

ABSTRACT

SENTENO, DOMINIC ANDREAS. Application of Unsupervised Machine Learning Algorithms to Neutron/Gamma Particle Identification using Organic Scintillators. (Under the direction of Dr. John Mattingly.)

Neutron and gamma particle identification (PID) has numerous applications in radiation detection, but PID distinguishing between low energy neutrons and gammas remains challenging. Some organic scintillators enable PID through pulse shape discrimination (PSD). Previous research has explored multiple PSD methods including the application of machine learning (ML). Some showed successful PID but not for low energy neutrons. This thesis presents an application of unsupervised ML to implement PID of neutrons at all energies in a gamma rich environment. Time-of-flight (TOF) experiments with a Na-22 source and a Cf-252 source were used to collect training and test data consisting of neutron and gamma pulses. The charge integration PSD method was applied to these measurements to establish a baseline for ML PID performance. In addition to the total charge and the ratio of delayed charge to total charge, quantiles measuring the time required to collect specified fractions of the total charge were used as features in the unsupervised ML algorithms. The K-means, MiniBatch K-means, Gaussian Mixture Model with Expectation-Maximization (GMM-EM), and Variational Bayesian Gaussian Mixture Model (VBGMM) algorithms clustered neutrons from a Cf-252 source and gammas from a Na-22 source the ML PID results were compared to ground truth visually and quantitatively. The quantitative comparison employed the true positive rate (TPR) and false positive rate (FPR) for each of the ML algorithms. The results show the GMM-EM algorithm with the quantiles, long gate integral, and PSD ratio features performs the best PID compared to the other methods. Therefore, the GMM-EM model should be the focus of future applications of ML techniques for sensitivity analysis and model comparison.

© Copyright 2021 by Dominic Andreas Senteno

All Rights Reserved

Application of Unsupervised Machine Learning Algorithms to Neutron/Gamma Particle
Identification using Organic Scintillators

by
Dominic Andreas Senteno

A thesis submitted to the Graduate Faculty of
North Carolina State University
in partial fulfillment of the
requirements for the degree of
Master of Science

Nuclear Engineering

Raleigh, North Carolina
2021

APPROVED BY:

John Mattingly
Committee Chair

Ge Yang

Xu Wu

DEDICATION

This work is dedicated to my wife, Allison, and two boys, William and Auston. This would not have been possible without your love and support. I love you, always.

BIOGRAPHY

Dominic Senteno received a Bachelor of Science in Mathematics from the United States Military Academy in 2009. Commissioned as an Engineer Officer, he completed assignments in Fort Bragg, NC, Fort Irwin, CA, and San Luis Obispo, CA including two deployments to Afghanistan. Upon completion of company command in Fort Irwin, he served as an assistant professor in the Military Science and Leadership Department at California Polytechnic State University. Upon completion of his teaching assignment, he transferred to functional area 52, Nuclear and Counterproliferation. He has been in the Army for nearly 12 years and is currently in the rank of Major. He is married with two kids ages 3 and 1.

ACKNOWLEDGMENTS

I would like to sincerely thank my advisor, Dr. John Mattingly, without your guidance, professionalism, and expertise I would not have been able to accomplish this. It was a pleasure learning from you and working with the RADIANS research team members. I would like to thank Dr. Mudit Mishra for his support with the time-of-flight experiment. I would also like to thank my committee members and instructors at North Carolina State University for supporting and challenging me in my life-long learning endeavor. A special thanks to LTC Joseph Dzvonic, Dr. Jim Prince, and Dr. Edward Fuselier for their support. Finally, I would like to thank my friends, family, and mentors for their support and motivation to persevere, achieve my goals, and have fun along the way.

TABLE OF CONTENTS

LIST OF TABLES	vi
LIST OF FIGURES	vii
Chapter 1: Introduction	1
1.1 Motivation.....	1
1.2 Prior work	1
1.3 Novel contributions.....	5
Chapter 2: Background	6
2.1 Organic scintillators	6
2.1.1 Gamma detection	8
2.1.2 Neutron detection.....	10
2.2 Fluorescence	13
2.2.1 Prompt fluorescence.....	16
2.2.2 Delayed fluorescence	16
2.3 Pulse shape discrimination (PSD).....	17
2.3.1 Charge integration method.....	18
2.3.2 Figure-of-merit (FOM)	21
2.3.3 Time-of-flight (TOF) method	22
2.4 Quantiles	24
Chapter 3: Implementation of pulse shape discrimination (PSD)	26
3.1 Time-of-flight (TOF) experiment	26
3.1.1 Record processing.....	29
3.1.2 TOF results.....	34
3.2 PSD calculation.....	36
3.2.1 Pulse alignment.....	37
3.2.2 Charge integration.....	38
3.2.3 Gaussian fits.....	40
3.2.4 Figure-of-merit (FOM) calculations	41
3.2.5 Optimization	42
Chapter 4: Machine learning (ML) application to particle identification	45
4.1 Development of ML models	45
4.1.1 K-means and MiniBatch K-means models	51
4.1.2 Gaussian mixture model (GMM) and GMM variants	54
4.2 Validation of ML models.....	63
Chapter 5: Conclusions and future work	67
5.1 Conclusions.....	67
5.2 Future work.....	68
References	70

LIST OF TABLES

Table 4.1 TPR and FPR with standard deviation for the primary algorithms considered based on PID clustering. GMM-EM with all features has best FPR and TPR results..65

LIST OF FIGURES

Figure 2.1	The three main types of photon interaction. The lines show boundaries where one becomes more dominant than another.....	9
Figure 2.2	Diagram of Compton scatter; photon scatters off a free electron. $E = h\nu$ where h is Planck's constant and ν is the photon's frequency.....	9
Figure 2.3	Diagrams of the center-of-mass and lab frame systems for elastic scattering of a neutron.....	11
Figure 2.4	Singlet and triplet levels of the π -electron state	15
Figure 2.5	The time dependence of light intensity for gamma rays, fast neutrons, and alpha particles in a stilbene organic scintillator	17
Figure 2.6	Example of the pulses detected from incident neutrons and gammas, with identification of the tail (red) and total area (blue + red)	18
Figure 2.7	PSD using the charge integration method for a measurement of Cf-252, gamma (γ) and neutron (n), source with EJ-309	20
Figure 2.8	An example of the histogram of the $\frac{Q_S}{Q_L}$ PSD parameter	21
Figure 2.9	Diagram of TOF setup.....	22
Figure 2.10	TOF histogram displaying the gamma (γ) and neutron (n) peaks from an experiment conducted at Lawrence Livermore National Laboratory (LLNL) with a Cf-252 source	23
Figure 2.11	Plot of cumulative integrals for the waveforms from (a) Cf-252 and (b) Na-22	24
Figure 2.12	Plot of quantile means with standard deviations for (a) Cf-252 and (b) Na-22	25
Figure 3.1	TOF experimental setup. Cf-252 is placed next to the trigger detector (left) 1 m from the study detector.....	27
Figure 3.2	Example waveforms from the same record [(a) trigger detector (CeBr ₃), (b) study detector (EJ-309)] with pulses in coincidence	28
Figure 3.3	Example waveform with features labeled	29
Figure 3.4	Original waveforms from records with varying amplitude and occurrence [(a) trigger detector (CeBr ₃), (b) study detector (EJ-309)].....	29

Figure 3.5	Baselined waveforms from records with varying amplitude and occurrence [(a) trigger detector (CeBr ₃), (b) study detector (EJ-309)].....	30
Figure 3.6	Inverted waveforms from records with varying amplitude and occurrence [(a) trigger detector (CeBr ₃), (b) study detector (EJ-309)].....	30
Figure 3.7	Histogram of standard deviation (σ). The boundaries for $1 < \sigma < 5$ are represented by the red dashed lines	31
Figure 3.8	Histogram of locations of maximum amplitude. The boundaries for $20 < l < 40$ are represented by the red dashed lines	32
Figure 3.9	2-D histograms of the effective pulse width vs. maximum amplitude. The trigger detector (CeBr ₃) before step five of the processing (a) and after (c). The study detector (EJ-309) before step five of the processing (b) and after (d)	33
Figure 3.10	2-D histograms of the effective pulse width vs. maximum amplitude after removing records with waveforms with multiple pulses [(a) trigger detector (CeBr ₃), (b) study detector (EJ-309)]	34
Figure 3.11	Histogram of the initial TOF a gamma only Na-22 source used to find the time delay	35
Figure 3.12	Histogram of the TOF for Cf-252	36
Figure 3.13	Waveforms from Cf-252 records	36
Figure 3.14	Waveforms with pulses aligned at maximum amplitudes.....	37
Figure 3.15	Reduced waveforms with pulses aligned at maximum amplitudes.....	38
Figure 3.16	2-D histogram showing PSD for a short gate fraction of 90%. Gammas (γ) are represented by the lower horizontal peak and neutrons (n) in the upper horizontal peak	39
Figure 3.17	The PSD histogram exhibiting a peak from gamma pulses (γ) and another peak from neutron pulses (n)	40
Figure 3.18	PSD histogram with double Gaussian curve fit.....	40
Figure 3.19	Comparison of the FOM by short gate fraction	41
Figure 3.20	(a) PSD histogram from optimized parameters and (b) PSD histogram with double Gaussian curve fit	44
Figure 4.1	PSD for (a) Cf-252 and (b) Na-22	48

Figure 4.2	PSD for Cf-252 only including records identified as neutrons through TOF. Red box represents gammas misidentified as neutrons	49
Figure 4.3	PSD of gammas (red) from Na-22 source and neutrons (blue) from Cf-252 source.....	50
Figure 4.4	PSD of gammas (red) and neutrons (blue) using a discrimination line.....	51
Figure 4.5	PID using (a) K-means and (b) MiniBatch K-means algorithms with $\frac{Q_S}{Q_L}$ ratio feature	53
Figure 4.6	PID using GMM-EM on the $\frac{Q_S}{Q_L}$ ratio with a tied covariance matrix	57
Figure 4.7	PID using GMM-EM on the $\frac{Q_S}{Q_L}$ ratio and Q_L with a tied covariance matrix	58
Figure 4.8	PID using GMM-EM using bottom-up slicing with a tied covariance matrix	59
Figure 4.9	PID using GMM-EM using quantiles, $\frac{Q_S}{Q_L}$ ratio, and Q_L with diagonal covariance matrix.....	60
Figure 4.10	PID using GMM-EM using quantiles only with a diagonal covariance matrix	61
Figure 4.11	PID using VBGMM using quantiles, $\frac{Q_S}{Q_L}$ ratio, and Q_L with diagonal covariance matrix.....	62
Figure 4.12	PID using VBGMM using quantiles only with diagonal covariance matrix	62

CHAPTER 1

Introduction

1.1 Motivation

Particle identification (PID) is the process of determining the incident radiation particle type from pulse-mode radiation detector signals.[1] The focus of this research is PID with organic scintillation counters. Historically scintillation counters have been used for both neutron and gamma detection.[2] The accurate identification of gammas and neutrons is applicable in research laboratories and in field environments; these applications include: detection of special nuclear materials (SNM); particle physics and astrophysics for background estimation and dark matter research; nuclear medicine for dosimetry; nuclear reactor instrumentation; and estimating plasma fusion yield and shielding characteristics.[3, 25, 31]

Since the origin of visual scintillation counting in 1908 (Crookes and Regener), significant advancements were made over the past century starting in the 1930s with the development of electronic detectors and counters and progressing to today's photomultiplier scintillation counters with analog-to-digital converters (ADCs).[2] Even with the improvement of detection devices and materials there still exist a challenge in discriminating between low energy neutrons and gammas. Different analog systems and analysis algorithms are used to implement PID; recently, machine learning (ML) has been applied to PID. This research further explores the application of ML to PID.

1.2 Prior work

As previously stated, the capabilities of scintillator counters and the pulse counting systems have steadily improved over the past century. As early as 1948, multiple studies on the decay time of fluorescence were performed by both chemists and physicists. Some attribute the

first exploration of scintillator counters and their properties for use in PID to Bailey and Rollefson [16] and Phillips and Swank [17]. Both groups conducted independent studies of decay times within scintillation detectors. In 1956, G.T. Wright published a study on the comparison of the decay times of photo-fluorescence and scintillation from electrons in scintillation counters. Wright found noticeable differences in the decay times of alpha and gamma particles.[18] The decay time properties of scintillation counters and how these properties assist in the discrimination between neutrons and gammas were further explored and reported by F.D. Brooks in 1959. Brooks demonstrated that the detected particles induce different fluorescence decay times.[19] Brooks continued his research using different scintillator detector materials and further identified the energy range over which discrimination between neutrons and gammas is possible.[20] Gatti and De Martini developed optimum linear filter circuits to conduct analog PID, the results of which are the baseline for comparison against several methods.[21] After the developments by Gatti and De Martini, in 1962, several methods for PID through pulse shape discrimination (PSD) were developed.

PSD methods may be categorized into three techniques: analog, time-domain, and frequency-domain. The common analog techniques include the zero-crossing method and charge integration method.[22] Common time-domain techniques include pulse gradient analysis, time of flight (TOF), triangular filtering, and trapezoidal filtering.[22] The common forms of frequency-domain techniques include frequency gradient analysis and discrete wavelet transforms.[22] Alexander and Goulding developed the zero-crossing method, or rise-time discrimination method, which is a distance-based method.[25, 38] The zero-crossing method transforms the pulse signal into a bipolar signal then integrates the time interval from the start of the pulse to the zero-crossing point. The pulses display small differences in decay time, such that

the zero-crossing point is different between particle types, which is used to identify the particle type of the radiation that created the pulse.[26, 29, 37] The charge integration method, or charge comparison method, is the method used in this research; it is discussed in greater detail in Chapter 2. The TOF technique was used in this research to provide the ground truth for the particle types; it is also discussed in greater detail in Chapter 2. Several novel ideas were explored based on these techniques to improve discrimination between neutrons and gammas by improving the ADC, circuitry, or the PSD method.[25, 28, 30] However, the charge integration method for PSD is one of the most common and most reliable methods for effectively identifying particle type.[22, 26] Several researchers use TOF to compare the results and provide the ground truth for the particle type based on a comparison of arrival times.[22, 31, 36] The figure of merit (FOM) is a common technique used to test the charge integration PSD method performance; it is discussed in greater detail in Chapter 2.[35]

The PSD methods discussed up to this point do not address the issue of identifying low energy neutrons. Some recent work has aimed to identify low energy neutrons using novel techniques like the reference-pulses PSD method or a slicing algorithm. Both methods use pulse shapes from the known high energy regions of gammas and neutrons to perform PID in the low energy region. The reference-pulses PSD method uses reference pulses to compare to the measured pulses to identify the type of particle that generated the pulse based on similarities in shape.[32] The slicing algorithm “slices” the PSD peaks into smaller samples to cluster and find the center of the data. Analysis of the sample’s clustering results are used to identify which points belong to different particle types by creating a discrimination line. In this research, the slicing algorithm is applied to the ML algorithms and discussed in Chapter 4.[3, 34]

At the turn of the 21st century, several researchers explored the PSD capabilities and advantages of improved technology and newly discovered mathematical methods and systems.[15] The development of ML has simplified the processing of data for PID and PSD. The research results of ML show the effectiveness of applying statistical learning to radiation detection. The minimal computational resources necessary for the implementation of ML to classify or cluster particle types also makes ML an appealing tool. ML is a form of statistical learning which may include supervised, unsupervised, or semi-supervised learning. Supervised learning is “a model trained with a dataset that contains both features and labels.”[15] Supervised models are associated with classification, while unsupervised learning clusters the data. Unsupervised learning, as the name describes, is “a model trained with a dataset that contains only features and no labels.”[15] Semi-supervised learning is a mixture of supervised and unsupervised learning that trains a model to classify or cluster with “partially labeled data”. [12, 44]

ML methods explored for PID include the classification or clustering techniques from artificial neural networks (ANNs), support vector machines (SVMs), principal component analysis (PCA), decision trees (DTs), C-means, K-means, Gaussian mixture models (GMMs), Gaussian mixture models with expectation-maximization (GMM-EMs), and Bayesian/Dirichlet process Gaussian mixture models (DPGMMs).[4-15] Some of the primary measures for comparison between ML methods are the execution time and fraction of correctly identified particles. From 2018 to present, researchers like B. Blair et al.[4], M. Gelfusa et al.[7], E. Doucet et al.[10], R. Wurtz et al.[12], and Z.H. Zhang et al.[14] have delved deeper into the use of ML to successfully characterize low energy neutrons and gammas.

1.3 Novel contributions

This research provides an in-depth analysis of the fundamental tools and technology required for PID through PSD. The offline digital processing of the data collected for PSD enables the optimization of parameters and exploration of features for different ML algorithms. The ML algorithms used enable the clustering of the records collected with a scintillator counter. Chapter 2 outlines the fundamental principles that drive the detection of gammas and neutrons in organic scintillators through fluorescence, then it describes the charge integration PSD method, FOM, and TOF. Chapter 3 covers the algorithms designed to process TOF data and the optimization tools implemented on the charge integration method and FOM for PSD. Chapter 4 provides an overview of the different ML algorithms used for PID, the tools used to measure performance, and the outcomes of the research. Chapter 5 shows some conclusions from this work and makes a few recommendations for future work.

CHAPTER 2

Background

2.1 Organic Scintillators

Organic scintillators are a group of radiation detectors that are excited at the molecular level by ionizing radiation. Organic scintillators can be made in a liquid, gaseous, plastic, thin film, glass, or crystalline state. Organic scintillators are unique in that they are made of aromatic hydrocarbon compounds; that is, they primarily consist of planar molecules made up of benzenoid rings. Due to their hydrogen content, organic scintillators are preferred for fast neutron detection because full energy transfer to the recoil proton from hydrogen-1 scatter is possible. Boron, lithium, cadmium, or gadolinium may be loaded into the solution to increase thermal neutron sensitivity in liquid and plastic organic scintillators. Hydrogen and carbon are poor absorbers of gamma particles. The solutions may be doped with lead to increase gamma sensitivity. There is a limit to how much neutron- or gamma-absorbing material may be added to the scintillator because its transparency is reduced with the addition of foreign material. To be a useful scintillator the organic material must be transparent to its own emission spectrum.[2, 39-41]

Organic scintillators are classified as unitary, binary, ternary, quaternary, or more complex compounds based on the solution mixture. Most organic scintillators are binary. The mixture in a binary organic scintillator consists of a solvent and solute combination. The energy deposition interaction of gammas and neutrons occurs primarily in the solvent, and fluorescence occurs in the solute to create the scintillation light.[2] An advantage organic scintillators possess that makes them desirable for radiation detection is a fast response time (<10 ns), which enables fast timing measurements.[40] A disadvantage of binary liquid scintillators is the phenomena of

“quenching”. Quenching is a reduction in the light output due to elements in the mixture either changing the optical properties or interfering with the interactions within the scintillator mixture.[39]

In this research, an EJ-309 liquid organic scintillator was used based on its gamma and fast neutron PSD capabilities. EJ-309 is a binary, undoped, xylene solution liquid organic scintillator. The chemical properties of EJ-309 make it a safe option for the field environment.[46] The EJ-309 liquid organic scintillator served as the study detector for the TOF experiment and provided the data for the PSD algorithm. Photon detection is not the primary use of organic scintillators due to their low efficiency, but some choices of mixtures, like EJ-309, are effective.

All organic scintillators operate on the same basic principles. The physical mechanisms that drive these principles exist in the singlet and triplet states within the π -bonds. A simple explanation is that the molecules are excited by ionizing radiation to upper excited singlet states, then the excited molecules dissipate some of their energy through vibration before reaching the energy minimum for the excited state and de-exciting back to the ground state. The difference between the excitation energy and de-excitation energy is what produces the scintillation light.[40] This oversimplification does not provide an explanation of the different types of light signal, called fluorescence, produced when radiation particles interact in the scintillator material. Fluorescence, both prompt and delayed, and the interactions that cause them are covered in greater detail in Section 2.2.

2.1.1 Gamma detection

This research focuses on gammas with energy ranging from 10 keV to 10 MeV. The light production in organic scintillators results from pair production, photoelectric effect, and Compton scatter.

