for use in a BDT trained on GMSB vs. QCD, and then a separate BDT trained on GMSB vs. γ + jets. This approach is not tested in the thesis research due to the additional signal samples needed — the same GMSB file cannot be used to train two BDTs.

Consider the approach of making a BDT (2 way classifier), then saving the discriminator values as an additional variable in the ntuple of the input ROOT files. The signal file could now be used to make another 2 way BDT, where the discriminator variable is cut on. Can the same signal file be used in both cases? *No — this original signal file will be biased.*

Consider $\vec{x}_1$ as the set trained on, with events t_1. $\vec{x}_2$ is a set not trained on, with events t_2. After the first training (GMSB vs. QCD), the discriminator for $\vec{x}_1$ will be very high (it is biased after being trained on). Therefore, given a discriminator value d_1:

$$p(d_1|\vec{x}_1) \neq p(d_1|\vec{x}_2) \tag{D.1}$$

And in particular,

$$p(d_1 > t_1|\vec{x}_1) > p(d_1 > t_1|\vec{x}_2) \tag{D.2}$$

Given this, then (up to normalization factor, application of Bayes' Theorem)

$$p(\vec{x}_1|d_1 > t_1) \propto p(d_1 > t_1|\vec{x}_1)p(\vec{x}_1) = p(d_1 > t_1|\vec{x}_1)p(\vec{x}) \tag{D.3}$$

Similarly for $\vec{x}_2$ (up to normalization factor)

$$p(\vec{x}_2|d_1 > t_1) \propto p(d_1 > t_1|\vec{x}_2)p(\vec{x}_2) = p(d_1 > t_1|\vec{x}_2)p(\vec{x}) \tag{D.4}$$

Then this means

$$p(\vec{x}_1|d_1 > t_1) \neq p(\vec{x}_2|d_1 > t_1) \tag{D.5}$$

Therefore, once a data set has been used in a BDT training, it will be biased based on the discriminator values from the first training. The probability that an event falls above the discriminator cut is not equal for the set trained on and another data set not trained on. Thus, a new simulation data sample has to be used in the second BDT training such that the model is not biased.

BIBLIOGRAPHY

[1] Stephen P. Martin. “A Supersymmetry primer”. In: (1997). [Adv. Ser. Direct. High Energy Phys.18,1(1998)], pp. 1–98. DOI: `10.1142/9789812839657_0001,10.1142/9789814307505_0001`. arXiv: `hep-ph/9709356 [hep-ph]`.

[2] Cristián Peña. “Searches for New Physics at the Compact Muon Solenoid Experiment and Precision Timing Calorimetry”. California Institute of Technology. PhD thesis. Nov. 2016.

[3] CMS. “Technical Proposal for a MIP Timing Detector in the CMS Experiment Phase 2 Upgrade”. In: *The Compact Muon Solenoid Experiment* CMS-TDR-17-006 (2017). Ed. by Lindsey Gray and T. Tabarelli de Fatis.

[4] Lindsey Gray. “Proposal for a MIP Timing Detector in the CMS Phase 2 Upgrade”. LHCC Meeting. Dec. 2017.

[5] Chris Neu on behalf of the MTD project. *MIP Timing Detector: Outcome of LHCC Review, Status and Plans*. 2018.

[6] CERN. “Overview of CMS Physics Goals and Detector”. In: *TWiki, CMS Public Web* (2015). URL: `https://twiki.cern.ch/twiki/bin/view/CMSPublic/WorkBookCMSExperiment`.

[7] “Performance of photon reconstruction and identification with the CMS detector in proton-proton collisions at $\sqrt{s} = 8$ TeV”. In: *Journal of Instrumentation* 10, P08010 (Aug. 2015), P08010. DOI: `10.1088/1748-0221/10/08/P08010`. arXiv: `1502.02702 [physics.ins-det]`.

[8] CMS Collaboration. “Particle-flow Reconstruction and Global Event Description with the CMS Detector”. In: *JINST 12 (2017) P10003* CMS-PRF-14-001, CERN-EP-2017-110 (2017).

[9] CMS Collaboration. “CMS Tracking Performance Results from Early LHC Operation”. In: *CMS-TRK-10-001* (2010). URL: `https://arxiv.org/pdf/1007.1988.pdf`.

