

# PULSE SHAPE DISCRIMINATION OF NEUTRONS AND GAMMA RAYS USING KOHONEN ARTIFICIAL NEURAL NETWORKS

Tatiana Tambouratzis<sup>1</sup>, Dina Chernikova<sup>2</sup> and Imre Pzsit<sup>2</sup>

<sup>1</sup>*Department of Industrial Management & Technology, University of Piraeus, address Street 107 Deligiorgi St., City Piraeus 185 34, country-region Greece*

<sup>2</sup>*Division of Nuclear Engineering, Chalmers University of Technology SE-412 96 Cityplace Gothenburg, country-region Sweden*

## Abstract

The potential of two Kohonen artificial neural networks (ANNs) - linear vector quantisation (LVQ) and the self organising map (SOM) - is explored for pulse shape discrimination (PSD), i.e. for distinguishing between neutrons (n's) and gamma rays ( $\gamma$ 's). The effect that (a) the energy level, and (b) the relative size of the training and test sets, have on identification accuracy is also evaluated on the given PSD dataset. The two Kohonen ANNs demonstrate complementary discrimination ability on the training and test sets: while the LVQ is consistently more accurate on classifying the training set, the SOM exhibits higher n/ $\gamma$  identification rates when classifying new patterns regardless of the proportion of training and test set patterns at the different energy levels; the average time for decision making equals  $\sim 100 \mu\text{s}$  in the case of the LVQ and  $\sim 450 \mu\text{s}$  in the case of the SOM.

## 1 Introduction

The potential of discriminating different types of particles in a detector signal has raised interest in a variety of scientific fields, including astrophysics, medicine, geology etc. The same topic has also attracted particular attention from the experts in the fields of nuclear safeguards, security and non-proliferation, as a means of:

- Enabling the timely and cost-efficient substitution of  $^3\text{He}$ -counters in nuclear safeguards and security applications. Although only sensitive to neutrons (n's), these counters have been particularly popular in detecting these particles. However, - due to the global shortage of  $^3\text{He}$  [1] - they will soon have to be replaced by other detectors.
- Implementing pulse shape discrimination (PSD), namely distinguishing between (and,

thus identifying) n's and gamma rays ( $\gamma$ 's). It should be mentioned that some of the detectors considered for the replacement of the  $^3\text{He}$ -counters are capable of simultaneously detecting n's and  $\gamma$  photons, but cannot distinguish between the two.

Among the detectors that might be used for PSD are liquid scintillators (LSs, such as EJ-309), and plastic scintillators (PSs, organic crystals). Although PSs are the most frequently applied in the field [2-6], they generally show poor performance on PSD when compared to LSs. As an example, the most recent experimental studies [7] indicate that the new PSs of type EJ-299-33 are inferior to the LSs of type EJ-309 in terms of both detection efficiency of n's, and quality of PSD.

The main challenge for PSD methods is their adequately fast operation as well as their sustained accuracy throughout the range of energy levels of

interest. In order to meet the imposed constraints, a number of analytical and computational intelligence (CI)-based approaches have been developed.

Analytical approaches to PSD include the charge comparison method (CCM [8]), the rise time method (RT [9]), the zero-crossing method (ZCM [10]),  $n$ - $\gamma$  model analysis (NGMA [11]), pulse gradient analysis (PGA [12]), simplified digital charge collection (SDCC [13]), and frequency gradient analysis (FGA [14]). The PSD results are data-dependent, whereby a consistent comparison between the aforementioned approaches cannot be reached. For instance, NGMA is reported as being more accurate than PGA, CCM, ZCM and RT in [15], while SDCC is reported as being better than PGA, CCM, and NGMA at  $n/\gamma$  discrimination in [16], and – finally - ZCM method is found superior to CCM in [17].

The existing CI-based approaches to date employ ANNs, and the back-propagation (BP) architecture [18] in particular, in its many forms, variants, and extensions that have been developed over the years (e.g. resilient BP [19], and scaled conjugate gradient [20]). Among the results reported in [21-30]:

- 95% accuracy is demonstrated for energy levels down to 100 keVee, although accuracy falls to 85% when the ratio of  $n$ 's to  $\gamma$ 's is changed by a factor of 10 [21];
- 98.5% accuracy, with each PSD decision implemented in about 2.9  $\mu$ s, is reached for a 60-20-20% training-validation-test set partitioning of the dataset [28];
- higher than 99% accuracy is attained at around 1,000 keVee for the related problem of  $^{12}\text{C}$  and  $^{13}\text{C}$  ion identification, with each decision realised in 20  $\mu$ s [29];
- about 84% accuracy is reported for energy levels below 400 keVee [30].

As for the analytical approaches discussed above, the results are data-dependent, hence not allowing a consistent comparison either between CI-based approaches or between analytical and CI-based approaches.