Pair production may occur when a photon with energy greater than two electron rest masses (1.022 MeV) interacts with the nuclear Coulomb field. The higher the energy of the photon, the greater the probability that pair production will occur. The likelihood of pair production also increases with atomic number Z as Z^2 . The incident photon is converted into the rest mass energy and kinetic energy of an electron and positron pair. The positron generally subsequently annihilates with an electron to produce two 511 keV photons.[2, 39-41]

The photoelectric effect is more probable for photons with lower energy. In the photoelectric effect, the photon is essentially absorbed by an orbital electron. Complete energy transfer from photon to the recoil electron occurs in photoelectric interactions. The energy of the recoil electron (photoelectron) is equal to the difference between the photon's energy and the binding energy of the original (K, L, or M) shell of the photoelectron. This leaves the atom in an ionized state with a vacancy in the original electron shell. The hole in the shell is quickly filled by either capturing a free electron or through the rearrangement of orbital electrons, which results in the emission of an x-ray or an Auger electron, respectively.[2, 39-41]

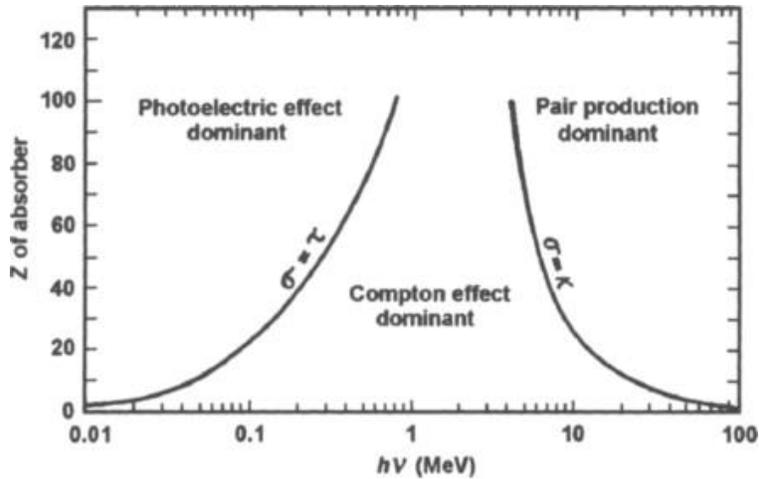


Figure 2.1. The three main types of photon interaction. The lines show boundaries where one becomes more dominant than another.[39]

The light produced from gamma interactions in organic scintillators predominately comes from Compton scatter because photoelectric absorption and pair production are improbable in low Z materials. Compton scatter is the most common interaction for photons in low-Z materials, and it is possible across the entire energy spectrum, as Figure 2.1 illustrates. Compton scatter is an interaction between a photon and an orbital electron; the energy of the photon is usually far greater than the binding energy of the electron, and, therefore, the electron may be treated as a free electron at rest, as Figure 2.2 illustrates. The incident photon collides with the free electron and the interaction results in part of the photon's energy being imparted on the electron in the form of kinetic energy.[2, 39-41]

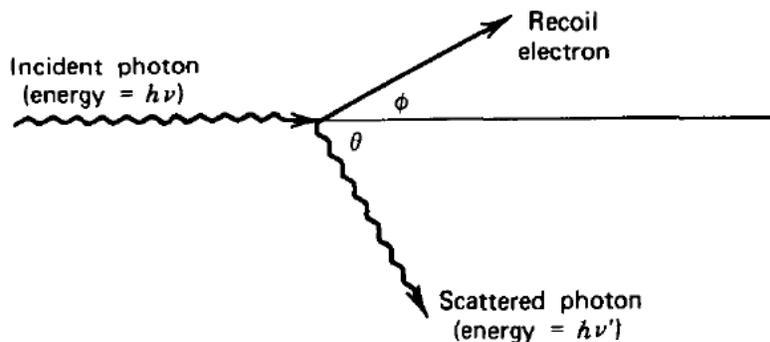


Figure 2.2. Diagram of Compton scatter; photon scatters off a free electron. $E = h\nu$ where h is Planck's constant and ν is the photon's frequency.[39]

Conservation of energy and momentum yield the kinematic relationship between the incident and scattered gamma energy E_γ and E'_γ given the scatter angle θ :

$$E'_\gamma = \frac{E_\gamma}{1 + (1 - \cos \theta) \frac{E_\gamma}{m_e c^2}} \quad (2.1)$$

The kinetic energy of the recoil electron $T_e = E_\gamma - E'_\gamma$:

$$T_e = \frac{(1 - \cos \theta) \frac{E_\gamma}{m_e c^2}}{1 + (1 - \cos \theta) \frac{E_\gamma}{m_e c^2}} E_\gamma \quad (2.2)$$

The photon never imparts all its energy the recoil electron. If $\theta = 180^\circ$, the photon has backscattered and $\cos \theta = -1$. This gives a scatter photon energy (E'_γ) at a minimum but greater than 0, and the kinetic energy of the recoil electron (T_e) is at a maximum, known as the Compton edge. As $\theta \rightarrow 0^\circ$, the photon transfers less energy and the electron recoil energy (T_e) goes to a minimum at zero. Compton scatter events with recoil electron energy between 0 and the Compton edge form the Compton continuum.[2, 39-41]

2.1.2 Neutron detection

Neutrons interact with nuclei via to two types of interactions: absorption and scatter. Absorption is generally more probable for slow neutrons while fast neutrons are more likely to scatter; this is not to say that the opposite is not possible, but it has a lower probability. Absorption occurs when a neutron is captured by the nucleus of an atom, then one or more particles are ejected. Common forms of absorption reactions include (n, p), (n, α), (n, 2n), (n, γ), and induced fission.[2, 39-41]

Scatter occurs when the incident neutron imparts some to all energy to a target nucleus, resulting in the neutron and the nucleus being scattered. Scatter interactions are either inelastic or elastic. Inelastic scatter occurs when the neutron is briefly captured by the nucleus distributing some of its energy to the nucleons before being scattered. This leaves the nucleus in an excited

state, which results in the release of the energy in the form of a gamma. The reaction for inelastic scatter is $(n, n'+\gamma)$. The elastic scatter interaction is a (n, n') reaction. For elastic scatter, the neutron and a nucleus interact like two hard spheres striking one another. The energy of the incident neutron and target nucleus change based on conservation of energy and momentum, and this change may be calculated in either the center-of-mass frame or lab frame. In the center-of-mass frame the target nucleus is in motion, while in the lab frame the nucleus is treated at rest as shown in Figure 2.3.[2, 39-41]

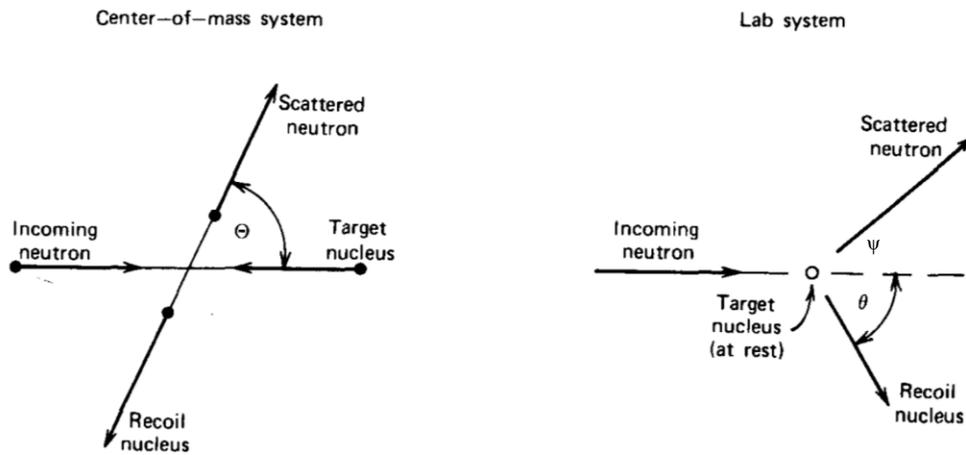


Figure 2.3. Diagrams of the center-of-mass and lab frame systems for elastic scattering of a neutron.[39]

Applying the conservation of energy and conservation of linear momentum for classical motion, the energy of the recoil nucleus (E'_N) in the center-of-mass frame is found to be

$$E'_N = \frac{2A}{(1+A)^2} (1 - \mu) E_n \quad (2.3)$$

where A is the mass number of the target nucleus, μ is $\cos \Theta$, where Θ is the scattered neutron angle in the center-of-mass frame, and E_n is the energy of the incident neutron. The center-of-mass frame may be transformed to the lab frame using the relationship

$$\cos \theta = \sqrt{\frac{1 - \cos \Theta}{2}} \quad (2.4)$$

where θ is the recoil nucleus angle in the lab frame. The conversion of Equation 2.3 using 2.4 shows the recoil nucleus energy in the lab frame to be

$$E'_N = \frac{4A}{(1+A)^2} (\cos^2 \theta) E_n \quad (2.5)$$

which may be used to find the maximum and minimum energy for the scattered nucleus by the recoil nucleus angle (θ). When $\theta = 90^\circ$ the energy of the recoil nucleus calculated using Equation 2.5 is zero; however, when $\theta = 0^\circ$ or 180° the energy transferred to the recoil nucleus is maximized. For hydrogen-1, the maximum energy for the recoil nucleus is 100% of the incident neutron energy, but for carbon-12 it is 28%. The probability distribution of the recoil nucleus energy and the probability distribution of the cosine of the scattered neutron angle are equal

$$P(E'_N)dE'_N = P(\mu)d\mu \quad (2.6)$$

where the isotropic probability distribution of the cosine of the scattered neutron angle is

$P(\mu)d\mu = \frac{d\mu}{2}$, $\mu \in [-1, 1]$. Equations 2.6 and 2.3 simplify to

$$P(E'_N) = \frac{A}{(1+A)^2} E_n \quad (2.7).$$

This is the probability distribution of the recoil nucleus across its entire energy range, from 0 to $\frac{4A}{(1+A)^2} E_n$. The average recoil nucleus energy is simply half of the maximum and expressed as

$$\overline{E'_N} = \frac{2A}{(1+A)^2} E_n \quad (2.8).$$

Organic scintillators are sensitive to fast neutrons primarily via proton recoil. Fast neutrons lose more energy from elastic scatter in materials with low atomic number than those with high atomic number. The fast neutrons elastic scatter off the hydrogen-1 and carbon-12 nuclei in organic scintillators with an average energy calculated using Equation 2.8. The average energy of the recoil nucleus is 50% of the incident neutron energy for hydrogen and 14% for carbon.

The recoil hydrogen-1 and carbon-12 recoil nuclei slow down through ionization and excitation of the other molecules in the scintillator, which in turn induces the emission of scintillation light.[2, 39-41]

2.2 Fluorescence

Fluorescence is a type of visible light (luminescence) observed in scintillation materials. Fluorescence is caused by charged particles, e.g. electrons, protons, and heavier ions exciting π -electrons from the ground state to an excited state, then de-excitation back to the ground state resulting in the release of photons (visible light). Recoil electrons from gamma interactions produce light that is essentially linear in intensity with energy disposition. The intensity of light production in response to protons from elastic scatter of neutrons is nonlinear with energy disposition. Compared to the recoil electron produced by gamma interaction, the recoil proton energy may also dissipate through lattice vibrations. The energy loss in lattice vibrations does not contribute to light production—this loss is called “quenching.”[39] Electrons and protons at the same energy do not produce the same amount of fluorescence; protons produce substantially less due to quenching. Gamma particle interactions can be identified through their short fluorescence. Relative to electrons, protons excite a greater density of π -electron triplet states. The triplet states decay more slowly than the singlet states, thus the scintillation time response is slower for the protons that result from neutron interactions. This phenomenon is delayed fluorescence.[2, 39-41]

The transfer of energy to populate the excited states occurs on the order of 0.2-0.4 ns. The subsequent decay of excited states to ground state occurs on the order of 2-3 ns. The categories of fluorescence are prompt and delayed; these categories are associated with the fall time of the π -electrons. The π -electrons have singlet and triplet states, see Figure 2.4. The energy

levels of the singlet states are labeled S_0, S_1, \dots . The energy levels of the triplet states are labeled T_1, T_2, \dots . Between the singlet and triplet states are vibrational states, which are labeled as S_{0x}, S_{1x}, \dots and T_{1x}, T_{2x}, \dots , respectively, where $x \geq 1$. The transitions between singlet states may occur in picoseconds to nanoseconds, while transition between triplet states are milliseconds. When π -electrons are excited to vibrational states, the de-excitation is energy loss through phonon production, not radiation release. Another type of luminescence is phosphorescence, which is slower and has longer wavelengths than fluorescence. As seen in Figure 2.4, phosphorescence only occurs when an excited singlet state (S_1, S_2, \dots) undergoes intersystem crossing to a triplet state (T_1, T_2, \dots), then de-excite to the ground singlet state (S_0). Phosphorescence is undesirable because it is long-lived and cannot be tied directly to any individual interaction.[2, 23]

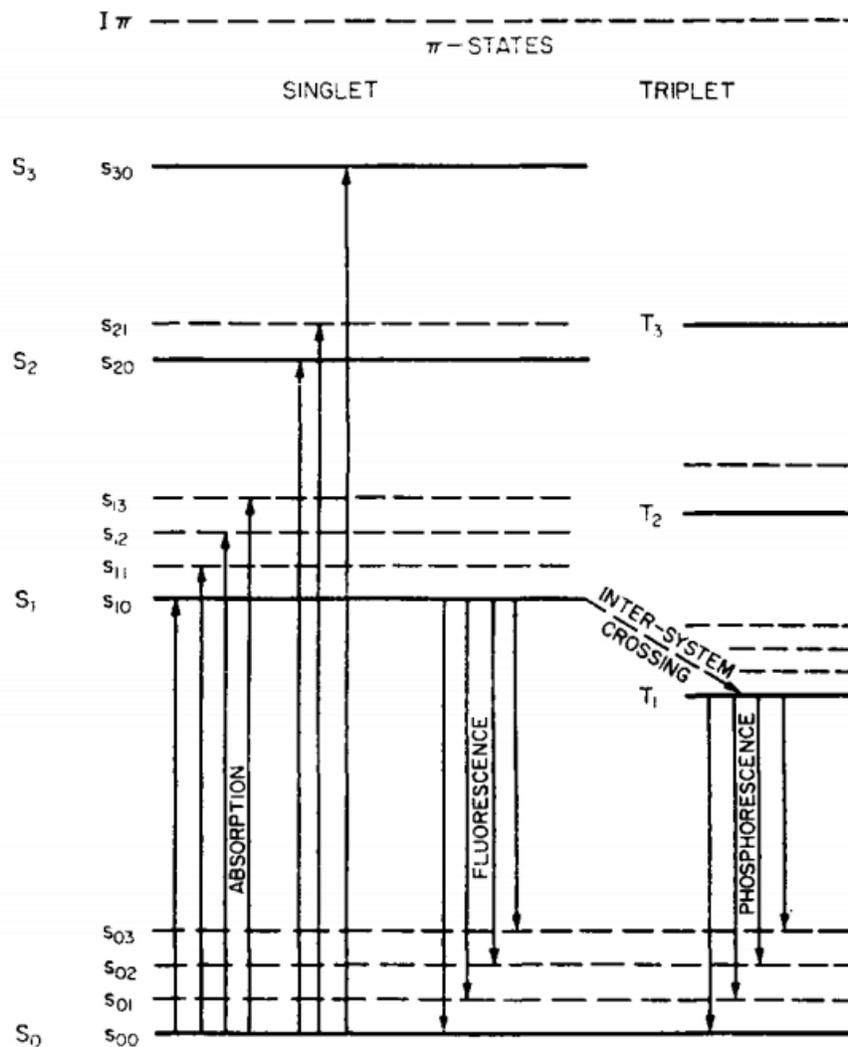


Figure 2.4. Singlet and triplet levels of the π -electron state.[2]

The luminescence produced in organic scintillators is in the form of visible-wavelength photons, which are converted to an electronic signal through a photomultiplier tube (PMT) or other photodetector. PMTs are composed of a photocathode to convert photons to electrons via the photoelectric effect. The electrons are directed through focusing electrodes to a series of dynodes, which re-emit multiple electrons for each incident electron. The electrons are collected on the anode as an electrical pulse.[2, 39-41]

2.2.1 Prompt fluorescence

Prompt fluorescence is the primary form of fluorescence. Prompt fluorescence is emitted when π -electrons de-excite from the excited state (S_{10}) to a vibrational level of the ground state (S_{0x}). If the π -electron falls to a vibrational level of the ground state, then it will subsequently vibrate down to the ground state (S_0) emitting heat. The π -electron is more likely to be excited to a higher energy level state above S_{10} , then the π -electron will quickly de-excite to the S_{10} excited state through internal conversion/vibration in the order of picoseconds. This transition is possible without the emission of fluorescence because the singlet states are closer together at higher energy states. The intensity of the light emitted through prompt fluorescence is expressed as

$$I = I_0 e^{-t/\tau} \quad (2.9)$$

where I_0 is the initial intensity, t is the time from the initial excitation, and τ is the fluorescence decay time in nanoseconds.[2, 23, 39]

2.2.2 Delayed fluorescence

The delayed fluorescence process is like phosphorescence because the π -electrons undergo intersystem crossing into the triplet states (T). The triplet states (T) have a long lifetime which leaves the π -electrons susceptible to thermal energy that may excite the π -electrons back to the lowest excited singlet state (S_{10}). From there the π -electrons undergo normal de-excitation through fluorescence. This results in the same emission spectrum as prompt fluorescence; however, it is characterized by a much longer emission time following excitation. Delayed fluorescence occurs for both gammas and neutrons but in greater quantities for the latter.[2, 23, 39]

2.3 Pulse Shape Discrimination

Pulse shape discrimination (PSD) is the process of discriminating different particle types through comparison of the prompt and delayed components of the anode pulse. The use of PSD is very common in neutron and gamma PID. The tail of a pulse is made up of two components: the fast component and the slow component. The fast component of the pulse is associated with prompt fluorescence. The slow component of the pulse is related to delayed fluorescence. As noted in Section 2.2, the recoil protons from neutron interactions are more likely to excite π -electrons to triplet states than the electrons produced from gamma interactions. The pulses of the neutrons and gammas display different time characteristics, the neutrons have slower fall times, as shown in Figure 2.5.[2, 18-19, 23]

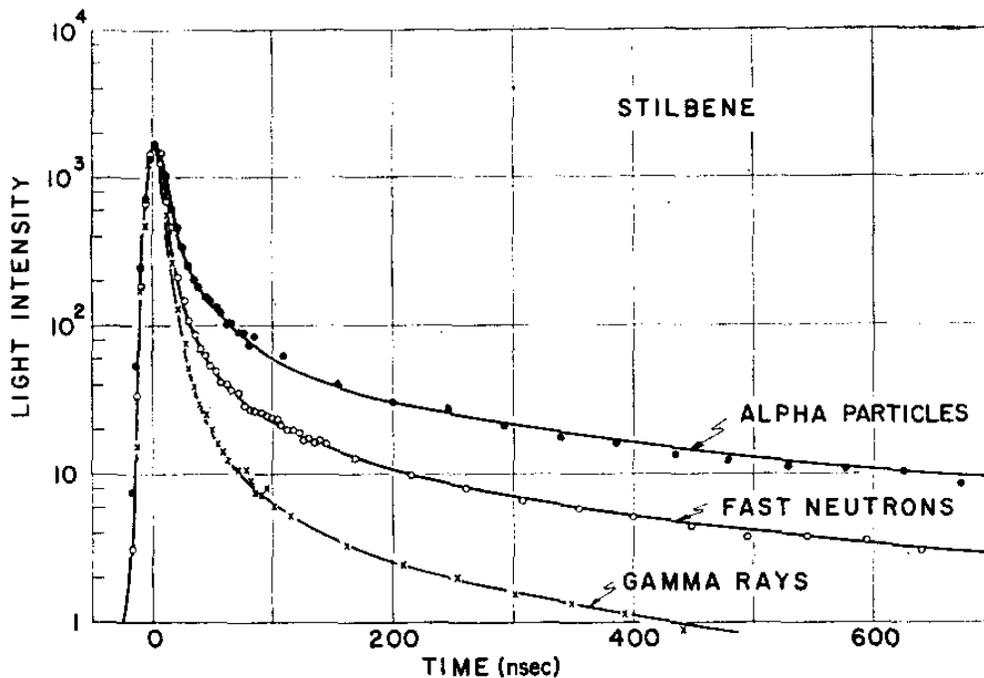


Figure 2.5. The time dependence of light intensity for gamma rays, fast neutrons, and alpha particles in a stilbene organic scintillator.[47]

Identification of the difference in the decay time that separates prompt and delayed fluorescence is exploited in PSD; it is illustrated in Figure 2.6. Prior research in the decay times of fluorescence materials paved the way for multiple methods of PSD, including the charge

integration method [33], optimum linear filter [21], pulse gradient analysis [22], zero-crossing method [29], and discrete wavelet transform [27]. The most common method is charge integration method, which is used in this research.

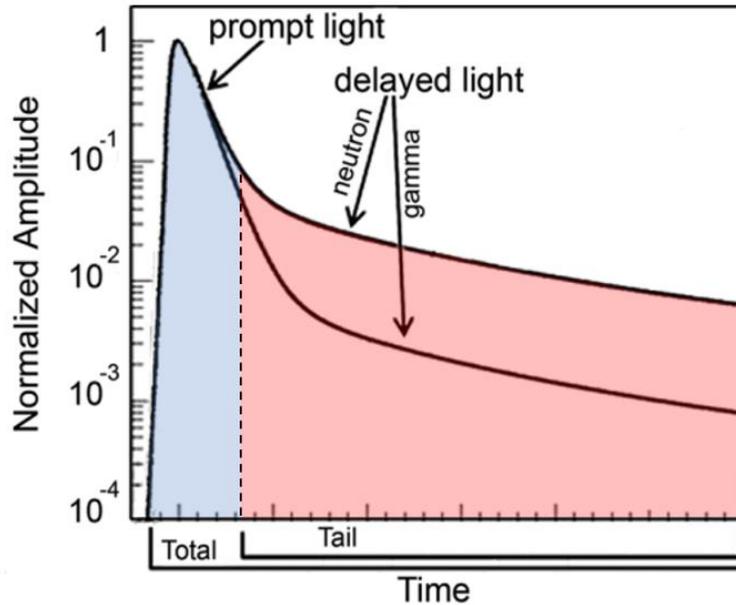


Figure 2.6. Example of the pulses detected from incident neutrons and gammas, with identification of the tail (red) and total area (blue + red).[48]

2.3.1 Charge integration method

The charge integration method is a PSD technique first introduced by Brooks in 1959.[19] I.A. Pawełczak et al. [33] and A.C. Kaplan et al. [34] demonstrated that pulses collected using an EJ-309 liquid organic scintillator processed through the charge integration method performs successful PSD. When radiation interacts with the detector medium and scintillation occurs, a pulse is formed from the charge collected on the anode. The charge of the pulse is proportional to the energy deposited by the incident particle. Application of the charge integration method integrates the pulse over two different areas and compares the integrals, displayed in Figure 2.6.

The first integral is taken over the entire duration of the pulse, known as the total charge or “long gate integral” (Q_L). The equation for the long gate integral is

$$Q_L = \int_{t_0}^{t_f} dt i(t) \quad (2.10)$$

where $i(t)$ is the anode current at time (t) in nanoseconds, t_0 is the start time of the pulse, and t_f is the end time of the pulse. The second integral is taken after the maximum pulse amplitude to the end of the pulse, known as the tail charge or “short gate integral” (Q_S). The short gate integral is

$$Q_S = \int_{t_i}^{t_f} dt i(t) \quad (2.11)$$

where t_i is the start time of the short gate, which is after the maximum pulse amplitude. The optimal selection of the start of the short gate provides the best PSD results. In this research, t_i was found by multiplying a fraction (f), called the “short gate fraction”, times the length of the long gate.

$$t_i = t_f - f(t_f - t_0) \quad (2.12)$$

The charge integration method compares the charge in the tail of the pulse to the charge in the total pulse, also known as the tail-to-total (TTT) ratio, charge ratio, or PSD ratio: [33]

$$PSD = \frac{Q_S}{Q_L} \quad (2.13)$$

The PSD ratio may be accumulated in a histogram as shown in Figure 2.7.[48] The delayed fluorescence properties that affect the tail of the pulse make differentiating pulses from different particle types possible. In the case of gamma and neutron radiation, neutrons exhibit higher tail charge than gammas at the same equivalent total charge.

