**B.Tech Project (PHN-400B) Report**

**Project Supervisor:**

**Anil K. Gourishetty**

**Group Members:**

**Karan (18122010)**

**Sutirtha Biswas(18122026)**

**Sankar Singh Meena(18122017)**

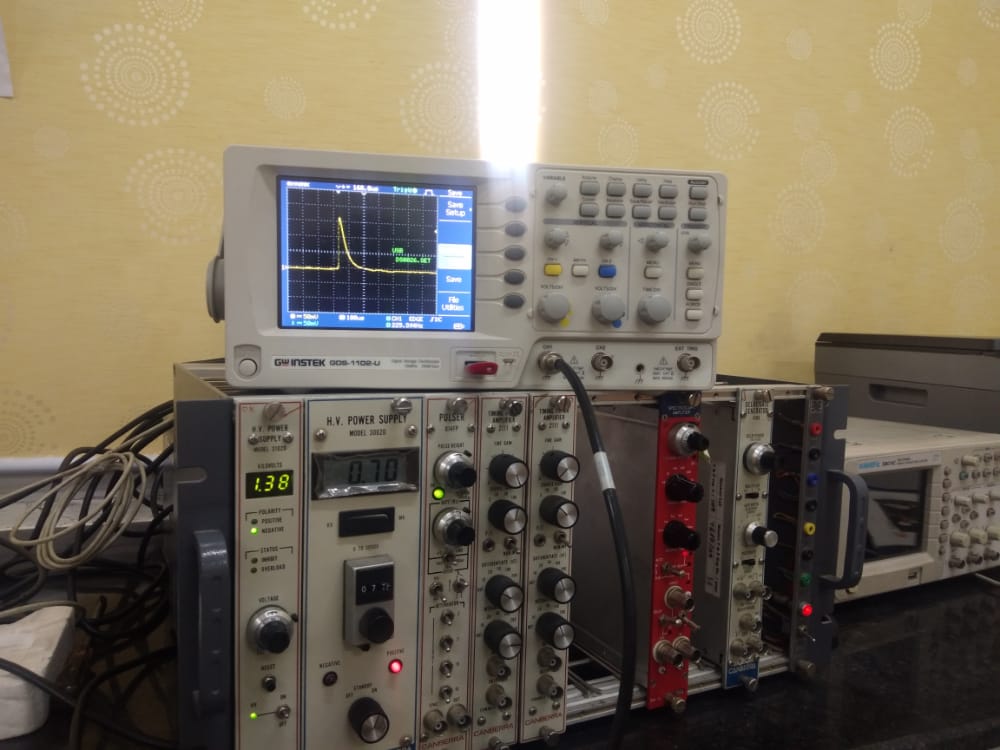
**Abstract:**

Radiations, specifically neutrons and gamma are acquired from different sources and are incident on various scintillators. The light output is collected in both analogue and digital format. Pulse Shape Discrimination is the process of analyzing different signal features which can lead to classifying a waveform as a neutron or a gamma. This work includes applying the conventional techniques to reproduce established results as well as introducing modern deep learning algorithms to increase the ability of discrimination. The conventional methods include extracting signal features like total area/ tail area, decay time, spectral energy density and calculating figure of merit. While new age AI techniques involve applying machine learning algorithms or neural networks to build an efficient classifier. Also a curved fitting algorithm mentioned in paper[1] has also been tested on the pulse data of CsI detector with AmBe source.

**Experimental Setup:**

The experimental setup consisted of an AmBe radiation source and detector with a cylindrical cell of 15.24cm in diameter and 7.62cm thick. The SMO10C Oscilloscope, with 1 GS/s sampling rate, 100 MHz Bandwidth was used to record the waveform. The module was installed on a PCI

