ALPHA/BETA PULSE SHAPE CLASSIFICATION IN A NOVEL SCINTILLATOR

by

Alexander Fay

A thesis submitted to the faculty of
The University of North Carolina at Charlotte
in partial fulfillment of the requirements
for the degree of Master of Science in
Electrical Engineering

Charlotte

2024

Approved by:

Dr. Samuel Shue

Dr. James Conrad

Dr. Ahmed Arafa

ABSTRACT

ALEXANDER FAY. Alpha/beta pulse shape classification in a novel scintillator.
(Under the direction of DR. SAMUEL SHUE)

The performance of existing methods for the digital and analog discrimination of alpha and beta particles in mixed radiation fields is examined for signals produced using a novel ruggedized scintillator. Leveraging frequency analysis of simulated signals, an experiment has been designed and used to obtain a representative sample of time domain electronic pulse data. The experiment was conducted at the Triangle Universities Nuclear Laboratory, where emissions from radioactive elements Americium 241, Plutonium 239, Cesium 137, Cobalt 60, Polonium 210, Strontium 90, and Thallium 204 were used to collect a diverse set of energies for alpha, beta, and gamma particles. Statistical signal processing techniques are utilized to determine optimal parameters for use in rise-time, pulse gradient, and time over threshold pulse shape discrimination techniques. Support vector machines, and k nearest neighbor classifiers are used to extend the discrimination performance by utilizing rise-time, pulse gradient, and time over threshold coupled with the peak height of the event. Rise-time discrimination is found to be 98.95 percent accurate for beta particles, and 99.29 percent accurate for alpha particles. Pulse gradient discrimination is found to be 99.92 percent accurate for beta particles, and 99.98 percent accurate for alpha particles. Finally, time over threshold discrimination is found to be 94.72 percent accurate for beta particles, and 99.50 percent accurate for alpha particles. Using classification methods and peak height, most methods produce an additional 1-2 percent accuracy in all cases.

All of the analysis is repeated after artificially down sampling the available data to mimic 50 MHz, 100 MHz, 200 MHz, 1 GHz, and 2 GHz sampling rates. Under this treatment, the best consistency is obtained with pulse gradient discrimination, which maintains greater than 99 percent accuracy even at 50 MHz, compared to rise-time, which falls to under 50 percent accuracy, and time over threshold which falls to under 80 percent accuracy at 50 MHz. The dataset is made available for more advanced study, and the gamma data is also available for researchers to attempt beta/gamma or alpha/gamma discrimination. While more features can be found for the pulses, and more advanced techniques can be used to obtain higher accuracy metrics, the benefit of the analysis completed here is its transferability to embedded and computationally constrained systems. Each method used here, with the exception of k nearest neighbors is capable of classifying an incoming pulse in a handful of clock cycles, compared to more complex systems that require more power and more computation time.

DEDICATION

This thesis is dedicated first and foremost to my fiancée Katy. The last two years of work undertaken for this degree would not have been possible without her constant support and encouragement. Whether it was picking up extra work around the apartment, or reminding me of what I am capable of, and what is really important, she was there for me every step of the way. I would also like to dedicate this thesis to my family, who have been instrumental in my growth into the person I am today, and who have fostered and encouraged my academic pursuits throughout my life. Finally I would like to dedicate this thesis to my cats, Bean and Sushi, who made sure I always had someone to run my ideas by, and who also made sure I never overworked myself, even if I wanted to.

ACKNOWLEDGEMENTS

The execution of this research was enabled by the guidance and discussion provided by Dr. Samuel Shue. Through his efforts, I have been able to elevate my research beyond what I would have expected myself to be able to achieve. I would also like to thank Dr. Ahmed Arafa for his instruction and guidance for theoretical and statistical approaches to signal processing. His teaching provided me with answers that I needed to advance the success and rigor of the methods applied here. Finally, I would like to thank Dr. Jim Conrad for his guidance in embedded systems, thanks to him I have a much better perspective on the operational constraints of embedded devices.

Thanks is also due to my coworkers at Corvid Technologies, who have sought to foster not just my professional growth, but my academic growth, and technical skill set. Dr. Sean Hunt in particular has spent a great deal of time helping me comprehend so much of the nuclear physics contained within, a field within which I have not spent much time before working with him.

Some material items obtained for the purposes of this research were funded by the Faculty Research Grant of the University of North Carolina at Charlotte. I would also like to acknowledge the financial support provided to me throughout the degree process by Corvid Technologies.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

ANN	An acronym for Artificial Neural Network
CTI	An acronym for Charge Tail Integration
FFT	An acronym for Fast Fourier Transform
KNN	An acronym for K Nearest Neighbors
PGA	An acronym for Pulse Gradient Analysis
PHA	An acronym for Pulse Height Analysis
PMT	An acronym for Photomultiplier Tube
PSA	An acronym for Pulse Shape Analysis
PSD	An acronym for Pulse Shape Discrimination
SiPM	An acronym for Silicon Photomultiplier
SVM	An acronym for Support Vector Machine
ToT	An acronym for Time over Threshold
TUNL	An acronym for the Triangle Universities Nuclear Laboratory

PREFACE

This thesis aims to provide a clear analysis of potential approaches to classifying the pulses produced by a new and novel scintillator, while providing theoretical rigor for the classifier design process. The techniques employed here are drawn from the existing literature in nuclear physics and classifiers, and are chosen to be feasible and practical for implementation in embedded and computation constrained environments. In particular, this thesis draws from foundational work in pulse shape discrimination, particularly in rise-time, pulse gradient, and time over threshold approaches. At both the experiment and analysis stages, motivation for the decisions made is derived through engagement with statistical signal processing theory. The results of this thesis should clearly inform the reader as to the best approach to be taken for their specific application involving this scintillator, while also providing a clear path forward for the furthering of this work.

The chapters of this thesis should build upon one another, beginning with an overview of the physics of scintillation, and its key importance across many fields. From there, analysis is conducted on the expected characteristics of the signals being investigated to drive experimental design and execution. This section also contains all necessary details for researchers to repeat and test the work conducted here. The methodology for handling the acquired data is covered in depth starting with the preprocessing techniques applied, and ending with a summarization of the supporting statistical theory. Finally, results are laid out and comparative analysis is applied to aid the reader in determining which of the assessed methods would perform best in the specific use case.

CHAPTER 1: INTRODUCTION

Particle radiation, whereby an unstable nucleus ejects energetic particles, plays a significant role in multiple fields, such as radioactive medicine, nuclear engineering, and the fundamental sciences. In medicine, lightly radioactive samples can be ingested by patients, with their radioactive emissions used to track its progress through their body, or strong radiation can be used to destroy harmful cells [1]. Radiation can even be used to create images of structures inside a patients body for the purpose of producing a clear diagnosis [2]. Nuclear engineering is a key component of power production, and defense [**?**]. Finally, particle radiation and its study is a key aspect of furthering investigation and validation of the standard model of physics [3]. The types of particle radiation are classified by the particle that has been ejected: α particles refer to the ejection of a He^{2+} nuclei, β particles refer to the emission of a high energy electron or positron, and γ particles refer to the emission of high energy photons.

Utilization and measurement of particle radiation is also a key component of radioactive hazard response, which can be due to both accidents, or adversarial attacks [4, 5]. While needs of specific applications vary, in many safety and hazard response situations, knowledge of the intensity of the particle radiation is just as important as information about the type of particle being measured. This is because different types of radiation interact with matter differently. α particles carry large amounts of energy, but due to their strong charge and larger mass, their energy is quickly attenuated in most mediums. This makes α radiation relatively safe so long as protective gear and precautions are employed. β radiation however, is less energetic, but also much less massive. It does not attenuate quickly through most mediums, and as such, it is generally better to employ dosage related precautions, such as minimizing exposure towards sources to as low as is reasonable. When responding to hazardous situations, or executing safety procedures related to radiation, handheld particle detectors are used to scan and classify the radiation seen in an affected area. Typically, due to the underlying physics related to the measurement of the particle radiation, two different handheld detector types are used. "Geiger counters" are used in areas where γ contamination is presumed to be present, and then α/β detectors are used in cases where a mixed radiation field is expected.

This thesis is concerned with the particular case of handheld α/β particle detectors and the advancement of their technology. Handheld particle detectors commonly use scintillation detection to measure and classify particle radiation. Scintillation is the emission of light following energy deposition by a radiating particle inside a special material, called a scintillator. The scintillator converts the energy into a pulse of light, and depending on the type of radiation that deposited the energy, the dynamics governing the light production are different. This difference in dynamics can give rise to differently shaped pulses when measured by a detector. Figure 1.1 shows an example of two different light pulses that were detected using a silicon photomultiplier (SiPM), with the left plot showing the pulse produced by a β particle, and the right plot showing an α pulse. The light pulses can be processed through use of pulse height analysis (PHA) and pulse shape analysis (PSA). In detectors where the signals produced by the scintillated light are powerful enough, PHA can be employed to very easily discriminate between particles by checking whether the max value of a pulse was above, or below a specific threshold. Given its simplicity, PHA can have a high error rate, and is only feasible in cases where the expected signal amplitudes are large. In scintillation detection, due to the absorption of light, imperfections in reflection of light, and other losses, it can be difficult to achieve the necessary signal strengths for PHA to be reliable. When this is the case, PSA is applied, whereby targeted filters are used to pull out features that are intrinsic to the "shape" of the pulse. While

PSA generally provides better accuracy in the classification of particle radiation, this benefit comes at the cost of more complicated implementation at either the hardware or software level.

Recently, a new ruggedized scintillator has been developed by Corvid Technologies which addresses the weakest component of many α/β detectors, the aluminized Mylar window. Most modern α/β detectors use a 4-6μm thick aluminized Mylar window which is easily damaged, often rendering the detector useless. However, due to intrinsic aspects of the ruggedized scintillators construction, it produces low amplitude signals that are incompatible with PHA. This thesis seeks to utilize existing methods of PSA, also called pulse shape discrimination (PSD) and apply them to the signals produced by this new scintillator. Three existing methods of PSD are drawn from the literature and examined in detail. Signal rise-time, pulse gradient analysis (PGA), and time over threshold (ToT) represent a gradient in the simplicity of implementation starting at rise-time as the simplest, to PGA as the most difficult and costly [6, 7, 8]. In order to produce the data required for an in depth assessment of the capabilities of the existing PSD methods, an experiment was designed and executed at the Triangle Universities Nuclear Laboratory (TUNL), and Corvid Technologies. In these experiments, ^{241}Am, ^{239}Pu, ^{137}Cs, ^{60}Co, ^{210}Po, ^{90}Sr, and ^{204}Tl were obtained to produce 140,000 pulses for the shape analysis. When broken down by particle type, approximately 60,000 pulses came from the α and β particles each, and about 20,000 were obtained from γ particles. The γ particle dataset will not be considered in this work, and is provided for potential follow-on research into potential differentiation between β and γ particle radiation. The effectiveness of the PSD developed here is also extended through the use of the classification techniques support vector machines (SVM) and k nearest neighbors (KNN).

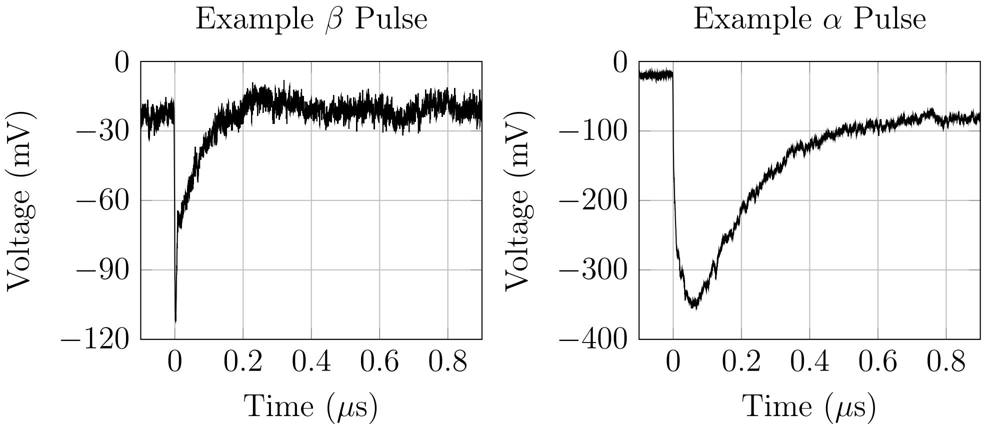

Figure 1.1: An example of two scintillations captured via photomultiplier, and resolved using a trans-impedance amplifier and a high speed oscilloscope. The left pulse originated from a β particle. Since β particles typically possess low energy, there is little to no delayed fluorescence, as can be seen by the clean exponential shape of the pulse. The right pulse originated from an α particle. α particles are much higher energy, and as such there is noticeable delayed fluorescence in the pulse shape.

1.1 Overview of Scintillation Detection

Scintillation detection has a long history in the physical sciences, which can be traced all the way back as far as the experiments of Rutherford and others [9]. The core principal of scintillation detection is that there are materials with unbound π electrons which can be excited by the electric fields intrinsic to α particles, which carry a positive charge, and β particles which can carry either a positive or negative charge. When excited, these unbound electrons are promoted into excited states, from which they eventually de-excite and return to the ground state. During the de-excitation process, each electron releases a photon with energy equal to that of the energy difference between the excited and ground states. This ensemble of photons which are emitted is referred to as "fluorescence," and the intensity of the emitted light over time can be modeled as an exponential function

$$I = I_0 exp(-\frac{t}{\tau}) \tag{1.1}$$

with I_0 the initial light intensity, t being time, and τ the intrinsic fluorescence decay time [10].

However, alternative paths do however exist, whereby an electron can oscillate for some nanoseconds between atomic states before fully returning to the ground state. During this oscillation period, photons of a different wavelength are emitted through a process known as “phosphorescence.” When the electron does finally return to the ground state and release a photon, this is called “delayed fluorescence,” and it is precisely this that causes the differences in light pulses that enable PSA and PSD to operate [11]. This delayed fluorescence can cause recorded light pulses to have shapes that are no longer perfectly exponential, and have decay times that are tens or hundreds of nanoseconds longer. Most importantly, the proportion of delayed fluorescence to general fluorescence is proportional to the energy loss of the particle along its path through the material [12]. This property essentially implies that the more energy a particle has, the wider its peak, and the slower its decay.

Aside from the delayed fluorescence, another driver of the resulting pulse shape is the scintillator itself. A single scintillating material, depending on its composition can be capable of scintillating α, β, and γ particles. Certain materials however, can be chemically or physically constructed such that they are statistically unlikely to scintillate radiations of different kinds. For example, thickness of a scintillator can determine the amount of energy a γ particle can deposit, and thus determine whether the resulting signal produces a measurable pulse [13]. If targeting different radiation than γ, the scintillator can then be made thin enough for most γ’s to be ignored. Other scintillators are constructed with band gaps that are large enough to ignore low energy radiation, which is useful for rejecting β particles and measuring only α particles, since there is a large disparity in carried energy. In α/β detection problems, the scintillator is usually composed of a stack consisting of an α sensitive layer such as silver activated zinc sulfide (ZnS:Ag), and a β sensitive layer such as EJ-212 which is a specially doped scintillating plastic. These different scintillating materials have their own light production dynamics driven by their chemical structure. This aspect further shapes the resulting light pulses produced by α and β particles.

These exponential light pulses are acquired using very sensitive sensors called photomultipliers, which generally use hundreds of volts to produce strong momentary currents which are resolved using a variety of circuits, such as a trans-impedance amplifier. An example of scintillation pulses produced by both an α and a β particle are displayed in Figure 1.1, where the concept of delayed fluorescence is visible when comparing the lower energy β particle scintillation with the higher energy α particle scintillation. A downside of light detecting electronics such as photomultipliers is the requirement that the detector be in total darkness. If exposed to ambient, or even low-light conditions, the photomultiplier can saturate and in many cases, destroy itself. This requires all scintillation detectors, handheld included, to have near perfect light-blocking mechanisms, usually achieved through a layer of material placed atop the scintillator, and sealed to the detector housing.

The full scintillation process from start to finish is shown in Figure 1.2 illustrating all of the components discussed above. It should be noted here that when light is emitted inside the scintillator, it does not have a coherent directionality. In principle, the resulting light emissions from the scintillator are released spherically, with only a fraction being released in the direction of the detection electronics. Due to this, many detectors, handheld or otherwise will employ varying degrees of optical engineering to redirect as much light as possible into the sensor, or sensors.

1.2 Motivation

The major motivating principle for this work is to assist in the ruggedization and advancement of modern scintillation detectors. A number of the target applications for handheld α/β detectors require portability and reliability in the field and potentially even in adversarial situations. This is a critical issue for modern scintillation

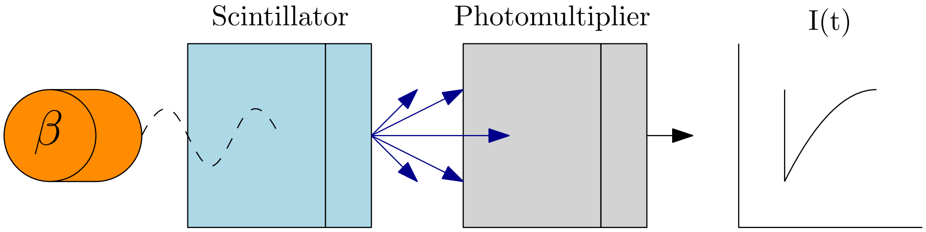

Figure 1.2: Diagram of the operating principle of a scintillation detector with a β particle source directed at a scintillator. Incident radiation impacts the scintillator and deposits its energy. The scintillator fluoresces, and photons are released. The photons are captured in the photomultiplier, which produces a current pulse that can be resolved by electronics.

detectors, mainly due to protective light blocking layers. A major constraint placed on the protective layers of α/β detectors is the rapid attenuation of α particles. β particles deposit very little energy in the materials they move through, however, α particles will deposit large amounts of energy, and quickly attenuate themselves into nothing in as few as millimeters of air. To accurately detect α particles without severely impacting the energy spectrum that can be measured, the protective layer must be kept as thin as possible. This leads to a conundrum, where the protective layer must be able to be sealed completely against outside light, while also being as thin as possible. Most modern scintillation detectors approach this problem using 2-4 μm of aluminized Mylar for their protective layer, which can be seen in Figure 1.3. However, Mylar this thin is incredibly fragile, and even light abrasion or contact with a blade of grass can be enough to puncture the layer and render the detector useless.