In this piece of research, the potential of two artificial neural networks (ANNs), inspired by Teuvo

Kohonen, namely linear vector quantisation (LVQ) [31] and the self organising map (SOM) [32], is investigated on the PSD problem. The MATLAB Neural Networks Toolbox [33] has been used for all the ANN implementations. The  $n$ - $\gamma$  dataset employed here is also used in [34-35], and contains PSD data from a wide range of energy levels, with a variable number of patterns collected at each energy level.

The evaluation of the two ANNs covers the effect that (a) the energy level, and (b) the relative size of the training and test sets, have on identification accuracy. The two ANNs demonstrate complementary discrimination characteristics for the training and test sets: while the LVQ is remarkably accurate at identifying  $n$ 's and  $\gamma$ 's of the training set, the SOM is more robust in its generalisation power that allows it to distinguish between novel (test)  $n$ - and  $\gamma$ -patterns. Despite these differences, both ANNs show high discrimination accuracy that is maintained even at the lowest energy levels.

The remainder of the paper is organised as follows: Section 2 describes the characteristics and pre-processing of the  $n$ - $\gamma$  dataset used here for PSD, with section 3 introducing the two ANN architectures; section 4 describes PSD data preparation as well as training and operation of the two ANNs; finally some conclusions are drawn, and some future directions are proposed in the Conclusion.

## 2 The Characteristics and Pre-Processing of the $n$ - $\gamma$ PSD Dataset

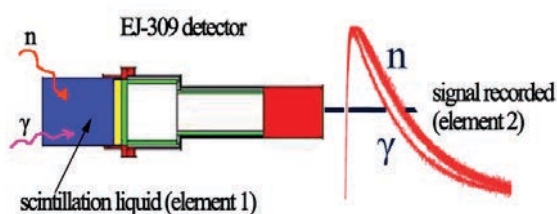
The  $n$ - $\gamma$  PSD dataset employed here contains more than 1,200 experimental data that have been generated by two sources, namely

- an Americium-Beryllium ( $\alpha,n$ )-source ( $^{241}\text{Am}$ -Be), which is considered as a source of both  $n$ 's and  $\gamma$ 's,
- a  $^{137}\text{Cs}$  calibration source, which is considered as a pure source of  $\gamma$ 's.

The data have been collected by an EJ-309 (D76 x 76 mm) LS at six different energy levels ranging from a high of 2,000 keVee to a low of 150 keVee. The settings and geometry of the experimental setup have already been described in [34].

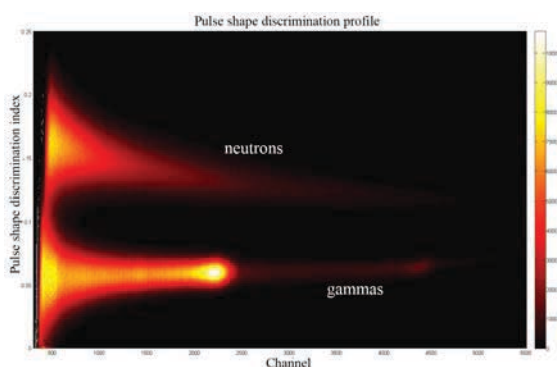
Neutrons and  $\gamma$ 's interact in different ways within the scintillation liquid (element 1 of Fig. 1) of the EJ-309 detector, i.e.

- n's interact mostly with medium via the mechanism of elastic scattering, resulting in the production of recoil protons (ionising particles);
- $\gamma$ 's interact mostly via the mechanism of Compton scattering, producing Compton electrons, which are also ionising particles.



**Figure 1.** Simplified illustration of the detection n's and  $\gamma$ 's in a scintillation detector

The difference in the masses of these ionising particles (namely  $\sim 10^{-27}$  kg for protons, and  $\sim 10^{-30}$  kg for electrons), gives rise to differences in the fraction of light that - due to delayed fluorescence - appears in the slow component of the recorded signal (element 2 of Fig. 1). It is this difference that constitutes the basis of PSD.



**Figure 2.** 2-D plot of pulse energy (x-axis) versus PSD index (y-axis); n's correspond to the higher horizontal peak, and  $\gamma$ 's to the lower horizontal peak, respectively

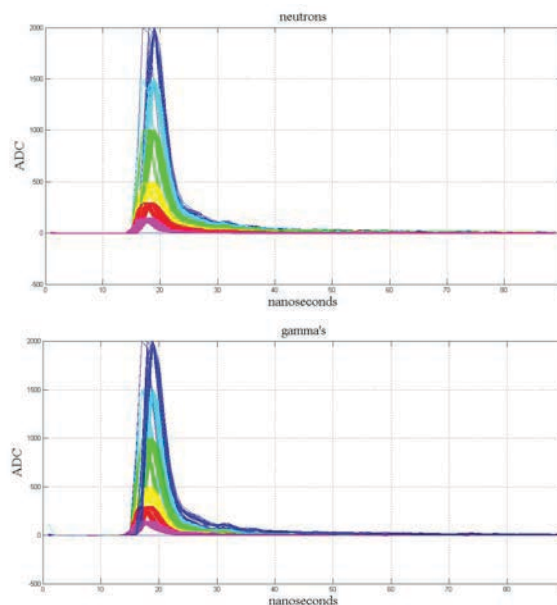
Various methods can be subsequently applied for separating between n's and  $\gamma$ 's. In the present experimental study, the CCM algorithm [8] is implemented for marking n's and  $\gamma$ 's. Subsequently, all the particles are classified as either n's or  $\gamma$ 's

depending on their PSD parameter, in the manner shown in Fig. 2.