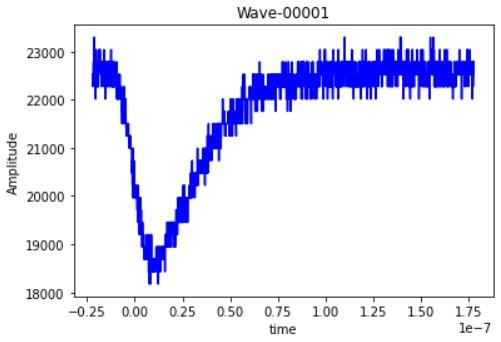
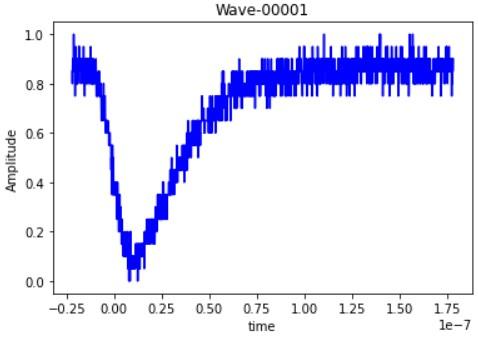
bus attached to PC with Windows 10 operating system.

**Data Preprocessing:**

Signals from CRO are of the format Amplitude vs Time. A few preprocessing steps must be carried out before relevant information can be extracted.

* **Normalization**: The waveforms differ in their max and min amplitudes. To bring them into a common parity the range of amplitudes is squeezed in between 0 and 1 using normalization. If the signal is represented by the array y = f(t), then the normalized signal array is given by y\_norm = (f(t) - min(f))/(max(f) - min(f))

Raw pulse Normalized pulse

* **Noise Filtration**: Raw signals obtained contain lots of noises that need some cleaning. Although lots of digital filters are available, for our present application Savitzky-Golay smoothing filters perform much better than standard averaging finite impulse response (FIR) filters, which tend to filter out most of the high frequency content of the signal along with the noise. This filter preserves the signal tendency and increases the precision of the data by applying convolution. The window slides across successive sub-sets of adjacent data points and fits it with a low degree [polynomial](https://en.wikipedia.org/wiki/Polynomial) by the method of linear least squares. We have chosen here a 3rd order polynomial and a kernel size of 41.

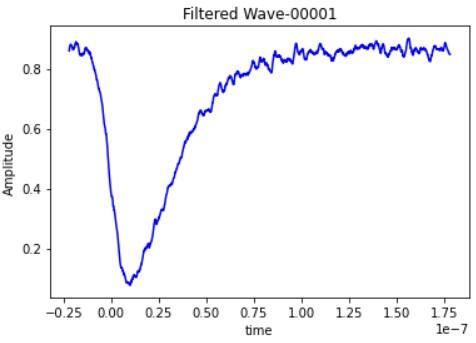


FIg.- Filtered wave

* **Clipping unwanted pulses**: No radiation source emits only neutrons and gammas solely, there are secondary radiations, low energy X-rays etc. These unwanted signals need to be clipped off before proceeding so that we get a dataset of pure neutrons and gammas. We train a Convolutional Neural Network classifier to distinguish required signals and unwanted radiations to achieve the purpose. Using only a few hundred training images the task has been accomplished successfully.

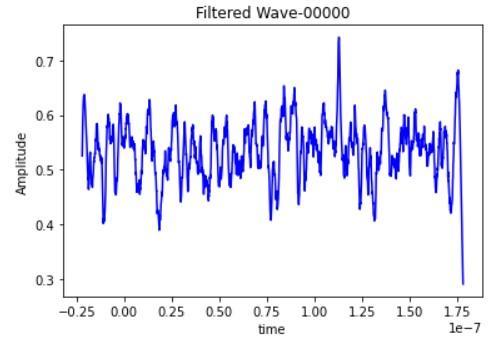
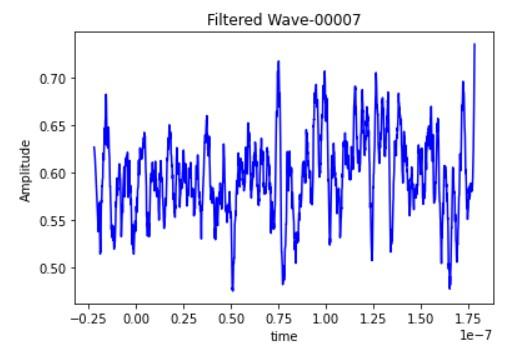
 