In response to this, Corvid Technologies has developed a novel, proprietary, ruggedized scintillator, which is shown in detail in Figure 1.4. The scintillator incorporates carbon doped composite materials that exhibit high degrees of light absorption and structural strength for the protective layer. Subsequent layers are also composed of ZnS:Ag powder suspended in epoxy for additional strength compared to similar scintillators where the adhesion of the scintillating power to the substrate is poor. In detectors that utilize aluminized Mylar, light emitted back towards the opaque

Figure 1.3: View of the aluminized Mylar screen on a popular Ludlum 43-93 model α/β detector. This screen is generally 2-4 μm thick, and very fragile, hence the protective mesh in front of the screen. Direct contact with most materials will easily puncture or shred the Mylar, rendering the detector useless.

layer can be reflected by the metallic surface, and eventually reach the detector. Due to the carbon doping of the Corvid scintillator protective layer, light emitted in the direction of the protective layer is absorbed rather than reflected.

The absorption of this light poses a major issue for widespread adoption of this new scintillator in handheld detectors. Most handheld detectors utilizing Mylar implement PHA as their method for classifying the produced light pulses. Figure 1.5 illustrates the issue preventing more widespread adoption of the Corvid scintillator. This issue motivated the search for suitable PSD methods that would be able to take the signals produced by this new, rugged scintillator, and discriminate the α particles from the β particles. Given that the target application is in handheld devices, particular importance is placed on PSD solutions that are feasible to implement on embedded hardware through either digital or analog electronics with classification latency as low as reasonably possible. Particular importance is placed on solutions with low computational requirements, as this allows the embedded processor to be cheaper, lower power, and have less complex software.

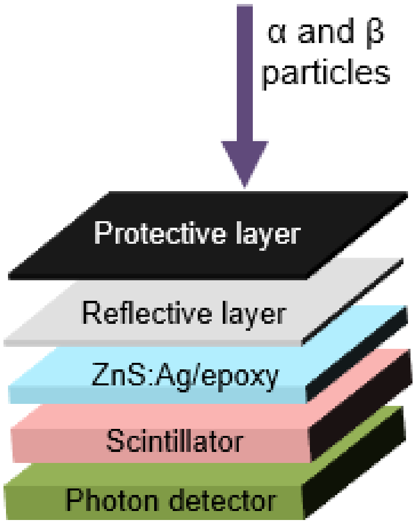

Figure 1.4: An exploded diagram of Corvid's proprietary scintillator showing the multi-layered construction. The protective layer's main objective is to stop ambient light from penetrating inside the detector, which would saturate the single photon counters. ZnS:Ag is a common α particle scintillator, and in this case EJ-212 is used as the β scintillator. The reflective layer is included to try and direct photons towards the detector, since the fluoresced light can be emitted in any direction.

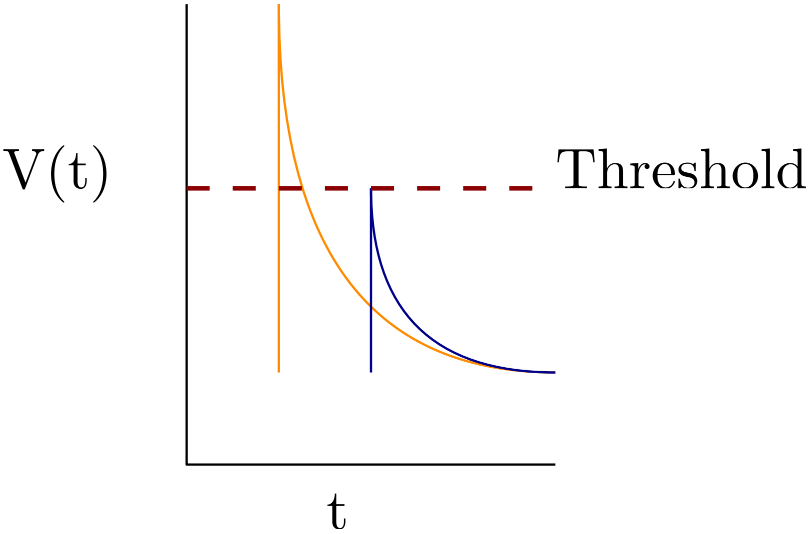

Figure 1.5: Demonstration of how PHA fails to classify radiation when scintillators have lower light output. The orange pulse represents a pulse similar to what is produced in a Mylar coupled scintillator, and the blue is the signal produced in the scintillator examined here for radiation of the same energy. Because the produced amplitude is much lower for all pulses, the classification accuracy is severely impacted.

1.3 Review of Pulse Shape Discrimination Techniques

PSD is a multi-disciplinary problem that has been approached by both engineers and physicists over a time period going as far back as the 1950s, with the study of pulse shape dependency on scintillating material going back even further [14]. The earliest investigations into PSD focused on the decay time, or decay constant that was observed on the exponential light pulse that a scintillator produced [15, 16, 17, 18]. The intention was to develop experimental setups that would allow for the consistent rejection of particles that were not the subject of a particular experiment, but could contaminate the readings if not ignored. By testing a wide variety of materials for different decay times, the earliest research sought to determine the right set of materials to use depending on which particles a researcher intended to reject. Limitations in the early nuclear electronics required that the rejection criteria was something easily handled in analog, with decay time and pulse height being the most feasible, and statistical methods existed to compensate for the effect these limited electronics had on the signal. In laboratory environments, with well designed experiments, these methods performed well at reasonably high energies, while low-energy event discrimination remained elusive [9, 19]. In the 1960s and beyond, PSD techniques began to evolve in complexity, with algorithms for pulse processing being derived that could be transformed easily into analog electronics.

Over this period, rejection techniques that utilized decay time remained popular [20], while new methods that were easier to reliably implement, and were more accurate emerged. The rise-time discrimination method employed by [6, 20, 21] centers on the time it takes for the integrated charge pulse to reach some determined fraction of its total magnitude. It has also been known to be called the zero crossing method, based on a popular implementation of the technique whereby an RC circuit is used to integrate the charge, and a CR circuit is used to differentiate the RC circuit. If the time constants are chosen judiciously, the time at which the derivative of the

RC circuit crosses zero matches well with the time at which the theoretical perfect integral would reach 90% or greater of its total amplitude [22]. This concept is shown in Figure 1.6, which compares the signal from the multiplier with a theoretical integral, the RC integral, and the CR differentiation, showing that the zero crossing occurs at approximately 90% of the rise. Such simple core circuitry contributed to its popularity in implementation, although because it is an approximation, it suffers particularly for the case of low energy events.

Proportional integral methods, of which Charge Tail Integration (CTI) is the most popular were employed in [23, 24, 25]. Proportional integral methods utilize the underlying physics involved in the delayed fluorescence and scintillator chemistry to greater degree by comparing the integrated charge during the "fast" component with the integrated charge over the full waveform, or over some different section of the waveform. As shown in [22] proportional integral methods had much higher performance compared to rise-time techniques, but were much harder to implement due to limitations in commercially available circuits at that time. Despite their relative simplicity, these PSD methods remained popular throughout the remainder of the century, and are still popular in low-cost electronics in the field of radiation detection and measurement. These methods also mark a departure in the literature from the original use case of PSD. Rather than being about rejection, these new methods were arising with the intention of accurately classifying the radiation, for use in building statistical models and achieving other research goals, although rejection remains an important aspect of some experiments [3]. Proportional integral methods in particular have become very popular for use in large fundamental science collaborations as readout electronics have improved. However, recently the availability of digitization methods has begun to introduce many new digital PSD approaches.

A modern, yet very simple PSD technique that utilizes a combination of analog and digital components is ToT PSD [8]. The theory behind this method relates back to the original work done on PSD, where the focus was on decay times. Since β-particle signals decay at a much faster rate than α particles, the time a signal spends above a given threshold should vary between the two particles. This measurement approximates the peak width of the signal, and lends itself to simple implementation with analog comparators and timers. This method also introduces a trade off, where the more thresholds are used, and the more widths obtained, the clearer a picture of the pulse shape is painted. However, more channels for timing involves more hardware, and more computational complexity. A useful digital PSD approach is that of pulse gradient analysis [7]. This technique was designed with waveform digitization in mind, and seeks to streamline picking out the "fast" and "slow" components of the waveform in a rapid and computationally efficient way. Each waveform is normalized such that it has a maximum magnitude of unity, and then a point is selected, generally within 20 ns, and its value recorded. These values themselves are the discrimination number, requiring no further computation. This method has been extended and verified experimentally alongside other PSD methods, and is very easy to implement in cases where the waveform can be digitized [26, 27]. As PSD applications have continued to shift in the direction of classification instead of rejection, the field of PSD has shifted towards utilizing the results of research into classifiers and artificial neural networks (ANN) as it has become easier to digitize pulse data [28, 29, 30]. Some researchers have even stretched the implementation of modern classification schemes even further to classify traditionally difficult data, such as when two events happen simultaneously, called pile up [31, 32].

1.4 Contribution

This thesis makes two main contributions to the state of the art for scintillation detection and pulse shape discrimination. First, a comprehensive dataset is provided for Corvid's new ruggedized scintillator which covers α, β, and γ particles from radioactive sources ^{241}Am, ^{239}Pu, ^{137}Cs, ^{60}Co, ^{210}Po, ^{90}Sr, and ^{204}Tl. While these signals

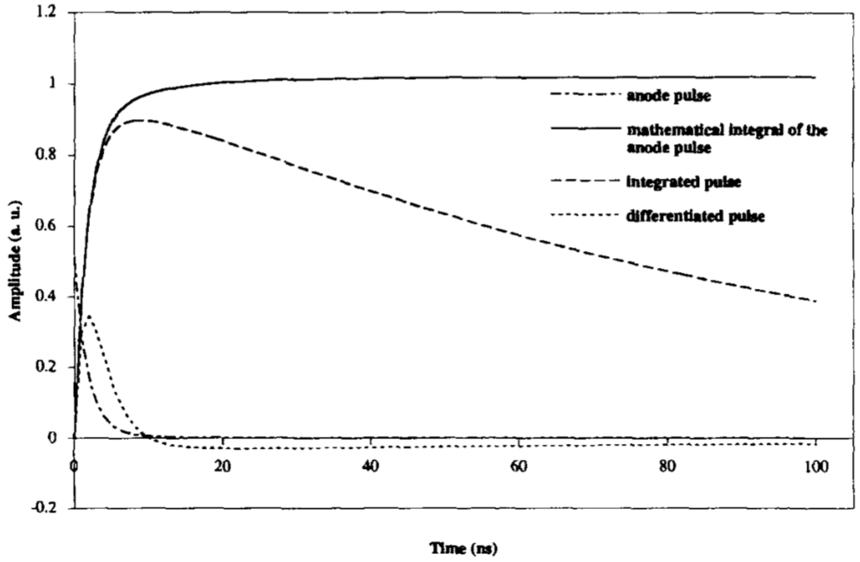

Figure 1.6: An example of "zero crossing" as an implementation of rise-time discrimination, reproduced from [22]. For properly chosen component values, the circuit differentiation of the integrator circuit will cross zero at approximately 90% of the signals peak.

were generated with Corvid's proprietary scintillator, the scintillating materials used inside the scintillator are commonplace, and the signal shapes themselves are representative of most α and β waveforms. The conditions under which this data was collected, and the procedure for replicating it are covered in detail as well to allow other experimenters to replicate this work.

Second, three different existing PSD techniques are evaluated, leveraging theoretical techniques from statistical signal processing to determine the optimal parameters to be used when implementing the PSD algorithms. This analysis is taken further as each PSD technique is used alongside the measured peak height, and classification techniques are employed to extend the accuracy of each analyzed method. These methods are then stretched to their limits by down sampling the waveform data to emulate lower sampling frequencies and lower bandwidths. The results of the effort are summarized in tables and plots, and both quantitative and qualitative analysis are provided to aid engineers and researchers in determining the ideal PSD method for their specific use case of the Corvid scintillator.

The overall structure of the thesis begins with the theoretical basis for, and construction of the experiment executed to produce the data in Chapter 2. The methodology developed to process the data, optimize the parameters of the classification techniques, and implement the SVM and KNN classifiers is covered in Chapter 3. The results of the developed classifiers are tabulated and analyzed, and potential sources of error that could not be controlled for are discussed in Chapter 4. Finally, the efforts undertaken in pursuit of this thesis goal, and its results are summarized, and potential future research in this subject is discussed in Chapter 5.

CHAPTER 2: DATA ACQUISITION

The overarching objective of acquiring this dataset was to provide the cleanest, most representative sample feasible of the light output produced by α and β pulses originating from this novel scintillator. A clean dataset facilitates the analysis completed here by enabling best case classification behavior to be evaluated. This analysis can be followed by investigation into the effects of degradation and corruption on the data, such as down sampling. Down sampling emulates the behavior of lower sampling rate ADC's by considering a smaller subset of the data. Although there are only a handful of particle radiation types, there are many different radioactive isotopes that emit them. Many of these isotopes have different emission energies, which can translate to different pulse shapes and amplitudes. To represent as wide a spectrum of the potential signals as possible, multiple different radiation sources were used across α, β, and γ species. The distance that the sources were from the scintillator was also varied to artificially create an energy gradient for the high energy α particles, which attenuate quickly through air. The source distancing scheme is shown in Figure 2.1, each spacer used should attenuate the energy of the α particle by approximately 10% . Figure 2.2 shows the empirical mean functions, $\mu(t)$ for a selection of different sources and distances illustrating the span of the collected data. It is also evident from Figure 2.2 that PHA is not suitable for processing this data, as the lowest energy α particles will be misinterpreted. The overall size and organization for the data included in this set are outlined in Table 2.1. The full dataset is available for download here [33]. Although γ events were not analyzed in this thesis, they are available in this dataset for future investigation into extensions on PSD algorithms that can be used with this scintillator.

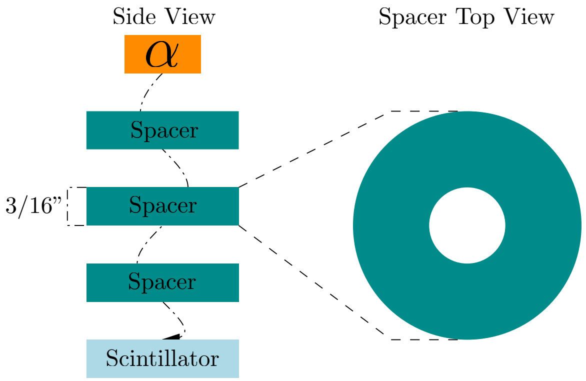

Figure 2.1: Setup for iterating α source distance. The spacers are made out of 3/16" acrylic, and have the center cut out so that the α particles attenuate through the air, rather than through the acrylic material itself. Each spacer corresponds to about 10% attenuation in particle energy, as received at the scintillator.

In low photon count applications main consideration for obtaining clean, discernible data is the speed of the resolving electronics. Referencing back to the dynamics of scintillation, much of what differentiates α and β components is contained in high frequency signal components. Much of the important data in the signal is contained in the “fast” components. Having an amplifier and digitizer with sufficient bandwidth and sampling rate to adequately resolve the “fast” component is key [34]. The data sheet for the AFBR-S4N66P014M SiPM used in this work to obtain the data establishes the rise times of the signal to be nanoseconds for single photon applications [35]. In Section 2.1, frequency domain analysis is conducted on a simulated single photon pulse. This analysis informs the equipment selected for use in the experiment, which is described in Section 2.2 along with locations at which data was taken, and the procedure followed.

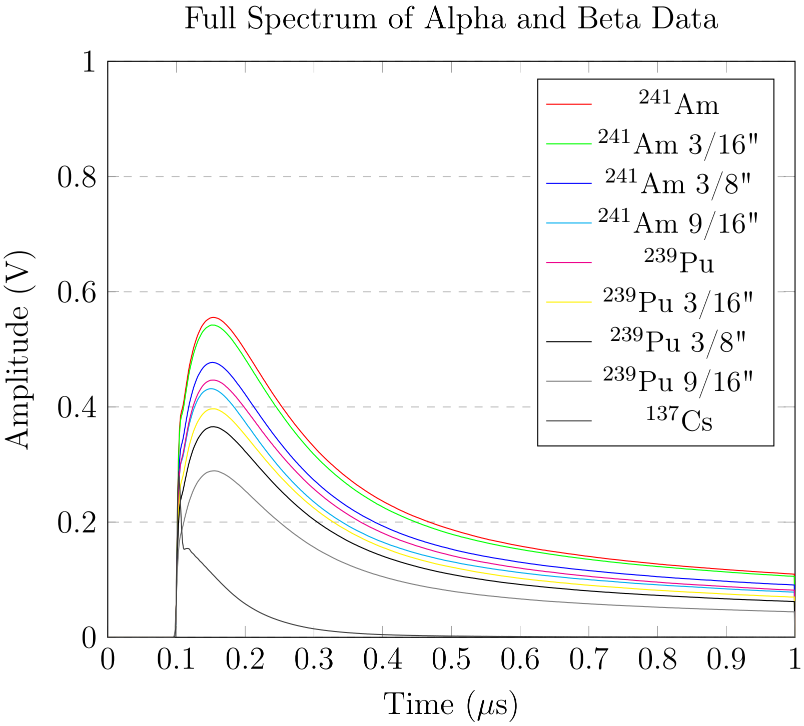

Figure 2.2: The empirical mean function $\mu(t)$ for the subsets of selected sources and distances. The potential span of shapes represented by the data obtained in this work is evident in the spread of the α mean functions. The waveforms are inverted here for presentation purposes. This figure also displays clearly why even in the ideal case, PHA is not a suitable analysis for this scintillator, as the lowest energy α particles will be misinterpreted.