The six datasets collected by the EJ-309 detector correspond to the six distinct energy levels shown in Table I, where the percentage of patterns in each energy level, and the proportion of n- and  $\gamma$ -curves available per energy level are also tabulated.

**Table 1.** Dataset Characteristics

Energy level (keVee)	(%) patterns	(%) N patterns	(%) $\gamma$ patterns
2,000	2.64	54.84	45.16
1,500	9.69	21.93	78.07
1,000	21	67.61	32.39
500	20.58	39.67	60.33
300	19.13	48.89	51.11
150	26.96	42.59	57.41

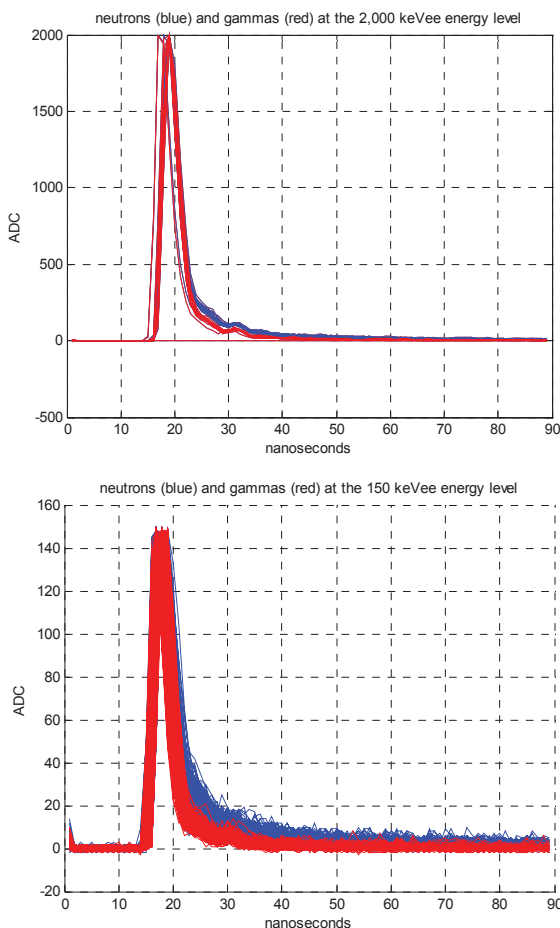


**Figure 3.** The shape of n- and  $\gamma$ -curves at the different energy levels, namely blue, cyan, green, yellow, red, and magenta for the 2,000, 1,500, 1,000, 500, 300 and 150 keVee energy levels

Figure 3 further illustrates the data obtained at each energy level, separately for n's and for  $\gamma$ 's. It can be observed that the two sets of curves demonstrate (a) decreasing and distinct magnitudes for declining energy levels, but (b) identical magnitudes at the same energy level. Hence, although the magnitude constitutes a distinguishing characteristic of the energy level, it cannot be used for differentiating

between n's and  $\gamma$ 's at the same energy level.

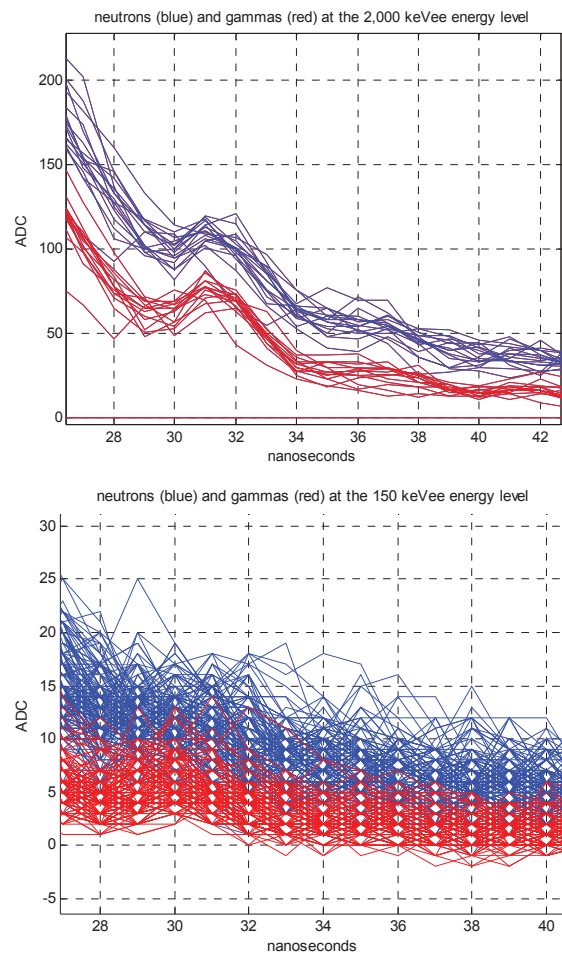
In fact, the variation between n-curves, as well as between  $\gamma$ -curves, is apparent at all energy levels. It is not uncommon for two n- (or for two  $\gamma$ -) curves at a given energy level to differ more than a pair of n- and  $\gamma$ -curves at the same energy level. This overlap becomes more pronounced for declining energy levels, a characteristic that - combined with the increasing ruggedness of the curves at these energy levels - is the main reason for the decreasing identification accuracy of PSD at declining energy levels. An illustration of the aforementioned observations is given in Fig. 4, for the energy levels of 2,000 and 150 keVee, respectively.