Fig.- Examples of unwanted pulses to be clipped

**Feature extraction:**

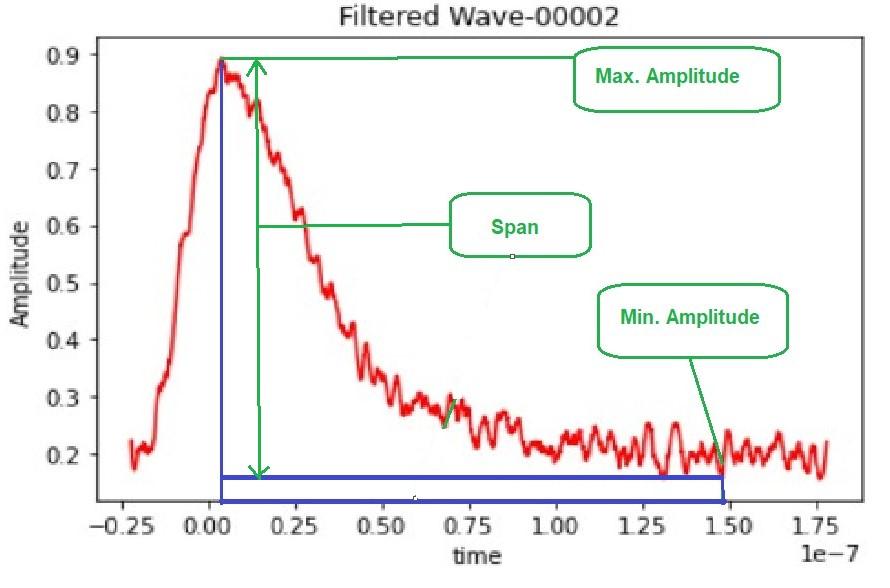
**Total Area/Tail Area:**

This method is also known as the Charge Comparison Method (CCM). The pulse is integrated over two different time intervals, a slow period, Q(slow) or the Tail Area and the total period, Q(tot) or the Total Area. The main disadvantage of the CCM is that the performance is highly dependent on the choice of the integration intervals. Here we have chosen the tail start area to be one-fourth the length of the pulse and tail end to be the end of the pulse.

**Decay Time:**

The time taken by the waveform to reduce to a certain fraction of its peak, is referred to as the decay time. The threshold is chosen as 36%. It's set optimally after a manual search as it provides the best results. The calculation is done as follows

1. Maximum amplitude is recorded
2. The base is non-zero, hence minimum amplitude is recorded. Span is defined as the range between the max and min amplitude.
3. Time taken to fall to 36% of span gives the decay time.

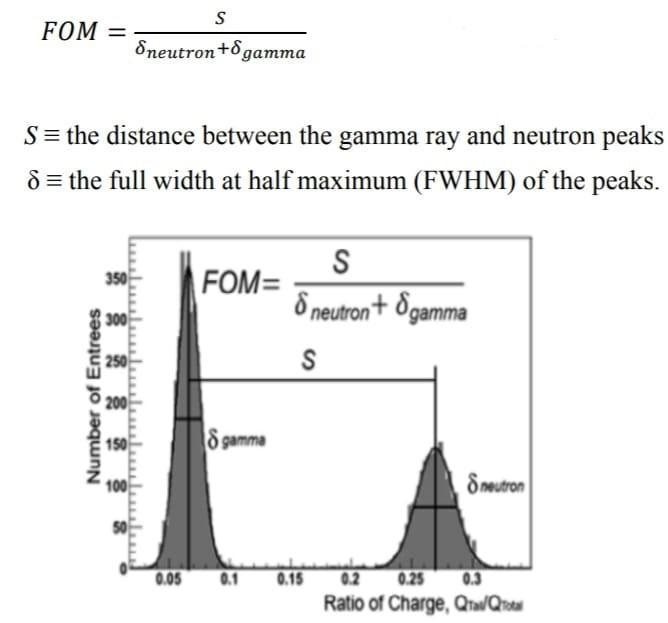


**Spectral Energy Density:** The energy spectral density (ESD) is the energy distribution of a signal as a function of frequency. Its integral is a measure of the total signal energy: E = ∫ |f(t)|^2, where f(t) is the input signal (amplitude as a function of time). The Fourier transform of the autocorrelation function of an energy signal is equal to the energy spectral density of the signal. Therefore, the ESD of such a neutron or gamma-ray pulse carries important information about the pulse, which is a useful parameter to carry out discrimination between a neutron and a gamma ray. We normalize the ESD metric in the following way,