Table 2.1: Number of pulses obtained for all sources and distances. Numbers are exact of this writing. The data is available online at [33].

Sources	Spacings			
α	0"	3/16"	3/8"	9/16"
^{241}Am	5,047	5,001	5,023	4,958
^{239}Pu	5,012	5,015	5,126	5,011
^{210}Po	5,011	5,010	5,054	5,505
Source	Spacings			
β	0"	3/16"	3/8"	9/16"
^{137}Cs	5,025	5,033	5,024	5,044
^{90}Sr	5,210	5,539	5,025	5,121
^{204}Tl	5,026	5,088	6,115	5,279
Source	Spacings			
γ	0"	3/16"	3/8"	9/16"
^{60}Co	5,056	5,021	5,056	5,021

2.1 Frequency Domain Analysis

As described previously, the single photon response of a SiPM has a rise-time measured in nanoseconds. To ensure that the signals are resolved accurately a simulated SiPM signal was programmed in MatLab using data from [35] which specifies an exponential pulse with rise-time from 0 V to 110 mV of 1 nanosecond. This simulates the SiPM's response to a handful of photons, and should provide an upper bound on the frequency content that we will require. This simulated signal can be Fourier transformed via a fast Fourier transform (FFT) algorithm to extract the frequency content of the signal. The simulated SiPM signal, and resulting FFT are plotted in Figure 2.3. The FFT shows that between the range of 100 MHz, and 1 GHz the frequency content of the signal rapidly falls off. This result is important in motivating both the design of the experiment, but also the values chosen for the down sampling later in this thesis. In order to get an ideal capture of the SiPM signal, the digitizer and amplifier should have bandwidth in excess of 500 MHz. Conservatively, an amplifier with 1 GHz bandwidth was selected to amplify the signal, which can allow for the obtained signals to be faster than simulated without distorting the collected data.

The requirements on the sampling rate of the digitizer used to convert the analog data to digital are driven by the Nyquist limit

$$f_{lim} = \frac{f_s}{2}, \tag{2.1}$$

with f_{lim} being the highest frequency that can be resolved without aliasing and f_s being the sampling frequency. From this, we can see that if we wish to be able to digitize the full range of signals that our amplifier can handle, a digitizer with at a minimum 2 GHz sampling rate is required. A digitizer chosen according to this specification will also far exceed the required sampling rate to extract the most relevant frequency content from the signal.

Broadcom Inc., who produce the SiPM used for this dataset also produce an evaluation board, the AFBR-S4E001, which incorporates an AD8099 amplifier. The AD8099 amplifier has been used before in multiple experiments involving SiPM's thanks to its gain bandwidth product of 3.9 GHz, and low noise performance [36, 37, 38]. With a gain bandwidth product of 3.9 GHz, the AD8099 amplifier on the Broadcom evaluation board is well suited for our experiment. For the high frequency digitizer, Corvid Technologies lent access to a Teledyne Lecroy Waverunner 8108HD, with 1 GHz analog inputs, 10 GHz sampling rate, and 12 bits of resolution. By the analysis completed here, this equipment should far exceed the minimum requirements for accurately digitizing these signals without impacting their shape.

2.2 Experimental Setup

The overall implementation of the experiment is shown in Figure 2.4, illustrating its relative simplicity. Light collection was conducted using a 6 mm by 6 mm Broadcom SiPM chip, connected to Broadcom's AFBR-S4E001 evaluation board [39]. To make the experiment easier to execute and work with, a "light-tight" enclosure was designed. The enclosure affixes a Corvid scintillator to the front of the SiPM, and

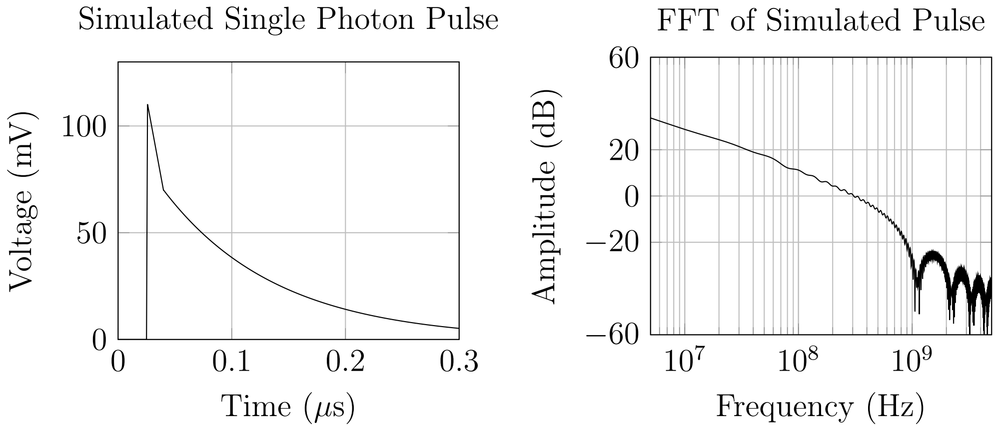

Figure 2.3: A simulated single photon capture in a SiPM, and its corresponding FFT. The left plot shows the simulated pulse that was produced to mimic that laid out in [35]. The FFT, shown right, has been cropped to specifically show the high frequency content, since the low frequency magnitude washes out the high frequency component. The FFT plot shows clearly that the high frequency content extends into the high MHz.

seals the surrounding area against ambient light. This allows the device to be used on the bench top, instead of inside of a special box. During initial checkout testing, where the scintillator was irradiated with high energy α particles, the output of the amplifier would saturate. This resulted in signal clipping that would cut off the top of the α particle peaks. While the β pulses were mostly unaffected, the clipping on the α particles created signal degradation that would have seriously impacted and complicated signal processing. To address this, small modifications were made to the amplifier gain on the evaluation board. The changes made to the gain also required the population of compensation circuits to prevent signal oscillation [38].

The AFBR-S4E001 amplifier board requires a bipolar power supply between +/- 5 and +/- 2.5 V, and the SiPM's require a reverse bias voltage between 38 and 50 V. The supply to the amplifier determines the dynamic range it can achieve, so to allow for the strongest pulses to be resolved without clipping, +/- 5 V was elected for the amplifier supply. The SiPM gain, noise, and correlated noise are essentially linear functions of reverse bias [35], and a reverse bias of 45 V represents a good trade off in terms of SNR. Two connections are made from the evaluation board to the oscilloscope: a pole-zero canceled timing channel, and a buffered output stage. The pole-zero canceled circuit filters the waveform, and as such is ill suited for PSA, but it has a very sharp rising edge, regardless of input magnitude. This makes this output a suitable candidate for triggering a scope acquisition. The buffered output channel produces the unfiltered waveform as it is produced by the SiPM without filtering, this output is saved and used for PSA in this dataset.

The “light tight” enclosure was designed in Solidworks® and 3D printed using poly lactic acid (PLA) filament. It is worth noting that even at high infills, 3D printed materials will not be single photon tight, due to the way 3D printers assemble the print. To account for this, rubber gaskets are utilized and compressed by the print to create a “torturous path” through which light is more likely to be absorbed than to reach the SiPM. A small acrylic square is placed between the scintillator and SiPM so that the scintillator lies flat atop the SiPM despite being larger in area. The spacers used here to stand off the radiation sources, shown in Figure 2.1, were laser cut out of 3/16" clear acrylic, and were made to have a hole through the middle so that the α particles moderate through air instead of acrylic.

To maintain consistency across the dataset, each source of radiation was tested according to the following procedure. First, with no source applied, live data was examined from the SiPM to check for suitable noise level. High noise levels generally indicated that a seal was not properly compressed around the SiPM. Once noise levels were confirmed, the chosen source was applied, and the waveforms were examined in real time. At this point, a trigger level was selected for the timing channel. Throughout all acquisitions in this dataset, the timing channel trigger level was able to be kept the same. Once the trigger was set, the oscilloscope began saving data, and the ambient temperature was recorded. Temperatures were recorded since SiPM’s have a gain function that is strongly correlated with operating temperature, and as such any notable change in temperature could impact the data. Once the oscilloscope had saved approximately 5,000 waveforms, acquisition was stopped, and a spacer was added before the process was repeated. Each source examined in this dataset was used with up to three spacers, corresponding to 3/16", 3/8", and 9/16" respectively.

A full image of the test setup as executed at TUNL is shown in Figure 2.5 within which the scope, power supplies, scintillator and enclosure can all be seen alongside the radioactive source. TUNL was selected as a testing site as they have access to radiation sources, such as ^{239}Pu and ^{241}Am that are not commercially available due to federal laws. All of the equipment used for the collection of the data is recorded in Table 2.2. Data was also collected at Corvid Technologies, using commercially available radiation sources ^{210}Po, ^{137}Cs, ^{90}Sr, ^{204}Tl, and ^{60}Co.

Table 2.2: List of equipment used to conduct the data acquisition for the dataset. Power supplies and oscilloscope were loaned by Corvid Technologies for the data collection efforts. The amplifier electronics and SiPM are available online from electronics distributors, and at Broadcom's website.

SiPM	AFBR-S4N66P014M
Amplifier Electronics	AFBR-S4E001
Bipolar Power Supply	GPE-2323
High Voltage Bias Power Supply	SPS-606
Oscilloscope	Teledyne Lecroy Waverunner 8108HD

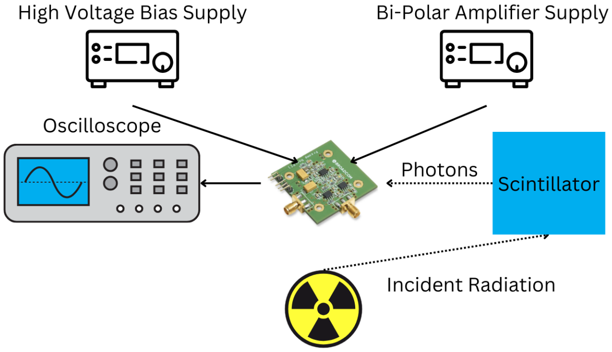

Figure 2.4: Block diagram illustrating the most important components involved in the experimental setup. Radiation is captured in the scintillator, where it is converted into a cascade of photons. The photons are captured by the SiPM attached to the evaluation board, and converted into an electrical pulse. The pulse is the digitized at 10 GHz and saved to a hard drive for offline analysis. The bipolar power supply provides the +/- 5 V power rails that the evaluation board requires. The high voltage bias supply provides 45 V to reverse bias the SiPM's.

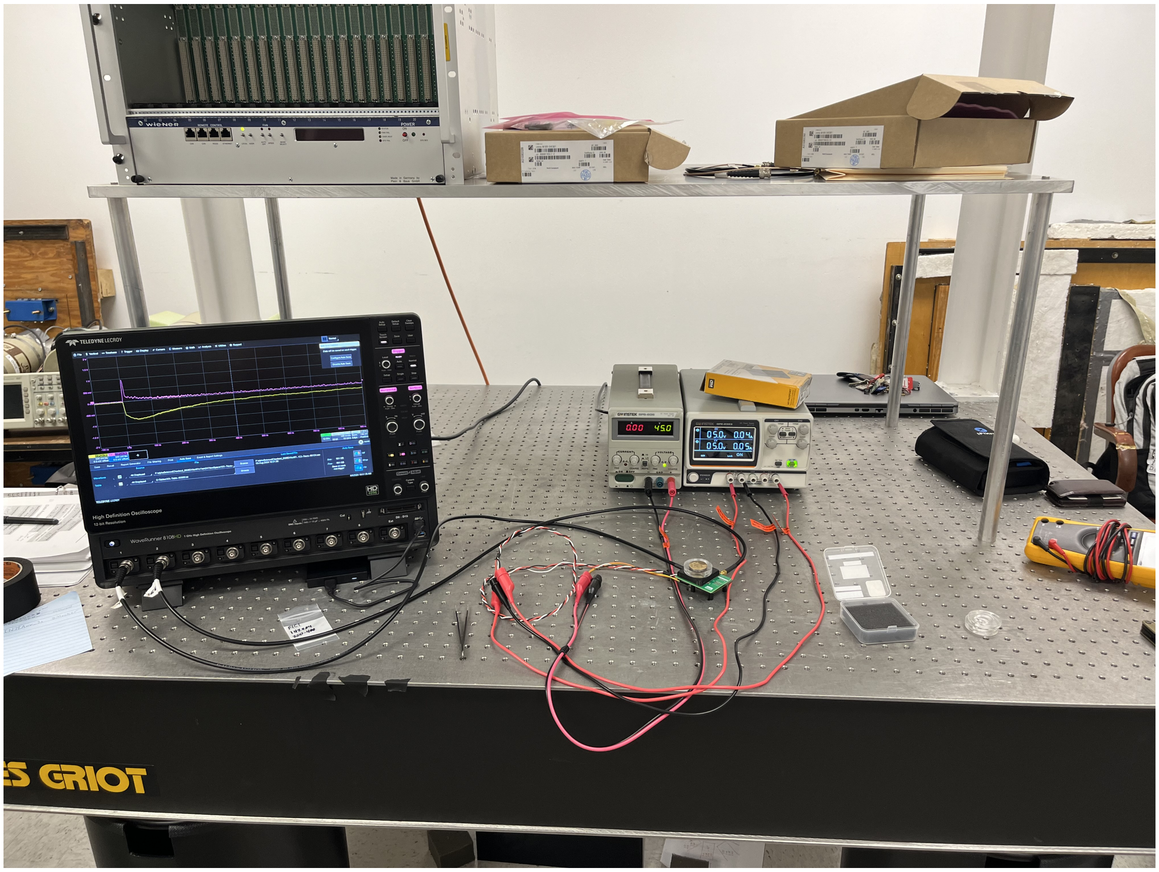

Figure 2.5: The test setup used at TUNL to acquire the pulse data. In this picture, both power supplies can be seen providing the requisite biases to the electronics, and the scope is reading out pulses. The purple channel is the output of the pole-zero circuit, and is used for timing. The yellow trace is the live pulse data that is being recorded for the shape analysis. The ^{241}Am source has two spacers between it and the detector, which can be seen in the middle of the table.

CHAPTER 3: CLASSIFIER DESIGN METHODOLOGY

This chapter elaborates on the methods and theory applied to achieve the goal of producing a robust and feasible PSD method for the unique signals produced by Corvid's scintillator. Before the data can be used for statistical and shape analysis, a few error control steps must be taken, as covered in Section 3.1. The PSD techniques that are used in this thesis are discussed in more detail from a technical and theoretical point of view in Section 3.2. The representative size of the dataset allows for the synthesis of the PSD techniques that we have discussed earlier in Section 1.3 with methods of statistical signal processing, covered in Section 3.3. Analytic techniques are used to obtain the optimal choice of both parameters used by the discrimination techniques, and the optimal decision making parameter in the case where a single method is used.

3.1 Pre-Processing

To avoid over fitting, before any processes are run, the data was sorted into training and testing datasets. From each entry in Table 2.1, approximately 2/3 was randomly assigned into the training set for the corresponding particle type and the remaining 1/3 sorted into the testing dataset. The resulting allocations are outlined in Table 3.1. This allows for the classifier to be developed off of one set of data, and have its accuracy evaluated on an independent dataset to ensure that the classification process is not overly biased to the specific dataset. With the dataset sorted other issues with the data such as pile-up, covered in Section 3.1.1, and offset, covered in Section 3.1.2, can be handled.

Table 3.1: Composition of data after being randomly sorted and distributed by radiation species into Training and Testing groups. The raw data that was collected according to source type and spacing was randomly shuffled and distributed into Training and Testing datasets to avoid over fitting, and to get a realistic idea of potential classifier performance. These numbers are exact as of this writing.

Source	α
Training	Testing
38,888	21,883
Source	β
Training	Testing
40,029	22,525
Source	γ
Training	Testing
8,688	5,852

3.1.1 Pile Up

“Pile-up” is the term given when two pulses occur very near each other in time, such that when the PSD algorithm filters the pulse, unexpected results are garnered. An example of such an event is provided in Figure 3.1. Pile up can be especially detrimental to algorithms like PGA, where the max value may occur far later in the waveform due to a second large pulse occurring, or if the second pulse happens in the region where the second sample is to be taken. Interestingly, some recent works have taken the approach of using ANN’s and advanced machine learning techniques to dynamically classify and de-convolve these events [32].

In the scope of this thesis, these pulses were identified and set aside before mixing all of the data from all of the sources. A script was written which looks for the occurrence of multiple “peaks” in a single waveform. In this context, a peak was identified when the rolling average over a window of 50 nS was significantly larger than the previous 50 nS. This window length needed to be tuned slightly for β particles due to their fast rise and decay times. Once identified, these pulses were copied over into a separate dataset for potential future use.

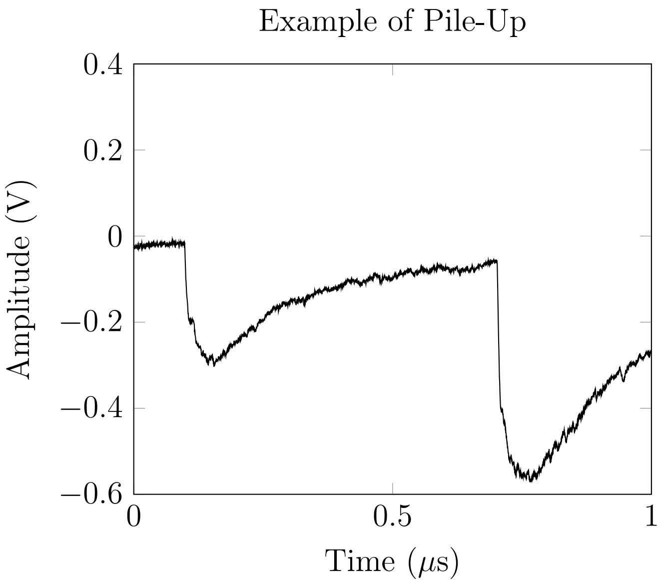

Figure 3.1: An example of a pile-up pulse that was removed from the dataset. While rare, these activities are more likely the higher the activity of the radiation source.