**Figure 4.** The n- and  $\gamma$ -curves (shown in blue and red, respectively) at the energy levels of 2,000 and 150 keVee; the increasing variability and overlap of the two kinds of curves is evident

Although the entire n- and  $\gamma$ -curves do not promote accurate PSD, a relatively small segment of the curves following the peak can still be used for

n/ $\gamma$  discrimination at all energy levels. This segment comprises the 12 points that correspond to sampling between 29 and 40 nanoseconds (as shown, for example, in Fig. 4), and is employed in the following for performing PSD. An example of the potential for n/ $\gamma$  discrimination within this time interval is shown in Fig 5, independently for the energy levels of 2,000 and 150 keVee. Further to facilitating PSD, using only a small segment of the curves also increases the computational efficiency of PSD.



**Figure 5.** Fig 5. The part of the n- and  $\gamma$ -curves (shown blue and red, respectively) offering optimal n/ $\gamma$  discrimination; illustration for the 2,000 and 150 keVee energy levels

A final observation (further detailed in [35]) is that the  $\gamma$ -curves are significantly more variable than the n-curves at all energy levels, even within the 29-40 nanosecond window. This renders  $\gamma$ 's more prone to misclassification than n's, especially at declining energy levels.

### 3 The Kohonen ANN Architectures Employed for PSD: LVQ and SOM

The two Kohonen ANNs are employed on the  $n$ - $\gamma$  PSD dataset, independently per energy level.

#### 3.1 The LVQ (or LVQ1)

Tailor-made to classification tasks, and widely applied to pattern recognition, classification, and data compression problems, the supervised LVQ [32, 36] comprises:

- the input layer, containing as many nodes as is the dimension of the patterns to be classified;
- a single hidden layer, whose nodes encode the codebook vectors (CVs), i.e. the prototype patterns<sup>1</sup>;
- the competitive (winner-take-all) output layer, with as many nodes as there are classes.

Full feed-forward connectivity is applied between the nodes of the input and the hidden layer; by contrast, the connectivity between the hidden and the output layer is limited to connections between all the CVs representing a given class and the output node of the class. This means that even though a class may be represented by one or more of CVs, each CV represents a single class.

LVQ training is performed iteratively via the modification of the CVs in such a manner as to ensure that all the training patterns are assigned to the output node that corresponds to their class. During a training iteration, each training pattern is presented - in turn - to the LVQ; the distance between the training pattern and the CVs is calculated<sup>2</sup>, and the “winner” CV (WCV, the one which is most similar to the training pattern) is selected. The WCV has its weight modified so as to become more similar to the training pattern if the class of the WCV coincides with that of the training pattern, or more

dissimilar to the training pattern if the class of the WCV is different from that of the training pattern.

CV modification is controlled by the learning rate. By gradually lowering the learning rate over successive training iterations, a progressive fine-tuning of the class boundaries is accomplished. Consequently, at the completion of training, not only are the training patterns assigned to their actual output node (class), but the CVs are assigned weights that fully cover the pattern space in a Voronoi-tessellation-like<sup>3</sup> manner.

Following training, each test pattern is presented to the trained LVQ, and directly assigned to the class represented by the most similar CV.

A significant number of variations and improvements to the original LVQ algorithm have been devised, e.g.

- LVQ2.1 [37], which uses pairs of WCVs that (a) are sufficiently near to the training pattern, but (b) belong to different classes, in order to fine-tune the class boundaries during training;
- LVQ3 [38], which further adds a stabilising term to LVQ2.1 in order to accomplish improved classification behaviour;
- the corresponding OLVQ1 and OLVQ3 (optimised LVQ1 and LVQ3, respectively, [39-42]), where each CV can be assigned its own learning rate.

Although these methodologies have slightly different properties from each other, their common aim is to optimise the class boundaries and/or to maintain the class distribution without either using an excessive number of CVs, or significantly slowing down convergence.

The LVQ and its variants have been found particularly adept at classification tasks concerning statistical data, outlier detection, and other classification/recognition problems with are characterised by complex boundaries between the classes to be distinguished.

<sup>1</sup>Choosing a suitable number of CVs, as well as appropriately initialising their weights promotes swift LVQ training as well as testing. A variety of methodologies exist to this end, including the one followed here, namely using an approximately equal number of training patterns from each class as initial weights for the codebook vectors.

