Investigating the use of machine learning classification techniques for neutron/gamma ray discrimination in plastic scintillators

Nick Ferguson

September 18, 2020

Dr Matthias Treder Dr G J Sykora

School of Computer Science and Informatics

Cardiff University

Abstract

The use of Gaussian mixture models to discriminate between neutron- and gamma ray-induced signals was investigated. Different feature sets were created based on quantities used in traditional analogue and digital signal processing techniques and used to inform a Gaussian clustering algorithm. Results of the clustering algorithm were validated against standard metrics. Gamma sensitivity of the detector was approximately 1×10^{-4}, two orders of magnitude above (worse) than traditional discrimination algorithms. Thermal neutron event selection proved however to be more selective in its choice of events than traditional methods, by choosing 2.4% fewer events, which translates to an exclusion of unwanted fast neutron events.

Acknowledgements

Thanks to Dr G J Sykora of the ISIS Detector group for sharing this project.

Contents

1 Introduction

The Science and Technology Facilities Council (STFC) is an non-departmental government body responsible for directing public money in to a number of world-leading scientific facilities around the UK[1]. Among these facilities are Daresbury Laboratory, home to the Hartree Centre for high-performance computing; the UK Astronomy Technology Centre, which designs and builds instruments for many of the world's major telescopes; and the Rutherford Appleton Laboratory (RAL), itself home to facilities such as the Central Laser Facility, RAL Space, and the ISIS Neutron and Muon Source.

The ISIS Neutron and Muon Source, shown in figure 1, is a world-leading facility for neutron and muon science[2]. Alongside muon science, the facility is dedicated to performing neutron scattering, a non-destructive set of techniques which give insights to the properties of materials on an atomic scale. Neutron beams are produced in a multi-stage process delivered to instruments housed in two Target Stations, the layout of which is seen in figure 2.

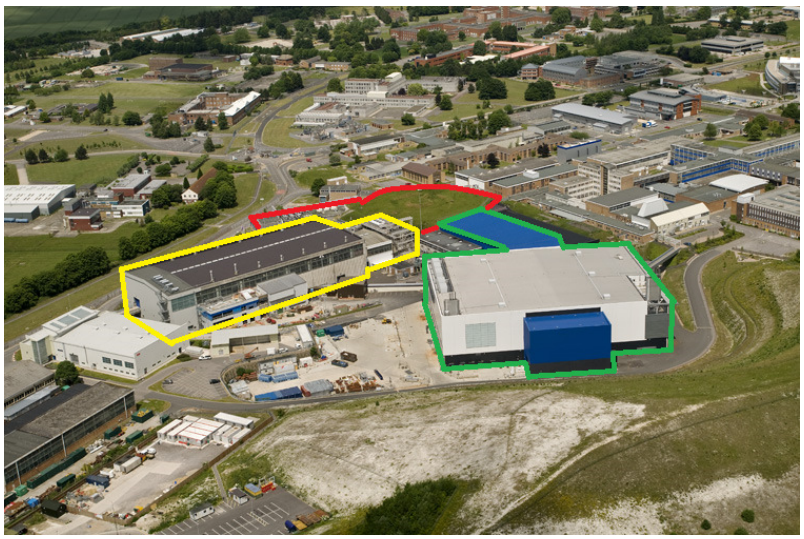


Figure 1: Aerial view of TS1 (yellow) and TS2 (green) with synchrotron mound and linac (red).

Firstly, an ion source produces two bunches of protons which are accelerated to an energy of 70MeV. The beam is then injected in to a synchrotron and resulting protons are guided in to a circular path by ten dipole magnets. Quadrupole magnets keep the beam focussed around the ring. After approximately 10,000 revolutions, the two proton bunches are approximately 100ns wide in time. This entire process is repeated 50 times

per second, resulting in an average beam current of 200μA. The beam is directed in to two beamlines, one of which is delivered to each target station after undergoing further focussing.

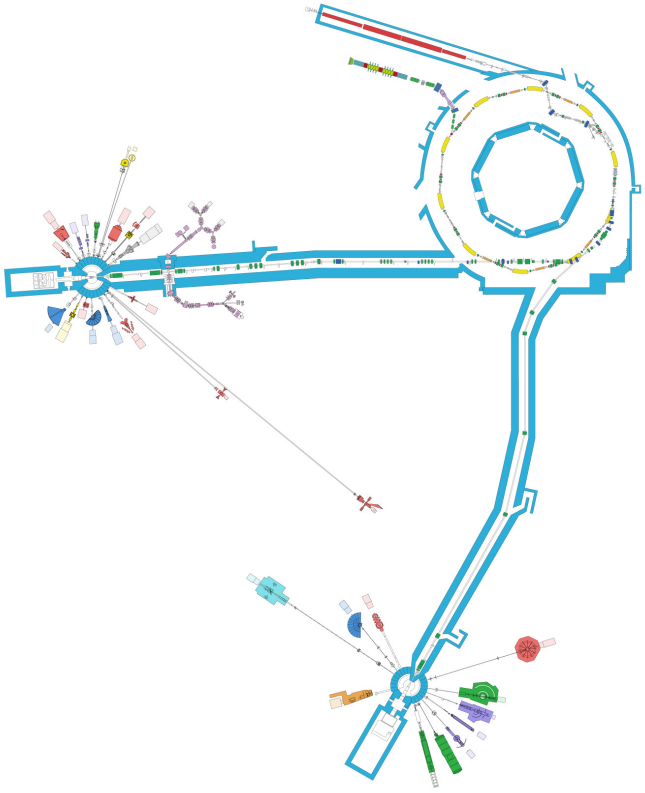


Figure 2: Layout of ISIS, showing ion source and linac (red line, top); synchrotron (yellow and green ring, middle); and proton beamlines to TS1 (left) and TS2 (bottom); with instruments positioned radially around the targets.

Neutrons are produced by *spallation*, during which the incoming protons physically eject neutrons from the target material. In Target Station 1, the proton beam is impacted on to a tantalum-clad tungsten target. Moderators slow down high-energy neutrons to useful energies. These neutrons are called *thermal* neutrons, and are of the most interest to neutron scattering experiments. However, the resulting energy distribution still includes some higher energy, *fast* neutrons. In Target Station 2, the production of long-wavelength neutrons is maximised for applications such as soft matter research. The instruments on ISIS are open to users around the world for neutron scattering experi-

ments[3]. Users can apply for beamtime on an instrument during one of the multiple user cycles which run throughout the year.

2 Aim and Objectives

The aim of the project is to investigate the feasibility and performance of a machine learning-based classifier for the purpose of classifying different types of signals from neutron detectors, with the goal of outperforming traditional methods. Specific objectives of the project include:

- Reviewing the current state of the art methods for utilising machine learning for the purpose of electronic signal processing
- Reviewing existing techniques for performing discrimination
- Investigating the feasibility of using machine learning-based techniques for classifying different types of electronic signals
- Comparing and contrasting the relative successes and failures of machine learning-based techniques and traditional analogue electronic processes in the context of signal processing
- Produce a deployable application which can be used by experimental scientists to utilise the final machine learning models.

3 Background

3.1 Scintillation-based detectors

Detectors on ISIS instruments consist primarily of ^{3}He-filled gas tubes or zinc sulphide-based scintillation detectors[3]. While the fine detail of the scintillation process will not be discussed, broadly speaking, scintillation-based detectors exploit the generation of photons due to captured particles. The duration of the scintillation process is affected by the type of particle that is absorbed: massive particles such as neutrons generate a large amount of photons and continue to release photons for a relatively long time after the particle has impacted the detector (afterglow). However, while gamma rays can be

as energetic as neutrons, the release of photons due to an incident gamma ray dies away very quickly after impact due to the physics of the scintillation material[4]. Scintillation photos pass through optical fibres and are collected by a photomultiplier tube (PMT), which produces an electrical signal proportional to the number of photons it receives. A family of techniques, broadly known as Pulse Shape Discrimination (PSD), form the basis for distinguishing between neutron- and gamma ray-induced pulses. These will be described later.