3.1.2 Offset Trimming

Due to variations in seal quality, SiPM gain, and other factors, the output from the SiPM, and thus the amplifier electronics was not always explicitly zero. An offset was present in nearly all of the data collected. Although the techniques applied here do not explicitly require handling off the offset, performance is improved through its mitigation, and both digital and analog methods of trimming it exist for practical implementation. To remove the offset, the average value of the first 50 nS of data was computed, and then subtracted from the waveform as a whole.

3.2 Discrimination Techniques

To review, the key issue in discriminating radiation in a scintillation detector is to determine the ratio of delayed fluorescence to immediate fluorescence. This ratio is represented as some feature that correlates with the underlying physics. In this chapter, the value for the features produced by these methods will be referred to as D. The rise-time technique accomplishes extraction of the fluorescence/delayed fluorescence ratio by measuring the elapsed time that a signal takes to go from 10%

of its maximum value to 90% of its maximum value. Although both the α and β waveforms have sharp rising edges, because the α causes delayed fluorescence, its sharp edge is followed by a slow climb to its peak value, compared to β particles which rise sharply all the way to their peak. As such, α particles should have a slower rise-time than β particles. In practice, this method is implemented using analog electronics, hence its popular nickname as the “zero crossing” method, in reference to its most popular implementation configuration. Given that the data obtained here is all digital, a digital version of rise-time is implemented here that maintains the core thesis of the idea. Additionally, due to the low peak voltage of the lowest energy β particles, and the noise on the signal, 20% to 80% rise-time was used here instead of the general 10% to 90%. A visualization of the rise-time technique is shown in Figure 3.2, showing a zoomed in rise of an α pulse to demonstrate the duration of time that is recorded.

PGA determines the delayed fluorescence by examining the other half of the signal; the decaying exponential. PGA begins by normalizing the waveform to its maximum value, such that the peak now has a value of unity, and all other points are a ratio of the original value to the peak value. From here, a point along the exponential decay of the waveform is selected and its value saved as D. This process is visualized in Figure 3.3, from which it can be seen that the exponential decay rate of the two different waveforms is proportional to the extracted value. This permits classification of particles to be done via this computationally fast approximation of the exponential decay rate.

Finally, ToT’s computation is similar to that of rise-time, however, it is extracting a different feature. It involves recording the amount of time that a signal spends above a threshold. This method is usually employed using analog circuity, as this task is easily done with modern high speed comparators. Generally, because of the long α decay rate, it performs well, except at low energies, due to the timings for both

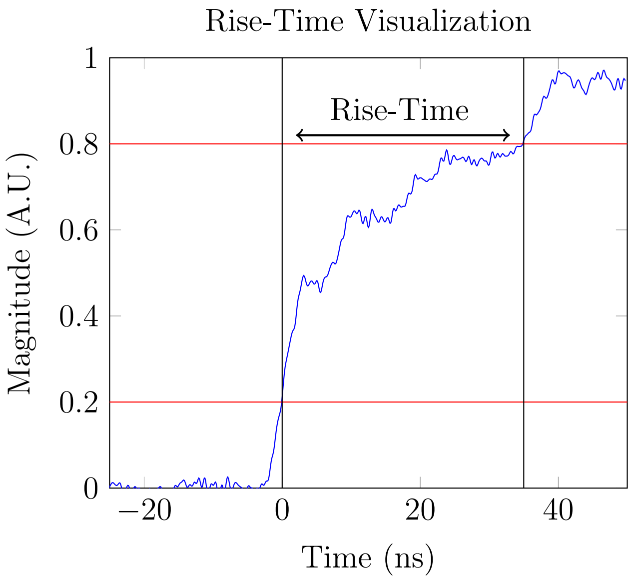

Figure 3.2: A visualization of the feature extracted by the rise-time technique. The α particle signal here has been normalized to its maximum value. Timing begins when the signal surpasses 20% of max, and end when it reaches 80% of max. For this particular α particle, we can read off that the rise-time is about 35 ns. This value is generally smaller for β particles, which have a much faster rise.

α's and β's being comparatively short. A visualization of how ToT is measured is shown in Figure 3.4 displaying the time that a β particle spends above the threshold. Since this is generally conducted in analog circuits, there are some considerations and adjustments we have to make for the digital case. Firstly, if the threshold is never reached, then a pulse will have been "missed", and these will have to be counted separately from the normal correctly and incorrectly classified counts. However, we want to miss as few as possible, so this will have to be accounted for when considering what threshold to use for the timing. This is discussed more in Section 3.3.2. Finally, because the data taken for this work has finite length, we also have to handle when a signal does not return under the threshold before the end of the available data. In these cases, it is easiest to substitute in the length of each waveform. This edge case handling does, however, affect the distributions of the D parameter which can complicate the theoretical treatment and implementation.

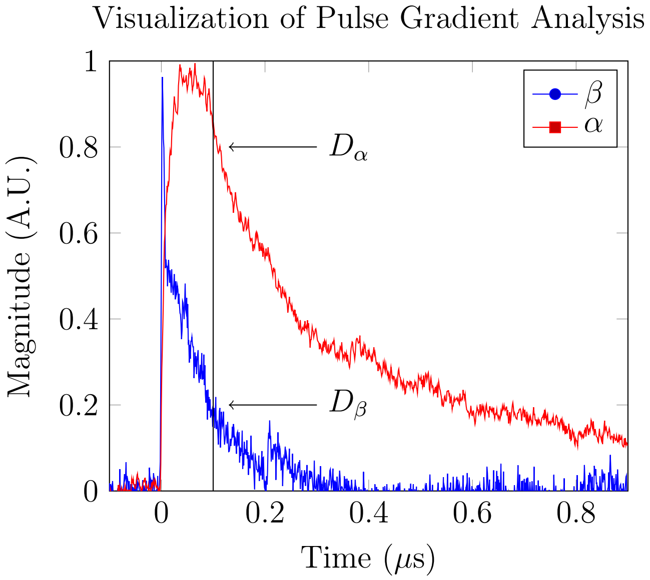

Figure 3.3: A visualization of PGA on both an α and a β particle simultaneously. Both the α and β waveforms have been normalized to their maximum value. Comparing the α and β pulses here, it is easy to identify how this value can be used to discriminate between the pulses.

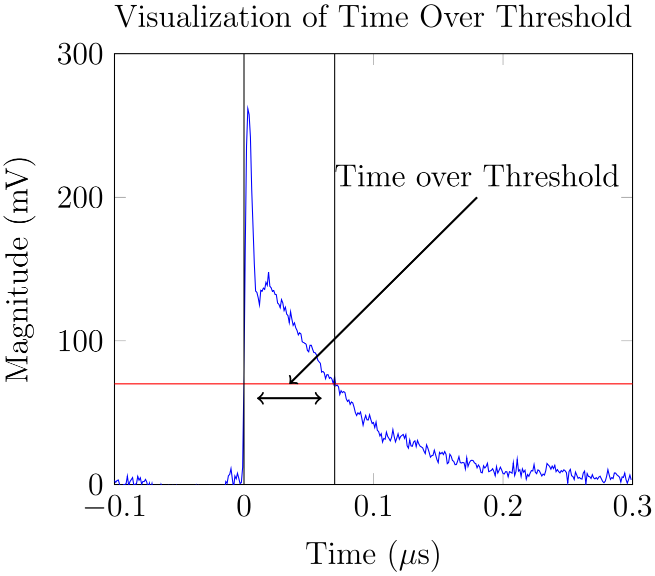

Figure 3.4: A visualization of time over threshold on a β pulse. Pulses are not normalized for this technique, as it is designed to be implemented in analog systems easily.

3.3 Statistical Signal Processing

The PSD techniques covered in Section 3.2 are often used to distill a large amount of data coming in from the waveform, into a much smaller form, D. This allows us to treat the problem of optimally discriminating between α and β particles as a binary detection problem for random processes and use techniques from binary detection and random process theory that obtain optimal results. Firstly, maximum a posteriori (MAP) rule, which is a technique from statistical signal processing and decision theory establishes a method for finding the optimal classification rule given a single feature of the data. The theory and application of MAP rule is covered in Section 3.3.1. Furthermore, both ToT and PGA have a hyper parameter. For ToT, it is the threshold above which to time the signal, and for PGA it is the point at which to sample the pulse gradient. Through an examination of the physics of the signal, and an examination of its mean and variance, which is covered in Section 3.3.2, we can optimize the hyper parameters to obtain maximum separation in the resulting distributions of D.

3.3.1 Map Rule

In order to properly apply the statistical techniques discussed here, we must first rigidly formalize problem. We define the waveforms produced by both α and β particles as random processes

$$\{\alpha_n, 0 \leq n \leq N\}, \tag{3.1a}$$

$$\{\beta_n, 0 \leq n \leq N\}, \tag{3.1b}$$

for α_n and β_n the value of an α and a β pulse respectively at sample n, and N the total number of samples taken. Given these definitions, we can now define any of our

PSD techniques as a functional of these processes

$$h(X) = D_X \tag{3.2}$$

where X is the vector of the realizations of the process from 0 to N coming from either an α or a β particle, and D_X is the scalar discrimination value resulting from the input process. By this definition, D can be modeled as a random variable with a probability distribution $f_D(d)$, where d is defined as a particular realization of D. By the definition in 3.2 it will have marginal distributions

$$f_{D|\alpha}, \tag{3.3a}$$

$$f_{D|\beta}, \tag{3.3b}$$

corresponding to the distributions of D produced from the an input of either process 3.1a or 3.1b. These marginals allow us to find the optimal cut off point through use of MAP rule. Formally, the MAP estimator is defined as

$$\hat{\theta}_{MAP}(d) \triangleq \underset{\hat{\theta}(d)\in\{\alpha,\beta\}}{\arg\max}\, P(\Theta = \hat{\theta}(d)|D = d). \tag{3.4}$$

In equation 3.4 $\hat{\theta}$ is the estimate of the originating radiation species, d is a realization of the discrimination parameter D, and Θ is the originating radiation species. From that definition, it follows that the MAP estimator is the estimator which maximizes the probability that the estimate is correct for each possible realization of D. Now, through application of Bayes rule to the right hand side of 3.4, which we are maximizing, we can obtain the following equations for our binary detection problem

$$P(\Theta = \alpha | D = d) = \frac{f_{D|\Theta}(d|\Theta = \alpha)}{f_D(d)} P_\alpha, \tag{3.5a}$$

$$P(\Theta = \beta | D = d) = \frac{f_{D|\Theta}(d|\Theta = \beta)}{f_D(d)} P_\beta. \tag{3.5b}$$

P_α and P_β are the probability of occurrence for α and β particles respectively. By setting 3.5a and 3.5b equal to one another, and assuming that we have equal likelihood of seeing either an α or a β particle, such that $P_\alpha = P_\beta$ we can find the value of d at which the probability is equal. This d value will correspond to a "tipping point" above which we should classify the measurement as one type, and below which we should classify it as the other. Applying this rule can be shown to theoretically minimize the average expected error that the classifier will incur. As an aside, by setting $P_\alpha = P_\beta$, we have reduced the MAP rule to the maximum likelihood, or ML rule. Referencing Table 3.1, we can see that we do not have equal amounts of α and β data. In order to satisfy the condition that $P_\alpha = P_\beta$, we must make sure that we use an equal amount of the α and β training data. This is a valid assumption to make in the general case, since in the context of using a scintillation detector, we do not have prior knowledge of the expected ratio of α and β radiation, so the best guess is for the events to be equiprobable.

To apply MAP rule to our dataset, we must run the data through the particular PSD algorithm, and generate a histogram of the resulting realizations of D. From here, the histogram can be fit to a distribution corresponding to both the α and β particles, and using the resulting distributions, the point of equality and optimal cutoff can be found. An example of this process is shown in Figure 3.5, where an example distribution of D for the PGA method is shown. Given the distributions, we can analytically or computationally solve for the point of equality, which we will denote as D^*.

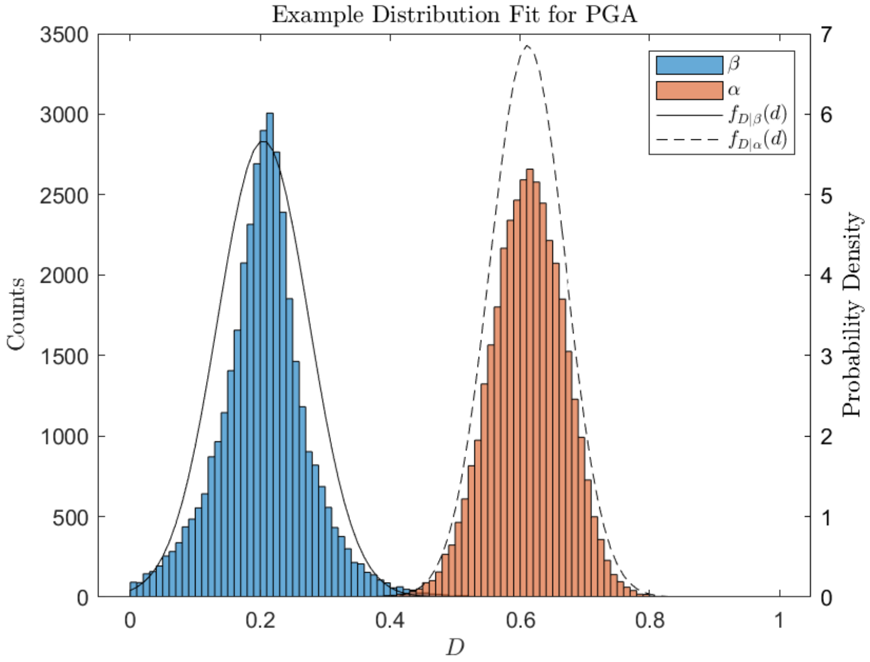

Figure 3.5: Example of D parameter histogram, with a PDF fit. The left axis corresponds to the histogram counts, and the right axis corresponds to the probability density at that point. The fits here are normal distributions, which have acceptable agreement with the data.

3.3.2 Parameter Optimization

Physically, the signals analyzed here can be described by two components. There is the exponential function behavior, which approximately follows equation 1.1 with the amplitude driven by I_0, which approximately corresponds to the particles energy, and there is random white noise. We can convert equation 1.1 into discrete time, $I_n = I_0 exp(\frac{nt_S}{\tau})$, with t_S being the sampling time. We can then define two functions B_n and A_n as

$$A_n = GI_0^{\alpha} exp(-\frac{nt_S}{\tau_{\alpha}}), \tag{3.6a}$$

$$B_n = GI_0^{\beta} exp(-\frac{nt_S}{\tau_{\beta}}). \tag{3.6b}$$

Here, G is the gain of the trans-impedance circuit in volts per amp, I_0^{α} adn I_0^{β} are the maximum current produced by a given α or β pulse respectively, τ_{α} is the decay constant for α pulses, and τ_{β} is the decay constant for β pulses. Here, A_n and B_n represent the noiseless voltage pulses produced by a scintillated event. If we then define the noise on the signal to be a random variable $Z \sim \mathcal{N}(0, N)$, for N the noise variance, α_n and β_n can be represented as linear combinations

$$\alpha_n = A_n + Z, \tag{3.7a}$$

$$\beta_n = B_n + Z. \tag{3.7b}$$

Examining equation 3.6 we observe that the only random component is the maximum voltage $V_0 = GI_0$. If this component is known, then A_n and B_n proceed deterministically. We can now treat V_0^{α} and V_0^{β} as random variables with arbitrary distribution.