<sup>2</sup>The distance expresses the degree of similarity between presented input vector and CVs; a small/large distance corresponds with a high/low degree of similarity, respectively. The Euclidean distance metric is the most widely used metric for calculating the similarity between input patterns and LVQ codebook vectors, and this metric is also used here.

<sup>3</sup>Without, however, necessarily approximating the density of the training patterns of each class.

### 3.2 The SOM

Similar in inspiration but distinct in terms of training and organisation, the unsupervised SOM [32,43-44] comprises only the first two layers of the LVQ, with its CVs arranged in a one- or two-dimensional hexagonal or rectangular grid, and demonstrating (like the LVQ) full connectivity to the input nodes. Due to the fact that the desired class output is unknown, the CVs are modified during training via the complementary mechanisms of competition and co-operation, resulting in a generally smooth map of the input space with (a) similar neighbouring CVs, and (b) demarcated areas on the CV grid corresponding to different classes. This is accomplished as follows:

- Prior to training, all the connection weights from the input layer to the CVs are randomly initialised to small values.
- During training
- The mechanism of competition is applied for each presentation of a training pattern. The nodes “compete” with each other in order to establish the WCV as the one with the smallest distance from the training pattern; subsequently, the WCV becomes modified so as to become more similar to the training pattern.
- The mechanism of co-operation is applied between neighbouring CVs. This is unlike LVQ training in that the weights of the CVs that are within the neighbourhood of the WCV are also modified by an amount that is proportional to their topological distance from the WCV. More specifically, the neighbourhood around the WCV equals the entire layer of CVs at the beginning of training; this neighbourhood progressively shrinks during training, and becomes confined to the direct neighbours of the WCV at the conclusion of training.

This increasingly competitive, and decreasingly co-operative (usually via an exponential decay), procedure creates a smooth, topologically ordered map with increased similarity between neighbouring CVs; in other words, during training, topologically proximal CVs acquire increasingly similar weights.

An appealing property of the SOM, which stems directly from the one- or two-dimensional grid-like structure of its CV layer, as well as its training procedure, is that a low dimensional and compressed mapping of the problem space is created. This mapping is rendered visible after training by the relative position of the CVs, even for high-dimensional and/or numerous pattern spaces.

Following SOM training, and prior to testing, each CV is assigned to a class; this is usually implemented by:

- either determining the class that is supported by the majority of training patterns that are assigned to the CV, and assigning this class to it,
- or clustering the CVs of the trained SOM into classes, and subsequently assigning the class of the cluster to each CV that belongs to the cluster.

Further to classification, the unified distance matrix (U-matrix [45]) can also be used for visualising the problem space in the one- or two-dimensional grid structure of its CV layer. A gray-scale U-matrix plot can also be created by plotting the distance between neighbouring CVs in shades of gray, thus demarcating clusters of proximal CVs as well as the boundaries between these clusters.

During testing, each test pattern is directly assigned to the class of the most similar CV.

A multitude of applications have been developed making use of the SOM properties, most of these applications falling into the general categories of high-dimensional input space visualisation, dimensionality reduction and data mining (e.g. [46]), and probability density function approximation and maximum likelihood classification of novel patterns (e.g. [47]). Additionally, a number of extensions to the original SOM have been put forward, for instance

- the time adaptive SOM (TASOM, [48]) which employs adaptive learning rates and neighborhood functions in order to render the mapping invariant to scaling, translation and rotation, and
- the growing SOM (GSOM) which starts off with a minimal number of CVs and appropriately creates new CVs in order to optimally represent the problem space [49].

## 4 ANN Training and Operation

### 4.1 Data Preparation

Identical training and test sets, with the training sets demonstrating equal numbers of  $n$ - and  $\gamma$ -patterns, are used by the two ANNs at each energy level. This practice is followed in order to avoid any bias that might occur from an unequal representation of the two classes when setting up the two ANNs<sup>4</sup>. The equivalent representation of the two classes in the training set leaves a variable number of  $n$ - and  $\gamma$ -patterns for testing, ranging from around 50% to around 75% of the corresponding datasets over the different energy levels; datasets with similar numbers of  $n$ - and  $\gamma$ -patterns demonstrate roughly equal sizes of training and test sets (around 50% each), while imbalanced datasets (especially at the energy levels of 1,500 and 1,000 keVee) correspond to significantly larger test sets.

**Table 2.** Proportion of Training Data

Energy level (keVee)	(%) training patterns
2,000	51.61
1,500	26.32
1,000	36.87
500	49.59
300	48.89
150	44.16

The proportion of patterns employed for training the two ANNs is shown in Table 2, independently for each energy level; the remaining proportion represents the relative size of the test sets. This training/test set partitioning is selected as it permits investigation of PDS prediction accuracy by the two ANNs for varying energy levels as well as for balanced and imbalanced datasets.

Aiming at extracting statistically valid results, the random creation of the training (and, consequently, of the test) sets is repeated 30 times per energy level. It is the averaged prediction results of these tests that are shown and discussed in the following two subsections.