ISIS and other facilities will soon generate higher neutron fluxes. When exposed to these higher rates, current scintillator-based detectors do not have the ability to distinguish between events located this close to each other in time, due to afterglow[5]. However, lithium-loaded plastic scintillators, particularly Eljen Technologies' EJ-270, a sample of which is shown in figure 3, are being investigated for their ability to cope with high fluxes due to their fast decay and PSD capabilities[6][7].

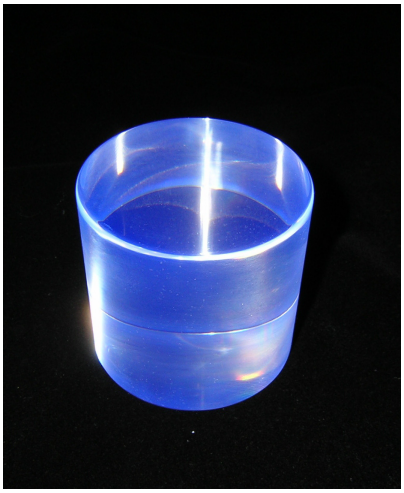


Figure 3: A sample of a lithium-loaded plastic scintillator emitting light due to incident ambient light[8].

3.2 Content of dataset

The training dataset used in this project contains approximately 28,000 events, each caused by either a neutron or gamma ray. The dataset was collected using a prototype detector containing a sample of EJ-270 as its scintillating material on an Americium-Beryllium source. Each event is a time series at 2ns intervals containing the amplitude of the pulse. Due to the PMT, what is effectively being measured is the light output of the

scintillator as a result of an impacting particle. There are 300 data points per pulse, covering a 150ns time period. Figure 4 shows an 80ns window, containing a typical neutron- and gamma ray-induced pulse. Both have had a background subtraction performed and are normalised to the peak in order to illustrate their different characteristics.

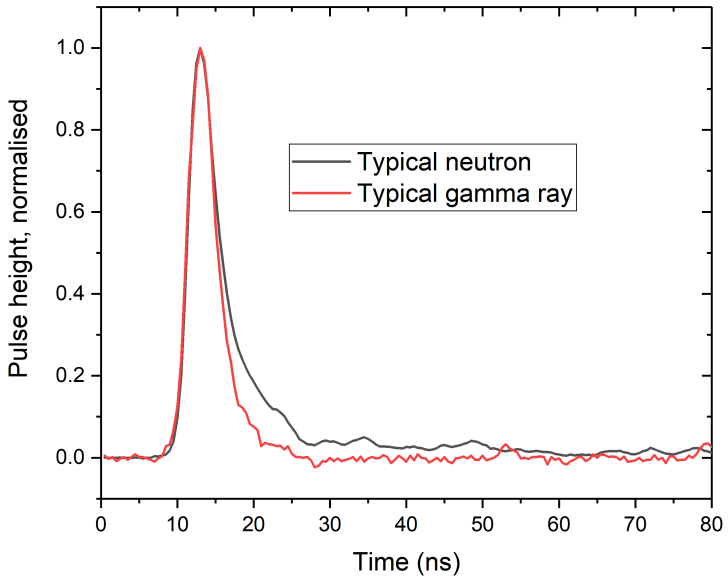


Figure 4: Typical neutron and gamma rays as they show up in the detector.

The gamma ray pulse was taken from a lower energy event, which explains why the background is noisier. The region of approximately 20-60ns is termed the *tail* of the pulse. The tail is not rigorously defined, but describes a general region of the pulse shortly after the peak, during which the detector material is still scintillating. The afterglow caused by the incident neutron is clearly visible in the region 30-50ns, whereas the signal has already returned to background in the gamma ray pulse. This illustration shows that the tail is a key distinguishing factor between neutron- and gamma ray-induced pulses. The relative portion of the pulse contained within the tail can be encoded in the *tail sum*, which is simply the sum of all data points within the tail. Figure 5 shows a scatter plot containing each pulse represented by it's peak amplitude (known from hereon as *pulse height*) along the x-axis and tail sum along the y-axis.

Considering the figure, one can visually separate three different regions. The region denoted A is the thermal neutron peak. The upper branch, denoted B, is the fast neutron region. The lower branch, denoted C, is the gamma ray region. A key point, and a problem which this study aims to address, is that not only are the boundaries between the regions diffuse and not fully resolvable, but each region does not exclusively contain

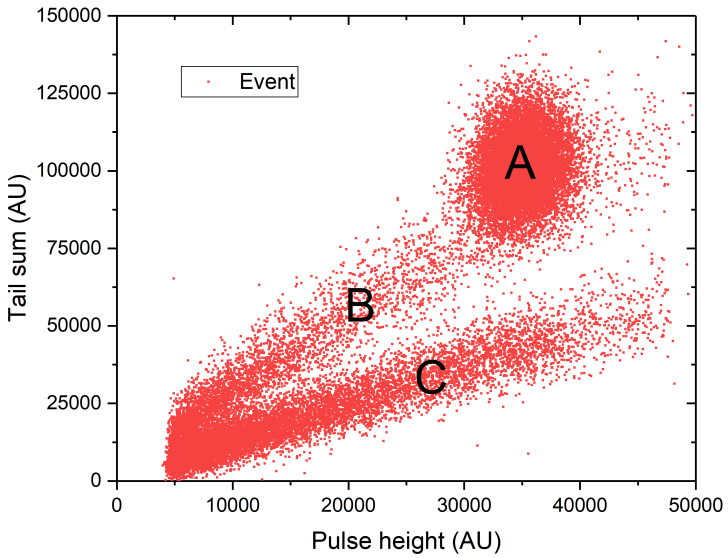


Figure 5: Each pulse in the dataset, represented by its height and tail sum

events of that type. For example, region A does not contain exclusively thermal neutron events, but certainly includes fast neutron events, as it sits on top of the branch of region B. Region A also likely contains some, but not many, gamma ray events.

This discrimination problem is part of a wider signal processing and data collection routine at ISIS. Events output by discrimination algorithms are passed to the data collection programs running on ISIS instruments. This is the data used to form the scientific analysis of experiments being performed by visiting scientists. As such, a key problem is not only discriminating between neutron and gamma ray events, but making sure that only thermal neutron events are selected, as opposed to fast neutron events. Traditionally, a bounding box is drawn around region A, and every event inside the limits will be classified as a thermal neutron. However, this results in many fast neutrons and gamma rays being classified as neutrons, which may contaminate scientific analysis.

3.3 Existing discrimination techniques

Distinguishing between neutron and gamma ray-induced signal is done via a broad range of techniques called Pulse Shape Discrimination (PSD). Methods belonging to this family are examined on a technique-by-technique basis. The use of machine learning in neutron/gamma ray discrimination is also surveyed, but on a study-by-study basis, because the algorithms and techniques used are more intrinsically linked.

3.3.1 Traditional analogue electronic discrimination

One of the most fundamental methods of PSD is the Charge Comparison Method (CCM). The CCM is a measure of the portion of light emitted in the different components of the scintillation process[9]. One way of implementing this method is to select two time windows, and for each pulse sum the amplitude values within each window and compare their relative areas. Values of the start and end point of these time windows, known from now on as *integration gates*, are optimised to provide the best figure of merit[10][11]. It has been established that the proportion of light released in the tail of the pulse is what distinguishes pulses caused by neutrons from pulses caused by gamma rays, so choosing time windows which are able to encode this difference is important.

Pulse Gradient Analysis (PGA) exploits fast and slow decay components of scintillation, similar to CCM[12]. However, instead of using integration gates based on these time intervals, the gradient of the pulse is taken at two points. D'Mellow used the peak of the pulse as the first point, and another sample 20ns after the peak. However, it is noted that properties of the experimental setup, such as type of scintillator and PMT may affect the optimal choice of the second time point.