This result is useful for optimizing the parameters pertaining to PGA. By normalizing the waveform to the maximum value, as is prescribed by the method, we remove the "randomness" from the equation, and our equations reduce to

$$a(n) = \frac{A_n}{V_0^\alpha} = exp(-\frac{nt_S}{\tau_\alpha}), \tag{3.8a}$$

$$b(n) = \frac{B_n}{V_0^\beta} = exp(-\frac{nt_S}{\tau_\beta}). \tag{3.8b}$$

Our random processes as defined in equation 3.7 are then reduced as well

$$\frac{\alpha_n}{V_0} = a(n) + \frac{Z}{V_0}, \tag{3.9a}$$

$$\frac{\beta_n}{V_0} = b(n) + \frac{Z}{V_0}. \tag{3.9b}$$

Which implies that the random processes of α and β, at any time n after normalization are normally distributed with means $a(n)$ and $b(n)$, and variances $N/{V_0^\alpha}^2$, and $N/{V_0^\beta}^2$ respectively. While it is not necessary, we make the assumption here that after normalization, $N/{V_0^\alpha}^2 = N/{V_0^\beta}^2 = 1$ to make the following math easier to follow. To find the ideal point to sample the gradient in PGA, we want to find the point n which produces the lowest probability of error. Using ML rule as derived above, our probability of error can be described as

$$P_e^* = Q(|D^* - a(n)|) + Q(|D^* - b(n)|) \tag{3.10}$$

where P_e^* is the minimum probability of error, $Q(x)$ is the Q function representing the probability of a normally distributed random variable being greater than x, and D^* being the point where the probability of being from either distribution is equal. Equation 3.10 can be understood qualitatively as the probability that an α particle is classified as a β particle, and vice versa. We begin by deriving an expression for

D^* using ML rule, and equation 3.9

$$\frac{1}{\sqrt{2\pi}} exp(-\frac{1}{2}(D^* - a(n))^2) = \frac{1}{\sqrt{2\pi}} exp(-\frac{1}{2}(D^* - b(n))^2), \tag{3.11}$$

dividing out the common factor and taking the natural log

$$-\frac{1}{2}(D^* - a(n))^2 = -\frac{1}{2}(D^* - b(n))^2, \tag{3.12}$$

then dividing both sides by $-\frac{1}{2}$ and taking the square root

$$D^* - a(n) = \overset{+}{-}(D^* - b(n)). \tag{3.13}$$

We now have to decide whether to take both sides having the same sign, or opposite sign. Choosing both sides to have the same sign however will remove D^* from the equation, and set $a(n) = b(n)$, which would imply that we have found the point that places the distributions atop one another. This solution will clearly not minimize the P_e. Instead, we take the left and right side of 3.13 to be of opposite sign and after simplification obtain this result

$$D^* = \frac{a(n) + b(n)}{2}. \tag{3.14}$$

We can now take this result and apply it to equation 3.10. Intending to minimize the probability of error, we arrive at the following

$$P_e^* = min[Q(|(\frac{a(n) + b(n)}{2}) - a(n)|) + Q(|(\frac{a(n) + b(n)}{2}) - b(n)|)]. \tag{3.15}$$

By simplifying, and utilizing the absolute value, we can rewrite the problem as

$$P_e^* = \arg\min_n [2Q(a(n) - b(n))]. \tag{3.16}$$

By considering that Q is a decreasing function, we can arrive at our final answer

$$P_e^* = \arg\max_n [a(n) - b(n)]. \tag{3.17}$$

It is important to acknowledge the limitations of this analysis. Equation 1.1 is an approximation, and does not explicitly address the intricacies of other phenomena such as delayed fluorescence. This extends to the steps taken in equation 3.8, by scaling with the max value, there are likely to still be elements present that are dependent on the dynamics of the specific particle being measured. However, compared to the dynamics described in equation 1.1 these components should be small, and not significantly impact our result. Additionally, the variance of Z is dependent on a wide variety of factors, and is very unlikely to be equal to 1 after scaling. However, even when the variances are not equal the result in equation 3.17 should still hold, since equation 3.14 should become a convex combination of $a(n)$ and $b(n)$ instead of a straight average. Even under this convex combination, the n that maximizes $a(n) - b(n)$ will still produce the lowest P_e.

Empirically, to obtain functions for $a(n)$ and $b(n)$ we recall that Z is zero mean and take the expectation of the normalized function

$$\mathbb{E}[\frac{\alpha_n}{V_0^\alpha}] = \mathbb{E}[a(n) + \frac{Z}{V_0^\alpha}] = a(n), \tag{3.18a}$$

$$\mathbb{E}[\frac{\beta_n}{V_0^\beta}] = \mathbb{E}[b(n) + \frac{Z}{V_0^\beta}] = b(n). \tag{3.18b}$$

This can be done computationally by averaging together the value of multiple waveforms at a single n, for all samples. An example of the resulting functions $a(n)$ and $b(n)$ is shown in Figure 3.6. Using these two curves, we can iteratively find the selection of n that maximizes the difference between them, and as such should minimize the probability of error when applying the PGA technique. An example of the dif-

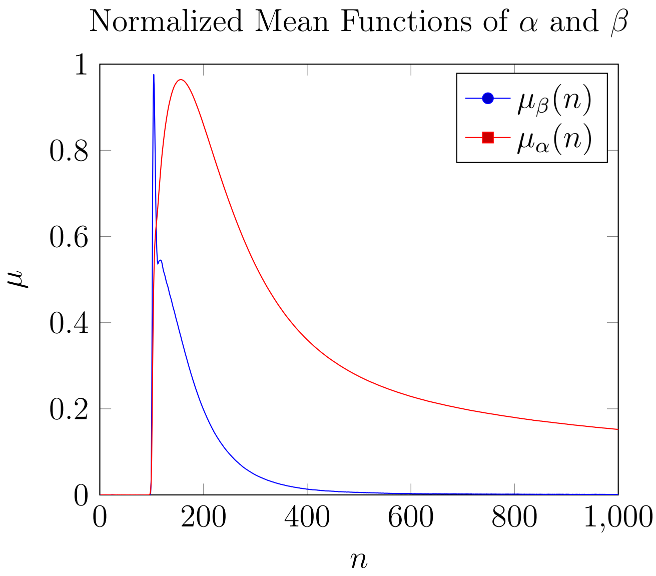

Figure 3.6: Normalized Mean Functions of α and β. If we assume that the noise is zero mean, then the mean function can be used to determine the optimal point to sample for the pulse gradient which will produce the best separation in the resulting marginals of D.

ference between $a(n)$ and $b(n)$ as a function of n is plotted in Figure 3.7, and the optimal choice of n can be determine to be approximately 80.

In the case of ToT, we can apply a similar analysis to that above, with some key differences. Unlike PGA, ToT does not work with normalized pulses, thus $A(n)$ and $B(n)$ are still random variables, driven by the energy distribution of the incident radiation. Energy distribution for α and β particles is dependent on the source of the radiation, and is impossible to be known beforehand. The best guess for the distributions of $A(n)$ and $B(n)$ then is to be uniform. Under this assumption, if we approximate the functional, $h_{ToT}(X) = D_X^{ToT}$ as a linear function of particle energy, then the choice of threshold that minimizes probability of error is the choice that maximizes the difference in D for the non-normalized mean functions. This difference can be found similarly to above and is shown as a function of n in Figure 3.8, where the best choice of threshold is approximately 70 mV.

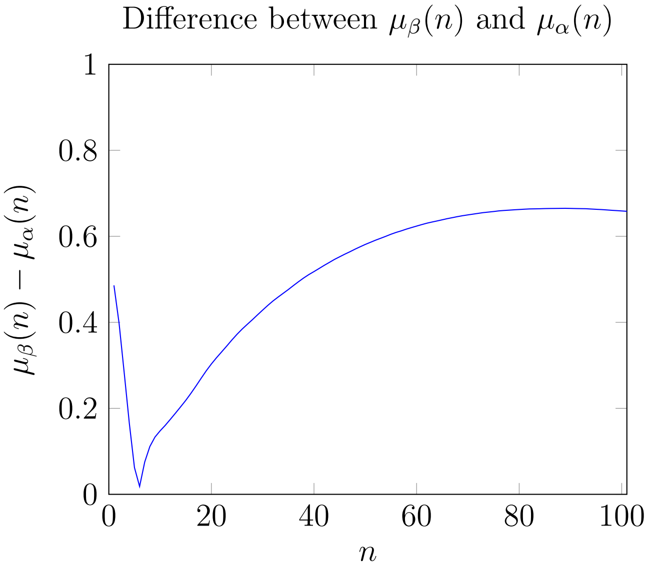

Figure 3.7: Difference between $\mu_\beta(n)$ and $\mu_\alpha(n)$. Since the mean functions should be analogous to $a(n)$ and $b(n)$, finding the maximum of this function should maximally separate the marginal distributions of D for given α and β.

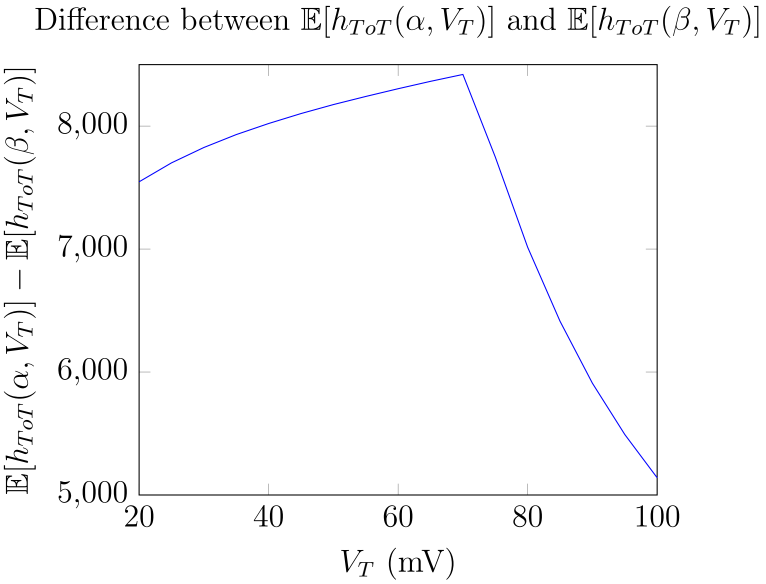

Figure 3.8: Difference between $\mathbb{E}[h_{ToT}(\alpha, V_T)]$ and $\mathbb{E}[h_{ToT}(\beta, V_T)]$. By finding the V_T that produces the most difference between D for μ_α and μ_β, we determine the optimal threshold for classifying.

3.4 Multi-Feature Classification

In addition to examining each of the above techniques for PSD, the ability of some well know classification techniques is used to classify the pulses the pulses investigated here, using both the feature extracted with one of the methods from Section 3.2, and the measured peak height. For each of the methods, an SVM and KNN classifier is trained according to the theory laid out in Sections 3.4.1 and 3.4.2 covering SVM and KNN respectively. The reader is directed to further reading for both methods.

3.4.1 Support Vector Machines

A support vector machine seeks to find what is called the "optimal hyperplane," which will maximize the margin of separation between two labeled sets of training data. While the support vector method has been extended to multi-dimensional data, in this case we are only dealing with two dimensions. The optimal hyperplane is found starting with the set of labeled training data

$$(y_1, X_1), ..., (y_l, X_l), y \in \{-1, 1\}, \tag{3.19}$$

which is said to be linearly separable if there exists a vector W and scalar b such that the inequalities

$$W \cdot X_i + b \geq 1, y_i = 1, \tag{3.20a}$$

$$W \cdot X_i + b \leq -1, y_i = -1. \tag{3.20b}$$

are met for each point in the training data set. To compare with our case, X_i is the ith vector of training data comprising peak height, and another feature, and y_i is -1 for β and 1 for α. Rewriting equation 3.20 more compactly we can obtain

$$y_i(W \cdot X_i + b) \geq 1, i = 1, ..., l. \tag{3.21}$$

The optimal hyperplane, defined as

$$W_0 \cdot X + b_0 = 0 \tag{3.22}$$

is the unique solution that separates the training data with maximal margin. In essence, the optimal hyperplane describes a direction $W/|W|$ which maximizes the distance between the the projections of the training vectors from the two different classes. This distance is derived to be given by

$$\rho(W_0, b_0) = \min_{X:y=1} \frac{X \cdot W}{|W|} - \max_{X:y=-1} \frac{X \cdot W}{|W|} \tag{3.23}$$

and utilizing 3.21 we can rewrite 3.23 as

$$\rho(W_0, b_0) = \frac{2}{|W_0|} = \frac{2}{\sqrt{W_0 \cdot W_0}}. \tag{3.24}$$

From this, the optimal hyperplane can be described as the one that minimizes 3.24 under the constraints of 3.21.

As a final note, if we define the “support vectors” as the vectors X_i that satisfy $y_i(W \cdot X_i + b) = 1$, then it can be shown [40] that a linear combination of the support vectors determines the optimal hyperplane

$$W_0 = \sum_{i=1}^{l} y_i a_i^0 X_i. \tag{3.25}$$

An example of SVM being used on a combination of pulse gradient, and peak height is shown in Figure 3.9 to display the training process. Because the data is not perfectly linearly separable, the hyperplane seeks to minimize the cost incurred from intersection with support vectors.

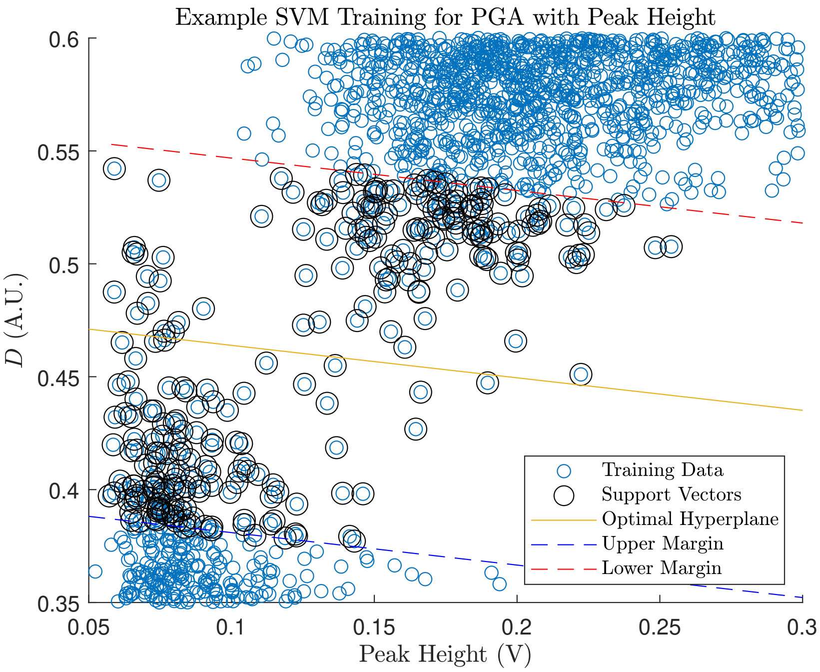

Figure 3.9: An example of SVM being trained on a training set composed of pulse gradient and peak height.

3.4.2 K Nearest Neighbors

Theoretically simpler than SVM, but possessing better performance for data with non-linear distributions, K nearest neighbors is also useful for this work. KNN operates on data, formatted as a set of pairs $(X_1, \theta_1), ..., (X_n, \theta_n)$ where X_i is a vector that takes values from the sample space Ω, containing the feature data, and θ_i is the population from which the data originated, and can take in practice take values $\{1, 2, ..., M\}$. When a new measurement of X is made without knowing its θ, we define a nearest neighbor to the new measurement as $X'_n \in \Omega$ if

$$\min d(X_i, X) = d(X'_n, X), i = 1, 2, ..., n. \tag{3.26}$$

The distance metric here $d(x_1, x_2)$ can be chosen according to the application. In our specific case, we use the standard euclidean distance. From the definition in 3.26 it follows then that we find the k nearest neighbors, i.e. the k X_i's in Ω that have the minimum distance to the newly measured point. We then look at the θ for each nearest neighbor, and if the majority belongs to a class, we assign that class to X [41, 42].

A visualization of this process, using data from the combination of the peak height feature with the rise-time feature is shown in Figure 3.10. The regions are colored according to the decision that would be made for a given point in that region, based on its nearest neighbors. Although the theory of this classifier is simple, it allows for complex decision spaces to be created that operate very well on data that may not be linearly separable, such as is the case with the rise-time data. An important caveat however, is that the KNN classifier is necessarily large, as it must hold the training dataset in memory to compare new data against the data that it learned. This can make it difficult to use on embedded sensors that have smaller available memory and computation power.

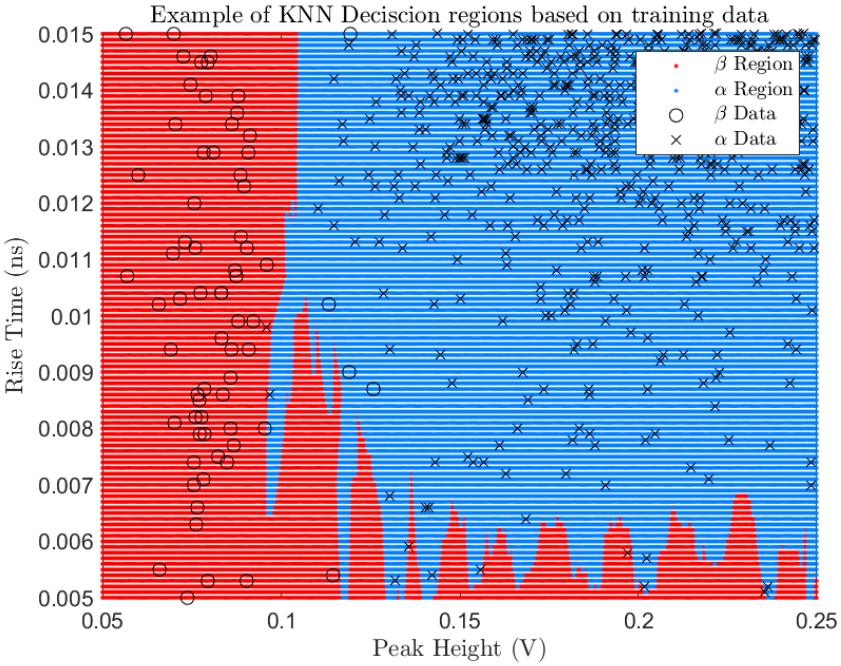

Figure 3.10: A zoomed in image of the decision boundary for the combination of rise-time discrimination data with peak height. The red regions represent areas where the nearest five neighbors will result in a β classification, while the blue represents the opposite. The distance metric used here is the euclidean distance.

3.5 Down Sampling

To test the transferability of these methods and analyses to lower bandwidth and sampling rate situations, down sampling of the waveforms in the dataset is employed. Referring back to the FFT in Figure 2.3, there was reason to believe that the frequency content of the produced SiPM signals would be quite pronounced between 500 MHz and 1 GHz. With the data in hand, this analysis was repeated, using the normalized mean functions $a(n)$ and $b(n)$ as described in the previous section. Figure 3.11 displays the normalized mean functions on the left, and the FFT of each of them on the right. The β pulse is shown to have larger amplitude at high frequency than expected, and in the range spanning 500 MHz to 1 GHz, there is a stark difference in the frequency content. This provides excellent motivation into the challenges that down sampling the data may pose. Based on the frequency content, the chosen frequencies to down sample at were 50 MHz, 100 MHz, 200 MHz, 1 GHz, 2 GHz, and 10 GHz, which was the native sampling rate. Down sampling was achieved by reduction of the number of points in each waveform when it was read in. For example, when emulating 50 MHz, every 200th sample of the native dataset was used.