### 4.2 ANN Training

In order to create general PSD prediction models, as well as to avoid the creation of ANNs that overfit the training sets (and, thus, whose performance cannot be generalised), the optimal LVQ and SOM architectures are established as being those that permit swift training and accurate PSD performance (during both training and testing) over all energy levels, with parameters such as the number of CVs being optimised via trial-and-error. The selected ANN architectures are:

- a 12-4-2 LVQ, trained in 100 cycles, and implementing each  $n/\gamma$  classification in 113  $\mu$ s;
- a 12-20 SOM with a 4 x 5 hexagonal-lattice CV layer, trained via incremental learning in 100 cycles, and implementing each  $n/\gamma$  classification in 441  $\mu$ s; classification is achieved by assigning a class to each CV according to (a) of Section 3.2.

**Table 3.** Prediction Accuracy on the Training Sets

Energy level (keVee)	LVQ	SOM
2,000	100	100
1,500	100	98.49
1,000	99.93	99.87
500	99.94	99.51
300	99.89	98.62
150	96.74	97.34

Table 3 depicts prediction accuracy of the two ANNs for the training sets of each energy level, and averaged over the 30 trials. Both ANNs are found particularly accurate at recalling the training sets.

<sup>4</sup>The problem of imbalanced datasets has been widely investigated in the related literature, with a variety of solutions proposed, e.g. oversampling the less represented class, undersampling the more represented class, a combination of the two, or variations to these approaches (for a review, the interested reader is referred to [50]).

**Table 4.** Percentage of n- $\gamma$  Missclassifications on the Training Sets

Energy level (keVee)	LVQ	SOM
2,000	n/a	n/a
1,500	n/a	0.45/99.55
1,000	18.32/81.68	2.25/97.75
500	21.04/78.96	5.39/94.61
300	23.57/76.43	10.72/89.28
150	29.17/70.83	16.37/84.63

The LVQ is found - in general - superior to the SOM at predicting the training set, even though a significant drop in LVQ performance is observed for the 150 keVee energy level, which renders the LVQ inferior to the SOM at that energy level. The SOM, on the other hand, although being slightly less accurate overall than the LVQ, demonstrates a milder decline in prediction accuracy at the lowest energy levels.

The excellent recall of the LVQ at the higher and medium energy levels can be explained by its inherent assignment of training patterns from the same class to the same CV during training. This works accurately as well as efficiently, especially for the energy levels that contain more clearly separated training patterns from the two classes. At the lowest energy level of 150 keVee, however, where a significant overlap between n- and  $\gamma$ -patterns is observed, prediction on the training set is not as successful: dissimilar training patterns of the same class are assigned to the same CV, while similar training patterns of different classes are assigned to different CVs. This assignment results in a representation of the class boundaries that disrupts accurate classification even for the known (training) patterns. These points are further supported by the findings of Table IV, which shows the occurrence of n- and of  $\gamma$ -pattern misclassifications, namely the mixed misclassification of both n- and  $\gamma$ -patterns from the energy level of 1,000 down to 300 keVee, with a declining bias towards misclassifying  $\gamma$ -patterns as n-patterns than of n-patterns as  $\gamma$ -patterns. The SOM, on the other hand, although generally not as accurate as the LVQ, shows a significantly clearer bias towards misclassifying  $\gamma$ -patterns as n-patterns; this bias ranges from almost 100% at the energy level of 1,500 keVee to

(almost) four times more probable to misclassify  $\gamma$ -patterns as n-patterns than to misclassify n-patterns as  $\gamma$ -patterns at the energy level of 150 keVee. In other words, although misclassifications are generally more frequent to the SOM, there is a significantly higher probability (and consistency over the declining energy levels) that the misclassification concerns a  $\gamma$ -pattern. This finding is corroborated by the more pronounced variability of the  $\gamma$ - than the n-curves, as already stated at the end of Section 2.

**Table 5.** Prediction Accuracy on the Test Sets

Energy level (keVee)	LVQ	SOM
2,000	100	100
1,500	99.26	99.52
1,000	97.39	99.86
500	98.63	99.07
300	97.56	98.31
150	92.11	96.75

### 4.3 ANN Testing

Both ANNs are, again, found capable of successfully generalising to novel (test) patterns. In contrast to the ANN results of Section 4.2, however, the SOM is found superior to the LVQ in classifying the test patterns at all energy levels, i.e. even at the higher energy levels where the LVQ was found more accurate at predicting the training patterns.

As shown in Table V, the higher discrimination/identification power of the SOM over the LVQ becomes increasingly evident as the energy level declines. In fact, when comparing Tables III and V, the SOM demonstrates a negligible difference in predicting n's and  $\gamma$ 's from the training and test sets at the same energy levels, respectively, whereas the LVQ shows an increasing deterioration in predicting test patterns for declining energy levels, which – as for the training sets – becomes more obvious for the lower energy levels.