Fourier-based methods are also popular. Similar to PGA, Frequency Gradient Analysis (FGA) exploits the difference between different components of the Fourier transform of pulses and has been shown to outperform PGA[13]. Rather than using raw pulse data, pulses are modelled as the superposition of two or three exponential decay functions[14]. Fourier Area Analysis (FAA) examines the area under the Fourier transformed pulse in order to perform discrimination and has been shown to outperform FGA and PGA[10].

In addition to the findings of this section, algorithms being investigated by Dr Sykora use custom figures for the window used to define the tail sum, as well as start and end points for long and short integration gates. These were used as starting points for features which informed the machine learning portion of the study.

3.3.2 Machine learning-based classification

Kaplan[15] proposes a classifier which utlises pure gamma ray sources and mixed neutron/gamma ray sources, similar to what is available at ISIS. Kaplan works with Tail-To-Total (TTT) figures, separating the neutron and gamma regions using a scatter plot with one axis containing the integral of the background-subtracted pulse and the other axis

measuring the fraction of the signal energy that the tail of the pulse contains. This is related to the CCM. Kaplan's classifier and the TTT method were tested on gold standard labels provided by a Time of Flight (TOF) labelling, which measures the time taken for the particle to arrive based on a trigger signal which activates when pulses of neutrons are generated. Gamma rays, travelling at the speed of light, arrive almost instantly, whereas slower, massive neutrons arrive afterwards. Pulse models were constructed based on a mixture of normal distributions in order to encode the different factors which influence the pulse shape. An Expectation-Maximisation algorithm is used to train the classifier, which outputs parameters which represent the pulse shapes with the highest likelihood for a pure and mixed source. In the testing phase, a score is generated for each pulse based on the pulses generated by the normal distribution modelling, using the parameters derived in the training phase. Findings from the study showed improved performance over a standard TTT method.

Yu et al.[16] adopted a supervised learning method using Support Vector Machines (SVMs). The CCM informs a figure of merit which is used to discriminate between neutrons and gamma rays. Obtain a training set for the SVM was performed by selecting regions of a scatter plot charting the energy of each event as a function of a figure of merit. The scatter plot separated neutron and gamma ray events reasonably clearly above a certain energy. Labels were assigned to these data points and used for training, while the rest of the events were used for testing. A Moment Analysis Method, described by Xie[17], is used to create feature vectors for each event. Only 3000 events of each type were used for training, and based on a test set containing 6,613 neutron events and 36,922 gamma ray events, they report an error of 0% and 0.04% respectively. Xie notes the advantage of the SVM method in that it can provide an accuracy score for each classification.

3.4 Unsupervised learning

A number of machine learning paradigms exist, the appropriateness of which depends on the problem at hand. In a *supervised* approach, a *learner* has access to the inputs and correct outputs of the system it is trying to create a model for. For example, consider task in which a learner which is required to determine whether or not a photo contains a cat or a dog. The learner will be fed examples of images which are labelled as containing

a cat or a dog. After a model is generated, based on this *training* data, it can be used on new, unseen images (*test* data) to predict a label for it, that is, whether it contains a cat or a dog. In contrast, *unsupervised* learning does not deal with training or test data. Instead, the learner aims to produce a representation of that data by grouping it in to useful subsets. The learner does not have access to what the correct outputs are[18].

Clustering is an unsupervised approach to machine learning which aims to identify meaningful subsets of data based on the parameters which make up each data point. Clustering algorithms use an iterative process, such Expectation-Maximisation[19] to arrive at an optimised solution. Mixture models are a type of unsupervised clustering concerned with resolving mixtures of distributions, which describe features of data[20] and are suited well to this project, as the dataset can be considered to be made up of a mixture of three distributions. A number of types of algorithm for performing clustering exist. Linkage-based clustering algorithms start with each point its own cluster, and then merges the closest clusters. Other algorithms, such as the k-means algorithm, define a cost function, which is to be minimised. Spectral clustering encodes the relationship between data points as nodes on a graph, connected by weights. The algorithm then seeks a partition of the graph between groups which have low weights and high weights[21].

3.5 Approach

While Yu performed labelling based on which region of a scatter plot the event was located in, this surely introduced contamination to the system due to the mixed nature of the data. Additionally, only events above a certain energy were labelled, causing the learning algorithm to only be exposed to a subset of the full range of possible neutron and gamma ray events. Kaplan's custom algorithm is a more sophisticated approach to creating labelled data by using time-of-flight information. However, that information is not available in this study. Additionally, Kaplan used a much more bespoke algorithm, but creating something in a similar vein was infeasible given the timeframe of the project.

In order to avoid contaminating the training dataset by using a substandard labelling approach, it was chosen to follow an unsupervised approach. A more bespoke approach such as Kaplan's may well have restricted the available feature sets, so a more generic, out-of-the-box approach allowed the scope of the study to remain wide, while allowing for a potential labelling, or other additions, later on.

4 Experimental

4.1 Model selection

The `sklearn.cluster` package[22] was chosen to perform the core clustering algorithms. It provides a number of different methods for clustering unlabelled data, including k-means clustering, the DBSCAN algorithm and its generalised form OPTICS, hierarchical clustering and Gaussian mixture models. Superficial exploratory analysis was performed in order to determine which clustering algorithm to use. This was validated through use of pulse height/tail sum plots such as in figure 5 as well as considering the mechanics of each algorithm. This approach was sufficient because the difference between algorithms was visible on a macroscopic scale, so thorough numeric validation was not deemed necessary.

k-Means clustering was ruled out as it aims to create clusters of equal variance, which is not something that is desired nor does it reflect the physics at the heart of the problem. The variance-equalising nature of k-means was clearly shown in a simple implementation, whereby cluster boundaries were clearly attempting to divide the data equally between clusters.The DBSCAN algorithm is able to discover clusters of arbitrary shape via initial discovery of high-density areas[23]. Additionally, there was potential for the output of the DBSCAN algorithm to contain a noise labelling. This was not desired, as noise reduction had already been performed to a sufficient level: we were happy for any event in the dataset to be labelled a thermal neutron, fast neutron or gamma ray. This can again be seen in figure 5 whereby signals which contained a pulse height below approximately 4500 were rejected. Similar results were obtained using OPTICS.

By making the approximation that the dataset consisted of a mixture of Gaussian distributions, the Gaussian mixture modelling algorithm suited well. Clusters formed by this algorithm are ellipsoidal in shape, which makes it suitable for fitting eccentric ellipsoids to both long arms of the fast neutron and gamma ray regions as well as the more circular thermal neutron peak. In addition to the data that is provided to the Gaussian clustering algorithm, a number of other parameters can be specified to the model - of particular importance is the number of components that the algorithm needs to create. Initially, both 2- and 3-component models were generated. The rationale behind generating 3-component models is clear: one component for each region of the pulse height/tail sum plot shown in figure 5. Searching for 2-component models was a