[10] CMS Collaboration. “Energy calibration and resolution of the CMS electromagnetic calorimeter in pp collisions at $\sqrt{s} = 7$ TeV”. In: *CMS-EGM-11-001* (Oct. 2013). URL: `https://arxiv.org/pdf/1306.2016.pdf`.

[11] Daniele Franci. “Search for SUSY events with off-time photons”. Sapienza Universita di Roma. PhD thesis. Oct. 2010.

[12] CMS Collaboration. “Search for long-lived particles in events with photons and missing energy in proton-proton collisions at $\sqrt{s} = 7$ TeV”. In: *ArXiv e-prints* (Dec. 2012). arXiv: `1212.1838 [hep-ex]`. URL: `http://adsabs.harvard.edu/abs/2012arXiv1212.1838C`.

[13] D. Contardo et al. *Technical Proposal for the Phase-II Upgrade of the CMS Detector*. Tech. rep. CERN-LHCC-2015-010. LHCC-P-008. CMS-TDR-15-02. Geneva, June 2015. URL: `https://cds.cern.ch/record/2020886`.

[14] European Organization for Nuclear Research - CERN Document Server. *Dark Matter*. 2018. URL: `https://home.cern/about/physics/dark-matter` (visited on 04/03/2018).

[15] G. Bertone, D. Hooper, and J. Silk. "Particle dark matter: evidence, candidates and constraints". In: 405 (Jan. 2005), pp. 279–390. DOI: `10.1016/j.physrep.2004.08.031`. eprint: `hep-ph/0404175`.

[16] Jennifer M. Gaskins. "A review of indirect searches for particle dark matter". In: *Contemporary Physics* 57 (Oct. 2016), pp. 496–525. DOI: `10.1080/00107514.2016.1175160`. arXiv: `1604.00014 [astro-ph.HE]`.

[17] Artur Apresyan et al. "Discussion". Feb. 2017.

[18] Cristián Peña and Dustin Anderson. "Discussion". Apr. 2018.

[19] CMS Collaboration. "Search for long-lived neutral particles decaying to quark-antiquark pairs in proton-proton collisions at $\sqrt{s} = 8$ TeV". In: *Physical Review D* 91 (2015), p. 012007.

[20] CERN. "Geant4 - A Simulation Toolkit". In: (2018). URL: `http://geant4.web.cern.ch`.

[21] CERN. "CMSSW Application Framework". In: *TWiki, CMS Public Web* (Nov. 2017). URL: `https://twiki.cern.ch/twiki/bin/view/CMSPublic/WorkBookCMSSWFramework`.

[22] "CMS Offline Software". In: (2018). URL: `https://github.com/cms-sw/cmssw`.

[23] CERN. "Full Simulation Offline Guide". In: *TWiki, CMS Public Web* (2017). URL: `https://twiki.cern.ch/twiki/bin/view/CMSPublic/SWGuideSimulation`.

[24] Torbjorn Sjostrand, Stephen Mrenna, and Peter Skands. "PYTHIA 6.4 Physics and Manual". In: *JHEP* 05 (2006), p. 026. DOI: `10.1088/1126-6708/2006/05/026`. arXiv: `hep-ph/0603175 [hep-ph]`.

[25] Christian Bierlich et al. "Present program version: PYTHIA 8.2". In: (2018). URL: `http://home.thep.lu.se/~torbjorn/Pythia.html`.

[26] Peter Z. Skands et al. "SUSY Les Houches accord: Interfacing SUSY spectrum calculators, decay packages, and event generators". In: *JHEP* 07 (2004), p. 036. DOI: `10.1088/1126-6708/2004/07/036`. arXiv: `hep-ph/0311123 [hep-ph]`.

[27] N. Desai and P. Z. Skands. "Supersymmetry and Generic BSM Models in Pythia 8". In: *European Physical Journal C, CERN-PH-TH/2011-210* 72, 2238 (Dec. 2012), p. 2238. DOI: `10.1140/epjc/s10052-012-2238-0`. arXiv: `1109.5852 [hep-ph]`.

[28] CERN. "ROOT Data Analysis Framework". In: (2018). URL: `https://root.cern.ch`.

[29] CERN. "Overview of ECAL Objects and Software". In: *TWiki, CMS Public Web* (2014). URL: `https://twiki.cern.ch/twiki/bin/view/CMSPublic/SWGuideEcalOverview`.