The results of the down sampling analysis will be particularly important for this application. A 10 GHz digitizer is neither financially nor physically feasible for use in a small handheld detector. Digitizers that operate at that speed often retail close to \$ 2,000, and most α/β detectors in the handheld market are sold for \$ 3,000. This would leave very little budget left over to be applied to critical electronics such as photomultipliers, or the processor, which would need to be able to run fast enough to handle such a fast digitizer. Conversely, a 50 MHz digitizer would be both cost, size, and power efficient in comparison. Approaching the problem from the analog perspective, electronics necessary to adequately resolve signals at 1 GHz and higher tend to be power hungry and sensitive. For embedded devices that need to be capable of operating for hours on battery power only, this could be a major impediment. If

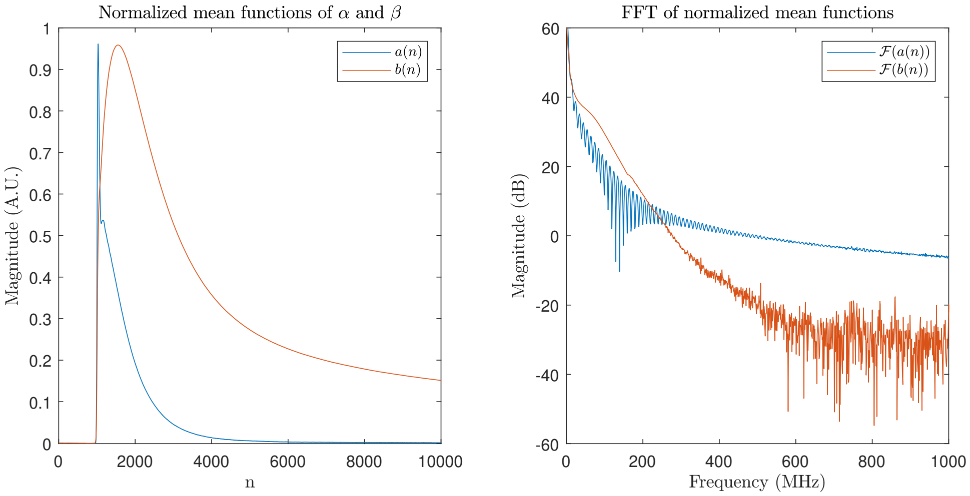

Figure 3.11: Frequency analysis of the normalized mean functions. The left plot shows the α and β normalized mean functions in terms of samples. The right plot shows the results of an FFT of these signals. Although there is a definite difference in the low frequency elements of the signal, the most apparent differences are seen in the high frequency band between 200 MHz and 1 GHz.

down sampling reveals that there are techniques that can work reasonably well at much lower bandwidths, compared to the wide bandwidth case, these will almost certainly present a more attractive option.

CHAPTER 4: CLASSIFICATION RESULTS

The methods and process for designing and testing the classifiers described in this chapter has been established in Chapter 3. In this chapter, we describe accuracy and other parameters for the single feature, and two feature classifiers, centered on rise-time, PGA, and ToT techniques. The set of optimal parameters related to measurement of pulse gradient and timing threshold are also reported, following the process laid out in Section 3.3.2. Potential errors in the experiment and design process are also discussed, along with steps taken to mitigate their impact.

The single feature classifier and its results are discussed in Section 4.1, and the results for the two feature classifiers is discussed in Section 4.2. The results of down sampling are discussed in Section 4.3. Finally, potential sources of error in the data, and the steps taken to mitigate or nullify them are established in Section 4.4.

4.1 Single Feature

As established in Chapter 3, approximately 2/3 of the available data was used for the purposes of fitting a probability distribution, and applying ML rule to determine D^*, the optimal cut off point for the one-dimensional classifier. The remaining 1/3 is used to test the effectiveness of the classifier and generate an accuracy metric based on the ratio of correct classifications to the total number of waveforms. Accuracy metrics are generated for both the α and the β particles, since the distributions for α and β particles need not be symmetric about D^*, and accuracies can be better or worse for one over the other. A summary of the achieved classification accuracy in the single feature case is shown in Table 4.1. Additionally, Table 4.2 contains the optimal paramters chosen in the single feature case to produce the accuracy numbers stated in Table 4.1, as well as the probability distribution fit parameters that were used to apply ML rule. Classification plots visualizing each of the different methods results, classification region, and accuracy are shown in Figures 4.1, 4.2, and 4.3.

Overall, PGA has the best classification performance for both α and β particles combined. In Figure 4.2 the separation in the distributions is very clear along the Y axis, compared with the separation achieved in Figure 4.1 and especially Figure 4.3. Rise-time achieves decent separation, although there are some definite outliers, and ToT appears separable but not for the one-dimensional discriminator which cannot take advantage of separability that could be achieved with two dimensions. PGA performs best here because it utilizes the digital nature of the data to a larger extent than the other methods, which were originally designed for analog.

It is important to consider here however, that PGA requires the most combined hardware and software compared to the other methods. Rise-time can be implemented using circuits as simple as CR/RC integrators and differentiators, or more rigorously using operational amplifiers. With the appropriate associated circuits, such as a time to digital (TDC) converter, the application's processor need only fetch the recorded time before a new event arrives. In general, dangerous to lethal radiation measurements correspond to event rates that are in the tens to hundreds of kHz. This means that modern micro-processors have nearly infinite time in between events to fetch the data and process it. The ToT method has a similarly simple implementation with comparators and a TDC, requiring minimal processor time or power.

Turning back to PGA, the price of its strong classification accuracy is that the signal must be digitized at 10 GHz, and the processor must be fast enough to receive all of this data, find the max value, and then conduct division on the point of discrimination to generate the PSD value. While modern microprocessors have made large gains in speed and floating point math, the requirements of PGA as compared to the other two methods present a very different class of processor as would suffice for rise-time

Table 4.1: Single feature discrimination results for all methods investigated. All methods perform well, being nearly 99% accurate except for in the case of ToT, which misses some particles due to the threshold issues.

Radiation Type	Rise-Time	PGA	ToT
β	98.95%	99.92%	94.72%
α	99.29%	99.98%	99.50%

or ToT. Viewing the problem through this lens, PGA stands to gain the most from the down sampling analysis. If PGA can still perform nearly as effectively at lower sampling rates, it makes a much stronger case for its usage in embedded and handheld detectors.

As is evident in Figure 4.3, ToT puts forth a unique issue when applying ML rule. Due to the handling of edge cases, discussed in Section 3.2, ToT has features that behave oddly. Notably, the line along the top of Figure 4.3, caused by the α particles which do not fall back below the threshold within the sampled window, is very difficult to fit a probability distribution to. In this case, we have just assigned the fall time to be the max length of the samples. Although it is not ideal, a normal distribution is used to fit the histogram, and verified by eye that the crossing point with the distribution of β particles is sensible. This worked well, as the distribution fit acceptably for the majority of the distribution, and failed at the edge. This is acceptable since the edge is far away from the region of interest for this work.

Table 4.2: Table of optimal parameters and probability distribution functions for the PSD features. ToT particularly produces an odd shape due to edge cases, the distribution fits to it are not ideal, posing an issue for MAP rule. Rise-time units are in μs unless otherwise specified. PGA's parameters are unit-less. ToT units are in tenths of a nanosecond (e-10) unless otherwise specified.

Rise-Time	
D^*	5.3 ns
β Distribution	Log-normal (-6.2482,0.3257)
α Distribution	Normal (0.0196,0.0052)
PGA	
D^*	0.4343
n^*	882
β Distribution	Normal (0.218637,0.0454226)
α Distribution	Normal (0.693336,0.0550436)
ToT	
D^*	143.4 ns
V_T^*	70 mV
β Distribution	Normal (398.5949,404.7976)
α Distribution	Normal (5830,2,626,000)

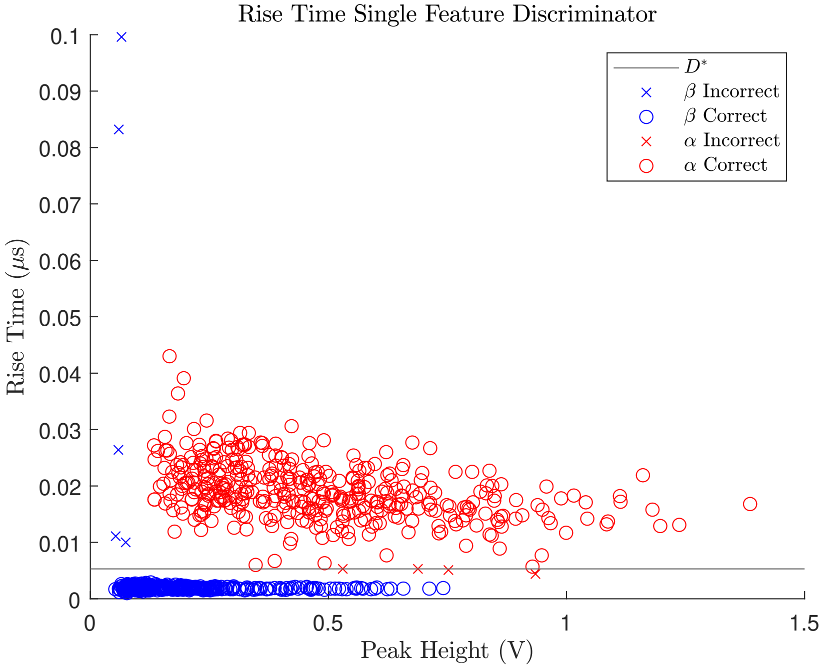

Figure 4.1: Plot of the single feature discriminator filter, using rise-time as the feature. Every 50th data point has been plotted for visibility purposes. Two separate distributions are clearly formed and the ML rule discriminator splits the distributions very well, ignoring the handful of outliers for the β particles.

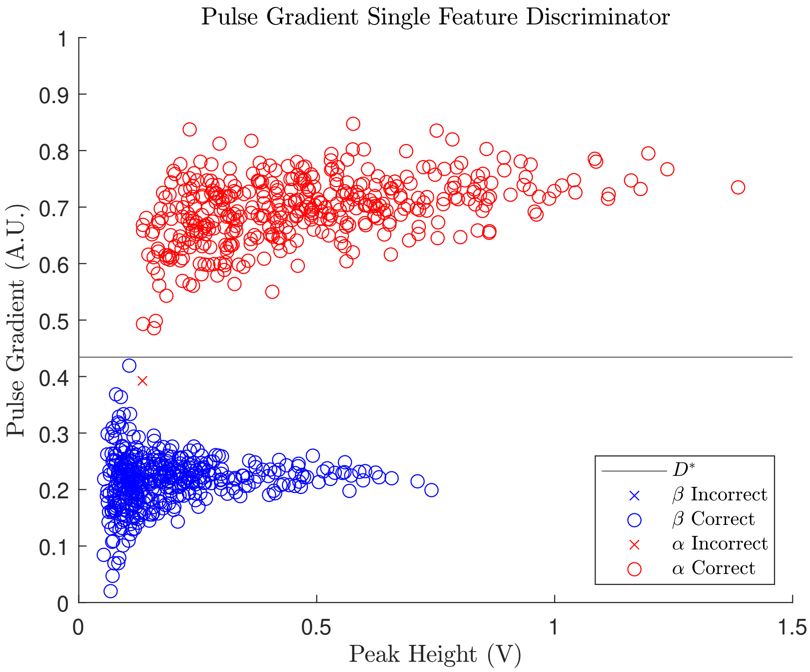

Figure 4.2: Plot of the single feature discriminator filter, using pulse gradient as the feature. Every 50th data point has been plotted for visibility purposes. Two distinct distributions are apparent from examination of the plot, and the D^* parameter splits them very well.

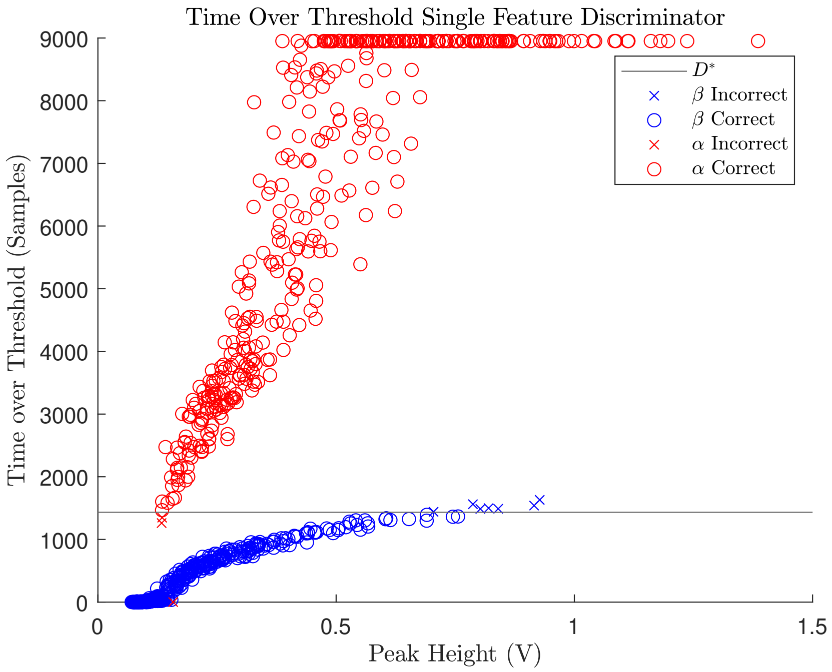

Figure 4.3: Plot of the single feature discriminator filter, using time over threshold as the feature. Every 50th data point has been plotted for visibility purposes. The clipping at the top is an artifact of α pulses that did not return below the threshold in the sampled window. While the distributions shown here are distinct, the single feature discriminator cannot take advantage, and there is overlap between the lowest energy α's and the highest energy β's

4.2 Multiple Features

Building on the results from the previous section, the results of the multi-feature classifiers are reported here. The method of implementation and parameters used for these techniques is covered in Sections 3.4.1 and 3.4.2. The resulting classification accuracies are outlined in Table 4.3. Parameters that were trained for both SVM and KNN are tabulated in Table 4.4 and 4.5 respectively. Partial visualizations of the classifier as used on the testing data set are provided for both methods across all three PSD techniques used here in Figures 4.4, 4.5, 4.6.

Between the two multi-feature classifiers used here, KNN has the best performance, although both classifiers improved accuracy. Despite its better performance it is important to consider the memory and computational requirements of KNN. In order to operate, the entire training dataset must be kept in memory and used as a lookup table to determine distances and neighbors. This is computationally expensive, and strains storage for embedded devices, such as handheld detectors. In comparison, SVM does not require the training data to be kept in memory, and computationally will only need to compare the newly measured data against the optimal hyperplane to derive a classification. This incurs only slightly more computational cost than the single feature case, making SVM an attractive route in all cases where it has comparable performance to KNN.

Some interesting cases arise from rise-time and ToT. SVM suffers when applied to the rise-time case, performing worse than the single feature threshold. This is likely due to the shape of the α and β distribution across the two features, shown in the top plot of Figure 4.4. The β particles have an L shaped distribution along the left side of the graph that is strongly non-linear. In attempting to fit to these "outlier" points, overall performance is negatively impacted. Since the probability of these points is low, they do not factor strongly into the ML rule calculation in he single feature case. It is possible that tuning of the cost function for the SVM training could improve its

Table 4.3: Classification using two feature classifiers. The comparatively low performance of ToT can be attributed to the missed pulses that go uncollected due to it's threshold technique.

Rise-Time	SVM	KNN
β	98.73%	99.84%
α	99.21%	99.59%
PGA	SVM	KNN
β	99.98%	99.99%
α	99.97%	99.96%
ToT	SVM	KNN
β	95.55%	95.34%
α	97.88%	99.51%

Table 4.4: Multi feature classification parameters for the application of SVM to each of the methods used here. These parameters can be used to implement the classifier in a system, or to replicate the experiment.

SVM	β_1	β_2	Bias
Rise-Time	1.5858	114.3963	-1.6089
PGA	1.7381	12.0921	-5.6928
Tot	-4.3797	0.0005	0.1387

result further in this case. The non linear shape however is inconsequential for KNN, which performs about a percentage point better in both cases, since KNN cares only about nearby points rather than overall distribution trends.

SVM also suffers on the α classification accuracy for ToT. Due to the clipping that occurs for α particles with long tails, along the top of the upper plot of Figure 4.6, the ToT value does not change over varying peak height due to the limited window of time that was sampled. In a real implementation, where a TDC or other analog implementation is used, this clipping will not occur, and the probability of an incorrect classification will be mitigated.

