Unsupervised Probability Density Expectation Maximization Algorithm for Particle Identification in Pulse-Shape Discrimination Applications

James B. Cole[a], Bryan V. Egner[b], Brian Frandsen[b], Darren E. Holland[b], James E. Bevins[b]

[a] Department of Physics
Bard College, Annandale-On-Hudson, NY
[b] Department of Engineering Physics
Air Force Institute of Technology, WPAFB, OH 45433

1 Introduction

Special nuclear material (SNM) emits both gamma rays and neutrons, and for nuclear security applications they are often encapsulated in shielding that moderates the neutron spectrum [1]. To characterize SNM, versatile detectors are required to separate the fast neutron, thermal neutron captures, and gamma-ray-induced signals, motivating the need for robust particle discrimination algorithms. Pulse-shape discrimination (PSD) is a common approach to discriminate particle types in organic scintillators, and the most common implementations are charge integration methods [2]. However, these approaches are limited in their ability to discriminate signals below a couple of hundred keVee.

One possible approach to improve discrimination below this threshold is the implementation of advanced approaches, such as Bayesian frameworks [3]. Alternatively, this work presents an unsupervised probability density expectation maximization (PDEM) algorithm for determining particle identification probabilities from one-dimensional charge-integration pulse-shape discrimination distributions. This is applied to data to extracting neutron capture events from a boron-loaded deuterated liquid scintillator in a mixed gamma and fast-neutron background.

2 Method

The proposed PDEM methodology determines the best fit to a set of experimental PSD data, $\{x_i\}_{i=1}^{N}$, using a weighted superposition of standard probability density functions (PDFs). It assumes that this fit accurately models the types of events in the charge integration PSD distribution (counts as a function of PSD parameter). The weighted superposition is given by

$$\rho(\mathbf{\Lambda}, x) = \sum_{c=1}^{C} w_c \rho_c(\lambda_c, x), \tag{1}$$

where the λ_c are the PDF parameters, w_c the weights, $\mathbf{\Lambda} = \{\lambda_c, w_c\}$, and ρ_c is the PDF for cluster c. Each PDF is defined by its own parameters λ_c (e.g. mean and standard deviation), and each particle type can follow independent distributions (e.g. Gaussian, Skew-Gaussian, Lorentzian, etc.) based on the detector settings and data cuts applied. The probability matrix P for N particles and C clusters in the PSD distribution is then given as

$$\mathrm{P} = \begin{bmatrix} P_{1,1} & P_{1,2} \cdots & P_{1,C} \\ P_{2,1} & & \vdots \\ \vdots & & \\ P_{N,1} & \cdots & P_{N,C} \end{bmatrix}, \tag{2}$$

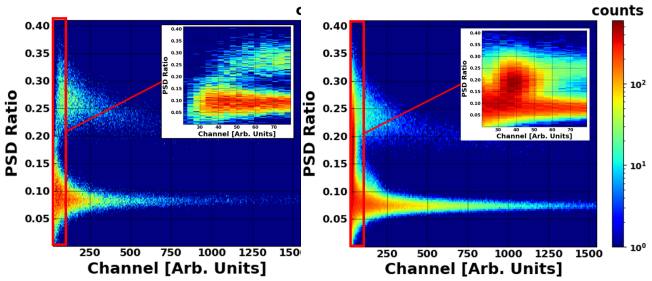
where $\mathrm{P}_{i,c} = w_c \rho_c(\lambda_c, x_i)/\mathcal{N}_i$. $\mathrm{P}_{i,c}$ is the probability relative (to other clusters) that event x_i is a member of cluster c. The $\mathrm{P}_{i,c}$ are thus fractional cluster memberships. $\mathcal{N}_i = \sum_{c=1}^{C} w_c \rho_c(\lambda_c, x_i)$ normalizes the row sums to unity to preserve the particle's probability to be in a cluster. The column sums n_c are probabilistic estimates of the cluster (particle) populations.

An iterative EM procedure is used to find the parameters, $\mathbf{\Lambda}$, that maximize $\sum_{i=1}^{N} \rho(\lambda_c, x_i)$ [4]. Taking the PDFs to be Gaussians, initial values of the cluster weights w_c^0, means μ_c^0, and standard deviations σ_c^0 ($\Lambda^0 = (w_c^0, \mu_c^0, \sigma_c^0)$) are estimated from the 2-D PSD histograms, Figure 1, and/or 1-D PSD histograms (data in Figure 2), using automated estimates from the peak locations, heights, and widths. This can be modified by the user to account for "hidden" peaks in the 1D data known to exist from the 2-D PSD data (e.g. Figure 1b).

The algorithm starts by computing $\mathrm{P}^0 = \mathrm{P}^0(\Lambda^0)$, and then sequentially updating the PDF parameters $n_c^{j+1}, w_c^{j+1}, \mu_c^{j+1}$, and σ_c^{j+1}. The iteration terminates when $\Delta^{j-1} - \Delta^j < g$, where Δ is the root mean square deviation of the model from the data and g is a small positive quantity.