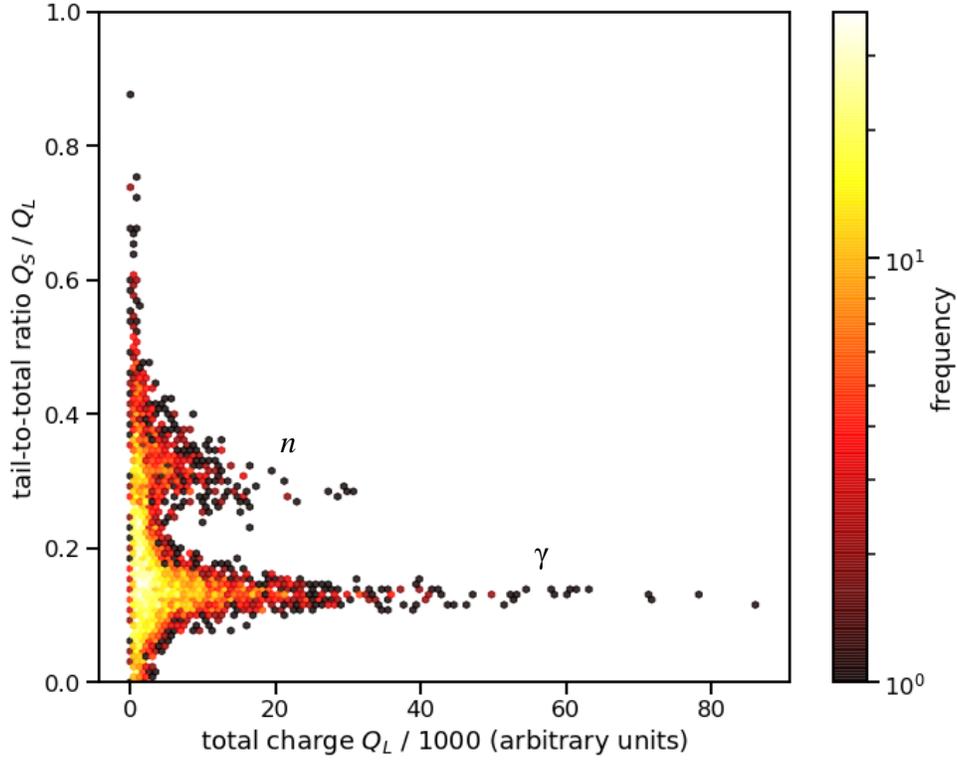


Figure 2.7. PSD using the charge integration method for a measurement of Cf-252, gamma (γ) and neutron (n), source with EJ-309.

The frequency distribution of $\frac{Q_S}{Q_L}$ exhibits two peaks, as seen in Figure 2.8. The peak at lower $\frac{Q_S}{Q_L}$ corresponds to gamma pulses, and the peak at higher $\frac{Q_S}{Q_L}$ corresponds to neutron pulses.

The PSD histogram is typically fit using two Gaussian distributions. The Gaussian function is

$$f(x) = a e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad (2.14)$$

where x is the $\frac{Q_S}{Q_L}$ ratio, a is the height of the frequency peak at the centroid, μ is the centroid of the peak, and σ is the standard deviation of the peak.[43] The double Gaussian model is

$$g(x) = a_\gamma e^{-\frac{(x-\mu_\gamma)^2}{2\sigma_\gamma^2}} + a_n e^{-\frac{(x-\mu_n)^2}{2\sigma_n^2}} \quad (2.15)$$

where the subscript “ γ ” denotes gamma pulses, and the subscript “ n ” denotes neutron pulses.[4]

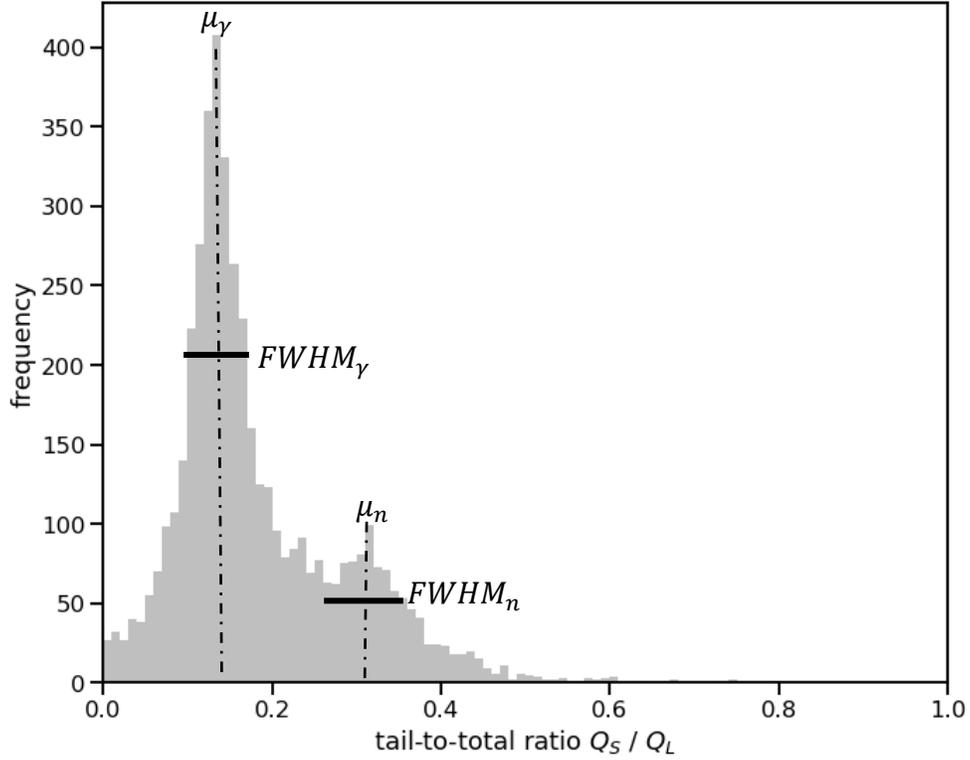


Figure 2.8. An example of the histogram of the $\frac{Q_S}{Q_L}$ PSD parameter.

2.3.2 Figure-of-merit (FOM)

The optimization of the start and stop times of the long gate and short gate integrals is important for effective PID using PSD. The figure-of-merit (FOM) is a measure of separation and determines the quality of charge integration PSD results. The FOM was first introduced by Winyard et al. [35] and is implemented across all PSD methods for PID. The higher the FOM, the better the PSD performance. The FOM is the absolute value of the difference between the gamma and neutron centroids for the PSD ratio divided by the sum of the full-width at half maximum (FWHM) of the gamma and neutron peaks, as illustrated in Figure 2.8. The FOM and FWHM are expressed as

$$FOM = \left| \frac{\mu_\gamma - \mu_n}{FWHM_\gamma + FWHM_n} \right| = \left| \frac{\mu_\gamma - \mu_n}{2.35 * (\sigma_\gamma + \sigma_n)} \right| \quad (2.16)$$

where μ_γ is the mean for the gamma peak, μ_n is the mean for the neutron peak, σ_γ is the standard deviation of the gamma peak, and σ_n is the standard deviation for the neutron peak.

$$FWHM = 2\sqrt{2 \ln(2)} \sigma = 2.35\sigma \quad (2.17)$$

One of the focal points of this research was creating an algorithm that optimizes the FOM based on the start time of the short gate integral relative to a fraction of the long gate integral.[35]

2.3.3 Time-of-flight (TOF) method

Time-of-flight (TOF) uses the travel time of the particles to identify the particle type. TOF is considered the “gold standard” technique for gamma and neutron (or other heavy particles) identification.[24, 36] The TOF method is often used as a comparison tool for different PSD methods. The results from the TOF method have been used in ML algorithms to calculate the true positive and false positive rates.[4-5, 12]

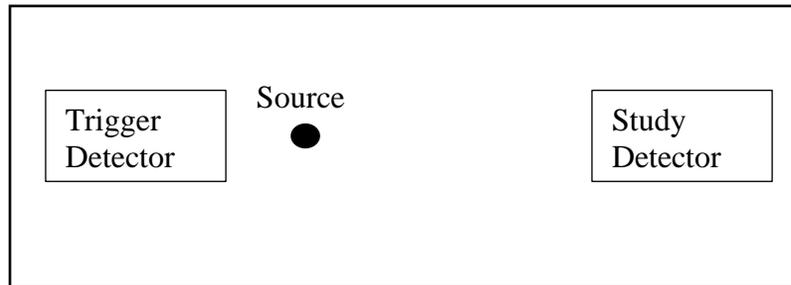


Figure 2.9. Diagram of TOF setup.

A simple experimental setup for the TOF method has a gamma and neutron source (frequently Cf-252), a trigger detector, a study detector, and data acquisition system, as Figure 2.9 illustrates. The fundamentals of the TOF method are based on the speed of the particles emitted by the source. The trigger detector selected may be sensitive to the detection of certain types of radiation (e.g., Cerium Bromide (CeBr₃) for gamma sensitivity). The study detector registers the pulses that are used for PSD. The two detectors have set spans of time, called gates, to collect the data. The trigger and study detectors collect detection events over the gates, and the

gates are compared to find pulses in coincidence. A pulse is considered coincident if a pulse from each detector occurs within the same gate. To calculate the flight time, the time of the maximum pulse amplitude of both trigger and study pulses are found and the difference is taken.

Analytically TOF is expressed as

$$t_{trig,max} = \frac{d}{dt} i_{trig}(t) \Big|_{t=t_{max}} = 0 \quad (2.18)$$

$$t_{study,max} = \frac{d}{dt} i_{study}(t) \Big|_{t=t_{max}} = 0 \quad (2.19)$$

$$TOF = t_{study,max} - t_{trig,max} \quad (2.20).$$

Gammas travel at the speed of light ($c \approx 2.998 \times 10^8$ m/s). The TOF is accumulated in a histogram that shows a narrow peak near 3.336 ns per meter of flight distance representing gammas and a subsequent, broader peak associated with neutrons, as illustrated in Figure

2.10.[24, 36]

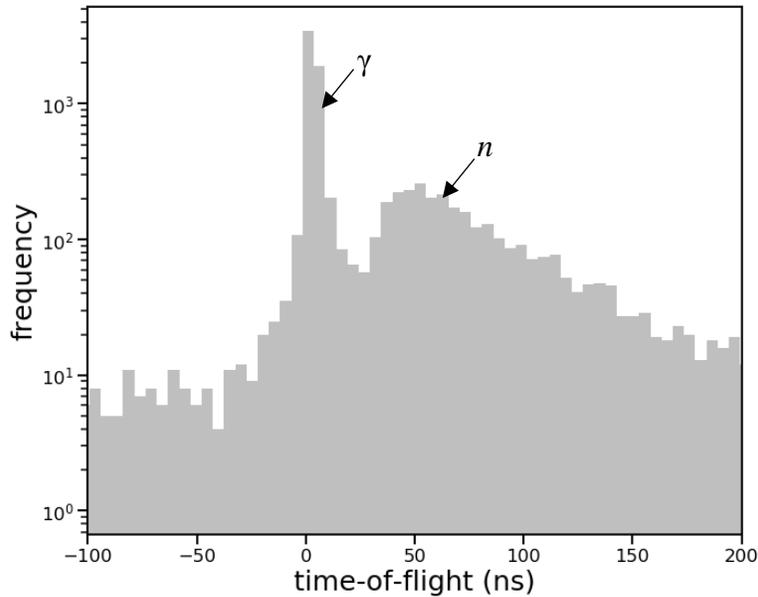


Figure 2.10. TOF histogram displaying the gamma (γ) and neutron (n) peaks from an experiment conducted at Lawrence Livermore National Laboratory (LLNL) with a Cf-252 source.[4]

2.4 Quantiles

A quantile is defined as the time required to accumulate a specified percentage of the total charge. The cumulative integrals over the samples of each record are displayed in Figure 2.11. The cumulative integrals were assigned as gammas and neutrons by using the mean values identified by the double Gaussian fit. The midpoint between the gamma and neutron peaks is used as the discrimination line to assign gamma and neutron labels to the quantiles. Using a discrimination line is rudimentary and does not accurately label each quantile with the ground truth, but the labels are used to calculate the quantile means and variances which are sufficient initial estimates for the constrained clustering ML algorithms.

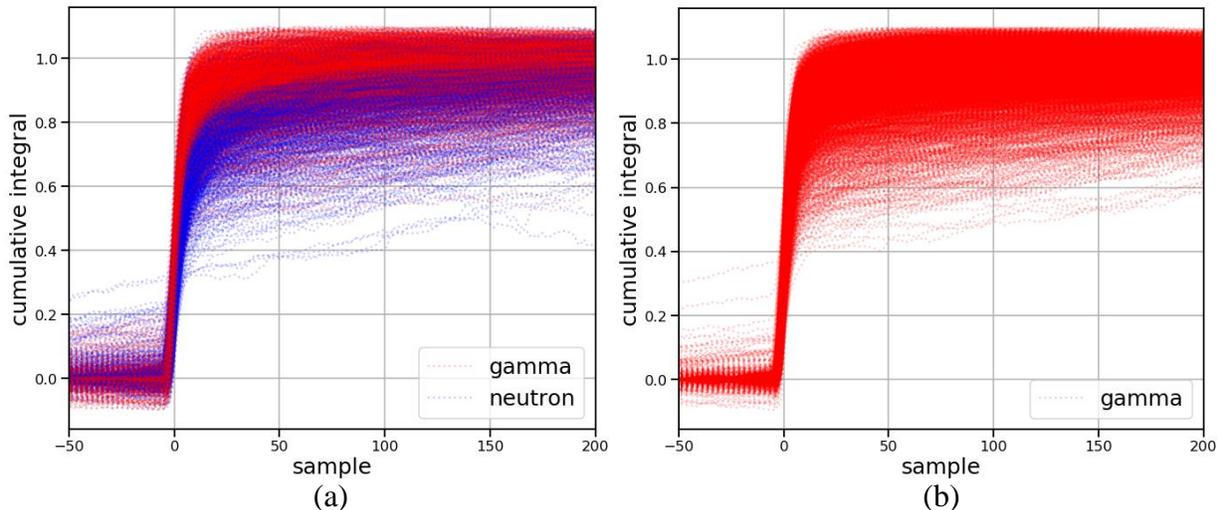


Figure 2.11. Plot of cumulative integrals for the waveforms from (a) Cf-252 and (b) Na-22.

The quantiles may vary in step size and range. For this research, the quantiles had a range from 10% to 90% with a step size of 10%. This results in nine quantiles per record. Each quantile designates the sample number at which the cumulative integral reached 10%, 20%, ..., and 90% of the total cumulative integral. The quantile means and variances were calculated by finding the average sample across all records at each quantile. The means and standard deviations of the quantiles for Cf-252 and Na-22 are presented in Figure 2.12.

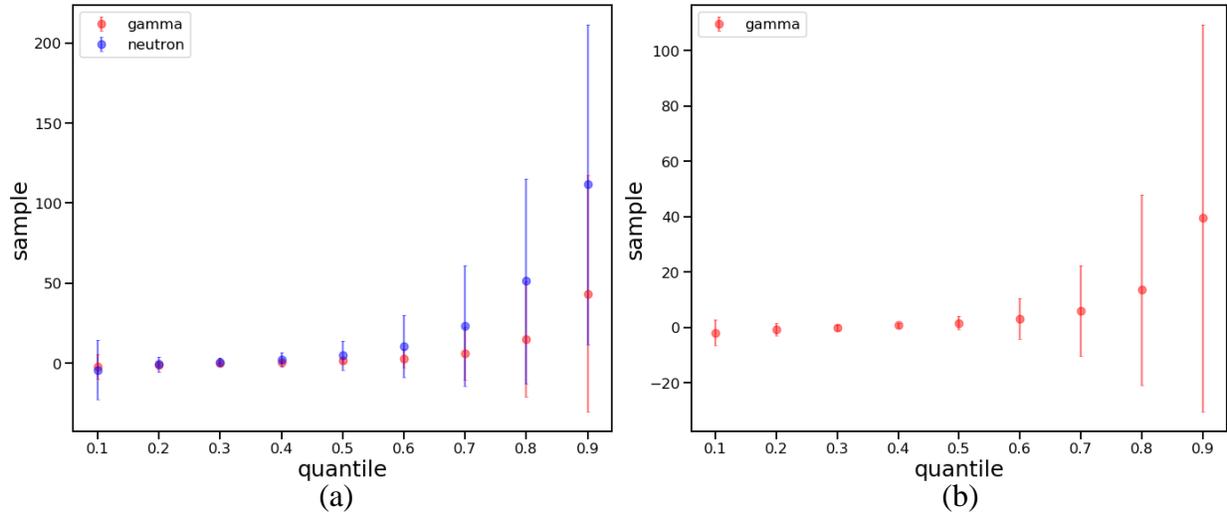


Figure 2.12. Plot of quantile means with standard deviations for (a) Cf-252 and (b) Na-22.

CHAPTER 3

Implementation of pulse shape discrimination (PSD)

3.1 Time-of-flight (TOF) experiment

A TOF experiment records coincident pulses; analysis of the arrival time of the pulses in the trigger and study detectors enable identification of particle type. TOF experiments are the “gold standard” for PID comparison and can simultaneously collect waveforms for PSD.[36] The TOF setup includes a source, a trigger detector, a study detector, and a digitizer. Two sources were used in the experiment: Na-22 and Cf-252. An Na-22 source emits a 1.274 MeV gamma and a 0.511 MeV gamma from the annihilation of the positron resulting from β^+ decay, which is the dominant form of decay for Na-22.[50] The Na-22 source was used to calibrate the detector time delay and provides the true gamma pulses for ML algorithm validation. Cf-252 is a spontaneous fission (SF) source of neutrons (3.77 per SF) and gammas (7.98 per SF). The probability of Cf-252 decaying by SF is approximately 3.09%, with the primary mode of decay being alpha decay (96.91%). The SF neutrons have an average energy of 2.3 MeV and the gammas have a range of 0.2-1.8 MeV.[49] The trigger detector was a CeBr₃ inorganic scintillator. CeBr₃ scintillators have high energy resolution and high intrinsic efficiency for high energy gammas.[60] The sensitivity of CeBr₃ makes it an ideal trigger detector for TOF experiments with sources that emit gammas. The study detector was an EJ-309 liquid organic scintillator. EJ-309 scintillators are sensitive to neutrons and gammas and possess excellent PSD properties.[33] The digitizer was a CAEN DT5730, 8 channel, 14-bit, 500 megasamples per second (MS/s) waveform digitizer. The CAEN DT 5730 digitizer is capable of digital waveform recording and collects waveforms from both the trigger and study detectors using a leading-edge trigger with 520 samples per record at 2 ns per sample.[61]

The setup for the TOF experiment used to collect the data for the validation of PID is shown in Figure 3.1. The pulses collected from the trigger detector and study detector are compared for coincidence, for both Na-22 and Cf-252. This process is discussed in greater detail later in this section.

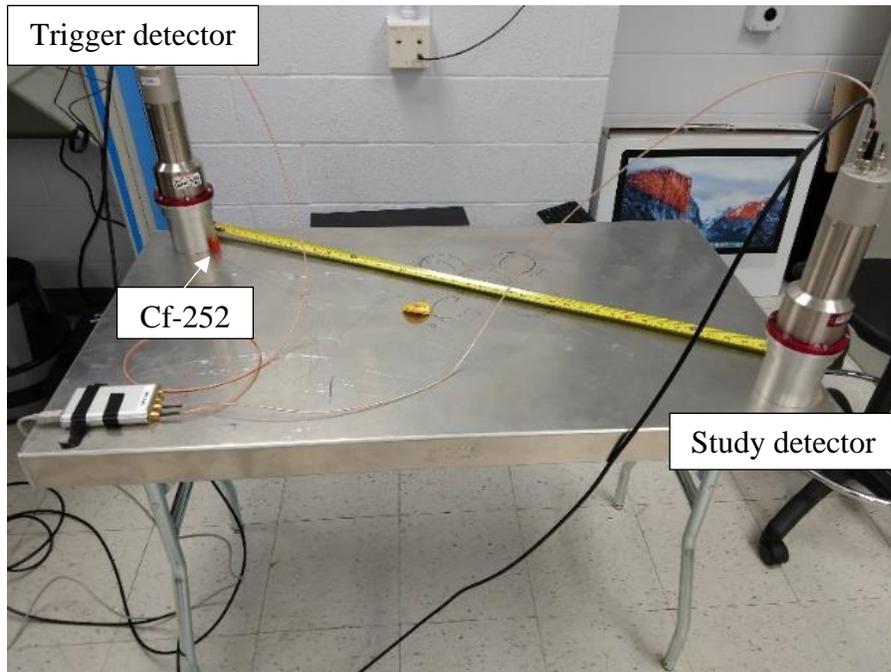


Figure 3.1. TOF experimental setup. Cf-252 is placed next to the trigger detector (left) 1 m from the study detector (right).

From the start of the record a waveform spans 1,040 ns. A record is triggered when the voltage exceeds a threshold, or lower-level discriminator (LLD). The method applied to identify the threshold used the average (μ_c) and standard deviation (σ_c) of the sample voltage for each recorded waveform from 0 to 40 ns, or samples 0 to 20. Samples 0 to 20 represent the pre-trigger region of a waveform, such that no pulse should be present in those samples. A pulse was determined to be present if the minimum voltage of the waveform was less than the average minus six times the standard deviation ($6\sigma_c$). Six sigma statistically represents 99.999% of the average voltage found in the pre-trigger region of the waveform. If a pulse exists in the same

record number from both the trigger and study detectors, then those pulses are in coincidence, as seen in Figure 3.2.

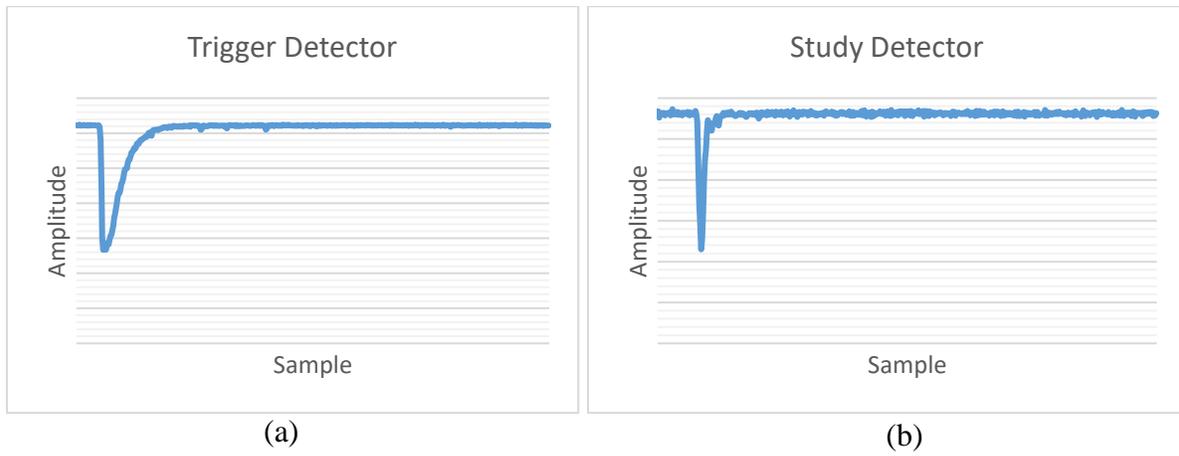


Figure 3.2. Example waveforms from the same record [(a) trigger detector (CeBr₃), (b) study detector (EJ-309)] with pulses in coincidence.