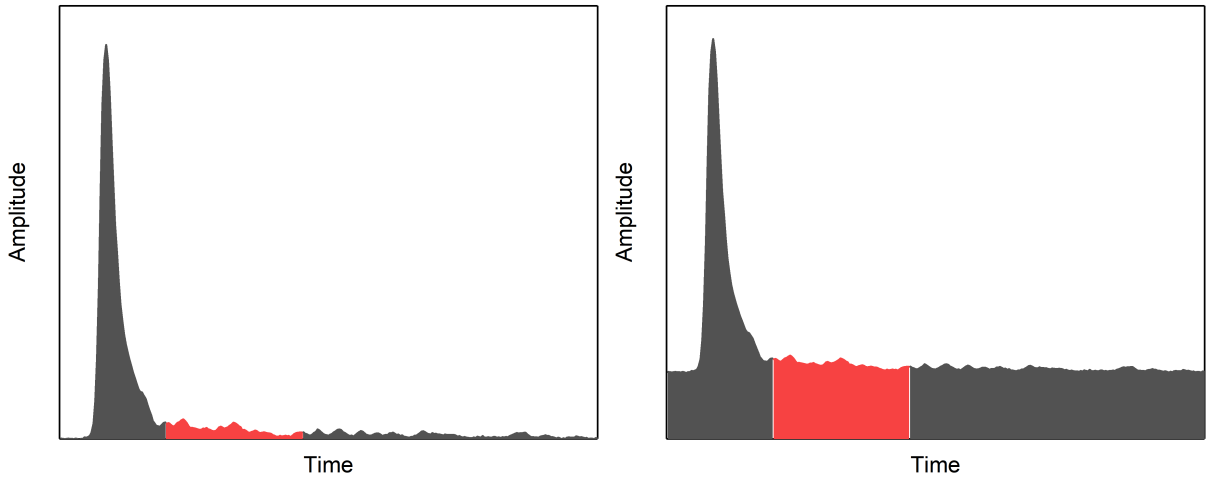
first attempt at creating a model which successfully localised the thermal neutron peak in one cluster, and "everything else" in the other - with the fast neutron and gamma ray regions merged. It soon emerged that artificially increasing the number of components to 4 provided interesting results. As will be demonstrated later, this forced the splitting of the fast neutron and gamma ray regions in to three clusters, which proved to isolate the thermal neutron peak from the upper end of the fast neutron branch better. Another parameter which could be passed to the clustering algorithm was an initial guess for the midpoint of each cluster. By selecting these to be roughly in the middle of each region, it was hoped that the algorithm would converge to a solution consisting of two ellipsoidal distributions capturing the fast neutron and gamma ray regions, and a more circular distribution fitted to the thermal neutron peak. However, despite multiple attempts, it was not possible to generate this result.

4.2 Feature creation and selection

One of the main tasks was selecting the information to pass to the clustering algorithm. It was important to provide enough features to enable a good-quality discrimination, but also to not overload the system with irrelevant data. The outcome of section 4.2.4 prompted returning to earlier feature sets to alter the way they were created. Different bounds for the integration gates were selected based on a visual assessment of what the normal distribution fits failed to capture. Additionally, a background subtraction was also performed in order to make any integration-based figures more polarised. As figure 6 shows, the proportion of the red area relative to the whole pulse area (grey and red areas combined) is a more representative when background subtraction has been performed.

4.2.1 Tail sums

As a first attempt at a clustering model, only two features were used - the pulse height and tail sum. Since ultimately this study will be compared to the results of Dr Sykora's work, the region used for the tail sum was kept the same to start with. The original tail sum was based on a value which maximised a figure of merit that was used in his approach, however it became clear that this value might not be the best approach for the clustering algorithm. 2, 3 and 4-component models were generated. In the 2-component model, it was found that the cluster containing the thermal neutron peak tended to leak too far



(a) With background subtraction.

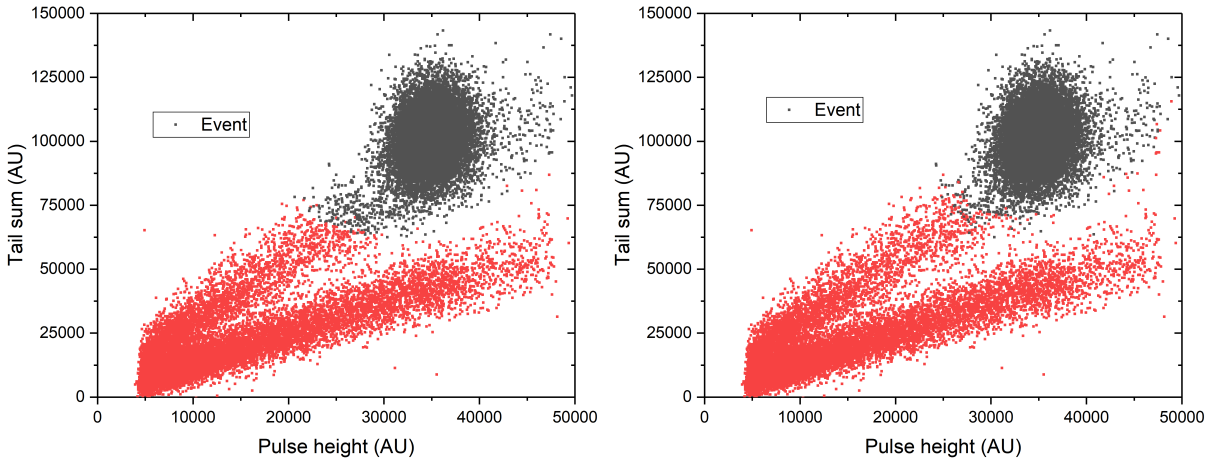
(b) Without background subtraction.

Figure 6: Results of running the clustering algorithm on different numbers of clusters with the pulse height and tail sum as features.

in to the fast neutron region, shown in figure 7a. A common theme with the 3-cluster approach was that it did not have the intended effect of separating the data in to one distribution per region. Instead, one distribution was generated centred on the densest area where the fast neutron and gamma ray branches merge. The second covered the less dense areas of both branches, while the third captured the thermal neutron peak well. By artificially increasing the number of clusters to 4, the fast neutron and gamma ray branches were split in to three clusters and the thermal neutron peak was better isolated from the fast neutron branch. This is achieving the intended effect of labelling neutrons in the fast neutron branch as fast neutrons rather than being labelled thermal neutrons. A modified tail window was also tested, based on a visual assessment of pulses, which led to much less leakage of the thermal neutron cluster in to the fast neutron branch, even in a 2-component model, as shown in figure 7b.

4.2.2 Tail-to-total

Kaplan's tail-to-total figure was trialled instead of using the tail sum. The tail-to-total values were obtained by dividing each pulse's tail sum integration by the total pulse integration. Since two values for the tail sum were investigated, this led to two versions for the tail-to-total value. Observing the 3-cluster models this time, shown in figure 8, the tail-total figures created using the original tail sum still leaked considerably in to the

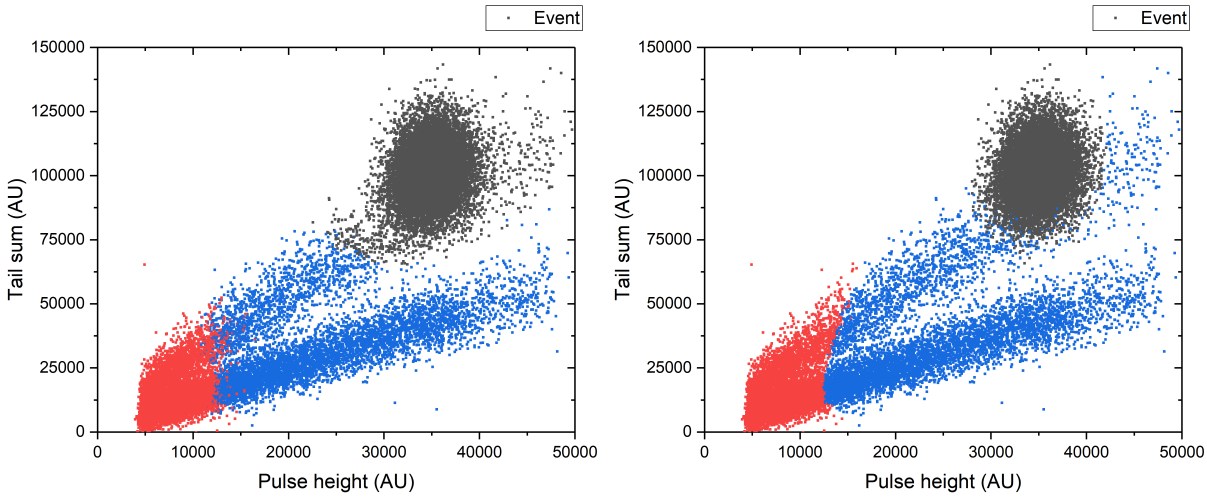


(a) Original tail sum values.