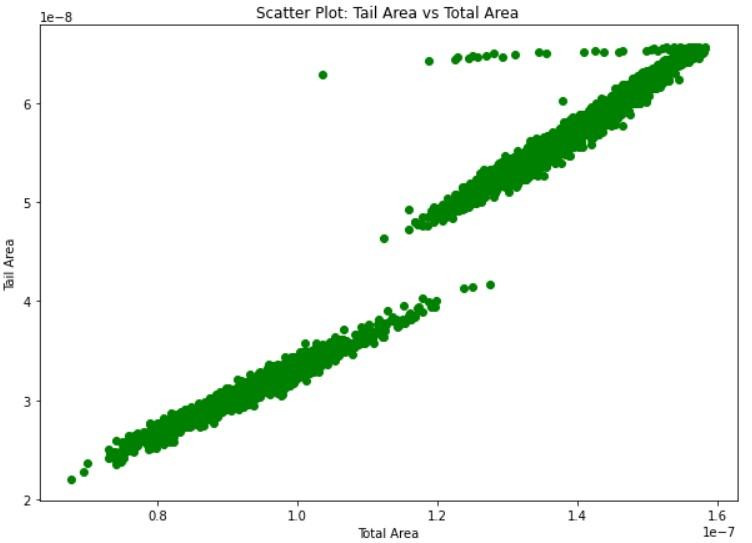
ESD (normalized) = ESD/max (ESD)

**Calculating Figure Of Merit:**

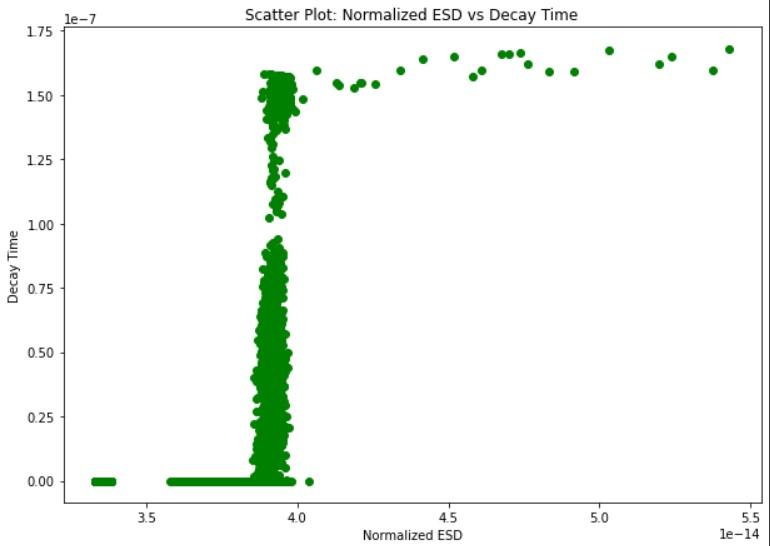
A figure of merit (FOM) is used to calculate the efficiency of the ability of discrimination between neutron and gammas. The FOM is calculated as the ratio of distance between gamma and neutron peaks and the sum of the full-width half maximum of the piles.