The TOF experiment originally had 253,792 records for Na-22 and 46,347 records for Cf-252 in coincidence. The records with coincident pulses were further processed through several steps:

1. Remove waveform baseline.
2. Invert waveforms.
3. Remove records with a pulse in the pre-trigger region of the waveform.
4. Remove records where the pulse occurs late in the waveform.
5. Remove records with saturated pulses.
6. Remove records with multiple pulses in the waveform.

A pulse with the features labeled is shown in Figure 3.3. After these steps, the remaining coincident pulses are plotted in a histogram of the difference in arrival time, a.k.a., the TOF. A single peak appears for the gammas detected for Na-22, and two peaks appear for Cf-252, one for gammas and the other for neutrons.

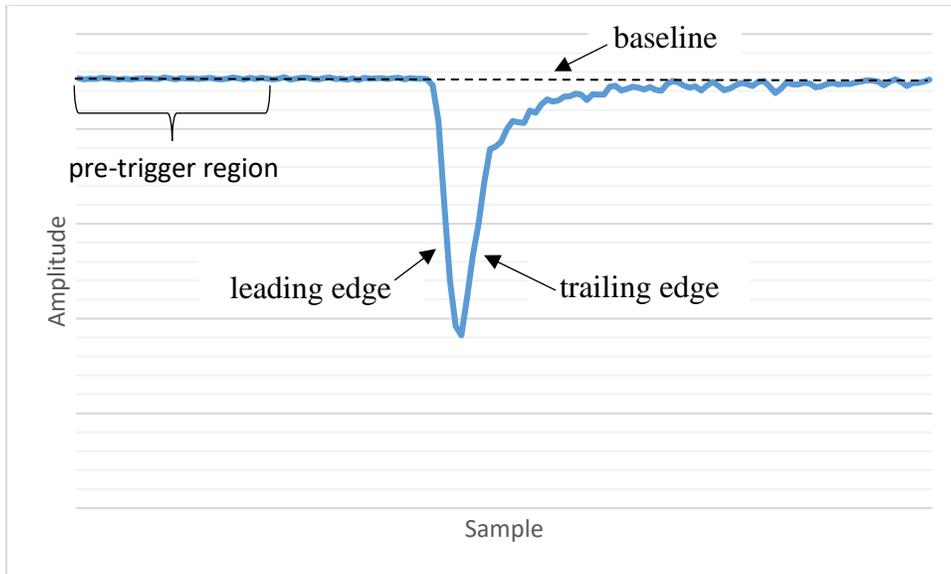


Figure 3.3. Example waveform with features labeled.

3.1.1 Record processing

The original waveforms displayed in Figure 3.4 shows the first 20 records without processing or alignment of the waveforms.

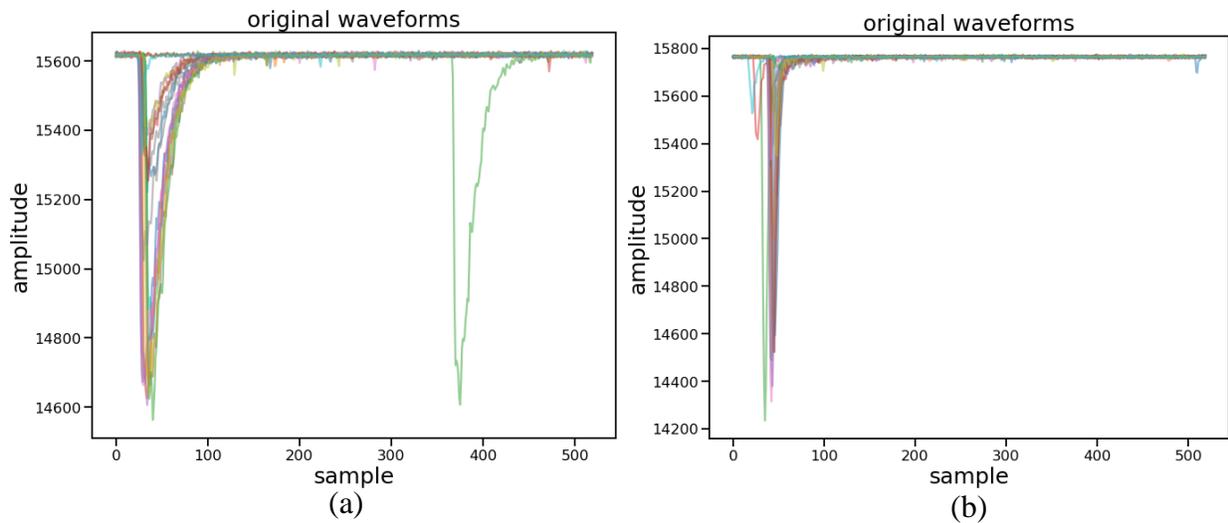


Figure 3.4. Original waveforms from records with varying amplitude and occurrence [(a) trigger detector (CeBr_3), (b) study detector (EJ-309)].

Baseline removal is the first step of processing. The baseline is found by calculating the average of the first 20 samples and subtracting that average from all sample values across the record.

Failure to remove the baseline of the records results in incorrect integral values. The pulses with baselines removed are presented in Figure 3.5.

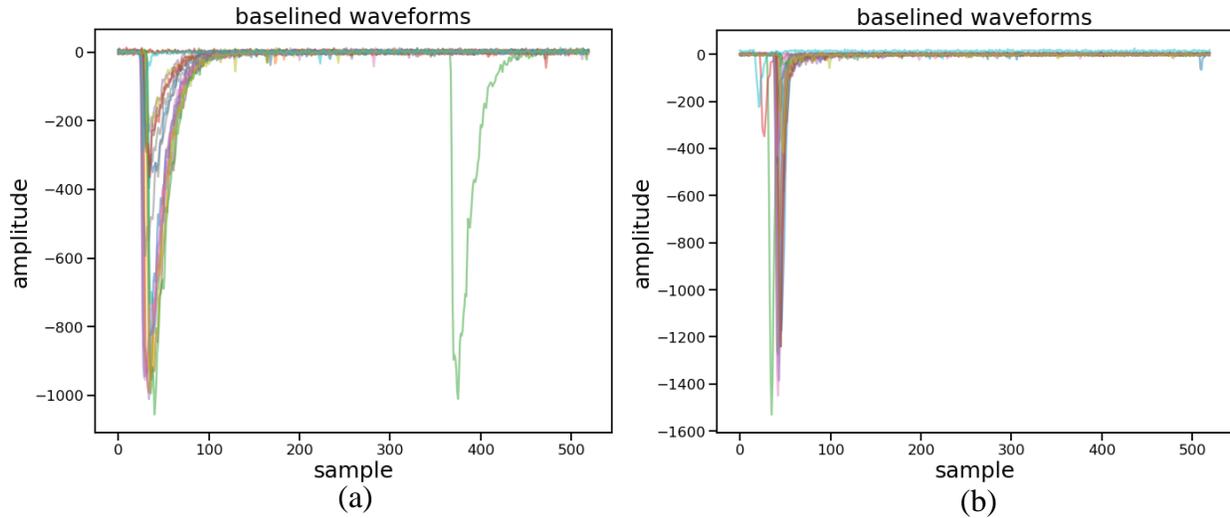


Figure 3.5. Baselined waveforms from records with varying amplitude and occurrence [(a) trigger detector (CeBr_3), (b) study detector (EJ-309)].

For the second step of record processing, the waveforms are inverted to make all values positive.

The inverted waveforms are displayed in Figure 3.6.

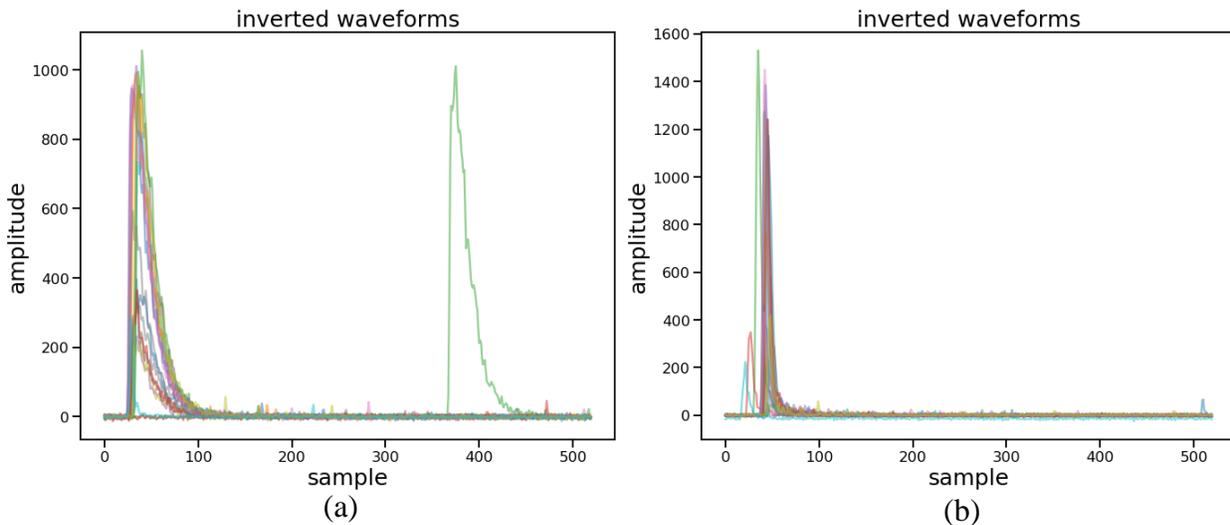


Figure 3.6. Inverted waveforms from records with varying amplitude and occurrence [(a) trigger detector (CeBr_3), (b) study detector (EJ-309)].

The third step is performed on the trigger detector records, then the associated study detector records are removed as well. The records removed are those with a pulse or tail of a pulse in the

pre-trigger region of the waveform. A pulse in the pre-trigger region is identified by taking the standard deviation of samples 0-20 of the waveform. The standard deviation for waveforms is low if the pre-trigger region does not contain a pulse. The threshold for the standard deviation was identified by histogramming the standard deviation of the baseline for all records. Figure 3.7 shows a peak at a standard deviation of approximately three. The threshold identified for the trigger detector was all standard deviation greater than one and less than five. The records with standard deviations outside of this range were removed.

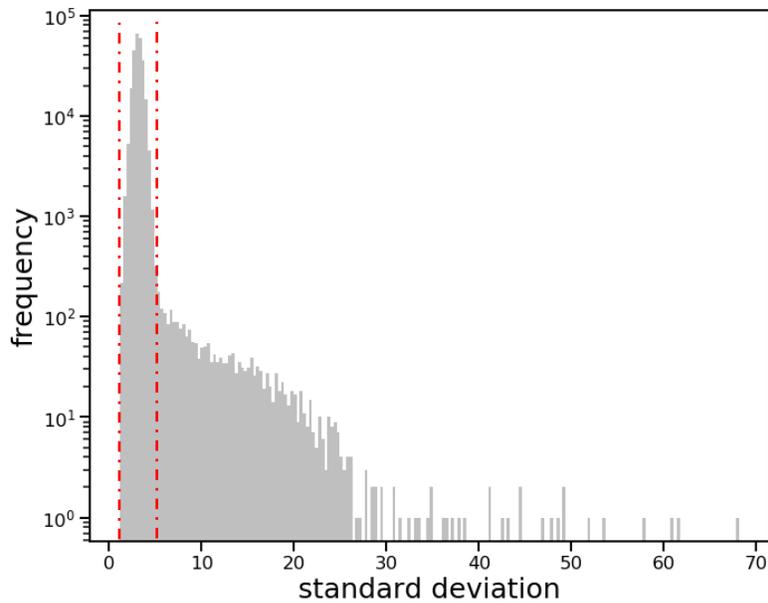


Figure 3.7. Histogram of standard deviation (σ). The boundaries for $1 < \sigma < 5$ are represented by the red dashed lines.

The fourth step of processing is the identification of records with late pulses. The pulse in the waveform of each record is expected to arrive within a set amount of time in the trigger detector. The previous step removed those records with pulses that arrived before 40 ns (or in the pre-trigger region). This step removes records with pulses arriving after 80 ns in the trigger detector. To identify the thresholds for the maximum amplitude location of the pulses, the sample value location (l) of the maximum amplitude in a waveform is found and recorded. Like the previous step, the maximum sample locations are histogrammed. The histogram was used to

identify the threshold for the maximum sample location. The records with maximum sample location outside of the thresholds in Figure 3.8 were removed.

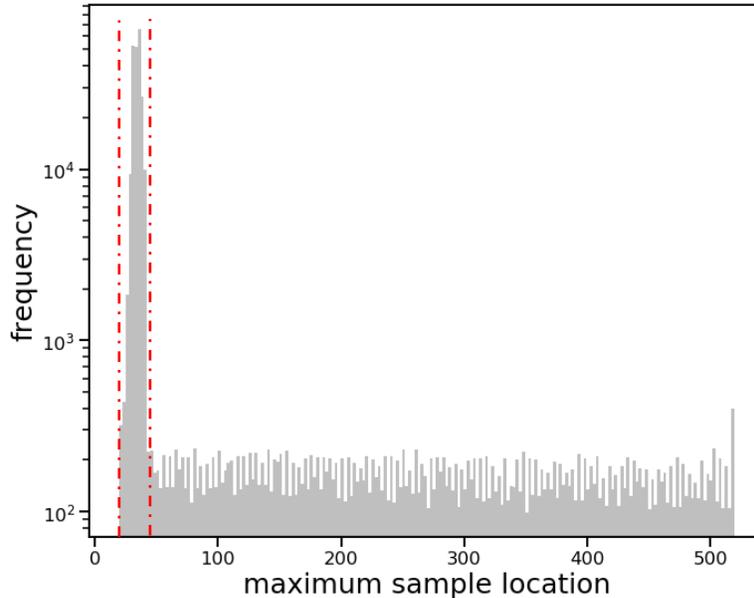


Figure 3.8. Histogram of locations of maximum amplitude. The boundaries for $20 < l < 40$ are represented by the red dashed lines.

The fifth and sixth steps are performed on both the trigger detector and study detector records. The fifth step identifies the records with waveforms displaying flat-top (i.e., saturated) pulses. The effective pulse width is the ratio of the area under the waveform to the maximum pulse amplitude. The area was found by integrating the waveform using the trapezoidal rule.[51] The effective pulse width and the maximum amplitude plots are displayed in Figure 3.9 (a) and (b). The flat-top pulses are easily identified as outliers along the maximum amplitude axis. The maximum amplitude threshold was set below the value of the highest maximum amplitudes to remove the outliers. Using a Gaussian fit, the concentration of the effective pulse width was used to identify a mean value of the distribution and a standard deviation. The range of effective pulse width used for the CeBr_3 trigger detector was 17 to 33, and 1 to 13 for the EJ-309 study detector. The reduced effective pulse width to maximum amplitude plot is displayed in Figure 3.9 (c) and (d).

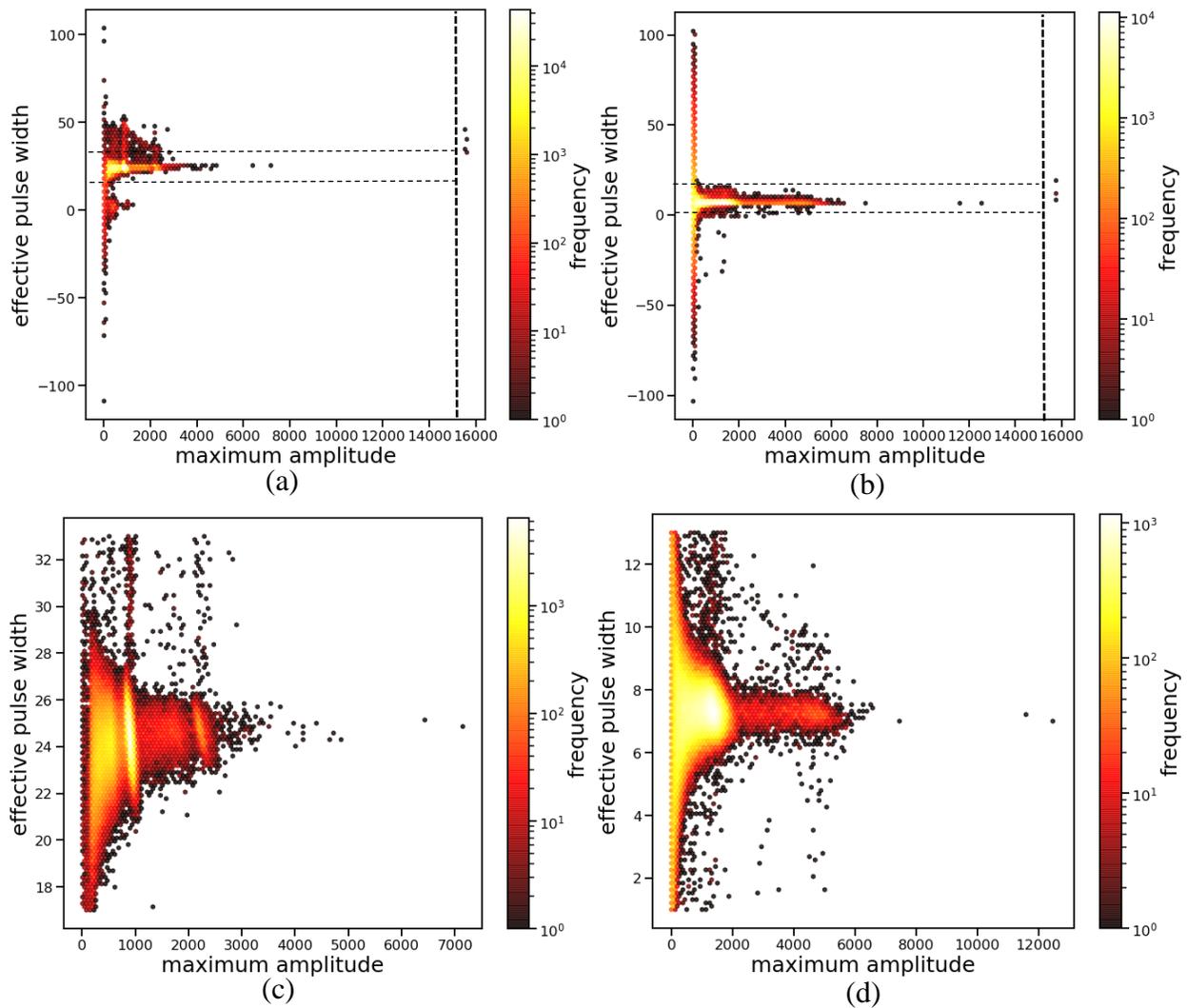


Figure 3.9. 2-D histograms of the effective pulse width vs. maximum amplitude. The trigger detector (CeBr₃) before step five of the processing (a) and after (c). The study detector (EJ-309) before step five processing (b) and after (d).

Although using the effective pulse width removes some records with additional pulses in the waveform, several waveforms with multiple pulses remained. Therefore, another evaluation of the number of pulses in each record was performed by visually identifying the waveforms with more than one pulse. The results of removing records with more than one pulse in the waveform are displayed in Figure 3.10. After the completion of record processing, the number of records was 142,502 for Na-22 and 5,774 for Cf-252.

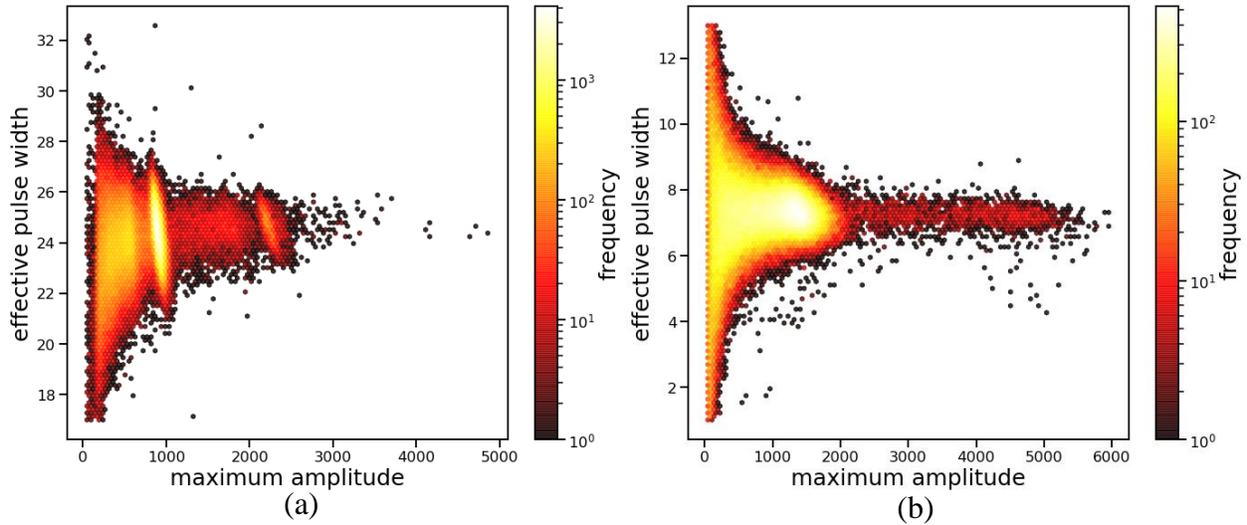


Figure 3.10. 2-D histograms of the effective pulse width vs. maximum amplitude after removing records with waveforms with multiple pulses [(a) trigger detector (CeBr_3), (b) study detector (EJ-309)].

3.1.2 TOF results

Only those records that pass each processing step were evaluated for TOF. The TOF calculation was performed as discussed in Section 2.3.4, the results of the TOF for Na-22 are presented in Figure 3.11. The Na-22 source was placed in the center between the trigger and study detectors, such that the maximum frequency peak of TOF is expected to have a centroid at zero. This is because Na-22 is a gamma-only source so coincident pulses should occur simultaneously in the trigger and study detectors. If the peak's centroid is not at zero it is a result of time delay between the CeBr_3 and EJ-309 photomultiplier tubes (PMTs). The time delay is a result of the charge collection time differences between the two detector systems and is due to the electron transit time in the PMTs. The time delay (τ) is found by identifying the maximum frequency in the TOF histogram, which occurs at 23.0 ns as seen in Figure 3.11. The time delay is then subtracted from the TOF originally calculated. The adjusted TOF results for Na-22 were saved with the processed records of the waveforms of the EJ-309 detector for future use.