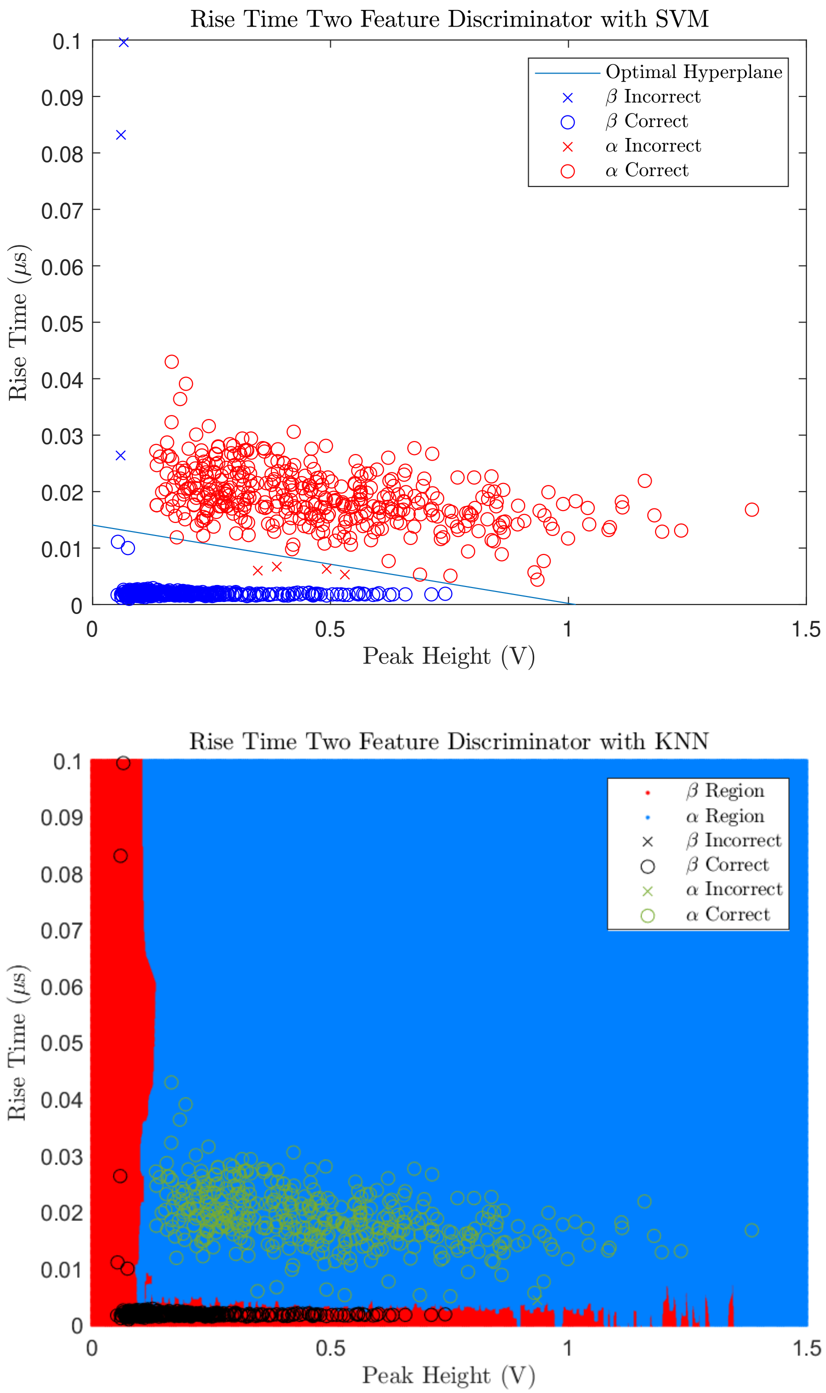

Figure 4.4: Multi-feature rise-time SVM and KNN discriminators. In both plots, only 400 data points out of the total 20,000 are shown for readability. In the top plot, the SVM classifier is shown as the light blue line. The blue population represents β particles, and the red population represents α particles. The circles represent correctly classified particles, and the crosses represent incorrectly classified particles. In the bottom plot KNN is shown. The blue region represents the area within which a new measurement is classified as an α particle, and the red region represents the area for β classification. β particles are shown in black, and α particles in green.

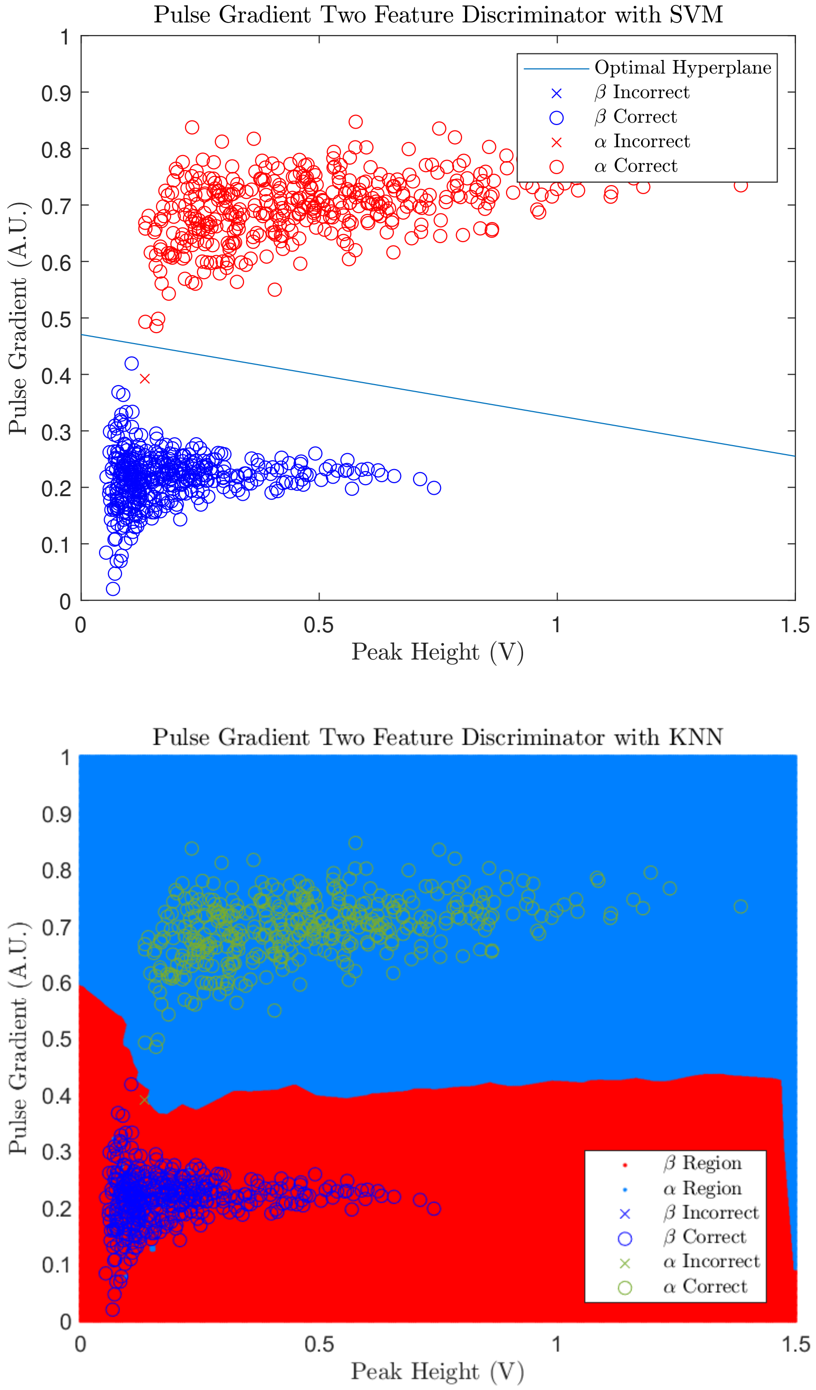

Figure 4.5: Multi-feature pulse gradient SVM and KNN classifiers. In both plots, only 400 data points out of the total 20,000 are shown for readability. The top plot shows the SVM classifier as a light blue line. α data is shown in red, β in blue. Circles represent correct classifications, crosses incorrect classifications. The blue region represents the area within which a new measurement is classified as an α particle, and the red region represents the area for β classification. Blue data points are β particles, green are α particles.

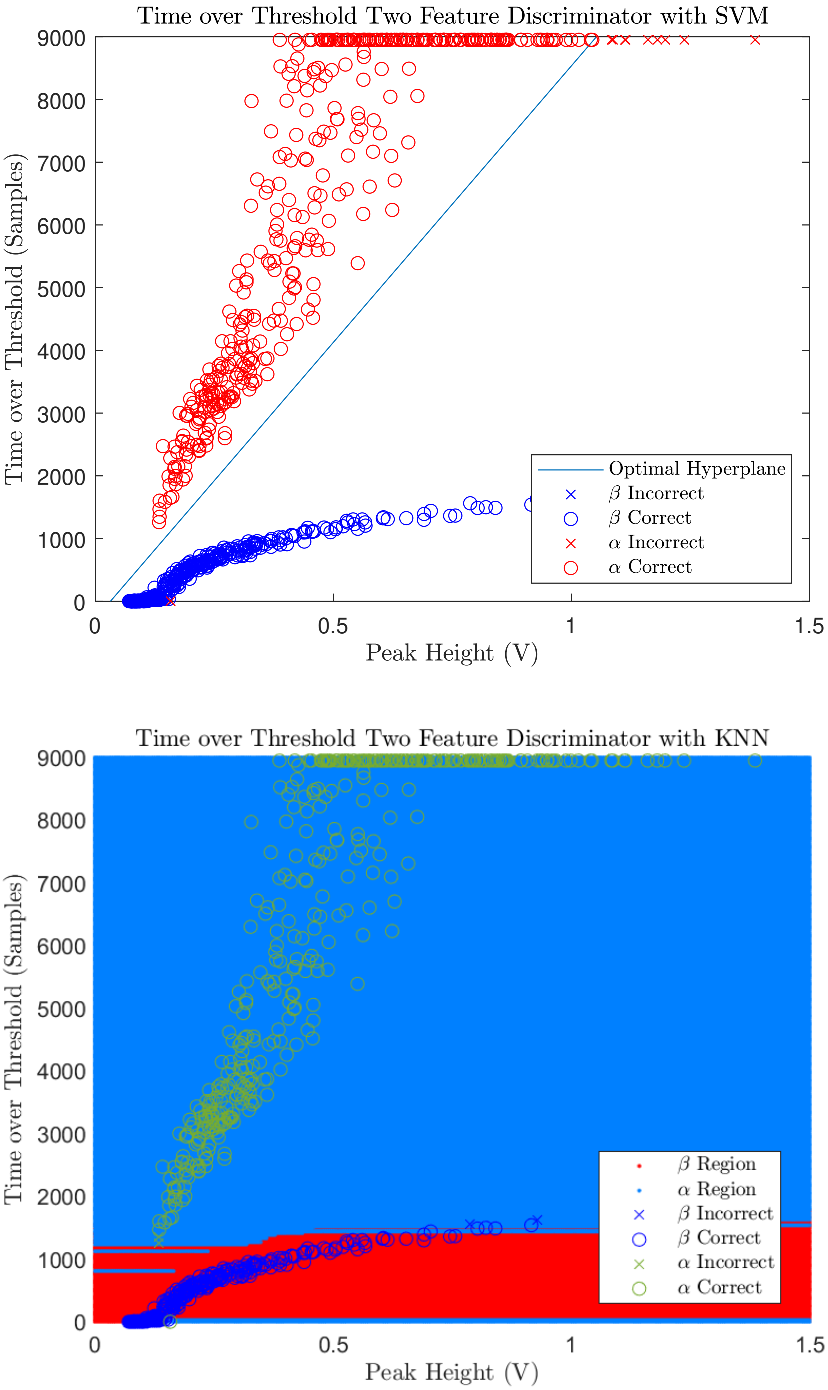

Figure 4.6: Multi-feature time over threshold SVM and KNN classifiers. In both plots, only 400 data points out of the total 20,000 are shown for readability. The top plot displays the SVM classifier as the light blue line splitting the two distributions. α data is plotted in red, β in blue. Crosses are incorrect classifications, circles are correct. The bottom plot shows the KNN classifier where the blue region represents the area within which a new measurement is classified as an α particle, and the red region represents the area for β classification. Blue data points are β particles, green are α particles.

Table 4.5: Multi feature classification parameters for the application of KNN to each of the methods used here. These parameters can be used to implement the classifier in a system, or to replicate the experiment. ToT has less samples than the other two methods because it cannot train on points that are missed by the rising trigger.

KNN	Samples	Neighbors
Rise-Time	76,000	5
PGA	76,000	5
Tot	74,129	5

4.3 Effects of Down Sampling

As covered in Section 3.5, these signals have strong frequency components in the range of 100 MHz to 1 GHz. Arbitrarily down sampling the data, which had a native sampling rate of 10 GHz, by taking every 5th, 10th, 50th, 100th, and 200th point, corresponds to effective sampling rates of 2 GHz, 1 GHz, 200 MHz, 100 MHz, and 50 MHz. Applying the same analysis applied in the previous sections to these cases allows for the robustness of the classifiers to be tested under more realistic conditions for a handheld, low power device. In particular in this section, changes to the distributions are assessed quantitatively, commentary is provided on which techniques are impacted the most, and the results are summarized. A compilation of the results for the down sampling analysis is provided in Table 4.6.

Although not all of the methods examined in this thesis are digital, the affect of down sampling mirrors the analog equivalent of reducing bandwidth. The affect of reducing sample rate and bandwidth on the signal size and shape is displayed in Figure 4.7. Comparing with the data in Table 4.6, the left plot of Figure 4.7 gives a visual representation of why rise-time's classification accuracy experiences a precipitous fall at low bandwidth. At 50 MHz, the rise-times are nearly identical, and the distributions lie essentially on top of each other. Figure 4.8 shows this evolution from 200 MHz to 100 MHz. Evaluating the results, proper execution of rise-time requires a bandwidth between 200 MHz and 1 GHz. Below this, accuracy is severely impacted. Another interesting result from the rise-time section of Table 4.6 is that

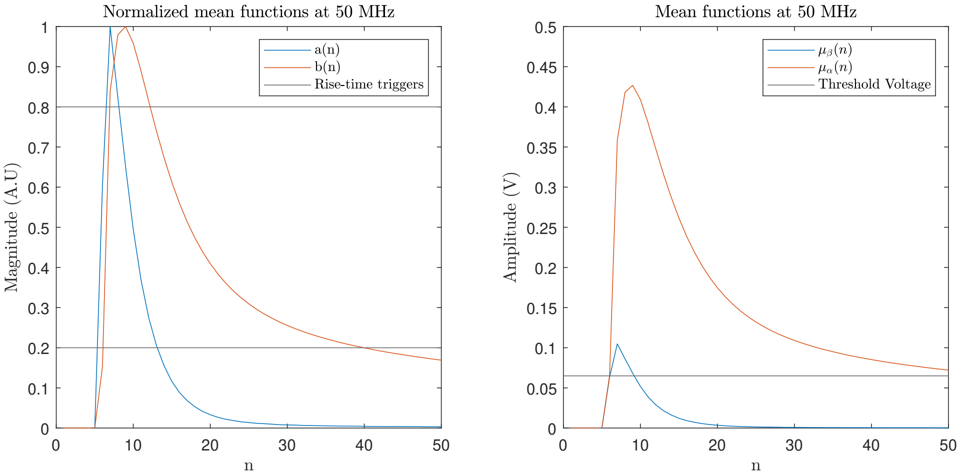

Figure 4.7: Normalized and non-normalized mean functions at 50 MHz. The left plot shows the normalized mean function for α and β particles. The rapid fall off in accuracy for rise-time can be clearly discerned. At 50 MHz, the rise-times are on average nearly indistinguishable. The right plot shows the non-normalized mean function, which explains the decaying β accuracy for ToT. Rather than the timings becoming tighter, the signal amplitude is severely decreased due to the low sampling rate. Many pulses come in below the starting threshold and are missed.

SVM and KNN are able to recover a significant amount of accuracy. The right plot of Figure 4.7 shows us how, as the SVM and KNN classifiers rely less and less on the rise-time data, and more on the peak height data as the rise-time feature degrades.

4.4 Sources of Error

This experiment has two unique sources of error that could not controlled for within the scope of its execution, cross-talk and gain drift. Cross-talk, covered in Section 4.4.1 has to do with the physics of scintillation detection, and impacts the ability to obtain pure datasets. Gain drift has to do with the strong correlation between SIPM gain and temperature. This is covered in Section 4.4.2.

4.4.1 Cross-Talk

As discussed in Section 1.1, the Corvid scintillator used here has two different radiation sensitive layers. Each layer is constructed out of materials that are sensitive to a specific radiation. However, both layers can also scintillate in the other as well.

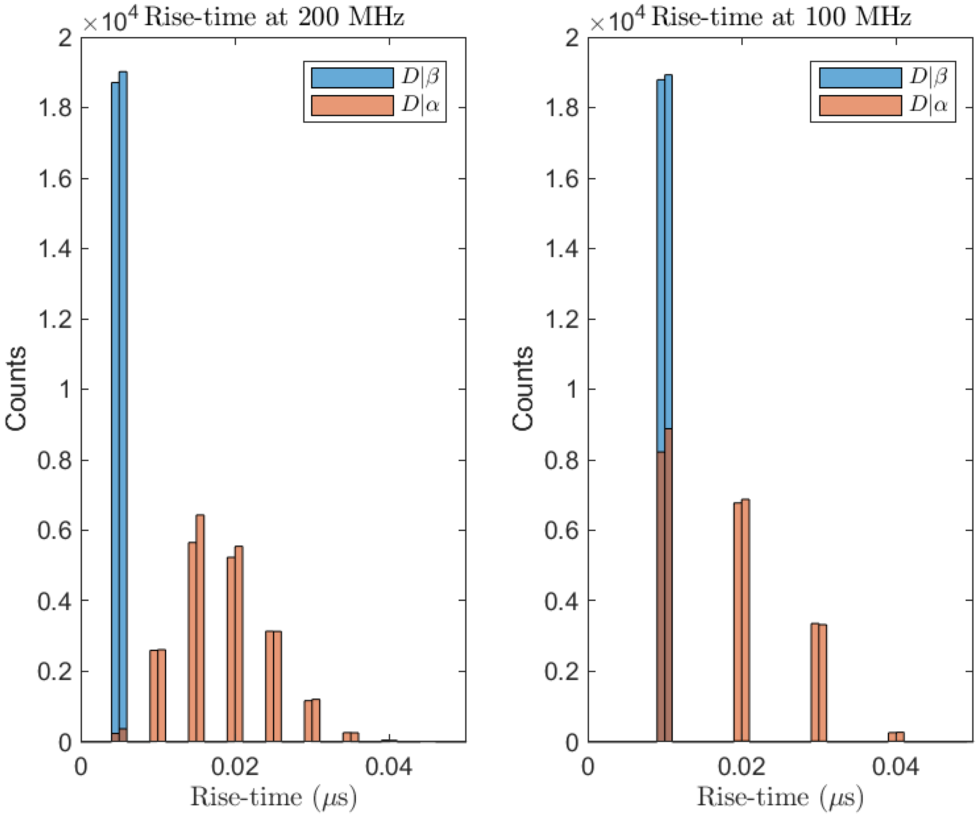

Figure 4.8: Degradation of the distribution of rise-time with decreasing frequency. On the left, the histogram of the rise-times at 200 MHz is shown, and on the right, the same distributions are shown at 100 MHz. The lower the bandwidth or sample frequency, the further into the β distribution the α distribution shifts.