(b) Modified tail sum values.

Figure 7: 2-component cluster modelling using both the original and modified tail sums.

fast neutron region, whereas the tail-to-total figures created using the modified tail sum proved to isolate the thermal neutron peak exceptionally well.



(a) Tail-to-total values created using the original tail sum region.

(b) Tail-to-total values created using the modified tail sum region.

Figure 8: 3-component cluster modelling using the pulse height and tail-to-total figures.

4.2.3 Integration gates

Findings from the survey of existing discrimination techniques showed that use of different integration gates was a common way to characterise the pulse (the CCM). It has been established that the length of short and long gates do not have rigid definitions, but are a result of optimising a figure of merit. In order to reduce the risk of losing information

by condensing these two values in to one, the raw values were used as features. While the values for the start and end point for both of these gates could still be optimised, this was not performed because the process would take up a not inconsiderable amount of time and any improvements likely marginal. Both gates opened at 0ns to ensure that the rise of the pulse was taken in to consideration. The short gate ran until 15ns, while the long gate ran until 67.5ns. Figure 9 shows a model pulse generated using Marrone's model[14] along with an approximation of the integration gates.

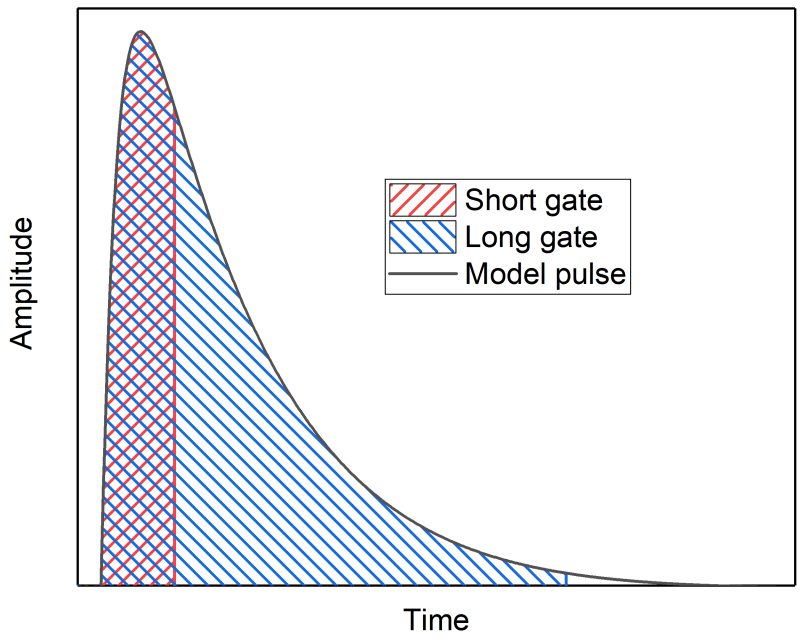
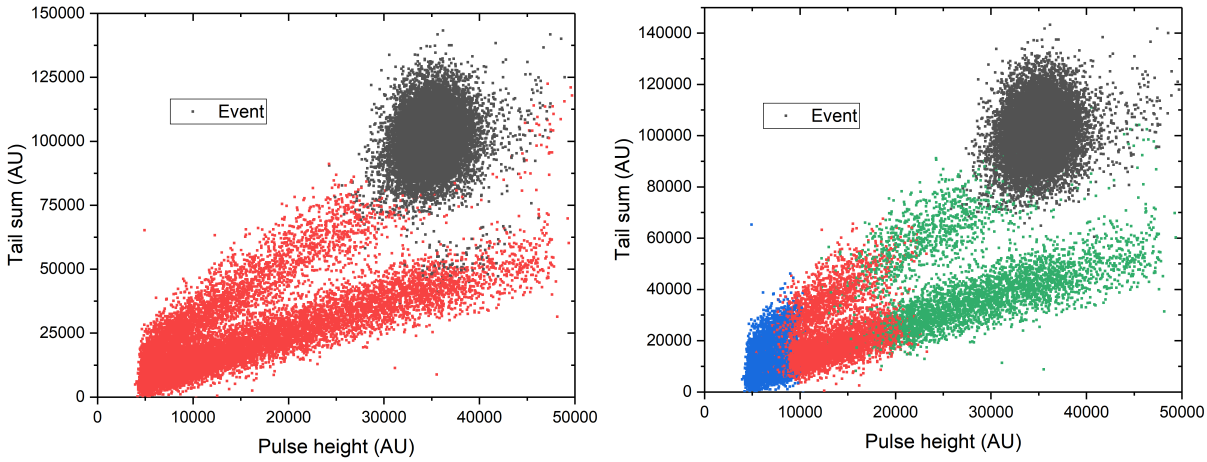


Figure 9: Demonstration of different integration regions on a model pulse.

As an exploration of these features, models were generated using only the short and long gate integration values as features. However, these exhibited a significant leakage of the thermal neutron cluster (shown in figure 10a) down in to the gamma ray branch which persisted in to 4-component models. Incorporating pulse height and the modified tail-to-total values for each pulse in to the feature set significantly improved the clustering, as shown in figure 10b.

Performance was comparable to the clustering obtained by using only the pulse height and tail-to-total values. However, some leakage was still present outside of the thermal neutron peak, particularly in to the fast neutron branch containing pulse heights of $x >$ 40,000.



(a) 2-cluster modelling using only short and long gates.

(b) 4-cluster modelling incorporating the pulse heights and modified tail-to-total values.

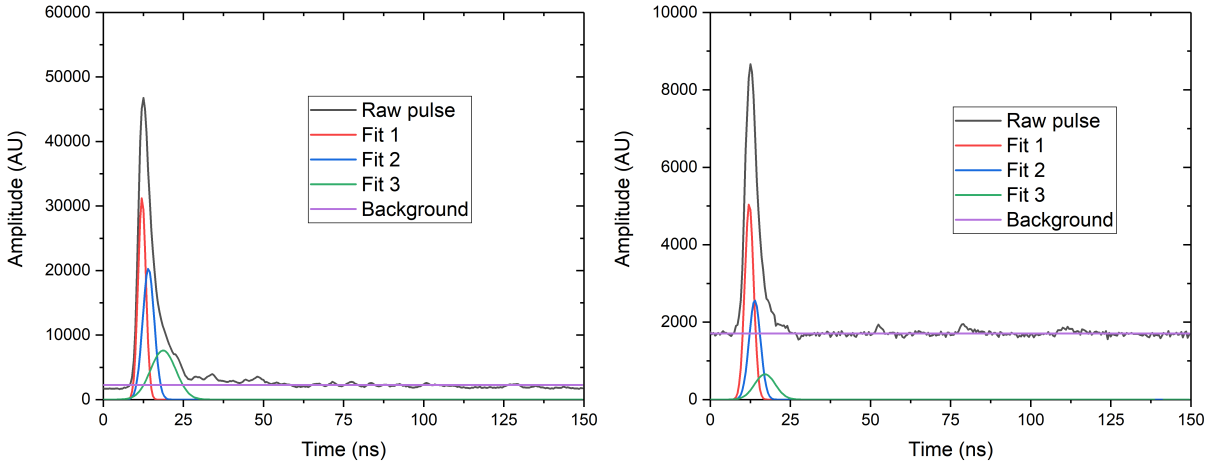
Figure 10: Incorporating the pulse height and tail-to-total values was necessary in order to obtain a good clustering.

4.2.4 Normal distribution modelling of pulses

In order try and better encode the overall shape of the pulse, modelling each pulse as the superposition of a number of normal distributions was explored. This was performed using `scipy.optimize.curve_fit`[24], a least-squares minimisation function, which takes a template function whose parameters are to be optimised, and a set of data to fit it to. The resulting parameters of these fits (amplitude, mean and standard deviation for each distribution) formed the features which were passed to the model.