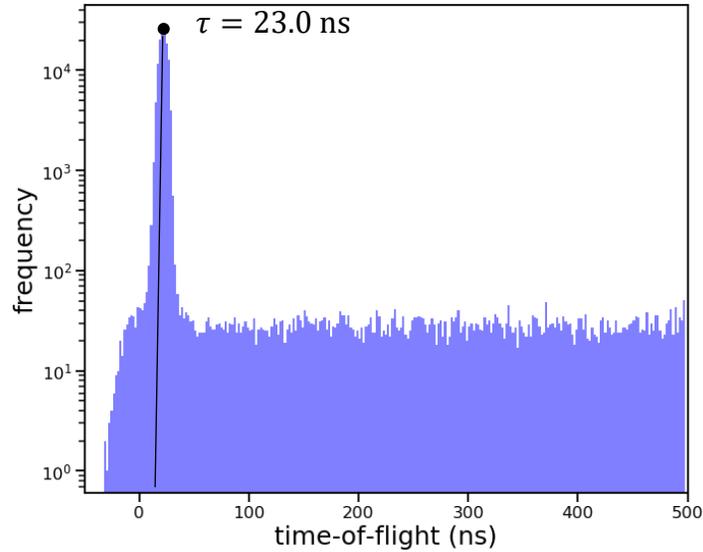


Figure 3.11. Histogram of the initial TOF a gamma only Na-22 source used to find the time delay.

The processing and TOF for the Cf-252 records were done similarly to the Na-22 records. The difference between Cf-252 and Na-22 are in the setup and TOF calculation. The Cf-252 source was placed next to the trigger detector, as illustrated in Figure 3.1. A separate time delay is not calculated for the Cf-252 source but the TOF calculation uses the time delay found using the Na-22 source. The results of the TOF are presented in Figure 3.12. The final TOF results are stored with the waveforms of the EJ-309 study detector records for comparison against the ML algorithms' performance.

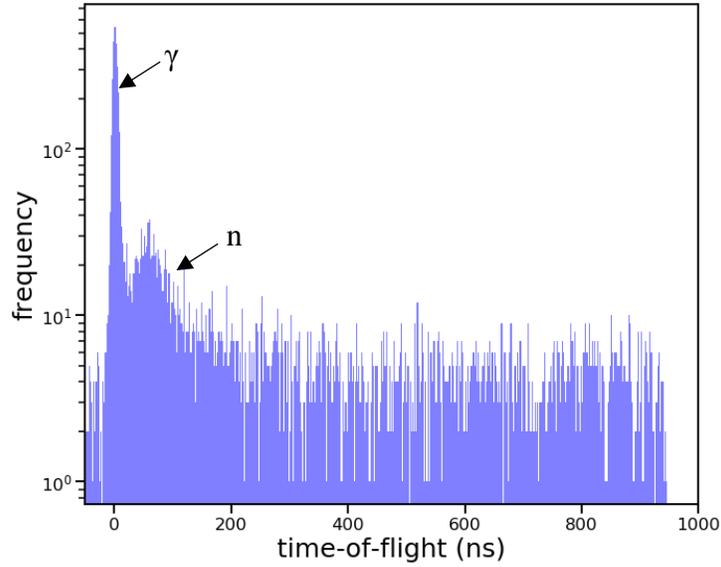


Figure 3.12. Histogram of the TOF for Cf-252.

3.2 PSD calculation

The application of PSD for the Cf-252 source involves additional steps. The records previously processed from the study detector are shown in Figure 3.13. Although PSD calculation is possible for the waveforms as they are presented, the repetitive model for PSD calculation is faster and has fewer issues with optimization when the pulses are aligned in time.

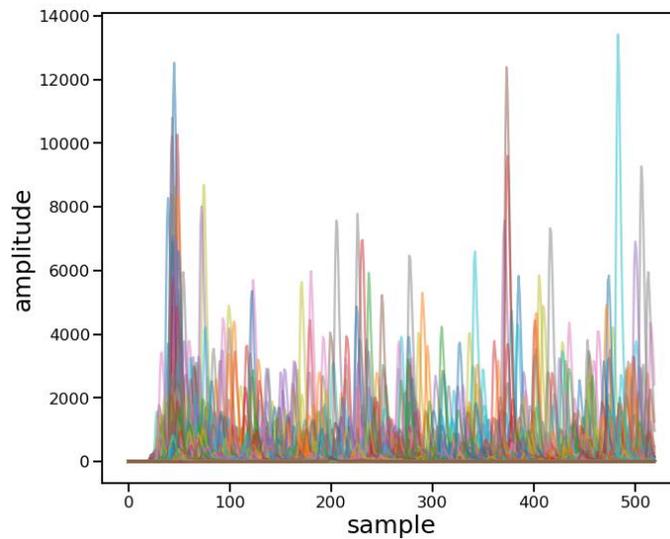


Figure 3.13. Waveforms from Cf-252 records.

3.2.1 Pulse alignment

The waveforms are shifted along the horizontal axis, such that the maximum amplitudes of the waveform occur at the sample equal to zero, as illustrated in Figure 3.14.

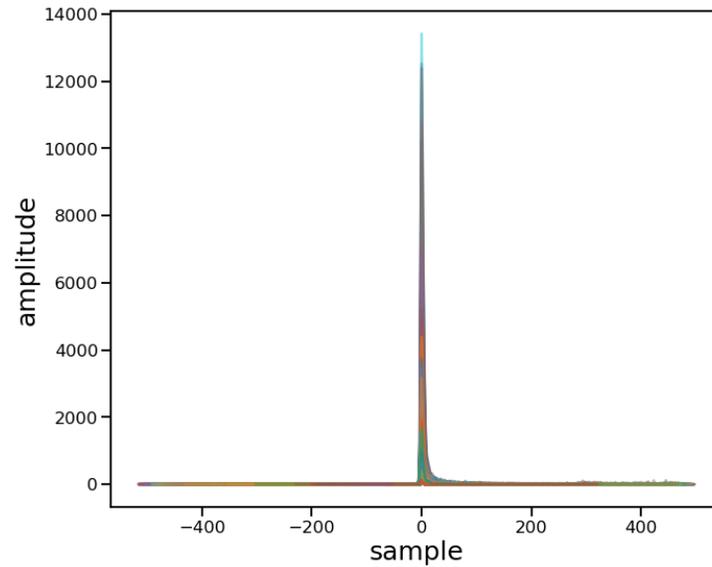


Figure 3.14. Waveforms with pulses aligned at maximum amplitudes.

The shift to align the amplitudes of each waveform does not address the issue of the length of the waveform following the pulse. To allow the model to integrate the waveforms effectively along the same range of samples, all records with waveforms that had less than 200 samples after the pulse were discarded. The remaining aligned pulses are displayed in Figure 3.15. This leaves 4,868 of the 5,774 records left for PSD calculation and optimization.

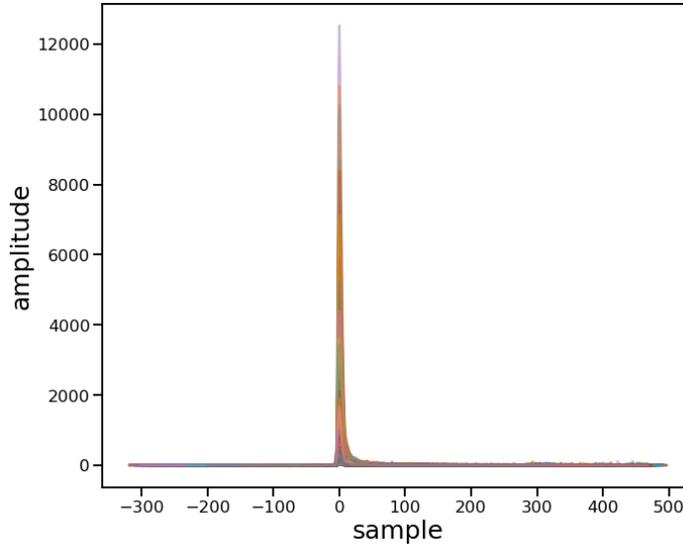


Figure 3.15. Reduced waveforms with pulses aligned at maximum amplitudes.

3.2.2 Charge integration

The charge integration method explained in Section 2.3.1 was applied to the remaining study detector records. To find the integral of the waveform, the trapezoidal rule and Gaussian quadrature were applied in this research. This is not to say other methods cannot be used, Nakhostin compared three different integration algorithms in research comparing the zero-crossing and charge integration methods.[26] Although the trapezoidal rule is effective for quick integration and optimization functions, it does not provide the most accurate integration for the waveforms. The Gaussian quadrature method provides the best results because it identifies the appropriate number of samples to integrate over the function that fits the waveform.[52]

Special attention was given to the selection of the start and stop times for short gate (Q_S) and long gate (Q_L) integrals. When applying the charge integration method, the stop times for both Q_S and Q_L are the same; they are typically at the end of the data set. To identify the optimal stop time a visual inspection of the waveform was done. The stop time selected was based on the sample number where all waveforms return to baseline. The start time for the Q_L was identified as the sample number where the rise of the pulse in the waveform occurs. The start time for the

Q_S can be varied and, for the model developed, was designed to be a fraction (short gate fraction) of the long gate pulse length. For the initial calculation of the Q_S , the short gate fraction of the long gate pulse length evaluated was 90%. The PSD plot may be viewed in two different ways. The first is presented in Figure 3.16.

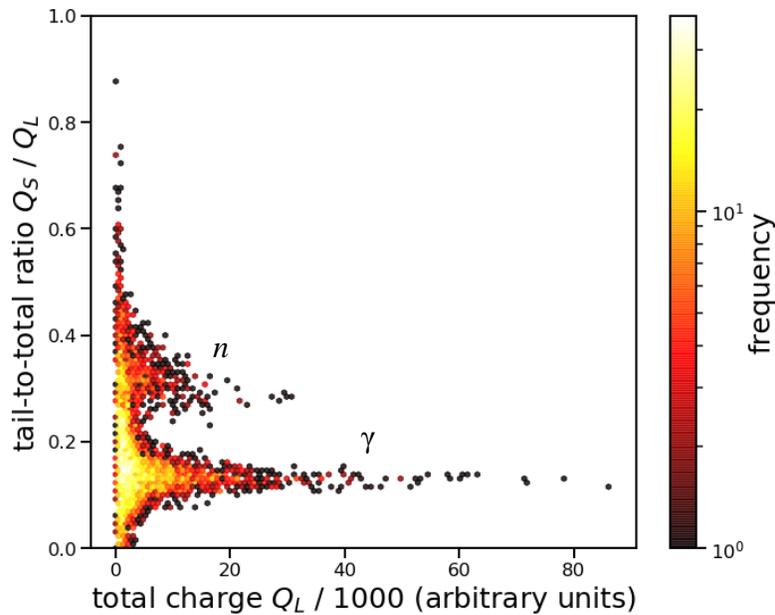


Figure 3.16. 2-D histogram showing PSD for a short gate fraction of 90%. Gammas (γ) are represented by the lower horizontal peak and neutrons (n) in the upper horizontal peak.

The second PSD plot is a histogram of the frequency of each $\frac{Q_S}{Q_L}$. The histogram plot exhibits easily distinguishable peaks for gammas and neutrons, as seen in Figure 3.17.

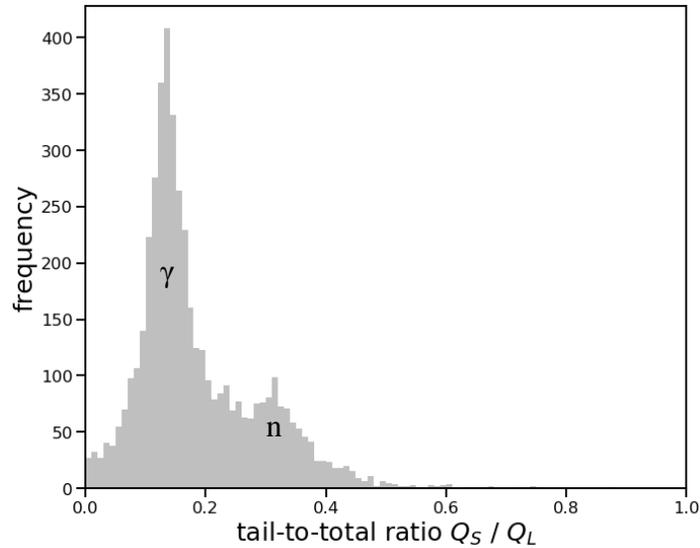


Figure 3.17. The PSD histogram exhibiting a peak from gamma pulses (γ) and another peak from neutron pulses (n).

3.2.3 Gaussian fits

The PSD histogram can be fit by the sum of two Gaussian functions. The fit for the double Gaussian function provides the heights of the gamma and neutron centroids of the frequency peaks, the $\frac{Q_S}{Q_L}$ value for the gamma and neutron centroids, and the standard deviations of each Gaussian. The double Gaussian curve fit to the PSD histogram plot is shown in Figure 3.18.

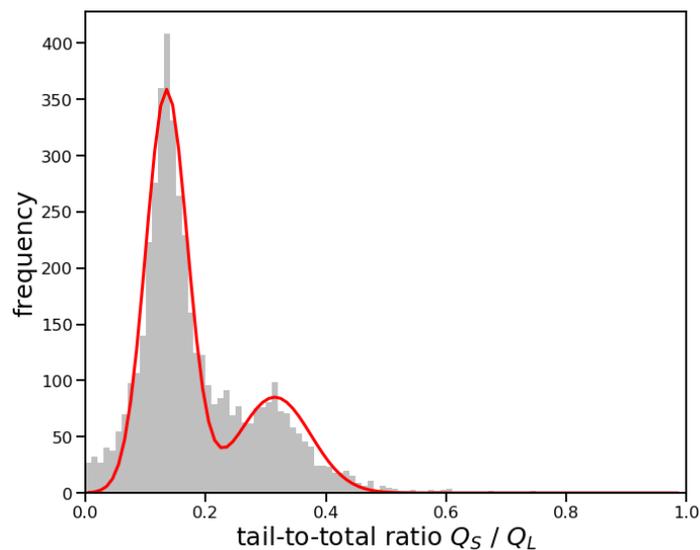


Figure 3.18. PSD histogram with double Gaussian curve fit.

3.2.4 Figure-of-merit (FOM) calculations

The FOM was used to identify the optimal short gate fraction. The FOM is:

$$FOM = \left| \frac{\mu_\gamma - \mu_n}{FWHM_\gamma + FWHM_n} \right| = \left| \frac{\mu_\gamma - \mu_n}{2.35 * (\sigma_\gamma + \sigma_n)} \right| \quad (3.1).$$

The parameters for the FOM are found using the results of the double Gaussian fit function. The gamma and neutron centroid $\frac{Q_S}{Q_L}$ values and standard deviations are the inputs for the means and standard deviations in the FOM equation. The optimal PSD is found at the maximum FOM. A FOM comparison was done manually by varying the short gate fraction to find the optimum is shown in Figure 3.19. The error associated with the FOM was calculated by applying propagation of uncertainty to the FOM equation. The optimal short gate fraction occurs at 91% with a FOM of approximately 0.8167 ± 0.0591 .

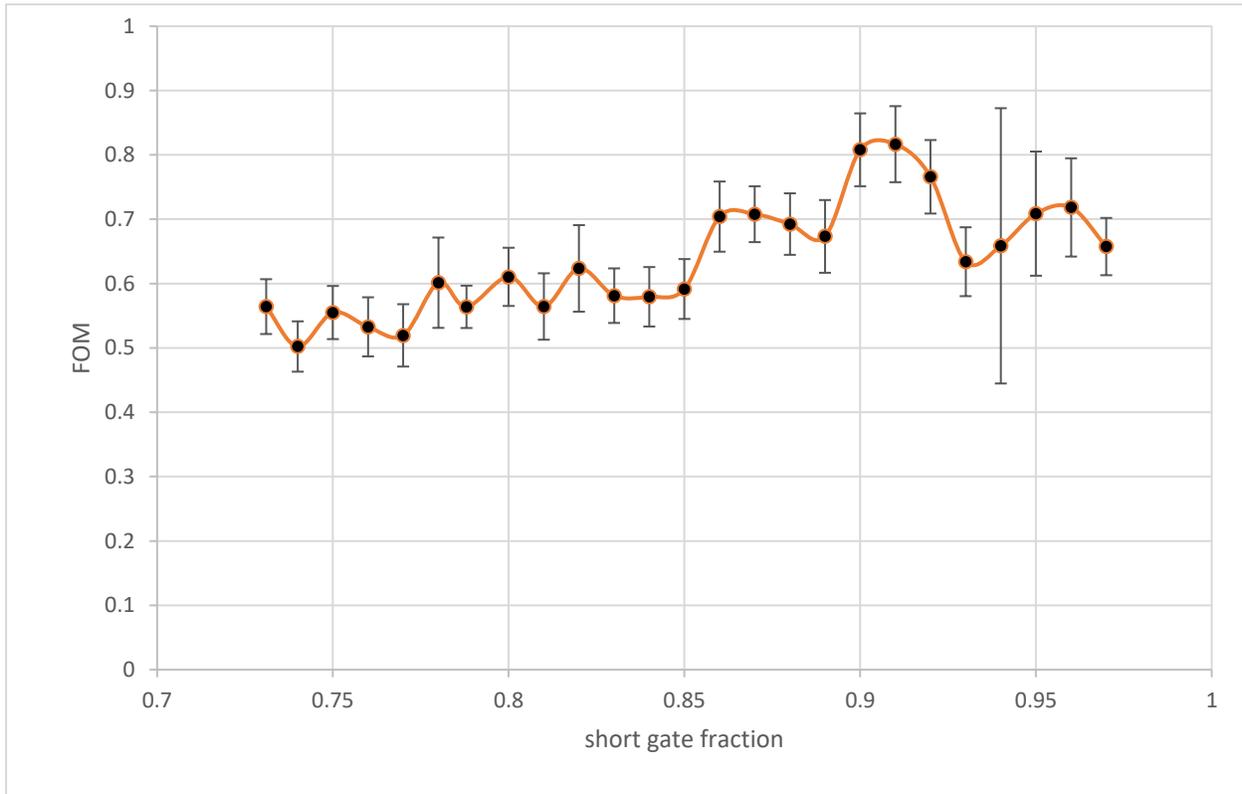


Figure 3.19. Comparison of the FOM by short gate fraction.

3.2.5 Optimization

Several optimization functions for minimization were analyzed to find the best method for maximizing FOM. As seen in Figure 3.19, the FOM has a defined global maximum and several local maxima at low values of short gate fraction. The optimization methods seek to find the minimum of the negative FOM. The optimization methods tested in this research include the Nelder-Mead, sequential least squares programming (SLSQP), limited-memory Broyden, Fletcher, Goldfarb, and Shanno with bounds (L-BFGS-B), brute force, and Brent's methods. The application of Brent's method provided the best results for optimizing the FOM.[53]

Nelder-Mead is an unconstrained method that is slower and meant for local optimization, so with bounds the method may not find the global minimum if other minima exist. Due to the upper and lower bounds of the short gate fraction, if the initial estimate for the short gate fraction is not close to the global minimum, then misidentification of the optimal FOM may occur. If the short gate fraction initial value is set below the optimum, then a different local minimum is found as the optimal FOM.[53-54]

L-BFGS-B is a bound-constrained quasi-Newton method. Like Nelder-Mead, quasi-Newton methods are good for identifying local minima. The L-BFGS-B method was selected because it uses bounds to constrain the algorithm. The L-BFGS-B method is designed to optimize over smooth functions with constraints. The BFGS algorithm finds the direction of descent using the gradient of the function. The issue with the use of L-BFGS-B method was the FOM vs. short gate fraction is not smooth; therefore, the L-BFGS-B method does not find the global minimum without limiting the bounds of the function to only include the global minimum.[53-56]

SLSQP is like the L-BFGS-B and may be applied to nonlinear functions. SLSQP is a constrained quasi-Newton method that optimizes using BFGS over the sequential least squares approximations of the function. The SLSQP method has some of the same issues as L-BFGS-B when the function being optimized is not smooth. Thus, the results of the SLSQP method are the same as the L-BFGS-B method.[53, 57-58]

The brute force method is a global optimization technique that may be bounded over a range of values. As the name implies, the brute force method calculates the value of the function at each point to find the global minimum. The application of the brute force method to find the optimal FOM has issues with efficiency and runtime errors for large, complex problems. These issues arise from an exponential growth of the number of points evaluated based on the step size between points.[53]

The final minimization algorithm examined for optimization of the FOM was Brent's method. Brent's method guarantees convergence by using a combination of secant, bisection, and inverse quadratic interpolation methods. The nature of Brent's method allows it to be used on constrained or unconstrained, linear or nonlinear, and smooth or oscillating functions. The secant method is the first and primary step of Brent's method. The secant step is a linear interpolation of the function that identifies two points and compares the values of the function at the points. The secant step continues until an optimal value is found. If the secant step does not find the optimum in a section, then the bisection step takes over. The bisection method guards against too little progress over several iterations. The bisection step reduces the step size by approximately half or bisecting it until the step size is less than or equal to two times the stopping tolerance, then performing the secant method with the new value. The stopping tolerance is chosen by the user as any value greater than zero, the default is 1×10^{-5} . The

smaller the stopping tolerance, the higher the precision of the optimization. The inverse quadratic interpolation is the final step of Brent's method and is performed to solve for the optimum FOM. An inverse quadratic is better for use than a quadratic because the inverse quadratic avoids having to solve for a square root of the secant function. The steps of Brent's method allow it to find the global minimum without finding local minima in error. After applying Brent's method to the FOM function over the range of short gate fraction (0.73 to 0.98) with a stopping tolerance of 1×10^{-6} , the optimal FOM and short gate fraction found were approximately the same as the values found in FOM calculations. The short gate fraction found was 90.71% with a FOM of 0.8284, and PSD plots for the optimal results are shown in Figure 3.20. The short gate fraction associated with the optimal FOM was used in the application of the ML algorithms.[42, 53]

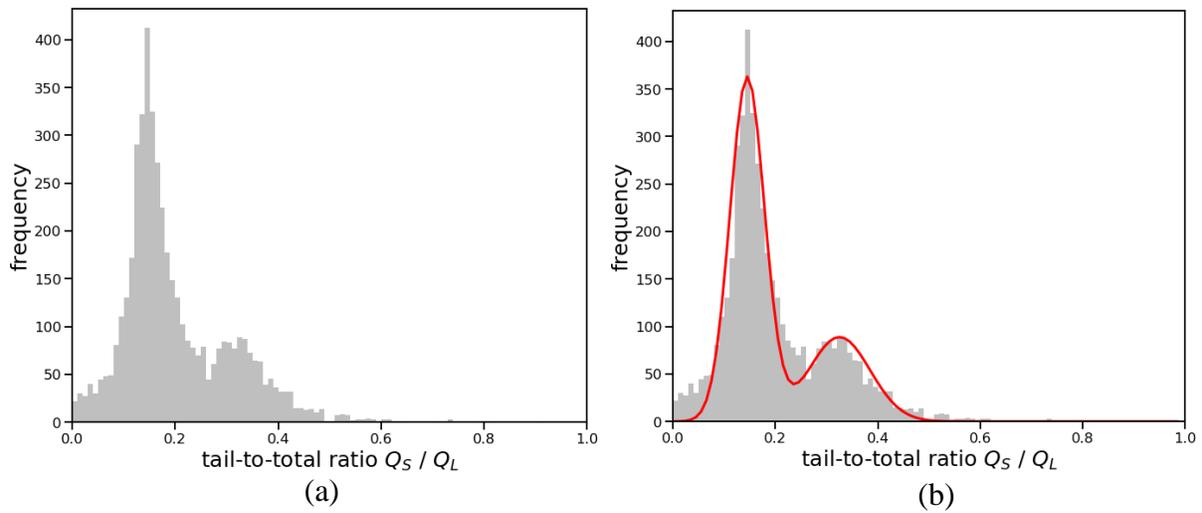


Figure 3.20. (a) PSD histogram from optimized parameters and (b) PSD histogram with double Gaussian curve fit.