**Table 6.** Percentage of n- $\gamma$ -Misclassifications on the Test Sets

Energy level (keVee)	LVQ	SOM
2,000	n/a	n/a
1,500	38.72/61.28	0.59/99.41
1,000	22.76/77.24	2.87/97.13
500	30.62/69.38	6.72/93.28
300	26.52/73.48	11.08/88.92
150	29.88/70.12	16.23/83.76

As far as misclassifications are concerned, the probability of a test  $\gamma$ -pattern (and, thus, also n-pattern) being misclassified remains at comparable levels with the corresponding probability observed for the training sets, at all energy levels, and for both the LVQ and the SOM.

Finally, it is worth noting that, when investigating the identity of the misclassified patterns, around 50% of the erroneously identified n- and  $\gamma$ - patterns are misclassified by both ANNs. This finding highlights the parallels between the two discrimination/identification approaches (for more details the interested reader is referred to [35]). For the remaining patterns, the enhanced performance of the SOM is due to the clustering of the inputs of similar training patterns independent of class, combined with the allocation of a class to each CV according to the class of the majority of training patterns that are assigned to it. Consequently, the SOM becomes less sensitive to misclassifying previously unseen patterns than the LVQ, especially at the more variable patterns and more complex boundaries between classes exhibited at declining energy levels.

#### 4.4 Further Observations - Comparisons

No obvious effect of the proportion of patterns used for training and testing the ANNs has been found on prediction accuracy of either the training or the test sets. It should be mentioned, however, that for the test sets only, and when the LVQ is employed for PSD, identification accuracy is higher at the 500 and 300 keVee energy levels (where the training sets constitute around 50% of the corresponding dataset) than for the higher energy level of 1,000 keVee (where the training set constitutes roughly a third of the corresponding dataset). It should be stressed, however, that further experi-

mentation is necessary in order to establish whether the relative sizes of the pairs of sets actually influence n/ $\gamma$  classification accuracy.

Taking into account the findings concerning the performance of CI-based existing methodologies to PSD described in the Introduction, the use of the two Kohonen ANN architectures (i.e. other than the PB architecture) seems to produce higher PSD accuracy at comparable energy levels, with the results being more robust to the ratios of n- and  $\gamma$ -patterns in the corresponding datasets. Still, a comprehensive and fair comparison can only be made by employing identical datasets, as well as training and test sets.

Although PSD is performed in the related literature at the level of  $\mu$ s per event, the results reported here are produced directly by MATLAB without a hardware realisation of the underlying code; as a result no direct comparisons can be made at this point.

## 5 Conclusion

Two Kohonen ANNs, namely the LVQ and the SOM, have been employed for performing n/ $\gamma$  PSD at an extensive range of energy levels. Both ANNs have been found (a) well-suited to the task, with their discrimination power falling only slightly for declining energy levels, and (b) independent of the relative sizes of the training and test sets.

Although the classification-oriented LVQ allows better recall of the training set, the SOM has been found superior overall: it is more accurate at classifying novel patterns (attributed to the generalisation capability of its clustering-oriented structure) at all energies and for all proportions of training and test set patterns tested.

It should also be reminded that the 12-4-2 LVQ, trained in 100 cycles, implements a n/ $\gamma$  classification in 113 $\mu$ s, whereas the 12-20 SOM with a 4 x 5 hexagonal-lattice CV layer, and trained via incremental learning in 100 cycles, provides the answer in 441 $\mu$ s. Since it is not clear whether the observed time difference can be related to the fact that the underlying algorithms have been implemented by the use of the MATLAB code (C++ shell), a hardware realisation of the MATLAB code will further demonstrate the appeal of the two Kohonen ANNs (as well as the actual differences between them in

terms of efficiency) for actual PSD applications.

Last, but not least, future research shall incorporate the identification of pile-up events.