**Scatter Plots and Bivariate Analysis Results:**

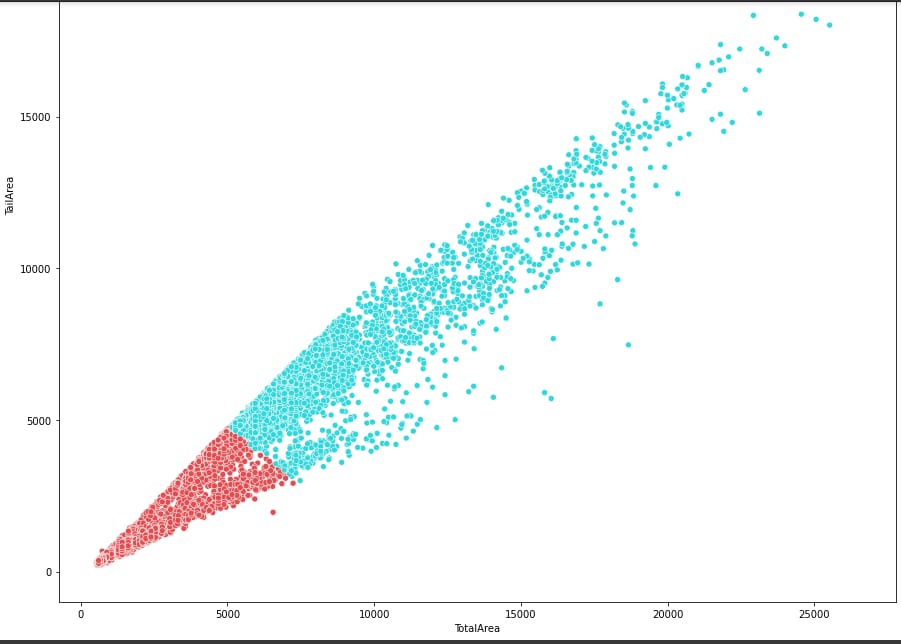






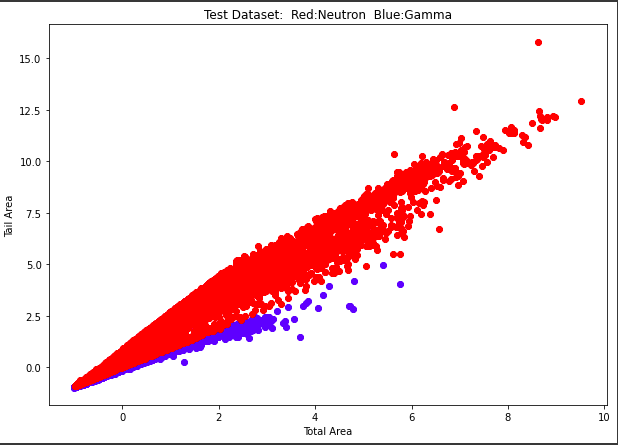
**Particle Discrimination using Supervised and unsupervised models**

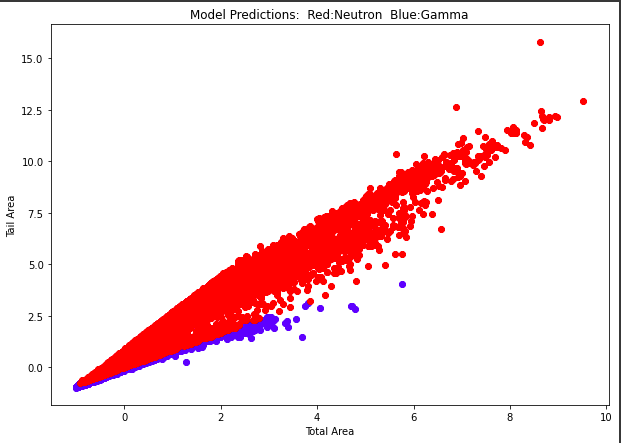
Various unsupervised ML models were tested on a labeled dataset from PuBe radiation source.The two features in the dataset were tail area and total area. The results were highly inaccurate due to the fact that unsupervised ML algorithms tend to find clusters in the dataset, and since the dataset we used had none, it failed to produce correct results.

**

*Clusters formed using K-Means Clustering Algorithm.*

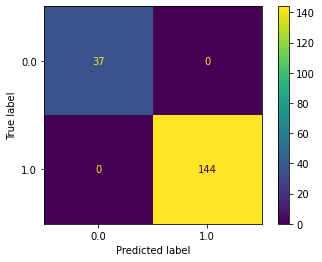
A Neural Network is trained on this dataset to find the performance of supervised algorithms. The accuracy of the neural network was considerably better given the fact that the dataset had a huge class imbalance.