Table 4.6: Table of results for down sampling all methods. Accuracies as separated by particle type. PGA performs the best, scoring better than 99% accuracy under all conditions and classifiers.

Rise-time	ML Rule	SVM	KNN
10 GHz (β)	98.95%	98.73%	99.84%
10 GHz (α)	99.29%	99.21%	99.59%
2 GHz (β)	99.22%	99.01%	99.84%
2 GHz (α)	99.35%	99.21%	99.61%
1 GHz (β)	99.38%	99.21%	99.83%
1 GHz (α)	99.28%	99.21%	99.62%
200 MHz (β)	99.35%	96.08%	99.60%
200 MHz (α)	98.41%	98.48%	98.91%
100 MHz (β)	99.37%	93.63%	97.14%
100 MHz (α)	55.50%	97.05%	98.28%
50 MHz (β)	99.75%	90.51%	89.07%
50 MHz (α)	16.06%	91.53%	90.91%
Pulse Gradient Analysis	ML Rule	SVM	KNN
10 GHz (β)	99.92%	99.98%	99.99%
10 GHz (α)	99.98%	99.97%	99.98%
2 GHz (β)	99.91%	99.96%	100%
2 GHz (α)	99.97%	99.96%	99.97%
1 GHz (β)	99.90%	99.95%	100%
1 GHz (α)	99.92%	99.96%	99.98%
200 MHz (β)	99.97%	99.92%	99.99%
200 MHz (α)	99.97%	99.95%	99.97%
100 MHz (β)	99.58%	99.77%	99.98%
100 MHz (α)	99.71%	99.82%	99.94%
50 MHz (β)	99.46%	99.78%	99.99%
50 MHz (α)	99.81%	99.88%	99.95%
Time over Threshold	ML Rule	SVM	KNN
10 GHz (β)	94.72%	95.53%	95.34%
10 GHz (α)	99.50%	97.88%	99.5%
2 GHz (β)	94.45%	95.19%	95.06%
2 GHz (α)	99.55%	99.67%	99.65%
1 GHz (β)	93.72%	94.46%	94.40%
1 GHz (α)	99.62%	99.74%	99.78%
200 MHz (β)	87.42%	87.79%	87.77%
200 MHz (α)	99.78%	99.94%	99.92%
100 MHz (β)	71.37%	71.60%	71.60%
100 MHz (α)	99.70%	99.93%	99.95%
50 MHz (β)	69.25%	69.74%	69.74%
50 MHz (α)	99.75%	99.93%	99.95%

This causes a problem, as the shapes of the pulses produced through scintillation, are jointly driven by delayed fluorescence and differences in decay times for excitations in the respective scintillators. Although it is not a common occurrence, it is possible for an α particle to scintillate in the β sensitive layer, and vice versa. Due to the physical construction of the scintillator, α particles that reach the β sensitive layer will have low energies, and therefore less delayed fluorescence. This means that an α particle scintillating in the β layer may be very difficult to differentiate from a normal β scintillation. On the other side, a β particle would require a very large amount of energy to bridge the energy gaps in the α sensitive layer, so a β scintillation in the α layer will look like a very low energy α pulse, with a tighter peak. In general, these features are difficult to near impossible to pick out without complicated algorithms.

In this thesis, cross-talk was handled through pre-processing of the dataset utilizing an un-optimized version of PGA on the collected data from each element used in this thesis. This approach netted a handful of pulses over the full 100,000 plus dataset. However, while PSD methods can generate strong classification results, no PSD approach is perfect, and there is likely to be some contamination in the training data, that would impact and “poison” the training. Thankfully, these events should have a low likelihood of occurrence, and are unlikely to have made a strong impact.

4.4.2 Gain Drift

SiPM gain has a strong temperature dependence, which requires most applications using them to either have an active cooling solution, or some way of measuring the temperature to calibrate for the changing gain. The SiPM used in these experiments has a temperature gain dependence of 1% per 5 °C [35]. Although this may not seem significant, over consumer, industrial, and military specification ranges this corresponds to as much as +/- 18% change in gain. The potential effects of this drift become much more pronounced when it is considered that typical SiPM gain is in the range of 10^6. A drift of 5% can correspond to 50,000. To control for and mitigate this issue in the case of unexpected results, temperature measurements were taken for every subset of data included in the overall dataset. Thankfully, these measurements indicated that all data was taken within +/- 5 °C of the rest, so individual datasets should vary by less than 1% from each other.

CHAPTER 5: CONCLUSIONS

In this thesis, multiple potential classifiers have been evaluated for use with the unique signals produced by a novel scintillator. Pulse gradient analysis in particular is a strong fit for implementation, as it can boast greater than 99% accuracy even at sampling rates as low as 50MHz using only one feature. Other methods investigated perform well, able to achieve greater than 90% accuracy at high sample rates using mutliple features, but struggle at lower sampling rates, falling to below 70% in some cases. Overall, pulse gradient analysis is a strong candidate for more strenuous testing and implementation.

Over the course of this thesis, an experiment was designed through examination of the theoretical dynamics of the signal in question. After examining the frequency domain of the signal, selecting suitable hardware, and developing a rigorous procedure, the experiment was run at TUNL. Utilizing a wide array of radioactive sources, a wide spectrum of energies for both α and β particles was obtained. In total, the resulting dataset comprised of greater than 60,000, each for α and β pulses, with approximately 20,000 γ pulses collected as well.

A rigorous process was developed for producing a pulse shape classifier designed specifically for the unique signals investigated here. MAP rule, which reduced to ML rule in the case of the training, was used to produce the optimal cut off point for classifying particles in the one-dimensional case. Theoretical analysis of the underlying signal structure was also done to determine the optimal parameter selection for PGA and ToT techniques. Further, SVM and KNN classifiers were reviewed from the literature and motivated for use on a combination of multiple features.

Finally, discriminating filter performance was examined across the single feature, multiple feature, and down sampled case. Through this examination a strong case was made for PGA as a discriminator for digital signals sampled at low sampling rates. Allowing for bandwidth of signals to be between 200 MHz and 1 GHz, analog methods also have a strong showing, as they can provide comparable results without as much computational overhead. Sources of noise unique to particle detection and particle physics were also covered, and mitigation methods were introduced and used to sanitize the dataset as much as was reasonable.

As a final note, before this work was begun, it was not expected that the signals would be so well distinguished from one another. It was also expected that more realistic data compared to what is available here would have been obtained. The dataset provided here represents ideal data, where the light is created in the scintillator physically near the detector. Signal to noise ratios suffer if the light is generated further away, as more photons are absorbed in transit. This thesis represents a first step towards the development of PSD techniques and implementations that robust in realistic detectors.

5.1 Future Work

Future work involving this dataset, scintillator, or classifier design methodology should investigate the γ dataset gathered here. Although the data was collected, timing constraints precluded an inclusion of analyzing the ability to discriminate between α, β, and γ particles in this thesis. Likely, much more advanced methods would need to be employed in this case, as β and γ particles are both known to interact with the same scintillator, through very similar processes. Distinguishing between the two would most likely require many more features, or specifically targeted features to produce good results. α/γ discrimination can be conducted already using the results of this thesis. In order to obtain statistically significant sample sizes for γ analysis it is likely that more data would need to be gathered from γ sources.

Another future avenue for investigation is corruption of the signals. To get the best sense of the shape of the signals, and what techniques would be pliable to use on them, ideal data was taken here. The sources were placed directly above the scintillator, which was directly above a SiPM. In practice, this is rarely the case. Often sources are coming in at random angles, and scintillations may not happen directly above a sensor, impacting the signal to noise ratio as photons can be absorbed through other surfaces, or inside the scintillator itself. It is likely very worth the effort to develop an experiment that will create degraded and low amplitude signals, to assess the accuracy of methods investigated here on these imperfect signals. This analysis will be more transferrable to field applications.

Another potential avenue of future research is the investigation of the affects of tiling on signal quality. Since SiPM sensors tend to be small (6 mm x 6 mm in this experiment), most applications that need to measure across a wide aperture will employ tiled arrays of sensors. In these cases, there is a capacitive effect that builds up due to intrinsic terminal capacitance in the SiPMs. This capacitance is likely to shape the resulting signals, potentially smearing out the fast rising β signals out over a longer period. Re-evaluating the effectiveness of methods used here on signals that have been stretched by this capacitive effect would be a very useful result for experimenters and engineers in multiple fields.

Finally, an interesting future analysis would be into whether or not cross-talk pulses can be accurately classified. It is possible that since the resulting pulses are a convolution of the scintillator dynamics, and the ratio of fluorescence to delayed fluorescence, cross-talk pulses may have their own unique shape that can be picked out and classified.

REFERENCES

[1] M. Marengo, C. J. Martin, S. Rubow, T. Sera, Z. Amador, and L. Torres, "Radiation safety and accidental radiation exposures in nuclear medicine," *Seminars in Nuclear Medicine*, vol. 52, no. 2, pp. 94–113, 2022. Radiation Exposure and Dosimetry.

[2] P. M. M. C. et al., "Precise positioning of gamma ray interactions in multiplexed pixelated scintillators using artificial neural networks," *Biomedical Physics Engineering Express*, vol. 10, p. 045038, jun 2024.

[3] H. B. et al., "Pulse-shape discrimination with the counting test facility," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 584, no. 1, pp. 98–113, 2008.

[4] Y. Morishita, A. Di Fulvio, S. Clarke, K. Kearfott, and S. Pozzi, "Organic scintillator-based alpha/beta detector for radiological decontamination," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 935, pp. 207–213, 2019.

[5] I. W. Croudace, P. E. Warwick, D. G. Reading, and B. C. Russell, "Recent contributions to the rapid screening of radionuclides in emergency responses and nuclear forensics," *TrAC Trends in Analytical Chemistry*, vol. 85, pp. 120–129, 2016. Trends in CBRN Measurements for safety and security.

[6] P. Sperr, H. Spieler, M. Maier, and D. Evers, "A simple pulse-shape discrimination circuit," *Nuclear Instruments and Methods*, vol. 116, no. 1, pp. 55–59, 1974.

[7] B. D'Mellow, M. Aspinall, R. Mackin, M. Joyce, and A. Peyton, "Digital discrimination of neutrons and γ-rays in liquid scintillators using pulse gradient analysis," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 578, no. 1, pp. 191–197, 2007.

[8] G. Bourlis, A. Tsirigotis, and S. Tzamarias, "Time over threshold electronics for an underwater neutrino telescope," in *2009 3rd International Workshop on Advances in sensors and Interfaces*, pp. 201–206, 2009.

[9] A. T. Krebs, "Early history of the scintillation counter," *Science*, vol. 122, no. 3157, pp. 17–18, 1955.

[10] J. Birks, *The Theory and Practice of Scintillation Counting, Electronics and Instrumentation.* Pergamon Press, 1964.

[11] F. Brooks, "Development of organic scintillators," *Nuclear Instruments and Methods*, vol. 162, no. 1, pp. 477–505, 1979.

[12] G. F. Knoll, *Radiation detection and measurement; 4th ed.* New York, NY: Wiley, 2010.

[13] A. Kumar and A. Waker, "An experimental study of the relative response of plastic scintillators to photons and beta particles," *Radiation Measurements*, vol. 47, no. 10, pp. 930–935, 2012.

[14] H. Kallmann and G. J. Brucker, "Decay times of fluorescent substances excited by high-energy radiation," *Phys. Rev.*, vol. 108, pp. 1122–1130, Dec 1957.

[15] R. B. Owen, "The decay times of organic scintillators and their application to the discrimination between particles of differing specific ionization," *IRE Transactions on Nuclear Science*, vol. 5, no. 3, pp. 198–201, 1958.

[16] F. Brooks, "A scintillation counter with neutron and gamma-ray discriminators," *Nuclear Instruments and Methods*, vol. 4, no. 3, pp. 151–163, 1959.

[17] F. D. Brooks, R. W. Pringle, and B. L. Funt, "Pulse shape discrimination in a plastic scintillator," *IRE Transactions on Nuclear Science*, vol. 7, no. 2/3, pp. 35–38, 1960.

[18] R. Monaghan and B. F. Wilson, "Dual phosphor detectors," *IRE Transactions on Nuclear Science*, vol. 7, no. 4, pp. 32–35, 1960.

[19] M. Roush, M. Wilson, and W. Hornyak, "Pulse shape discrimination," *Nuclear Instruments and Methods*, vol. 31, no. 1, pp. 112–124, 1964.

[20] E. Gatti, C. Cottini, S. Donati, V. Svelto, and F. Vaghi, "Particle identification by pulse shape discrimination," *Energia nucleare*, vol. 17, p. 60, 1970.

[21] F. T. Kuchnir and F. J. Lynch, "Time dependence of scintillations and the effect on pulse-shape discrimination," *IEEE Transactions on Nuclear Science*, vol. 15, no. 3, pp. 107–113, 1968.

[22] G. Ranucci, "An analytical approach to the evaluation of the pulse shape discrimination properties of scintillators," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 354, no. 2, pp. 389–399, 1995.

[23] J. Heltsley, L. Brandon, A. Galonsky, L. Heilbronn, B. Remington, S. Langer, A. Vander Molen, J. Yurkon, and J. Kasagi, "Particle identification via pulse-shape discrimination with a charge-integrating adc," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 263, no. 2, pp. 441–445, 1988.

[24] D. W. Jones, "A new circuit for pulse shape discrimination," *IEEE Transactions on Nuclear Science*, vol. 15, no. 3, pp. 491–499, 1968.

[25] J. Kasagi, T. Murakami, and T. Inamura, “Use of charge integrating adcs with pulse shape discriminators for neutron-gamma discrimination,” *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 236, no. 2, pp. 426–427, 1985.

[26] M. Aspinall, B. D’Mellow, R. Mackin, M. Joyce, Z. Jarrah, and A. Peyton, “The empirical characterization of organic liquid scintillation detectors by the normalized average of digitized pulse shapes,” *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 578, no. 1, pp. 261–266, 2007.

[27] M. Aspinall, B. D’Mellow, R. Mackin, M. Joyce, N. Hawkes, D. Thomas, Z. Jarrah, A. Peyton, P. Nolan, and A. Boston, “Verification of the digital discrimination of neutrons and γ rays using pulse gradient analysis by digital measurement of time of flight,” *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 583, no. 2, pp. 432–438, 2007.

[28] M. Gelfusa, R. Rossi, M. Lungaroni, F. Belli, L. Spolladore, I. Wyss, P. Gaudio, and A. Murari, “Advanced pulse shape discrimination via machine learning for applications in thermonuclear fusion,” *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 974, p. 164198, 2020.

[29] D. B. et al., “Optimized α/β pulse shape discrimination in borexino,” *Phys. Rev. D*, vol. 109, p. 112014, Jun 2024.

[30] S. L. et al, “Csi(tl) pulse shape discrimination with the belle ii electromagnetic calorimeter as a novel method to improve particle identification at electronâpositron colliders,” *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 982, p. 164562, 2020.

[31] G. Liu, M. Aspinall, X. Ma, and M. Joyce, “An investigation of the digital discrimination of neutrons and γ rays with organic scintillation detectors using an artificial neural network,” *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 607, no. 3, pp. 620–628, 2009.

[32] N. M. Michels, A. J. Jinia, S. D. Clarke, H.-S. Kim, S. A. Pozzi, and D. D. Wentzloff, “Real-time classification of radiation pulses with piled-up recovery using an fpga-based artificial neural network,” *IEEE Access*, vol. 11, pp. 78074–78083, 2023.

[33] A. Fay. https://www.kaggle.com/datasets/alexfay/ruddeized-alpha-beta-scintillator. Ruggedized Alpha Beta Scintillator Pulses.

[34] P. P. Caló', F. Ciciriello, S. Petrignani, and C. Marzocca, "Sipm readout electronics," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 926, pp. 57–68, 2019. Silicon Photomultipliers: Technology, Characterisation and Applications.

[35] Broadcom Inc., *AFBR-S4N66P014M NUV-MT Silicon Photomultiplier.*

[36] D. R. Schaart, S. Seifert, H. T. van Dam, M. R. de Boer, R. Vinke, P. Dendooven, H. Lohner, and F. J. Beekman, "First experiments with labr 3: Ce crystals coupled directly to silicon photomultipliers for pet applications," in *2008 IEEE Nuclear Science Symposium Conference Record*, pp. 3991–3994, IEEE, 2008.

[37] S. Sajedi, N. Zeraatkar, M. Taheri, S. Kaviani, H. Khanmohammadi, S. Sarkar, H. Sabet, and M. Ay, "A generic, scalable, and cost-effective detector front-end block for pet," in *2017 IEEE Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC)*, pp. 1–3, IEEE, 2017.

[38] Analog Devices, *AD8099: Ultralow Distortion, High Speed, 0.95 nV/âHz Voltage Noise Op Amp.*

[39] Broadcom Inc., *AFBR-S4E001 NUV-MT SiPM Evaluation Kit.*

[40] V. V. C. Cortes, "Support-vector networks," *Mach Learn.*

[41] J. H. E. Fix, "Nonparametric discrimination: Consistency properties," tech. rep., USAF School of Aviation Medicine, 1951.

[42] T. Cover and P. Hart, "Nearest neighbor pattern classification," *IEEE Transactions on Information Theory*, vol. 13, no. 1, pp. 21–27, 1967.