CHAPTER 4

Machine learning (ML) application to particle identification

4.1 Development of ML models

The primary objective for the application of ML algorithms was to implement accurate PID of low energy neutrons in a gamma rich environment. The ML algorithms used in past research include classification, clustering, regression, and dimensionality reduction techniques. The research and development of different statistical ML algorithms for PID include:

- Artificial Neural Network (ANN) [8-10, 13]
- Support Vector Machine (SVM) [6-7, 11, 14]
- Principal Component Analysis (PCA) [12, 14]
- Decision Tree (DT) [11]
- C-means [15]
- K-means [10, 15]
- Gaussian Mixture Model (GMM) [4-5, 7, 12, 15]
- Gaussian Mixture Model with Expectation-Maximization (GMM-EM) [5, 12]
- Dirichlet Process Gaussian Mixture Model (DPGMM) [4]

ML algorithms employ data features and model parameters. For this application, features are categorical and/or numerical characteristics of a digitized pulse/record. The features used in this research include $\frac{Q_S}{Q_L}$ ratio, total charge/long gate integral, and quantiles. The parameters used in this research are statistical hyperparameters. Parameters are properties of the ML model that may be user-specified, or components of the model construction that are not learned from the data. Each ML algorithm has different parameters that may be used to construct the model.[59]

The desired result is to train an ML algorithm and create a model to implement PID without complete knowledge of the source and to minimize misidentification of gammas and neutrons. In this research, clustering ML algorithms were used to find an effective way to conduct PID without labeled features. Semi-supervised methods were also tested. Both supervised and unsupervised methods may be transformed to semi-supervised methods. Supervised methods transform to semi-supervised classification methods and unsupervised methods transform to constrained clustering. Constrained clustering differs from normal unsupervised methods in that it includes some information about the clusters before processing the data in an ML algorithm.[44-45, 59]

Clustering is an unsupervised ML technique that groups data based on the similarities in the features evaluated. The clustering algorithms explored in this research include:

- K-means
- MiniBatch K-means
- GMM-EM
- Variational Bayesian Gaussian Mixture Model (VBGMM)

Each of the ML algorithms were implemented with the features $\frac{Q_S}{Q_L}$ ratios, long gate integrals (Q_L), quantiles, or a combination of these features. Constrained clustering was evaluated in this research. However, constrained clustering did not provide reasonable results for all the algorithms tested; this is discussed in the following sections.

To achieve constrained clustering, the unlabeled data was separated into sections through “slicing,” i.e., segmenting the data by total charge. The concept of slicing the PSD results showed significant PID capabilities in the research by Polack et al. [3] and Kaplan et al. [34]. There were two approaches taken to slice the unlabeled data. The first used a defined total charge

boundary at specified values that were not equally spaced (e.g., $Q_L = 2, 5, 10, 15, 30,$ and 60) found through visual inspection of the records. The second separates the unlabeled data by a percentage of the population over equal spacing (e.g., 0-100% in 20% increments). The second slicing method provided the best results and did not require adjustments for different data. The algorithms used in this research used both a top-down and bottom-up slicing approach for comparison. Slicing was achieved by separating the population into equally sized samples (e.g., 0-100% in 20% increments). The ML algorithm sequentially trained each slice by using the initial weights and means from the PSD calculation in Chapter 3 for the first slice, then the next slice used the weights and means from the results of the prior slice as initial parameters. Bottom-up slicing started at the lowest slice along the total charge axis, 0-20% of the maximum total charge, and moved up the total charge axis in 20% increments. Top-down slicing was achieved by starting at the maximum total charge, the 80-100% slice, then moved down the total charge axis in 20% increments.[44]

The $\frac{Q_S}{Q_L}$ ratios and Q_L from both the Cf-252 and Na-22 sources were found using the methods discussed in Chapters 2 and 3. All 5,774 Cf-252 records and 10,000 Na-22 records from the TOF calculation were processed through the PSD algorithm and are displayed in Figure 4.1.

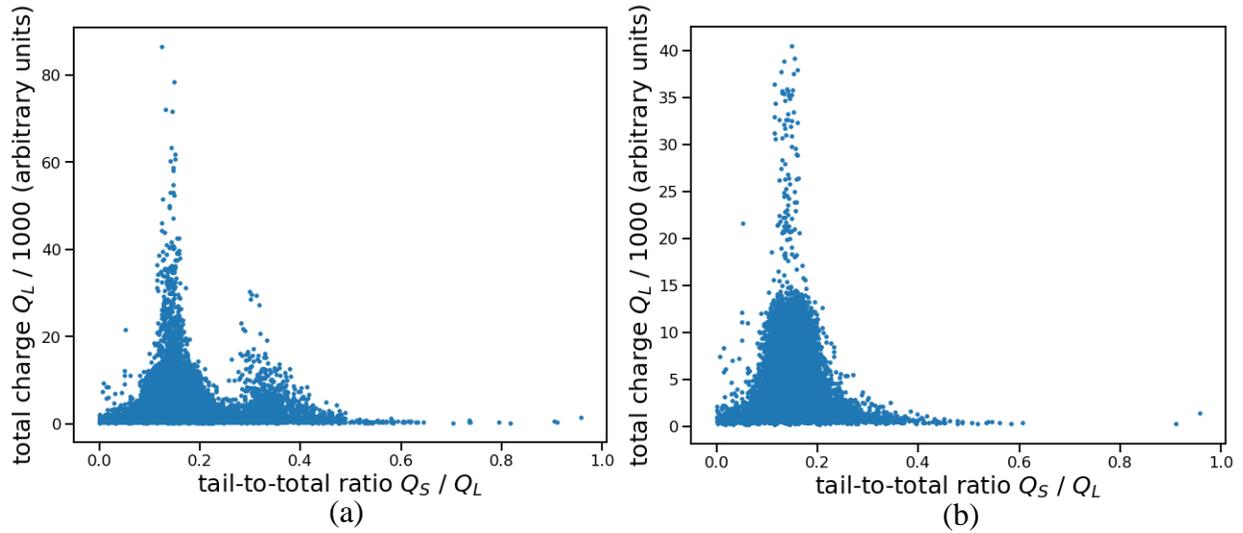


Figure 4.1. PSD for (a) Cf-252 and (b) Na-22.

The Na-22 source $\frac{Q_S}{Q_L}$ ratios and Q_L are used to simulate a gamma rich environment. The $\frac{Q_S}{Q_L}$ ratios and Q_L associated with the Cf-252 source were reduced to simulate a neutron only source. This was done by identifying those records associated with neutrons using TOF discrimination. Reviewing the TOF histogram displayed in Figure 3.11, the gamma records and neutron records are identified by the two peaks. The Cf-252 records associated with neutrons fall in the peak between $50 \leq \text{TOF} \leq 150$ ns. Removing the records that do not have a TOF within the neutron range reduces the PSD for the Cf-252 source to the histogram shown in Figure 4.2.

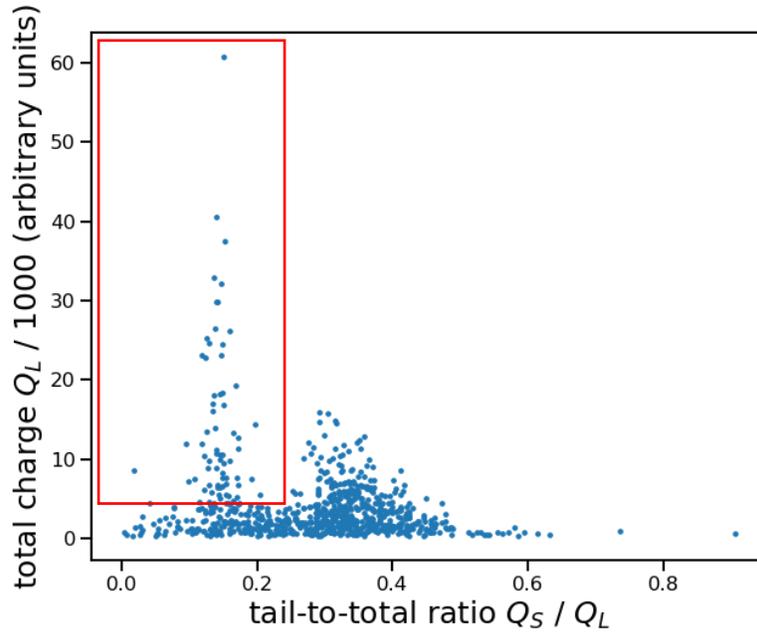


Figure 4.2. PSD for Cf-252 only including records identified as neutrons through TOF. Red box represents gammas misidentified as neutrons.

Note that ideally only a single Gaussian neutron peak would appear, but without a pure neutron source in a controlled environment misidentification of gammas in the TOF distribution will occur primarily due to accidental coincidence.[5] The results in Figure 4.2 are used to identify the true neutrons for validation of the ML algorithm results. The red box in Figure 4.2 outlines records that were identified as neutrons through TOF analysis but are in fact gammas. Approximately 5% of neutrons are misidentified gammas. The misidentified gammas are a result of accidental coincidence caused by gammas from a spontaneous fission different from the one detected in the trigger detector.[31-32, 37] A Cf-252 source produces on average 7.77 gammas and 3.77 neutrons per spontaneous fission (i.e., 62% gammas and 38% neutrons).[6, 12]

The records identified as neutrons from the TOF analysis of the Cf-252 source were combined with records from the Na-22 (gamma only) source to form Figure 4.3. The records from Na-22 source are labeled as gammas and the records from the Cf-252 source are labeled as neutrons.

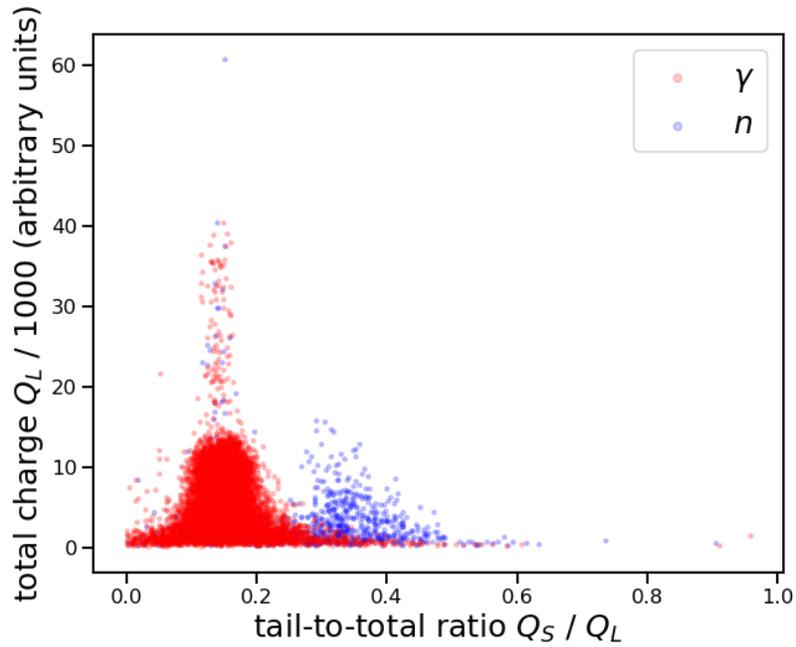


Figure 4.3. PSD of gammas (red) from Na-22 source and neutrons (blue) from Cf-252 source.

The use of a discrimination line for PID is used in this research as a baseline comparison against the ML algorithm results. The discrimination line was drawn at the midpoint between the mean gamma and mean neutron centroids found from the double Gaussian fit discussed in Chapter 3. The clustering of neutrons and gammas using a discrimination line is presented in Figure 4.4.

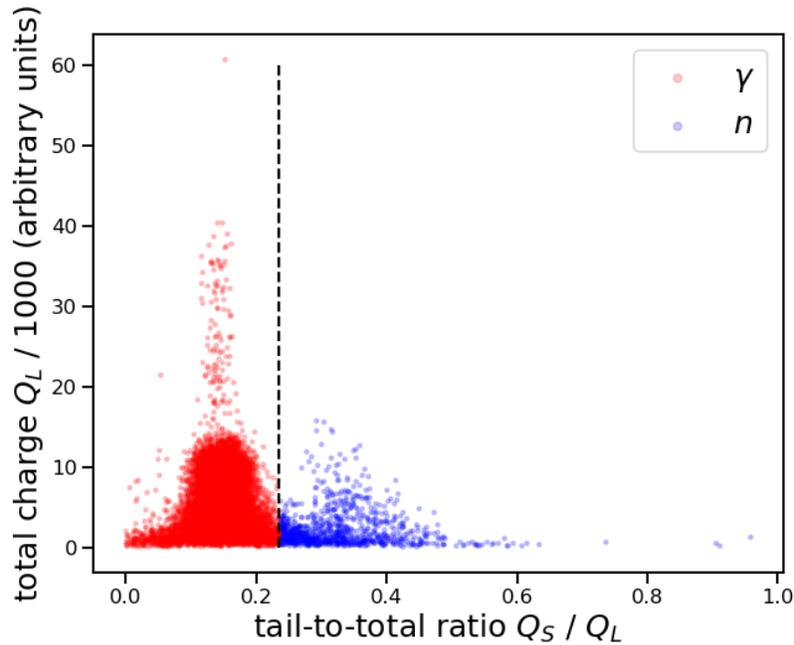


Figure 4.4. PSD of gammas (red) and neutrons (blue) using a discrimination line.

The final preparation of the data separated the population into two datasets: 70% in a training dataset and 30% in a test dataset. The training dataset was used to train the ML algorithms. The test dataset was used to test the predictability of the model. Both are included in the final results for the models and validation calculations.

4.1.1 K-means and MiniBatch K-means models

The K-means algorithm is a clustering technique that seeks to identify a centroid of each cluster to minimize the distance between the associated points and the centroid of the cluster. K-means uses Lloyd's algorithm.[59] The process of clustering starts with the input of the number of clusters desired. Two options for producing initial centroids exists: one generates the centroids at random, and the other generates centroids at a maximum distance apart; the latter was applied in this research. The initial centroids are generated and the minimum distance from each point to each centroid is calculated. The points are then clustered with the centroid that is the point's nearest neighbor. The average of the clustered points is taken, and the average value becomes the

new centroid. The minimum distance calculation and clustering process is repeated until the new average value is equal to the prior average value within a user-specified tolerance.[45]

The MiniBatch K-means algorithm operates under the same principles as the K-means algorithm. The difference between MiniBatch K-means and K-means is how the data are processed. In each iteration of the K-means algorithm, the entire population is evaluated and split into the number of clusters specified. Meanwhile, the MiniBatch K-means algorithm separates the population into several smaller samples or “miniature batches” prior to identifying the initial centroids.[59] The process selected for the initial centroid identification was the same for MiniBatch K-means and K-means by generating centroids at a maximum distance apart. The initial centroids are used to cluster the first sample, and the average of the clusters is calculated and used as the new centroid. Another sample is clustered using the new centroid and the process is repeated through clustering and calculating a running average for the next centroid. The process is repeated until the entire population is clustered or the change in centroid values is within a user-specified tolerance. Observations and research comparing the MiniBatch K-means and K-means algorithms show that MiniBatch K-means is faster but sometimes yields poorer results.[59]

The K-means and MiniBatch K-means algorithms were used in unsupervised form. The number of features used as inputs to both the K-means and MiniBatch K-means algorithms were limited and did not include labeled data. The primary parameter for the K-means and MiniBatch K-means algorithms is the desired number of clusters, which were assigned to particle type, i.e., there were two clusters, neutron and gamma. The $\frac{Q_S}{Q_L}$ ratio feature with training and test datasets showed similar results in both K-means and MiniBatch K-means algorithms, shown in Figure 4.5. The PID results for the K-means appear the same as the discrimination line results because

the K-means algorithm finds cluster centroids that are nearly the same as the means of the gamma and neutrons found using the double Gaussian fit. The MiniBatch K-means results were similar to the K-means with an increase in both cluster centroid values.

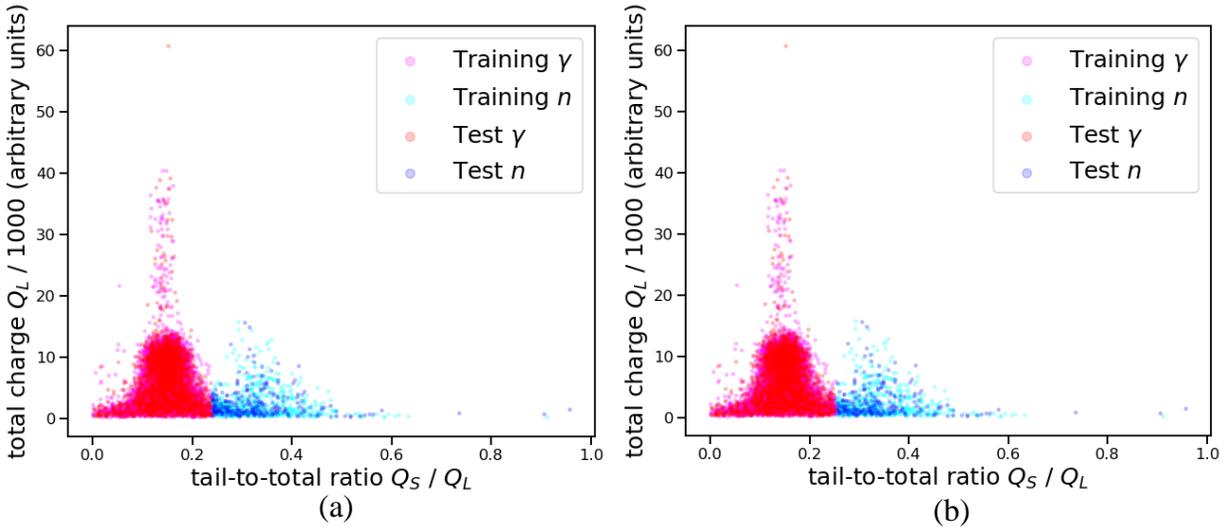


Figure 4.5. PID using (a) K-means and (b) MiniBatch K-means algorithms with $\frac{Q_S}{Q_L}$ ratio feature.

When evaluating the $\frac{Q_S}{Q_L}$ ratio and Q_L features, the K-means and MiniBatch K-means algorithms do not cluster the PID as anticipated. Instead of two clusters appearing side-by-side, the clusters appear stacked. This observation shows that these algorithms selected the centroid primarily based on Q_L instead of $\frac{Q_S}{Q_L}$ ratio. This is a result of the algorithm finding centroids that cluster and calculate minimum distance between the Q_L values because it has a wider range. Similar results occur when applying K-means and MiniBatch K-means algorithms to only Q_L . The K-means and MiniBatch K-means algorithm results including the Q_L feature do not exhibit good PID so they are excluded from the final evaluation at the end of this chapter.

The impact the Q_L had on the K-means and MiniBatch K-means algorithms was reduced by adding the quantiles feature. Although adding another feature removes the issue that occurs with the Q_L carrying too much weight and causing poor PID, there is significant

misidentification of neutrons as gammas at higher total charge in the neutron peak. These poor PID results made further validation of this ML model unnecessary because the desired result is an ML algorithm that performs good PID at all energies.

The slicing algorithm was applied to the K-means and MiniBatch K-means with varying features; however, due to the dependence the slicing algorithm has on Q_L , the PID results were poor. This showed that for PID the K-means and MiniBatch K-means perform better with fewer features and without slicing.