**Noise Removal Using Machine Learning Algorithm:**

The unwanted noise data in our dataset must be removed for further analysis. Instead of using a signal processing algorithm we trained a Random Forest Classifier which can classify an unwanted noise pulse from a clean pulse. The ML model is trained on a dataset of 500 pulses and is validated on 200 different pulses. The confusion matrix below shows that the model is capable of accurately classifying the noised pulses.



*Confusion matrix of the prediction of Random Forest Classifier on Test Dataset*

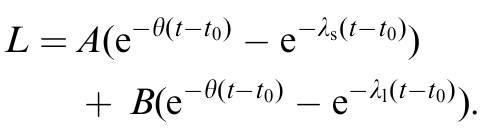
**Pulse shape analysis using 1-parameter fit algorithm**

**Introduction :**

In this project the shapes of pulses produced by g-rays and neutrons have been studied for the CsI detector. A 1-parameter ﬁtting algorithm is used, to gather knowledge on the radiation class and energy. The ability of this method in energy resolution and n/g classification is then evaluated.

1. **Shape Analysis:**

In the reference [1] an equation is derived to interpolate the pulse shape of different particles. The equation consists of 6 distinct variables : A,B,t0,, and .



Here A,B are the normalization variables, t0 is the time for beginning of the signal, , and are the decay constants of the detector which won’t change with different radiations . The use of variable t0 is useful, because the start of the signal is helpful in signal interpretation.

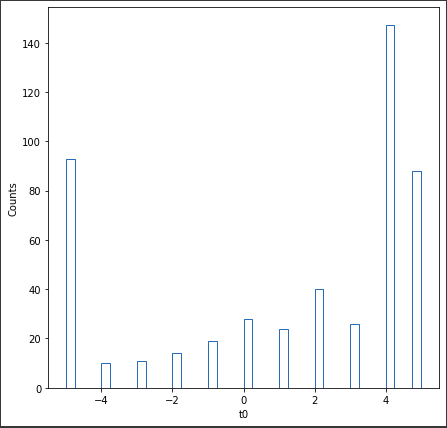
A ﬁt according to Eq. (1) with six variables is computationally intensive . To simplify the ﬁtting process, reduce the number of free variables by finding their single best values or developing relations between them. Thus the following process is :

(1) 500 events are ﬁtted with Eq. (1), and optimal values are calculated for all 6 variables.

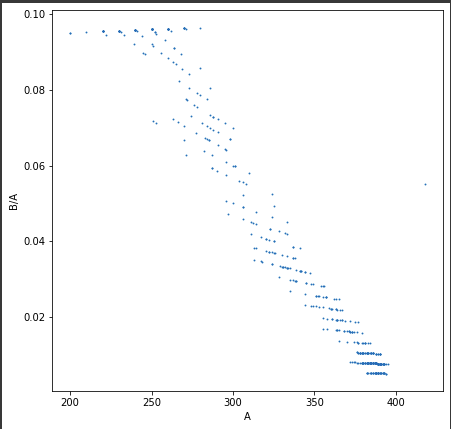
(2) The distribution of the three decay constants (, and ) is found, and their average is used .

(3) Fixing the 3 exponential decay constants ,we fit the curves only to calculate t0, A and B..

(4) An average value is chosen for t0 ; and a ratio between A and B is found, thus all pulses can be ﬁtted with only 1 normalization constant.

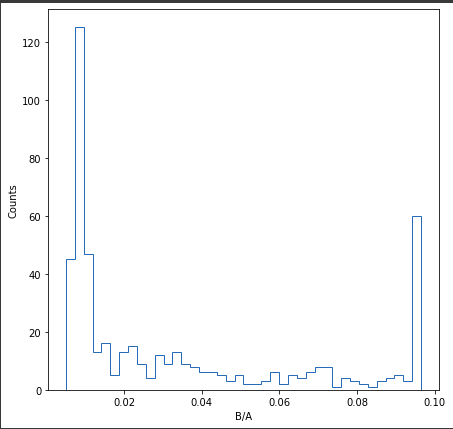


*The frequency plot for the optimal t0 parameter is plotted, and bases on that 4 and -4 are considered as t0 values for gamma and neutron respectively.*