Considering the physical origins of scintillation, fitting a single distribution to each component of the scintillation process should yield around 5 different distributions, which was the initial number fitted. However, even when the parameters' upper and lower bounds were constrained, frequently the fit function would fit 4 distributions, but often only 3. While good fits decompose the bulk of the pulse in to distinct distributions and also fit to the background, they fail to capture the information in the tail. The poor 5-distribution fits often lacked any contribution for 2 out of its 5 fitted distributions. In order to achieve a more consistent and reliable fitting outcome, it was decided to reduce the number of distributions fitted to 3, plus a constant background as shown in figure 11. A constant background was fitted as opposed to a Gaussian background because, since the background and tail portions of the pulse are so shallow, the mean and standard deviation

of these parameters varied considerably, too much for any meaningful contribution.



(a) Note how the tail is still not being fit.

(b) A typical 3-distribution fit.

Figure 11: More consistent 3-Gaussian fits plus a uniform background.

After restricting the number of distributions to 3, ensuring that all distributions were fitted in the same order for each pulse was imperative, otherwise the clustering algorithm would be comparing different parts of the pulse from one pulse to the next. For example, should this constraint *not* be kept, and the distributions were fit in a random order, the clustering algorithm could be comparing the distribution fit to one pulse's rise to the distribution fit to the next pulse's tail. In other words, for a fit comprising 3 distributions with parameters $(a_1, m_1, s_1, a_2, m_2, s_2, a_3, m_3, s_3)$, we require that $m_1 < m_2 < m_3$ for each pulse. This was ensured, as much as it could be, by providing the fitting function with an initial guess for each parameter, as well as upper and lower bounds. While the 3-distribution method proved the most consistent and reliable, its ability to encode the very shallow afterglow region of the tail was still limited.

Despite the effort recorded here, clustering results for this feature set were the worst of all. The algorithm arrived at a solution placing the vast majority of events in one cluster, and only a few fringe events in other clusters, shown in figure 12.

5 Results and discussion

As mentioned earlier, the neutron/gamma ray discrimination process is part of a wider data analysis pipeline. Data output from detectors and their discrimination algorithms goes on to be used in the experiment that is being performed on the instrument, so

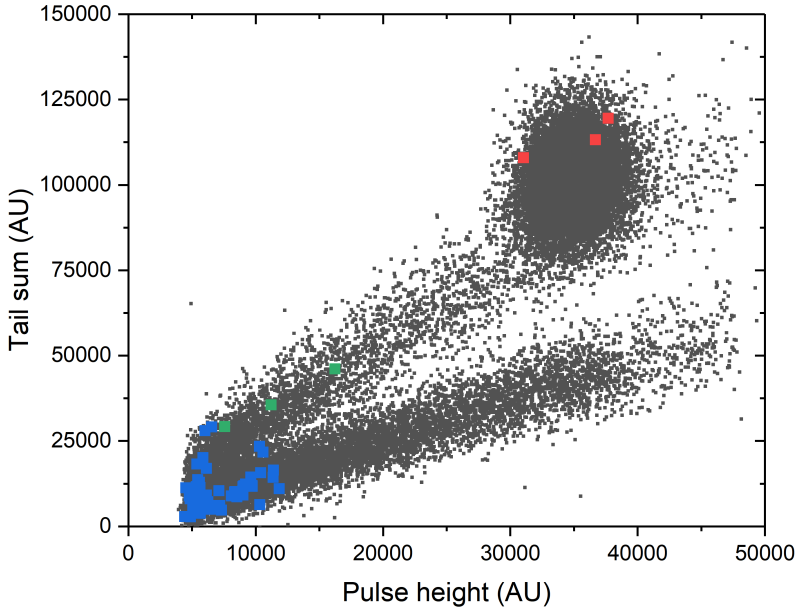


Figure 12: Clustering using Gaussian fits of pulses

ensuring each event is classified correctly is critical. The models' performance were numerically assessed by two metrics. The *gamma sensitivity* is the measure of the proportion of gamma ray events are labelled as being caused by neutrons. Minimising this quantity is important in order to reduce the number of gamma rays that go on to inform the scientific analysis of the experiment, as they are simply noise, but labelled as useful data. The second metric is relates to the discrimination between the thermal neutron peak and the fast neutron and gamma ray branches. Not only is it important that gamma rays are not labelled as thermal neutrons, but it is important that fast neutrons are not labelled as thermal neutrons either. This second metric measures the proportion of events which the model labels as thermal neutrons, relative to the traditional method of thermal neutron labelling. Minimising this quantity is also of interest, as this represents a more selective model. Data from two feature sets were not numerically validated: the normal distribution fits to each pulse and short and long integration gate models. Feature sets in this section are referred to by an abbreviation according to table 1.

5.1 Gamma sensitivity

As has been established, the dataset was taken on a single source consisting of mixed thermal neutrons, fast neutrons and gamma rays. While it was not possibly to obtain pure neutron data, a pure gamma ray source exists which enables the gamma sensitivity

Abbreviation	Long name
PH	Pulse height
TS	Tail sum (original)
NTS	Tail sum (new)
TTT	Tail-to-total using TS
NTTT	Tail-to-total using NTS
SGI	Short gate integration
LGI	Long gate integration

Table 1: Abbreviations of feature names.

to be validated. Gamma ray data was collected from a cobalt-60 source and passed through the model, which enabled a label to be predicted to be made for each event. Only gamma rays with a pulse height in the range $0 < x < 50,000$ were used in the testing of the model. This is because the energy spectrum of the gamma ray source was different to that of the source on which the training dataset was taken, and contained events with larger pulse heights, which the model was not trained on. Figure 13 shows the full gamma ray dataset, along with the cutoff point. Label predictions were only made for events in the red region.

Table 2 shows the gamma sensitivity values for different component numbers and feature sets. Out of the approximately 21,000 events which were able to test the gamma sensitivity, the best results were achieved using a 3-component clustering and the feature set consisting of the pulse heights, where only 2 events were labelled as neutrons. Data from two features sets was not numerically validated: the normal distribution fits to each pulse and short and long integration gate models. While there is not much variance between the best results, it becomes clear that selecting fewer high-quality features is more important than both the number of clusters that the model searches for and the number of features. Even the 2-cluster model using the fundamental feature set [PH, TS] proved create a better model than one which contain more features. The [PH, NTTT, SGI, LGI] was known to contain good-quality features (as shown in the [PH, NTTT] results), and the [SGI, LGI] feature sets produced acceptable results, but the combination of these two

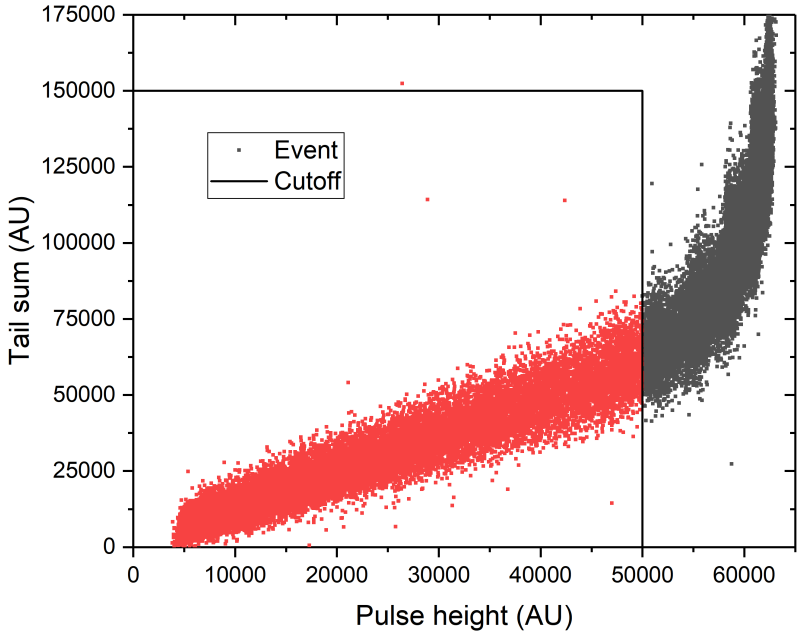


Figure 13: Each pulse in the pure gamma ray dataset, with the restricted area shown.

did not increase the quality of the clustering.

n	**PH, TS**	**PH, NTS**	**PH, TTT**	**PH, NTTT**	**SGI, LGI**	**PH, NTTT SGI, LGI**
2	6.69×10^{-4}	1.00×10^{-3}	6.69×10^{-4}	3.34×10^{-4}	2.02×10^{-2}	3.19×10^{-1}
3	6.69×10^{-4}	1.43×10^{-3}	6.69×10^{-4}	$\mathbf{9.55\times10^{-5}}$	8.83×10^{-3}	3.20×10^{-3}
4	2.44×10^{-3}	1.43×10^{-3}	2.39×10^{-3}	4.78×10^{-4}	2.34×10^{-3}	2.34×10^{-3}

Table 2: Gamma sensitivity of different feature sets for n clusters.