[30] CMS Collaboration. "The Electromagnetic Calorimeter Technical Design Report". In: CMS TDR 4, CERN/LHCC 97–33 (1997). Ed. by Felicitas Pauss.

[31] L. Garren et al. "Monte Carlo Particle Numbering Scheme". In: *Particle Data Group* (2009).

[32] G.F. Giudice and R. Rattazzi. "Theories with gauge-mediated supersymmetry breaking". In: *Physics Reports* 322.6 (1999), pp. 419–499. ISSN: 0370-1573. DOI: `https://doi.org/10.1016/S0370-1573(99)00042-3`. URL: `http://www.sciencedirect.com/science/article/pii/S0370157399000423`.

[33] R. Barbiera et al. "R-Parity-violating supersymmetry". In: *Physics Reports* 1-195 (2005). URL: `https://ac.els-cdn.com/S0370157305003327/1-s2.0-S0370157305003327-main.pdf?_tid=9b77776c-fbc6-4f7c-a4e5-73bc7f8fc635&acdnat=1524439399_78b705073fccb4a6fe13492fc94bd208`.

[34] Stephen Myers. "The LEP Collider, from design to approval and commissioning". In: *CERN-91-08* (1991). Delivered at CERN, 26 Nov 1990.

[35] Microsoft Azure. "Boosted Decision Tree Regression". In: (Jan. 2018). URL: `https://docs.microsoft.com/en-us/azure/machine-learning/studio-module-reference/boosted-decision-tree-regression`.

[36] Thomas M. Mitchell. *Machine Learning*. 1st ed. New York, NY, USA: McGraw-Hill, Inc., 1997.

[37] Thomas G. Dietterich. "An Experimental Comparison of Three Methods for Constructing Ensembles of Decision Trees: Bagging, Boosting, and Randomization". In: *Machine Learning* 40 (2000), pp. 139–157.

[38] Ron Kohavi and Clayton Kunz. "Option Decision Trees with Majority Votes". In: *In Proceedings of the Fourteenth International Conference on Machine Learning* (1997), pp. 161–169.

[39] XGBoost. *Introduction to Boosted Trees*. 2016. URL: `http://xgboost.readthedocs.io/en/latest/model.html`.

[40] Wikipedia contributors. *Logistic function — Wikipedia, The Free Encyclopedia*. [Online; accessed 21-April-2018]. 2018. URL: `https://en.wikipedia.org/w/index.php?title=Logistic_function&oldid=836519677`.

[41] Tom Fawcett. "An introduction to ROC analysis". In: *Pattern Recognition Letters* 27 (2006), pp. 861–874.

[42] Jerome H. Friedman. *Greedy Function Approximation: A Gradient Boosting Machine*. IMS 1999 Reitz Lecture. 1999.

[43] CERN. *Cut Based Photon ID for Run 2 - Working points for 2016 data for 80X and later samples and Selection implementation details for SPRING16*. 2018. URL: `https://twiki.cern.ch/twiki/bin/viewauth/CMS/CutBasedPhotonIdentificationRun2`.

[44] Adi Bornheim et al. "Displaced Photon Searches with CMS in LHC Run II". In: (Long-lived Exotica WG Meeting). Nov. 2017. URL: `https://indico.cern.ch/event/681779/contributions/2793288/attachments/1560769/2456940/DelayedPhoton_17Nov2017.pdf`.

[45] Si Xie. "Precision Timing for Future Collider Experiments". PowerPoint from Harvard LPPC Seminar. Nov. 2015.

[46] B. Bilki. "CMS Forward Calorimeters Phase II Upgrade - 16th International Conference on Calorimetry in High Energy Physics". In: *IOP Publishing* Conference Series 587 (2015).

[47] William R. Leo. *Techniques for Nuclear and Particle Physics Experiments - A How To Approach*. Springer-Verlag, 1987.

[48] Torbjorn Sjostrand. "Pythia 8 Index - Program Overview - SUSY". In: (). URL: `http://home.thep.lu.se/~torbjorn/pythia81html/Welcome.html`.

[49] G. F. Giudice and R. Rattazzi. "Theories with gauge mediated supersymmetry breaking". In: *Phys. Rept.* 322 (1999), pp. 419–499. DOI: `10.1016/S0370-1573(99)00042-3`. arXiv: `hep-ph/9801271 [hep-ph]`.