4.1.2 Gaussian mixture model (GMM) and GMM variants

The GMM algorithm uses the probabilistic properties of the Gaussian distribution to cluster the data. The GMM algorithm may be applied to univariate or multivariate datasets to generate a finite number of Gaussian clusters. The GMM parameters include the number of desired clusters, the covariance matrix type, and the features for unsupervised execution of the GMM algorithm. In an unsupervised format, the GMM algorithm operates like K-means algorithm by identifying the mean, variance, and weight parameters of each Gaussian distribution by assigning initial approximations of the parameters and calculating the centroids based on the features. The GMM algorithm recalculates the mean, variance, and weight after the initial clustering and repeats the process with the new parameters until the parameter values of the previous step are the same or within a user-specified tolerance. The GMM algorithm is also capable of semi-supervised execution. If the initial mean, variance, and weight parameters of the desired clusters are known, then these values are added as parameters of the model. Using labeled information accelerates and improves the performance of the GMM algorithm.[44, 59]

The expectation-maximization (EM) algorithm may be included to improve the processing speed and optimization of the GMM model. The EM algorithm is implemented in

most GMM algorithms including the one used in this research. The EM algorithm consists of two steps to optimize the parameters of the GMM. The expectation step computes the probability that the samples of a dataset would have come from the estimated parameters found by the GMM algorithm. The maximization step optimizes the probability found in the expectation step by modifying the parameters. The EM algorithm continues until convergence of the parameters is achieved.[44, 59]

The VBGMM algorithm is a variant of the GMM algorithm. The variational Bayesian estimation of the GMM uses variational inference and the Dirichlet process for clustering. The Dirichlet process is a method that finds the prior probability distribution using the probability distributions of the features. The Dirichlet distribution is a multivariate beta distribution used in variational inference. Variational inference is similar to the EM algorithm in terms of performance, but instead of maximizing the probability that samples came from the estimated parameters, it seeks to maximize the conjugate prior or posterior Dirichlet distribution. The VBGMM algorithm is most effective for models that have more features. The weight of the concentration prior parameter addresses the unique difference between the VBGMM and GMM algorithms. The weight of the concentration prior determines the number of features that influence the clustering. The higher the weight of the concentration prior, the greater the number of features that have an influence in the VBGMM algorithm.[4, 59]

The GMM algorithm variants all use a covariance matrix parameter. The options for the covariance matrix type are full, tied, diagonal, or spherical. The selection of the covariance matrix type is dependent on the relationships of the clusters and the features. For a full covariance matrix each cluster has its own general covariance matrix. The clusters and features have some independence and do not need to have the same shape, so each cluster has its own

mean and variance vectors. A tied covariance matrix means the clusters share the same general covariance matrix, i.e., the clusters and features have some dependence and have the same shape. The diagonal covariance matrix treats the features within each cluster as uncorrelated. The diagonal covariance matrix clusters have similar properties to a full covariance matrix except the clusters may have non-circular shapes. A spherical covariance type gives each cluster its own single variance; the clusters are independent with a circular shape. The full, tied, and diagonal covariance matrices were tested in research by Blair et al. [4], where they found that the tied covariance matrix in a GMM displayed the best clustering by particle type. All covariance types were tested for each GMM algorithm variant.[59]

The data evaluated by GMM algorithms and variants are processed and labeled based on the features and parameters given. The GMM-EM algorithm labels the data with quantitative values representing the particle type. Misclustering occurs when the algorithm mislabels the clusters or a portion of the clusters. Misclustering may occur on an iteration basis, which may be fixed by running the algorithm again. Misclustering may also be more permanent due to the features or parameters used, which requires a change to the cluster labels. Special attention must be given to evaluate the cause of misclustering when it occurs.

In this research, the GMM algorithm appeared to be obsolete to the GMM-EM with no improvement to the results outside of computational time. The first GMM-EM clustering used the $\frac{Q_S}{Q_L}$ ratio feature and the parameter of covariance matrix type, initial weights, and initial means. The initial means are extracted from the double Gaussian fit results found through processes described in Chapter 3. The initial weights used were 62% for gammas and 38% for neutrons to match what is expected from a Cf-252 source. To maintain accurate proportionality of neutrons to gammas, the first application of the GMM-EM algorithm only included 1,273

samples from the Na-22 source with 780 samples from Cf-252 source. To be more representative of a gamma rich environment, the number of Na-22 source samples was increased, and the initial weights were adjusted accordingly. The number of gammas from the Na-22 source was increased until maximizing the initial weights at 92% for gammas and 8% for neutrons. However, adjusting the initial weights did not change the weights found by the GMM-EM algorithm. The covariance matrix was varied between tied, full, diagonal, and spherical types for comparison, and like the findings of Blair et al. [4], the tied covariance matrix exhibited the best PID clustering, as seen in Figure 4.6. These results are expected because the clusters share the same overall mean and variance with the $\frac{Q_S}{Q_L}$ ratio feature.

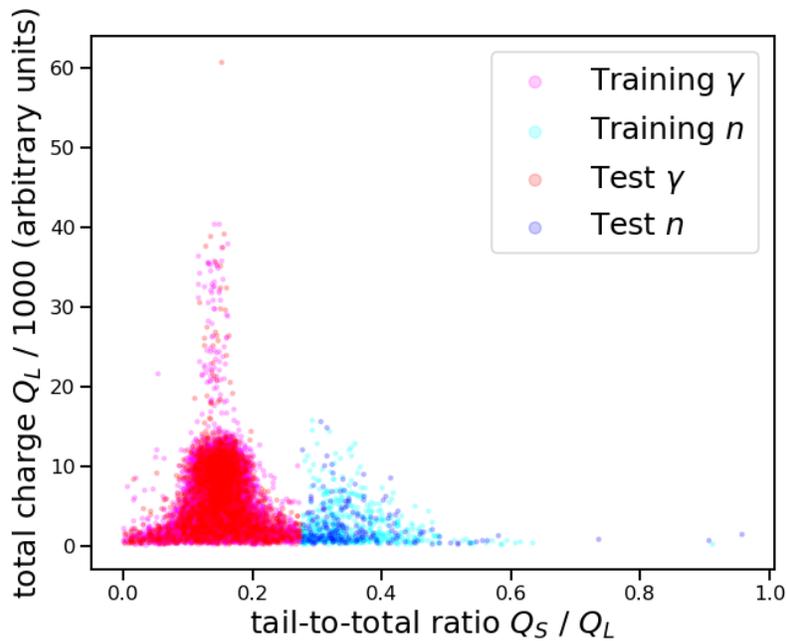


Figure 4.6. PID using GMM-EM on the $\frac{Q_S}{Q_L}$ ratio with a tied covariance matrix.

The GMM-EM algorithm was then applied to the Q_L feature, but the results were similar to the K-means and MiniBatch K-means where the clusters appeared to be stacked instead of side-by-side. The features $\frac{Q_S}{Q_L}$ ratio and Q_L were also examined using the GMM-EM algorithm with the same parameters of initial weight and initial mean, while varying the covariance matrix type. The

full, diagonal, and spherical covariance types failed to provide adequate clustering. The PID results of the tied covariance matrix application are presented in Figure 4.7. When compared to the discrimination line PID in Figure 4.4, there is a noticeable shift in PID with an increase in the number of neutrons falsely labeled as gammas. There is a slight improvement in the PID of neutrons at low total charge when the Q_L feature is added. Unlike the K-means and MiniBatch K-means, the GMM-EM algorithm did not degrade when more features were added to it. This is because GMM-EM has parameters to help regulate the impact of additional features.

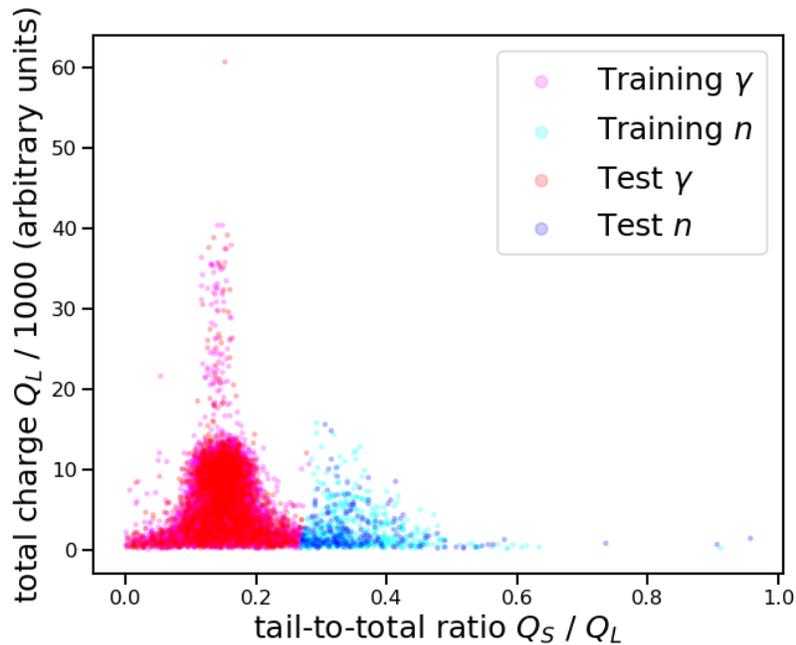


Figure 4.7. PID using GMM-EM on the $\frac{Q_S}{Q_L}$ ratio and Q_L with a tied covariance matrix.

The next GMM-EM algorithm included slicing the data by Q_L (total charge). A comparison of the bottom-up and top-down approaches, with the features of $\frac{Q_S}{Q_L}$ ratio and Q_L , showed that the PID of the bottom-up approach was better. All covariance types of the bottom-up and top-down approaches for GMM-EM were tested. The tied covariance matrix of the bottom-up slicing exhibited the best clustering; it is shown in Figure 4.8. The top-down slicing did not improve with different covariance matrices. Top-down slicing was also tested in the other

GMM algorithms with varying features and covariance matrices, but it did not improve the results.

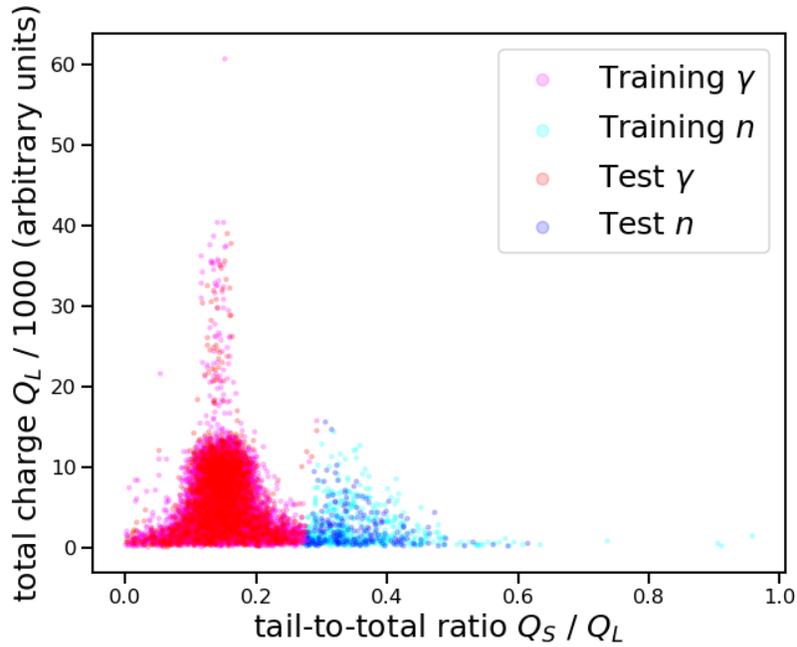


Figure 4.8. PID using GMM-EM using bottom-up slicing with a tied covariance matrix.

The quantiles feature was introduced to the GMM-EM algorithm. The quantiles feature is a multivariate feature; for each record it has nine distinguishable quantiles (i.e., 10-90% in 10% increments). The applications included:

1. All features (quantile, $\frac{Q_S}{Q_L}$ ratio, and Q_L).
2. Bottom-up slicing with all features.
3. Quantiles as the only feature.
4. Bottom-up slicing with quantiles as the only feature.

The evaluation of all features including the quantiles only feature yielded good PID results. The clusters had correct PID and there was a good mixture at low total charge, as seen in figures 4.9 and 4.10.

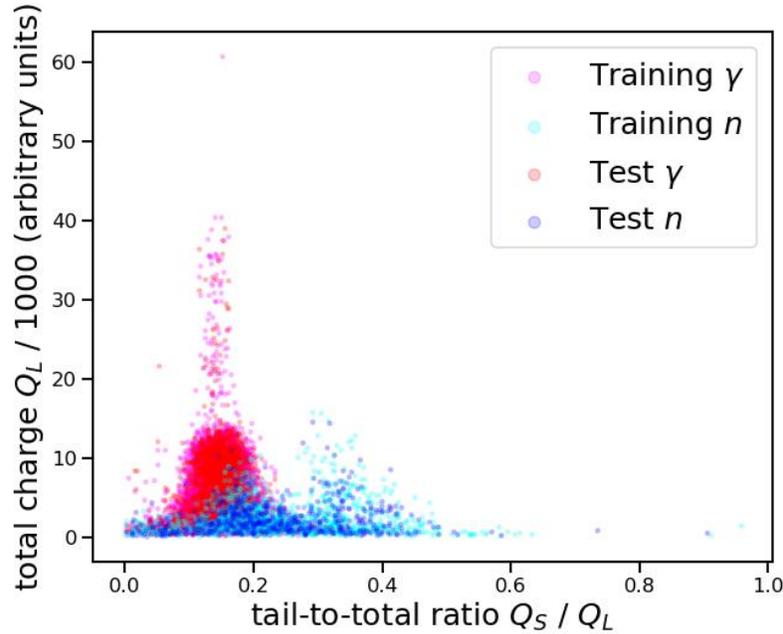


Figure 4.9. PID using GMM-EM using quantiles, $\frac{Q_S}{Q_L}$ ratio, and Q_L with diagonal covariance matrix.

The tied covariance matrix for the GMM-EM algorithm with quantiles was not the best choice. The quantiles feature adds dimensionality to the clusters that requires separate covariance matrices. The quantiles feature is independent of the $\frac{Q_S}{Q_L}$ ratio and Q_L features. For the GMM-EM algorithm, the quantiles, $\frac{Q_S}{Q_L}$ ratio, and Q_L features need either a full or diagonal covariance matrix function, while the quantiles only feature requires a diagonal covariance matrix. The full covariance matrix parameter in the GMM-EM algorithm applied to quantiles, $\frac{Q_S}{Q_L}$ ratio, and Q_L features required a change to the labels to avoid misclustering. Therefore, a diagonal covariance matrix was used for the GMM-EM algorithm with the addition of the quantiles feature. The diagonal covariance matrix is the optimal choice of the covariance matrix types for the application of GMM-EM with quantiles. This is due to the relationship that the quantiles have with the particle type; the covariances of the quantiles for each particle type are independent.

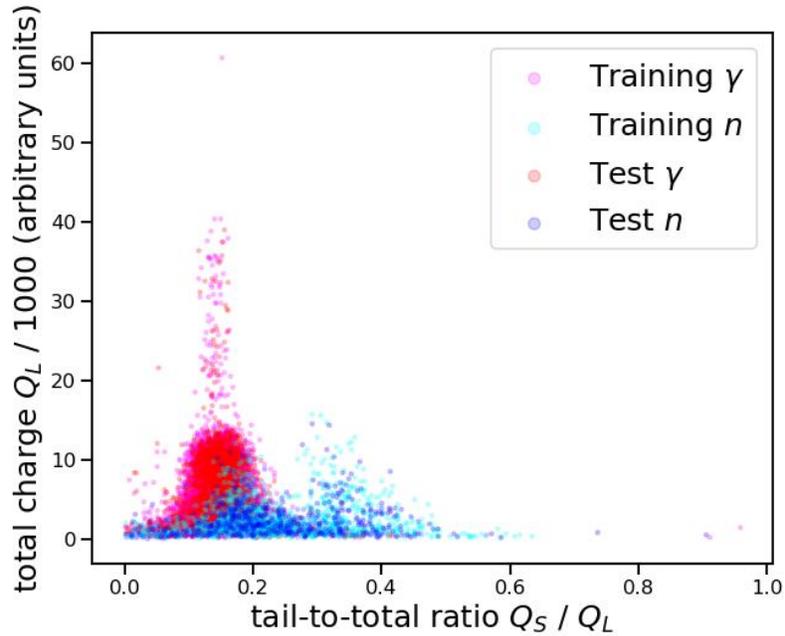


Figure 4.10. PID using GMM-EM using quantiles only with a diagonal covariance matrix.

Bottom-up slicing was also applied to the GMM-EM algorithm with quantiles and with all features, but the clustering did not have the desired PID. Bottom-up slicing caused a large group of misidentified neutrons on the righthand side of the gamma peak.

As discussed, the VBGMM algorithm performs well with multiple features, so it was not used to evaluate the $\frac{Q_S}{Q_L}$ ratio and Q_L features. The VBGMM algorithm was applied to all features and quantiles only feature with and without slicing. The VBGMM algorithm with all features and quantiles feature without slicing provide good clustering results with a diagonal covariance matrix are presented in figures 4.11 and 4.12.

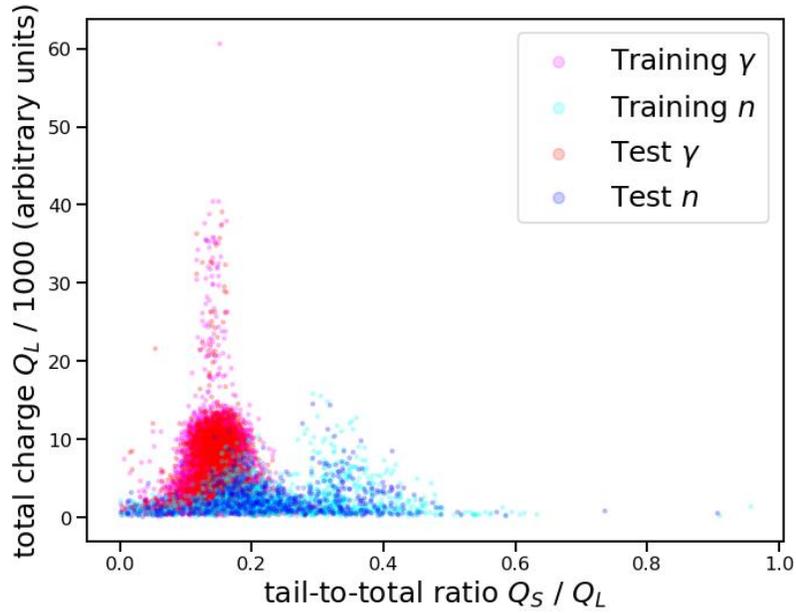


Figure 4.11. PID using VBGMM using quantiles, $\frac{Q_S}{Q_L}$ ratio, and Q_L with diagonal covariance matrix.

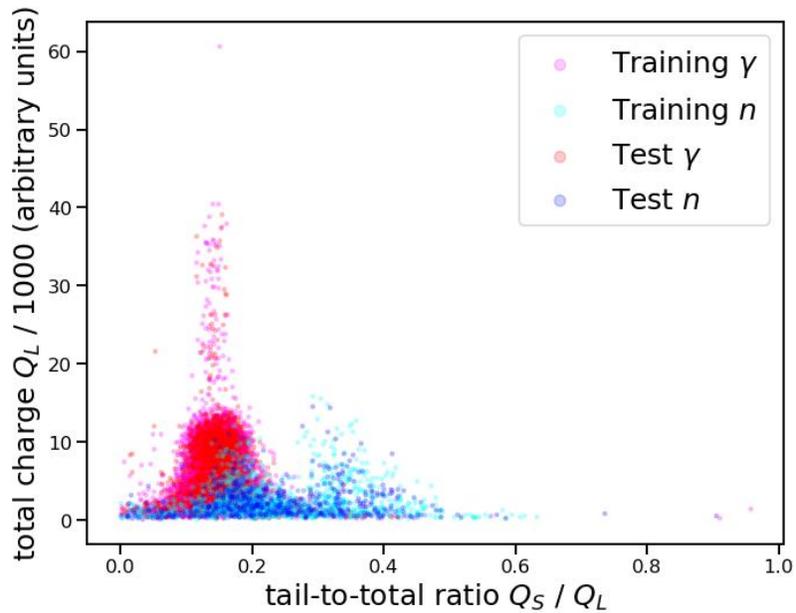


Figure 4.12. PID using VBGMM using quantiles only with diagonal covariance matrix.

The use of bottom-up slicing with VBGMM had similar results to the bottom-up slicing with GMM-EM with the quantiles feature. The PID clustering results with slicing for the quantiles only feature and all features show a large group of misidentified neutrons on the

righthand side of the gamma peak cluster. Because of the poor PID exhibited by the slicing algorithm it was not considered further with VBGMM.

To evaluate the quantiles impact on the PID results for the GMM-EM and VBGMM algorithms, the number of quantiles necessary for good PID was tested by removing some quantiles. Removing any of the quantiles (e.g., 10%, 90%, or both 10% & 90%) degraded the PID results of the GMM-EM and VBGMM algorithms. The identification of the optimal quantiles was beyond the scope of this research and is recommended for future work.

The GMM-EM and VBGMM algorithms evaluated in this section required further validation of the results. The PID clusters appear to fall within the correct areas for most applications of the ML algorithms, but a comparison to the ground truth is necessary to validate the clustering results. The best visual PID results at low total energy appear to occur with the GMM-EM and VBGMM algorithms with the $\frac{Q_S}{Q_L}$ ratio, Q_L , and quantiles features, and with the quantiles only feature without using a slicing algorithm.

4.2 Validation of ML models

The ML models for each of the algorithms and iterations with different features and slicing were validated. The validation involves calculating the true positive rate (TPR) and false positive rate (FPR) and comparing the TPR and FPR for each PID clustering result. The TPR is the number of identified neutrons per true neutron. The FPR is the number of neutrons identified as gamma per true gamma. The true neutrons are labeled using the TOF calculation for the Cf-252 source that are presented in Figure 4.2. The true gammas are labeled using the records from the Na-22 source. All the records from Na-22 are gammas, while a small portion of the true neutrons from the Cf-252 source are also gammas ($\approx 5\%$).[4-5, 12]

The TPR is calculated by finding the number of clustered neutrons that have the same $\frac{Q_S}{Q_L}$ ratio and Q_L as the true neutrons. After finding the number of clustered neutrons that match the true neutrons, the matched clustered neutrons are counted and the ratio of the number of clustered neutrons to the number of true neutrons is taken. This ratio is the TPR. The FPR is a ratio calculated in a similar way. The number of clustered gammas that have the same $\frac{Q_S}{Q_L}$ ratio and Q_L as the true neutrons is found. If the clustered gamma matches a true neutron, then it is considered misidentified. The FPR is calculated as the ratio of the number of misidentified neutrons to the number of true gammas. The rates follow a binomial distribution so the variance for the TPR and FPR is calculated using Equation 4.1.

$$Var = \frac{rate(1-rate)}{N} \quad (4.1)$$

where the rate is either the FPR or TPR, and N is the sample size of the true gammas or neutrons, respectively.[12, 24]

The results of the TPR and FPR calculations for the discrimination line and the algorithms with good visual PID clustering are included in Table 4.1. The GMM-EM algorithms are calculated with a tied covariance matrix unless quantiles are included in the features, in which case a diagonal covariance matrix was used. The VBGMM algorithms were calculated with a diagonal covariance matrix.

Table 4.1. TPR and FPR with standard deviation for the primary algorithms considered based on PID clustering. GMM-EM with all features has best FPR and TPR results.