Key to improving the fitting is considering which gamma ray events were labelled as neutrons. This is shown in figure 14. The event located at (28896,114264) contains an anomalously large tail sum for a gamma ray of that pulse height. While there could be a case for removing this event by classing it as an outlier, it is no doubt caused by a gamma ray and therefore there is a possibility of observing more events like these in the detector. In fact, gamma rays in this region which are classed as neutrons are the most critical events to discriminate between, as these are gamma rays which the system otherwise considers to be neutrons.

For each event mis-labelled as a neutron, the gamma sensitivity increases by 4.78 ×

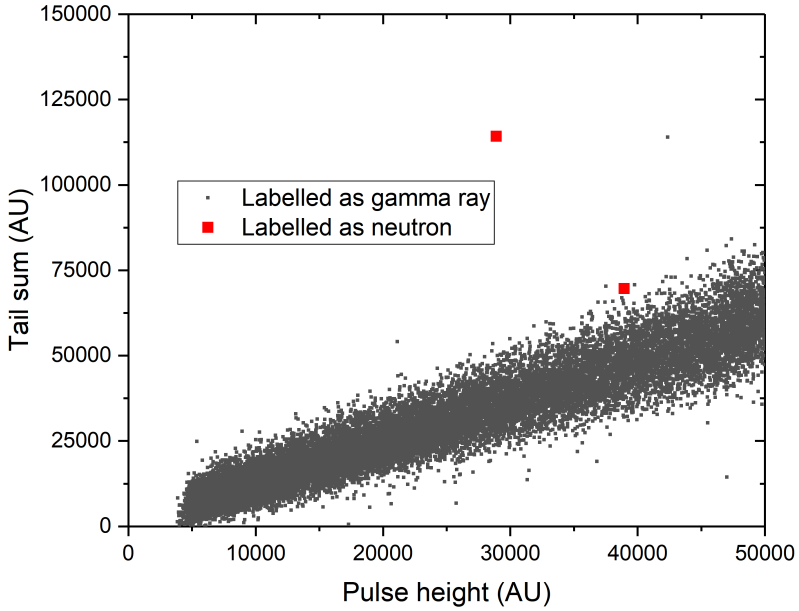


Figure 14: Caption

10^{-5}. By approximating the uncertainty in the gamma sensitivity as the precision of the measurement, this gives an error of $\pm 50\%$ for the best result. This is proportionally large compared to the absolute value, but reducing this and obtaining more accurate values is entirely possibly by using a larger test dataset.

5.2 Thermal neutron peak

The traditional method employed to select thermal neutron events is to simply create a bounding box around the thermal neutron peak and label everything inside that a thermal neutron. However, this causes fast neutron and likely some gamma ray events to be labelled as thermal neutrons, which is not a desirable outcome. Gaussian clustering presents a more selective approach as events within the thermal neutron peak can easily be extracted from the model. Traditionally, any event with a pulse height between the values of $28,000 < x < 42,000$ and tail sum within $70,000 < y < 135,000$, shown in figure 15 would be labelled a neutron.

When assessing the number of neutrons that each model placed in the thermal neutron peak, the bounding box was also applied such that an event was only considered a thermal neutron if both the model placed it in the cluster which contains the thermal neutron peak, *and* it lied in the region $28,000 < x < 42,000$ and $70,000 < y < 135,000$. This enables comparison to the traditional method, as well as being necessitated by some 2-

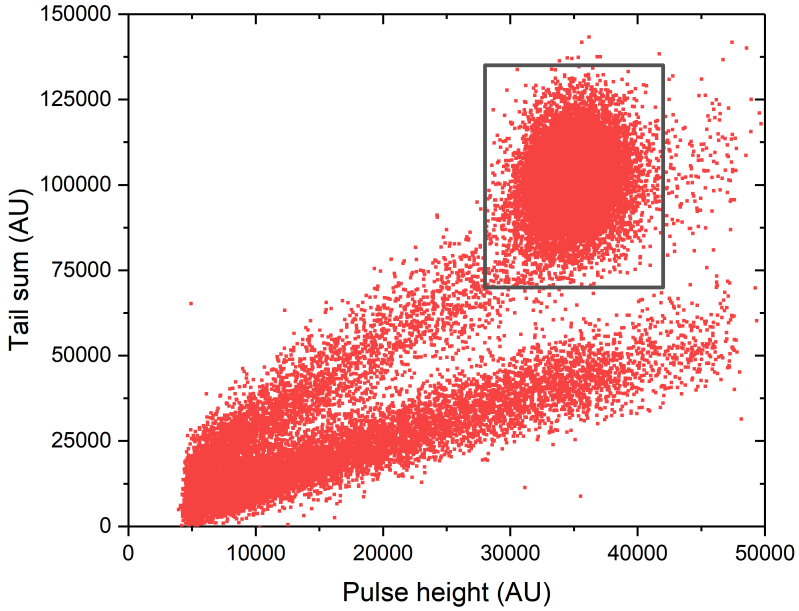


Figure 15: Bounding box for thermal neutron classification

component models' whose thermal neutron cluster leaked too far in to the fast neutron branch. Since we are looking for a more refined approach to thermal neutron selection, the lower amount of thermal neutrons that a model selected, the better. While this means more time is required to collect the same amount of data, slower run times but a higher quality of thermal neutron discrimination are preferable to faster run times and poorer discrimination. Table 3 shows the proportion of thermal neutrons that were labelled as such by the model compared to the number of neutrons within the bounding box. A value of 1 indicates that the model was no more selective than the bounding box method.

n	**PH, TS**	**PH, NTS**	**PH, TTT**	**PH, NTTT**	**PH, NTTT SGI, LGI**
2	1.000	0.998	1.000	0.986	1.000
3	1.000	0.994	1.000	0.986	0.996
4	0.989	0.988	0.988	0.997	0.993

Table 3: Proportion of thermal neutrons labelled as such by the model.

6 Conclusion and further work

An unsupervised learning approach has been successfully implemented to perform discrimination between neutron- and gamma ray-induced pulses in a detector. While the gamma sensitivity of the model is around two orders of magnitude higher than using traditional methods, a larger test dataset may improve this value. Selection of thermal neutron events has been shown to be 2.4% more selective than a traditional bounding box approach.

Perhaps the biggest change which could be made to the study is to find a way to label the data. There are a few options for this. Firstly, manually labelling each data point is technically one option, but this would be extremely labour-intensive and distinguishing between thermal and fast neutrons would be difficult and not entirely accurate. An approach similar to Kaplan, using time-of-flight data seems like the best "fast" way to generate labels. However, while ISIS does have time-of-flight data available, gamma rays are generated by neutrons colliding with beamline equipment throughout their travel time such that they are present throughout the neutron pulse's arrival, not just concentrated at the start when the neutron pulse is generated. Should a reliable (enough) labelling method be found, this enables the possibility of more sophisticated supervised learning techniques, such as SVMs.