3 Experimental Data

The PSD response for a 1-inch right circular cylinder ^{10}B-enriched deuterated toluene liquid scintillator [5] irradiated by a ^{252}Cf source was measured to test the proposed method. For the measurement, the scintillator cell was coupled to a $1\frac{1}{8}$ inch Hamamatsu R6094 photomultiplier tube, biased using an 8-channel CAEN DT8033N high-voltage power supply, and connected to an 8-channel 500 MS/s CAEN DT5730 digital pulse analyzer. Full digitized radiation-induced waveforms were collected using CAEN Multi-PArameter Spectroscopy Software (COMPASS), which also can record 2 ps-precision event time-stamps when operating in constant fraction discriminator mode. This mode was used for neutron time-of-flight (nToF) measurements to provide tagging validation of fast neutron and gamma-ray events. 2-D PSD histograms for two particle (fast neutrons and gamma rays) and three particle (alphas from ^{10}B(n,α) capture, fast neutron, and gamma rays) are shown in Figure 1. For the initial validation studies presented here, only Figure 1a contains time-tagged fast neutrons and gamma-rays; the data in Figure 1b do not.



(a) (b)

Fig. 1: 2-D PSD validation histograms with different particle types (a) fast neutrons and gamma rays and (b) alphas, fast neutrons, and gamma rays.

4 Model Validation

The proposed methodology was tested on the tagged experimental data shown in Figure 1a using Gaussian and truncated-Gaussian PDFs to compute the gamma and neutron populations, where truncated-Gaussian PDFs were found to be more representative of the data due to the data cut applied at PSD = 0. Comparisons to the known populations are given in Table 1, for which a Kolmogorov-Smirnov (KS) Two-Sample Test p-value of 0.96 was obtained (model was drawn from same distribution as the data).

Table 1: True vs PDEM populations estimates.

	Population	PDEM
fast n	55,752	54,020
γ	17,885	18,617

The estimated populations were 3 and 4% off from the true populations for neutrons and gamma rays, respectively. This is due to the fact that the data are not truly Gaussian and exhibit slight skewness in the distributions.

The PDEM algorithm was then used to analyze the data shown in Figure 1b (also grey histogram in Figure 2), using the *a priori* information that there are three particle types. The best fit parameters are given in Table 2. Initial particle populations were estimated using $N_c = Nw_c$, where $N = 316,145$ is the total number of events in the distribution. The normalized histogram of the data and the best fit truncated-Gaussian PDFs, with a KS Two-Sample Test p-value of 1.0 indicating high statistical likelihood that model was drawn from same distribution as the data, are shown in Figure 2.

5 Conclusions

Reasonable estimates of the particle populations were obtained using the outlined PDEM algorithm when compared to tagged fast neutron and gamma ray data. Initial estimates of three-particle type data were obtained that demonstrate the capability to develop models to separate events in highly overlapped, low-energy deposition regions enabling better quantification of interaction type in PSD-capable organic scintillators. Future work will implement a flexible Beta distribution and develop tagged data to validate performance across multiple interaction types and levels of data mixing.

Table 2: PDEM 3-cluster distribution parameters.

	pop.	w_c	μ	σ
γ	152,113	0.481	0.094	0.033
alpha	131,338	0.415	0.197	0.044
fast n	32,694	0.103	0.253	0.047

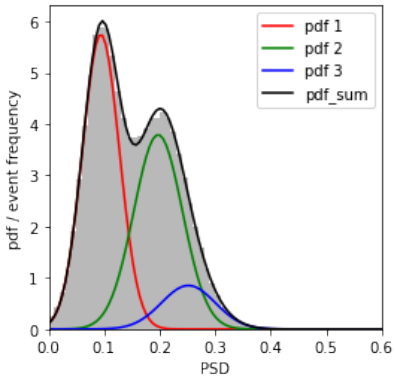


Fig. 2: One-dimensional normalized PSD data from Figure 1b (grey histogram); γ distribution (red), alpha (green), fast n (blue).

Acknowledgment

This work is supported by the Defense Threat Reduction Agency under grant HDTRA-1033292. The views expressed in this article are those of the authors and do not necessarily reflect the official policy or position of the United States Air Force, the Department of Defense, or the United States Government.

References

[1] C. C. Lawrence, et al., Warhead verification as inverse problem: Applications of neutron spectrum unfolding from organic-scintillator measurements, J. Appl. Phys 120 (6) (2016).

[2] I. A. Pawełczak, , et al., Studies of Neutron-Gamma Pulse Shape Discrimination in EJ-309 Liquid Scintillator Using Charge Integration Method, Nucl. Instrum. Methods Phys. 711 (2013) 21–26.

[3] M. Monterial, , et al., Application of Bayes' theorem for pulse shape discrimination, Nucl. Instrum. Methods Phys. 795 (2015) 318–324.

[4] N. Gershenfeld, Mathematical Modeling, Cambridge University Press, 2011.

[5] B. V. Egner, et al., Characterization of a boron-loaded deuterated liquid scintillator for fast and thermal neutron detection, Nucl. Instrum. Methods Phys. 996 (4) (2021).