<i>ML Algorithm</i>	FPR	TPR
<i>Discrimination line</i>	0.0252 ± 0.0016	0.6949 ± 0.0165
<i>K-means</i>	0.0255 ± 0.0016	0.6910 ± 0.0165
<i>MiniBatch K-means</i>	0.0276 ± 0.0017	0.6654 ± 0.0169
<i>GMM-EM w/ $\frac{Q_S}{Q_L}$ ratio</i>	0.0303 ± 0.0018	0.6333 ± 0.0173
<i>GMM-EM w/ $\frac{Q_S}{Q_L}$ ratio and Q_L</i>	0.0300 ± 0.0018	0.6359 ± 0.0172
<i>GMM-EM w/ slicing</i>	0.0307 ± 0.0018	0.6282 ± 0.0173
<i>GMM-EM w/ all features</i>	0.0234 ± 0.0016	0.7167 ± 0.0161
<i>GMM-EM w/ quantiles only</i>	0.0270 ± 0.0017	0.6731 ± 0.0168
<i>VBGMM w/ all features</i>	0.0237 ± 0.0016	0.7128 ± 0.0162
<i>VBGMM w/ quantiles only</i>	0.0273 ± 0.0017	0.6692 ± 0.0168

The desired results are a TPR close to 100%, and an FPR value close to 0%. A good baseline for comparison is the FPR and TPR for the PSD clustered by the discrimination line. The discrimination line and the K-means algorithm have similar TPR and FPR values, which is a result of the K-means algorithm clustering based on a centroid value similar to the mean values used to calculate the discrimination line. MiniBatch K-means performs slightly worse than K-means which is expected based on the algorithms' sampling methods. The GMM-EM algorithm with the $\frac{Q_S}{Q_L}$ ratio feature, $\frac{Q_S}{Q_L}$ ratio and Q_L features, and slicing results appear to have similar PID properties, but compared to the discrimination line results the TPR and FPR are worse. The application of quantiles only leads to results with higher FPR and lower TPR for both the GMM-EM and VBGMM algorithms, with GMM-EM performing slightly better. The best results for TPR and FPR come from the applications of GMM-EM and VBGMM algorithms with $\frac{Q_S}{Q_L}$ ratio,

Q_L , and quantiles features. The GMM-EM algorithm only slightly outperforms the VBGMM algorithm, but the results are within the standard deviation for both the TPR and FPR. Although GMM-EM and VBGMM algorithms differ in their methods, the algorithms seek to find a GMM that clusters using the same features which give the algorithms similar results.

The impact of the accidental coincidences on the FPR and TPR was evaluated by removing most of the gammas misidentified as true neutrons from the TOF of Cf-252. This was done by removing the Cf-252 source records that are highlighted in the red box in Figure 4.2. The removal of the gammas misidentified as true neutrons that are correctly identified as gammas in the algorithm was expected to improve the TPR and FPR. The identification of the misidentified gammas for removal was done by selecting those records less than the midpoint between the means of the two peaks found using a double Gaussian fit found in Section 3.2.4. To keep more potential neutron records, only the records with a total charge greater than a selected threshold were removed. Visually, all PID clusters appear to have more accurate distributions with the removal of the misidentified gammas. The TPR and FPR also improved. The improvement ranged from a 9-11% increase for TPR and 1-2% decrease for FPR. The removal of misidentified gammas does not improve the performance of one clustering algorithm over the others, but it does improve the clustering results for each clustering algorithm. The GMM-EM algorithm with the quantiles, $\frac{Q_S}{Q_L}$ ratio, and Q_L features still provides the best PID results with $TPR = 0.8294 \pm 0.0145$ and $FPR = 0.0121 \pm 0.0011$.

CHAPTER 5

Conclusions and future work

5.1 Conclusions

The use of ML algorithms for PID of gammas and neutrons in a gamma rich environment has several applications including nuclear security and safeguards, nuclear medicine, and particle physics. This research demonstrates that the use of ML for PID improves performance over more traditional PID methods. In this research a TOF experiment was used to collect digitized pulses from a CeBr₃ inorganic scintillator and an EJ-309 liquid organic scintillator. After parsing and processing the pulse records, standard techniques were used to calculate TOF, perform charge integration PSD, optimize the PSD FOM, and calculate the quantiles feature. The TOF setup use a CeBr₃ trigger detector and EJ-309 study detector to reduce the number of accidental coincidences caused by neutrons while collecting all gamma and neutron records for PSD. The parsing and processing of records removed additional accidental coincidences. The long gate (Q_L = total charge) and short gate (Q_S = delayed charge) integrals of the records for the application of charge integration method was selected using the FOM. Optimization found a maximum FOM of 0.8284 with a short gate fraction (fraction of the long gate integral) of 90.71%. The results of the optimized FOM were used in the charge integration PSD method to extract the $\frac{Q_S}{Q_L}$ ratio and Q_L features for the ML algorithms. The charge integration method results were also used to generate the quantiles feature used in the ML algorithms.

The K-means, MiniBatch K-means, GMM-EM, and VBGMM ML algorithms were evaluated for PID performance without labeled features. The K-means and MiniBatch K-means algorithms performed well with fewer features, but did not show good PID at low total charge. The PID performance at low total charge was better in the GMM algorithm variants. The GMM-

EM and VBGMM algorithms with the $\frac{Q_S}{Q_L}$ ratio, Q_L , and quantiles features exhibit good visual PID at low total charge when compared to the other ML algorithms tested. The ML algorithms' true positive rate (TPR) and false positive rate (FPR) match the visual PID clustering results. The best model was produced by the GMM-EM algorithm, which exhibited TPR and FPR within the standard deviation of the VBGMM algorithm. The TPR and FPR may be improved by removing some of the misidentified gammas from the true neutron records from the Cf-252 source.

Applying the GMM-EM algorithm with the $\frac{Q_S}{Q_L}$ ratio, Q_L , and quantiles features to the reduced dataset of Cf-252 neutrons and Na-22 gammas had the best validation results with a TPR = 0.8294 ± 0.0145 and FPR = 0.0121 ± 0.0011 , substantially better than standard PID using a discrimination threshold.

5.2 Future work

Proposed future work includes an experiment at the Triangle University National Laboratory (TUNL) tandem Van de Graaff accelerator to acquire measurements of a pure neutron beam. Although the TOF calculation to identify the true neutrons from a Cf-252 source provides a good method for PID it is not absolute as seen with the accidental coincidences in Figure 4.2. A pure neutron source measurement will better train the model and can be used to validate the PID without having accidental coincident gammas intermixed with the true neutrons. The evaluation of supervised or semi-supervised ML techniques with the same features should be compared to the ML algorithms used in this research. The supervised and semi-supervised ML techniques require the labeling of some of the data. An in-depth review and analysis of the features and parameters should also be done, including new features and parameter adjustments. The analysis of features should include sensitivity analysis to identify the amount of influence each feature has on the PID models. The review of the features should include the exploration of

the quantiles' step sizes. The number of clusters should also be tested. The number of clusters parameter may be increased to four or more based on separating the gammas and neutrons by total energy. The final recommendation for future work is to apply the clustering algorithms at lower total energy ranges. This will allow for the evaluation and comparison of the TPR and FPR at low total energies to those from the total energy spectrum.

REFERENCES

- [1] C.G. Tully, *Elementary Particle Physics in a Nutshell*, Princeton: Princeton University Press, 2011.
- [2] J.B. Birks, *The Theory and Practice of Scintillation Counting*, New York: Macmillan, 1964.
- [3] J.K. Polack, M. Flaska, A. Enqvist, C.S. Sosa, C.C. Lawrence, and S.A. Pozzi, “An algorithm for charge-integration, pulse-shape discrimination and estimation of neutron/photon misclassification in organic scintillators,” *Nuclear Instruments and Methods in Physics Research, Section A: accelerators, spectrometers, detectors and associated equipment*, vol. 795, pp. 253-267, 2015.
- [4] B. Blair, C. Chen, A. Glenn, A. Kaplan, J. Ruz, L. Simms, and R. Wurtz, “Gaussian mixture models as automated particle classifiers for fast neutron detectors,” *Stat Anal Data Min: The ASA Data Sci Journal*, vol. 12, pp. 479-488, 2019.
- [5] L.M. Simms, B. Blair, J. Ruz, R. Wurtz, A.D. Kaplan, and A. Glenn, “Pulse discrimination with a Gaussian mixture model on a FPGA,” *Nuclear Instruments and Methods in Physics Research, Section A: accelerators, spectrometers, detectors and associated equipment*, vol. 900, pp. 1-7, 2018.
- [6] T.S. Sanderson, C.D. Scott, M. Flaska, J.K. Polack, and S.A. Pozzi, “Machine learning for digital pulse shape discrimination,” *IEEE Nuclear Science Symposium and Medical Imaging Conference Record*, pp. 199–202, 2012.
- [7] M. Gelfusa, R. Rossi, M. Lungaroni, F. Belli, L. Spolladore, I. Wyss, P. Gaudio, A. Murari, and JET Contributors, “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, 164198, 2020.
- [8] A. Grobov, “Machine learning approach to pulse shape discrimination in liquid noble gas detectors,” *Journal of Physics: Conference Series*, 1390, 012110, 2019.
- [9] P. Holl, L. Hauertmann, B. Majorovitis, P. Schulz, M. Schuster, and A.J. Zsigmond, “Deep learning based pulse shape discrimination for germanium detectors,” *The European Physical Journal C*, 79, 450, 2019.
- [10] E. Doucet, T. Brown, P. Chowdhury, C.J. Lister, C. Morse, P.C. Bender, and A.M. Rogers, “Machine learning n/γ discrimination in CLYC scintillators,” *Nuclear Instruments and Methods in Physics Research, Section A: accelerators, spectrometers, detectors and associated equipment*, vol. 954, 161201, 2020.

- [11] Z. Zhang, C. Hu, Y. Zhang, B. Liao, J. Zhu, X. Fan, Y. Li, and L. Zhang, "The combined application of principal component analysis and decision tree in nuclear pulse shape discrimination," *Nuclear Instruments and Methods in Physics Research, Section A: accelerators, spectrometers, detectors and associated equipment*, vol. 943, 162425, 2019.
- [12] R. Wurtz, B. Blair, C. Chen, A. Glenn, A.D. Kaplan, P. Rosenfield, J. Ruz, and L.M. Simms, "Methodology and performance comparison of statistical learning pulse shape classifiers as demonstrated with organic liquid scintillator," *Nuclear Instruments and Methods in Physics Research, Section A: accelerators, spectrometers, detectors and associated equipment*, vol. 901, pp. 46-55, 2018.
- [13] E.K. Mace, J.D. Ward, and C.E. Aalseth, "Use of neural networks to analyze pulse shape data in low-background detectors," *Journal of Radioanalytical and Nuclear Chemistry*, col. 318, pp. 117-124, 2018.
- [14] Z.H. Zhang, C.Y. Hu, X.Y. Fan, B. Liao, Y.T. Li, J.J. Zhu, and L.Q. Zhang, "A direct method of nuclear pulse shape discrimination based on principal component analysis and support vector machine," *Journal of Instrumentation*, vol. 14, P06020, 2019.
- [15] R.E. Wurtz, "Consistent Principles for Particle Identification by PSD Systems," *Lawrence Livermore National Laboratory*, LLNL-CONF-781720, 2019.
- [16] E.A. Bailey and G.K. Rollefson, "The Determination of the Fluorescence Lifetimes of Dissolved Substances by a Phase Shift Method," *The Journal of Chemical Physics*, vol. 21, 1315, 1953.
- [17] H.B. Phillips and R.K. Swank, "Measurements of Scintillation Lifetimes," *Review of Scientific Instruments*, vol. 24, 611, 1953.
- [18] G.T. Wright, "Scintillation Decay Times of Organic Crystals," *Proceedings of the Physical Society, Section B*, vol 69, 358, 1956.
- [19] F.D. Brooks, "A Scintillation Counter with Neutron and Gamma-ray Discrimination," *Nuclear Instruments and Methods*, vol. 4, pp. 151-163, 1959.
- [20] F.D. Brooks, R.W. Pringle, and B.L. Funt, "Pulse Shape Discrimination in a Plastic Scintillator," *IRE Transactions on Nuclear Science*, vol. 7, 2-3, pp. 35-38, 1960.
- [21] E. Gatti and F. De Martini, "A New Linear Method of Discrimination Between Elementary Particles in Scintillation Counters," *Vienna: IAEA, Nuclear Electronics*, vol II, pp. 265-276, 1962.
- [22] M.J.I. Balmer, K.A.A. Gamage, and G.C. Taylor, "Comparative analysis of pulse shape discrimination methods in a ${}^6\text{Li}$ loaded plastic scintillator," *Nuclear Instruments and Methods in Physics Research, Section A: accelerators, spectrometers, detectors and associated equipment*, vol. 788, pp. 146-153, 2015.

- [23] F.D. Brooks, "Development of Organic Scintillators," *Nuclear Instruments and Methods*, vol. 162, pp. 477-505, 1979.
- [24] A.D. Kaplan, B. Blair, C. Chen, A. Glenn, J. Ruz, and R. Wurtz, "A neutron-gamma pulse shape discrimination method based on pure and mixed sources," *Nuclear Instruments and Methods in Physics Research, Section A: accelerators, spectrometers, detectors and associated equipment*, vol. 919, pp. 36-41, 2019.
- [25] M. Amiri, V. Prenosil, F. Cvachovec, Z. Matej, and F. Mravec, "Quick algorithms for real-time discrimination of neutrons and gamma rays," *J Radioanal Nucl Chem*, 303, pp. 583-599, 2015.
- [26] M. Nakhostin, "Recursive algorithms for digital implementation of neutron/gamma discrimination in liquid scintillation detectors," *Nuclear Instruments and Methods in Physics Research, Section A: accelerators, spectrometers, detectors and associated equipment*, vol. 672, pp. 1-5, 2012.
- [27] W.G.J. Langeveld, M.J. King, J. Kwong, and D.T. Wakeford, "Pulse Shape Discrimination Algorithms, Figures of Merit, and Gamma-Rejection for Liquid and Solid Scintillators," *IEEE Transactions on Nuclear Science*, vol. 64, 7, 2017.
- [28] M. Nakhostin, "A General-Purpose Digital Pulse-Shape Discrimination Algorithm," *IEEE Transactions on Nuclear Science*, vol. 66, 5, pp. 838-845, 2019.
- [29] M.L. Roush, M.A. Wilson, and W.F. Hornyak, "Pulse Shape Discrimination," *Nuclear Instruments and Methods*, vol. 31, pp. 112-124, 1964.
- [30] S. Marrone, D. Cano-Ott, N. Colonna, C. Domingo, F. Gramegna, E.M. Gonzalez, F. Gunsing, M. Heil, F. Kappeler, P.F. Mastinu, P.M. Milazzo, T. Papaevangelou, P. Pavlopoulos, R. Plag, R. Reifarh, G. Tagliente, J.L. Tain, and K. Wisshak, "Pulse shape analysis of liquid scintillators for neutron studies," *Nuclear Instruments and Methods in Physics Research, Section A: accelerators, spectrometers, detectors and associated equipment*, vol. 490, pp. 299-307, 2002.
- [31] M. Flaska and S.A. Pozzi, "Optimization of an Offline Pulse-Shape Discrimination Technique for the Liquid Scintillator BC-501A," *Oak Ridge National Laboratory*, 2006.
- [32] S.D. Ambers, M. Flaska, and S.A. Pozzi, "A hybrid pulse shape discrimination technique with enhanced performance at neutron energies below 500 keV," *Nuclear Instruments and Methods in Physics Research, Section A: accelerators, spectrometers, detectors and associated equipment*, vol. 638, pp. 116-121, 2011.
- [33] I.A. Pawelczak, S.A. Ouedraogo, A.M. Glenn, R.E. Wurtz, and L.F. Nakae, "Studies of neutron- γ pulse shape discrimination in EJ-309 liquid scintillator using charge integration method," *Nuclear Instruments and Methods in Physics Research, Section A: accelerators, spectrometers, detectors and associated equipment*, vol. 711, pp. 21-26, 2013.

- [34] A.C. Kaplan, M. Flaska, A. Enqvist, J.L. Dolan, and S.A. Pozzi, “EJ-309 pulse shape discrimination performance with a high gamma-ray-to-neutron ratio and low threshold,” *Nuclear Instruments and Methods in Physics Research, Section A: accelerators, spectrometers, detectors and associated equipment*, vol. 729, pp. 463-468, 2013.
- [35] R.A. Winyard, J.E. Lutkin, and G.W. McBeth, “Pulse shape discrimination in Inorganic and Organic Scintillators. I,” *Nuclear Instruments and Methods*, 95, pp. 141-153, 1971.
- [36] K.P. Lennox, P. Rosenfield, B. Blair, A. Kaplan, J. Ruz, A. Glenn, and R. Wurtz, “Assessing and minimizing contamination in time of flight based validation data,” *Nuclear Instruments and Methods in Physics Research, Section A: accelerators, spectrometers, detectors and associated equipment*, vol. 870, pp. 30-36, 2017.
- [37] W.L. Bryan, C.L. Britton, J.T. Mihalcz, J.S. Neal, S.A. Pozzi, and R.W. Tucker, “Fast Neutron-Gamma Pulse Shape Discrimination of Liquid Scintillation Signals for Time Correlated Measurements,” *2003 IEEE Nuclear Science Symposium. Conference Record*, vol. 2, pp. 1192-1195, 2003.
- [38] G. Ranucci, “A Review of the Statistical Foundations of Classical Pulse Shape Discrimination Techniques in Scintillation Applications,” *IEEE Symposium Conference Record Nuclear Science 2004*, vol. 2, pp. 804-809, 2004.
- [39] G.F. Knoll, *Radiation Detection and Measurement*, 4th edition, Hoboken: John Wiley & Sons Inc, 2010.
- [40] N. Tsoufanidis and S. Landsberger, *Measurement & Detection of Radiation*, 4th edition, Boca Raton: CRC Press, 2015.
- [41] J.K. Shultis and R.E. Faw, *Fundamentals of Nuclear Science and Engineering*, 3rd edition, Boca Raton: CRC Press, 2017.
- [42] R.P. Brent, *Algorithms for Minimization without Derivatives*, New Jersey: Prentice-Hall, pp. 47-60, 1973.
- [43] J. Unpingco, *Python for Probability, Statistics, and Machine Learning*, Switzerland: Springer Nature, 2nd edition, p. 83, 2019.
- [44] X. Zhu and A.B. Goldberg, *Introduction to Semi-Supervised Learning*, Morgan & Claypool, 2009.
- [45] W. Lee, *Python Machine Learning*, Indianapolis: John Wiley & Sons Inc., 2019.
- [46] *Data Sheet: Neutron/Gamma PSD Liquid Scintillator EJ-301, EJ-309*, Eljen Technology, 2018, accessed at https://eljentechnology.com/images/products/data_sheets/EJ-301_EJ-309.pdf.

- [47] L.M. Bollinger and G.E. Thomas, "Measurement of the Time Dependence of Scintillation Intensity by a Delayed-Coincidence Method," *The Review of Scientific Instruments*, vol. 32, pp. 1044-1050, 1961.
- [48] N. Zaitseva, A. Glenn, L. Carman, R. Hatarik, S. Harnel, M. Faust, B. Schabes, N. Cherepy, and S. Payne, "Pulse Shape Discrimination in Impure and Mixed Single-Crystal Organic Scintillators," *IEEE Transaction on Nuclear Science*, Vol. 58, 6, pp. 3411-3420, 2011.
- [49] B.C. Anderson, K.E. Holbert, and H. Bowler, "Design, Construction, and Modeling of a ^{252}Cf Neutron Irradiator," *Science and Technology of Nuclear Installations*, vol. 2016, 9012747, 2016.
- [50] Y. Hino and Y. Kawada, "Absolute ^{22}Na Radioactivity measurement by gama efficiency vairaiton of $4\pi\beta\text{-}\gamma$ coincidence method," *IAEA Japan*, M. Nakazawa (Ed.), KEK-PROC—94-7, pp. 202-209, 1994.
- [51] E.W. Weisstein, "Trapezoidal Rule," From *MathWorld--A Wolfram Web Resource*, <https://mathworld.wolfram.com/TrapezoidalRule.html>, accessed 16 June 2020.
- [52] E.W. Weisstein, "Gaussian Quadrature," From *MathWorld--A Wolfram Web Resource*, <https://mathworld.wolfram.com/GaussianQuadrature.html>, accessed 16 June 2020.
- [53] P. Virtanen, R. Gommers, T.E. Oliphant, M. Haberland, T. Reddy, D. Cournapeau, E. Burovski, P. Peterson, W. Weckesser, J. Bright, S.J. van der Walt, M. Brett, J. Wilson, K.J. Millman, N. Mayorov, A.R.J. Nelson, E. Jones, R. Kern, E. Larson, C.J. Carey, Í. Polat, Y. Feng, E.W. Moore, J. VanderPlas, D. Laxalde, J. Perktold, R. Cimrman, I. Henriksen, E.A. Quintero, C.R. Harris, A.M. Archibald, A.H. Ribeiro, F. Pedregosa, P. van Mulbregt, and SciPy 1.0 Contributors, "SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python," *Nature Methods*, 17(3), pp. 261-272, 2020, accessed at <https://docs.scipy.org/doc/scipy/reference/optimize.html>.
- [54] C. Audet and W. Hare, "Nelder-Mead," *Derivative-Free and Blackbox Optimization*, Springer: Springer Series in Operations Research and Financial Engineering, pp 75-91, 2017.
- [55] D.G. Luenberger and Y. Ye, "Quasi-Newton Methods," *Linear and Nonlinear Programming*, Springer: International Series in Operations Research & Management Science, vol. 228, pp. 285-317, 2016.
- [56] R.H. Byrd, P. Lu, J. Nocedal, and C. Zhu, "A Limited Memory Algorithm for Bound Constrained Optimization," *SIAM Journal on Scientific Computing*, vol. 16, 5, pp. 1190-1208, 1995.
- [57] C. Zhu, R.H. Byrd, P. Lu, and J. Nocedal, "Algorithm 778: L-BFGS-B: Fortran Subroutines for Large-Scale Bound-Constrained Optimization," *ACM Transactions on Mathematical Software*, vol. 23, 4, pp. 550-560, 1997.

- [58] Z. Fu, G. Liu, and L. Guo, "Sequential Quadratic Programming Method for Nonlinear Least Squares Estimation and Its Application," *Hindawi Mathematical Problems in Engineering*, vol. 2019, 3087949, 2019.
- [59] Scikit-learn: Machine Learning in Python, Pedregosa *et al.*, JMLR 12, pp. 2825-2830, 2011.
- [60] F.G.A. Quarati, P. Dorenbos, J. van der Biezen, A. Owens, M. Selle, L. Parthier, and P. Schotanus, "Scintillation and detection characteristics of high-sensitivity CeBr₃ gamma-ray spectrometers," *Nuclear Instruments and Methods in Physics Research, Section A: accelerators, spectrometers, detectors and associated equipment*, vol. 729, pp. 596-604, 2013.
- [61] User Manual UM3148, DT5730, 8-Channel 14-bit 500MS/s Digitizer, Rev. 1, Dec. 2014, Italy, CAEN.