The addition of different feature sets to the study is also of interest. Implementing features based on PGA, FGA and FAA may be of interest. Also optimisation of the different windows for the tail as well as long and short integration gates should be performed. Another feature which Dr Sykora trialled in his algorithms was a "time to 10%". This also attempts to encode the tail in a value, and takes the time value at which the signal has dropped from its peak, to 10% of its peak value. Additionally, based on the failure of modelling pulses based on a mixture of normal distributions, modelling pulses based on Marrone's method may be worth trying.

Being able to bind the shape of the fitted clusters to the shape of the regions which make up figure 5 would surely result in a fit which is both improved in terms of its gamma sensitivity and thermal neutron classification. This may yet be possible with the algorithm built in to `sklearn`, but a more custom approach may be required.

There is always the potential for using different feature sets in the algorithm. One option is utilising Marrone's fitting approach to fit a number of exponential functions to each

pulse and use the resulting parameters as features, similar to the Gaussian distribution fitting approach.

It emerged that the possiblity of producing a standalone app was not possible. Not only would this have required a significant amount of time which could be spent working on the algorithms, but the algorithm itself requires improvement before moving on to the next stages of implementing in standard detector development workflows.

7 Reflection

This project is a significant bridge between the world of applied physics and that of machine learning. As such, knowledge of the workflow of each discipline is required and can be difficult to reconcile. My background is in physics and during my degree I was able to undertake a placement year at the ISIS facility. At this point I was aware of the potential of the application of machine learning to this problem, so by the time this project came about, I was already familiar with the problem, the physics and their context.

Additionally, having not done a computer science degree, my coding practices are not as polished and streamlined as someone who has. While I am comfortable enough programming at this level, the project evolved in such a way that I began to regret designing the codebase in the way I did, and some time was used to go back and refactor code rather than working on improvements which would lead to better numeric results.

I would have liked to take a more bespoke approach to the creation of the clustering routine, rather than simply using a pre-built routine such as the one provided by `sklearn`. However, my knowledge of the mathematics of Gaussian mixture models is not sufficient. The `sklearn` package was however extremely easy to work with and enabled maximum time to be put in to trialling different feature sets rather than working on the minutae of a custom algorithm.

References

[1] Accessed 2020-08-10. URL: https://stfc.ukri.org.

[2] Accessed 2020-08-10. URL: https://www.isis.stfc.ac.uk/Pages/home.aspx.

[3] N.J Rhodes, E.M Schooneveld, and R.S Eccleston. "Current status and future directions of position sensitive neutron detectors at ISIS". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 529.1 (2004). Proceedings of the Joint Meeting of the International Conference on Neutron Optics (NOP2004) and the Third International Workshop on Position-Sensitive Neutron Detectors (PSND2004), pp. 243–248. ISSN: 0168-9002. DOI: https://doi.org/10.1016/j.nima.2004.04.152.

[4] Glenn F. Knoll. *Radiation Detection and Measurement*. Wiley, 2010. ISBN: 9780470131480.

[5] I. Stefanescu et al. "Neutron detectors for the ESS diffractometers". In: *Journal of Instrumentation* 12.01 (Jan. 2017), P01019–P01019. DOI: 10.1088/1748-0221/12/01/p01019.

[6] Natalia Zaitseva et al. "Pulse shape discrimination with lithium-containing organic scintillators". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 729 (2013), pp. 747–754. ISSN: 0168-9002. DOI: https://doi.org/10.1016/j.nima.2013.08.048.

[7] Nerine J. Cherepy et al. "Bismuth- and lithium-loaded plastic scintillators for gamma and neutron detection". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 778 (2015), pp. 126–132. ISSN: 0168-9002. DOI: https://doi.org/10.1016/j.nima.2015.01.008.

[8] Accessed 2020-09-18. URL: https://eljentechnology.com/18-products/plastic-scintillator.

[9] Gioacchino Ranucci. "An analytical approach to the evaluation of the pulse shape discrimination properties of scintillators". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated*

Equipment 354.2 (1995), pp. 389–399. ISSN: 0168-9002. DOI: https://doi.org/10.1016/0168-9002(94)00886-8.

[10] M.W.J. Hubbard, M.P. Taggart, and P.J. Sellin. "Exploration of Fourier based algorithms and detector designs for pulse shape discrimination". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 930 (2019), pp. 64–73. ISSN: 0168-9002. DOI: https://doi.org/10.1016/j.nima.2019.03.020.

[11] Matthew J.I. Balmer, Kelum A.A. Gamage, and Graeme C. Taylor. "Comparative analysis of pulse shape discrimination methods in a 6Li loaded plastic scintillator". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 788 (2015), pp. 146–153. ISSN: 0168-9002. DOI: https://doi.org/10.1016/j.nima.2015.03.089.

[12] B. D'Mellow et al. "Digital discrimination of neutrons and γ-rays in liquid scintillators using pulse gradient analysis". English. In: *Nuclear Instruments & Methods in Physics Research. Section A: Accelerators, Spectrometers, Detectors, and Associated Equipment* 578.1 (July 2007), pp. 191–197. ISSN: 0168-9002. DOI: 10.1016/j.nima.2007.04.174.

[13] G. Liu et al. "A Digital Method for the Discrimination of Neutrons and γ Rays With Organic Scintillation Detectors Using Frequency Gradient Analysis". In: *IEEE Transactions on Nuclear Science* 57.3 (2010), pp. 1682–1691.

[14] S Marrone et al. "Pulse shape analysis of liquid scintillators for neutron studies". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 490.1 (2002), pp. 299–307. ISSN: 0168-9002. DOI: https://doi.org/10.1016/S0168-9002(02)01063-X.

[15] Alan D. Kaplan et al. "A neutron-gamma pulse shape discrimination method based on pure and mixed sources". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 919 (2019), pp. 36–41. ISSN: 0168-9002. DOI: https://doi.org/10.1016/j.nima.2018.11.136.

[16] Xunzhen Yu et al. “Neutron–gamma discrimination based on the support vector machine method”. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 777 (2015), pp. 80–84. ISSN: 0168-9002. DOI: https://doi.org/10.1016/j.nima.2014.12.087.

[17] Xufei Xie et al. “Digital discrimination of neutrons and gamma-rays in organic scintillation detectors using moment analysis”. In: *Review of Scientific Instruments* 83.9 (2012), p. 093507. DOI: 10.1063/1.4754633. eprint: https://doi.org/10.1063/1.4754633.

[18] Stuart Russell and Peter Norvig. *Artificial Intelligence: A Modern Approach*. 3rd. USA: Prentice Hall Press, 2009. ISBN: 0136042597.

[19] A. P. Dempster, N. M. Laird, and D. B. Rubin. “Maximum Likelihood from Incomplete Data via the EM Algorithm”. In: *Journal of the Royal Statistical Society. Series B (Methodological)* 39.1 (1977), pp. 1–38. ISSN: 00359246. URL: http://www.jstor.org/stable/2984875.

[20] J.-M. Marin, K. Mengersen, and C.P. Robert. “Bayesian Modelling and Inference on Mixtures of Distributions”. In: *Handbook of Statistics* 25 (2005). cited By 225, pp. 459–507. DOI: 10.1016/S0169-7161(05)25016-2.

[21] Shai Shalev-Shwartz and Shai Ben-David. *Understanding Machine Learning: From Theory to Algorithms*. Cambridge University Press, 2014. DOI: 10.1017/CBO9781107298019.

[22] F. Pedregosa et al. “Scikit-learn: Machine Learning in Python”. In: *Journal of Machine Learning Research* 12 (2011), pp. 2825–2830.

[23] Martin Ester et al. “A density-based algorithm for discovering clusters in large spatial databases with noise.” In: *Kdd*. Vol. 96. 34. 1996, pp. 226–231.

[24] Pauli Virtanen et al. “SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python”. In: *Nature Methods* 17 (2020), pp. 261–272. DOI: 10.1038/s41592-019-0686-2.