*Scatter plot of A vs B/A. The lower part of the plot*

*Represents the gamma particles and the spread above 0.04 Represents the neutron particles.*

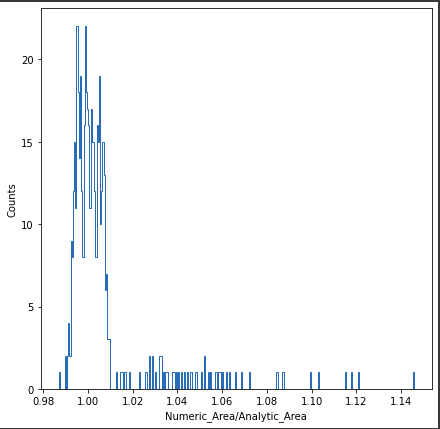


*Histogram plot for optimal B/A values. The left side peak represents*

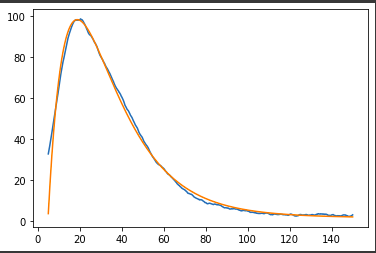
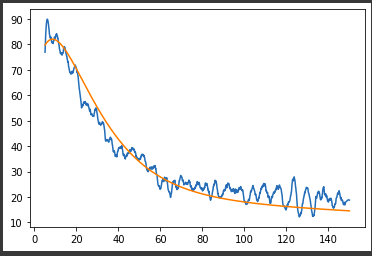
*The gamma particles while the spread after 0.04 represents the neutron particles.*

1. **Energy Resolution and n/y discrimination:**

As a performance evaluation of the ﬁt, we have measured the energy spectra obtained by the integration of Eq. (1) up to 150 ns to the result of numerical integration of the actual pulse signals till 150ns. The plot of curve fit for both particles and the event-by-event ratio between the light outputs is plotted below. Both plots shows that the ﬁt generates signals of significant quality.



*Count plot of ratio of Numeric Area/Analytical Area*

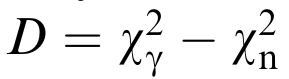
*Pulse plot for a gamma particle pulse. The blue Pulse plot for a neutron particle. The blue*

*Line indicate the original data and the Line indicate the original data and the*

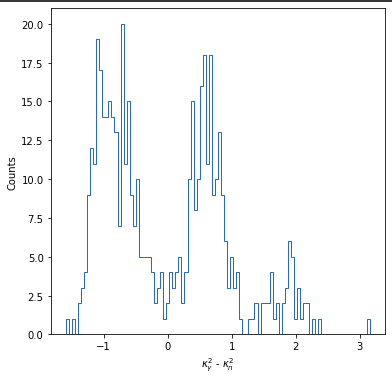
*Orange line indicates the plot by our Orange line indicates the plot by our*

*Equation derived above. Equation derived above.*

The neutron-gamma classification is done by using the chi-square test. The 6 variable equation is reduced to just 1-variable equation by fixing values of 5 variables for the gamma and neutron particle type. Then the normalization variable **A** is optimized for each pulse and chi-squared error is calculated.



The normalized error is supposed to be negative for the gamma particles and positive for the neutron particles. The count plot for this error below indicates that the separation between these two particle types is reasonably well using this 1-variable fitting algorithm.



*Kai-Squared plot for neutron and gamma pulses. The peaks on positive*

*Side indicates neutron pulses and peaks on negative side represents*

*Gamma particles.*

**Conclusion:**

The pulse data from AmBe source has been collected and cleaned using signal processing algorithm and machine learning algorithm. Once the pulse shape is known for gamma and

neutrons, 1-variable ﬁts can be applied to get knowledge about particle type and

energy . The accuracy of this method is checked by comparing the energy spectrum

and n/g discrimination ability obtained due to ﬁt by those obtained by numerical integra-

tion of signal. The performance of this method, in energy resolution and particle classification can be compared to well known methods in the field of nuclear physics.

**References:**

* S. Marrone et al. / Nuclear Instruments and Methods in Physics Research A 490 (2002) 299–307