## References

- [1] R. T. Kouzes, The 3He Supply Problem, Report 413 No. PNNL-18388, Pacific 414 Northwest National Laboratory, Richland, WA, 2009
- [2] A. Enqvist, placeI. Pzsit, S. Avdic, Sample characterization using both neutron and gamma multiplicities, *Nuclear Instruments & Methods A*, vol. 615, pp. 62-69, 2010
- [3] V. L. Romodanov, V. K. Sakharov, A. G. Belevitin, V. V. Afanas'ev, I. V. Mukhamad'varov, D. N. Chernikov, Computational-experimental studies of a facility for detecting fissile materials in airports, *Atomic Energy*, vol. 105, pp. 118-123, 2008
- [4] Y. N. Barmakov, E. P. Bogolyubov, O. V. Bochkarev et al., 2011, System of combined active and passive control of fissile materials and their nuclide composition in nuclear wastes, *International Journal of Nuclear Energy Science and Technology*, vol. 6, pp. 127-135, 2011
- [5] D. Chernikova, V. Romodanov, V. Sakharov, A. Isakova, Analysis of  $^{235}\text{U}$ ,  $^{239}\text{Pu}$  and  $^{241}\text{Pu}$  content in a spent fuel assembly using lead slowing down spectrometer and time intervals matrix, *Journal of Nuclear Materials Management*, vol.40, pp. 9-18, 2012
- [6] D. Chernikova, V. Romodanov, V. Sakharov, Development of the neutron-gamma-neutron (NGN) approach for the fresh and spent fuel assay, the 53rd Annual Meeting of the Institute of Nuclear Materials Management, Orlando, Florida USA, July 2012
- [7] S. A. Pozzi, M. M. Bourne, S. D. Clarke, Pulse shape discrimination in the plastic scintillator EJ-299-33, *Nuclear Instruments and Methods in Physics Research Section A*, vol. 723, 21 pp. 19-23, 2013
- [8] F. D. Brooks, A scintillation counter with neutron and gamma-ray discriminators, *Nuclear Instruments and Methods*, vol. 4, pp. 151-163, 1959
- [9] F. T. Kuchnir, F. J. Lynch, Time dependence of scintillations and the effect on pulse-shape discrimination, *IEEE Transactions on Nuclear Science*, vol. 15, pp. 107–113, 1968
- [10] P. Sperr, H. Spieler, M. R. Maier, D. Evers, A simple pulse-shape discrimination circuit, *Nuclear Instruments and Methods A*, vol. 116, pp. 55-59, 1974
- [11] S. Marrone, D. Cano-Ott, N. Colonna, C. Domingo, F. Gramegna, E. M. Gonzalez, F. Gunsing, M. Heil, F. Kappeler, Pulse shape analysis of liquid scintillators for neutron studies, *Nuclear Instruments and Methods in Physics Research A*, vol. 490, pp. 299-307, 2002
- [12] B. D. Mellow, M. D. Aspinall, R. O. Mackin, M. J. Joyce, A. J. Peyton, Digital discrimination of neutrons and gamma-rays in liquid scintillators using pulse gradient analysis, *Nuclear Instruments and Methods A*, vol. 578, pp. 191-197, 2007
- [13] D. I. Shippen, M.J. Joyce, M. D. Aspinall, A wavelet packet transform inspired method of neutron-gamma discrimination, *IEEE Transactions on Nuclear Science*, vol. 57, pp. 2617 – 2624, 2010
- [14] G. Liu, M. J. Joyce, X. Ma, M. D. Aspinall, A digital method for the discrimination of neutrons and gamma rays with organic scintillation detectors using frequency gradient analysis, *IEEE Transactions on Nuclear Science*, vol. 57, pp. 1682 – 1691, 2010
- [15] D. Wolski, M. Moszynski, T. Ludziejewski, A. Johnson, W. Klamra, O. Skeppstedt, Comparison of n- $\gamma$  discrimination by zero-crossing and digital charge comparison methods, *Nuclear Instruments and Methods A*, vol. 360, pp. 584-592, 1995
- [16] I. V. Muhamadyarov, Nondestructive testing and detection of fissile and radioactive materials in systems with pulsed neutron sources and digital processing of the experimental data, PhD Dissertation, Russian State Library Electronic Catalogue (OPAC), 2009
- [17] K. A. A. Gamage, M. J. Joyce, N. P. Hawkes, A comparison of four different digital algorithms for pulse-shape discrimination in fast scintillators, *Nuclear Instruments and Methods A*, vol. 642, pp. 78-83, 2011
- [18] R. E. Rumelhart, G. E. Hinton, R. J. Williams, Learning representations by back propagating errors, *Nature*, vol. 323, pp. 533-536, 1986
- [19] M. Riedmiller, H. Braun, A direct adaptive method for faster backpropagation learning: the rprop algorithm, *Proceedings of the IEEE Conference on Neural Networks*, San Francisco, CA, March 28<sup>th</sup>-April 1<sup>st</sup>, 1993, pp. 586–591, 1993
- [20] M. Moller, A scaled conjugate gradient algorithm for fast supervised learning, *Neural Networks*, vol. 6, pp. 525-533, 1993

- [21] Z. Cao, L. F. Miller, M. Buckner, Implementation of dynamic bias for classifiers, *Nuclear Instruments and Methods A*, vol. 416, pp. 438-445, 1998
- [22] B. Esposito, L. Fortuno, A. Rizo, Neural neutron/gamma discrimination in organic scintillators for fusion applications, *Proceedings of the 2004 IEEE International Joint Conference on Neural Networks*, Budapest, Hungary, July 25-29, 2004, vol. 4, pp. 2931-2936, 2004
- [23] P. Guazzoni, F. Previdi, S. Russo, M. Sassi, S. M. Saravesi, Pulse shape analysis using subspace identification methods and particle identification using neural networks in CsI(Tl) scintillators, *Proceedings of the IEEE Nuclear Science Symposium and Medical imaging*, Puerto Rico, October 23th -29th, 2005, vol. 3, pp. 1341-1345, 2005
- [24] D. Wisniewski, M. Wisniewska, P. Bruyndonckx, M. Krieguer, S. Tavernier, O. Devroede, C. Lemaitre, J. B. Mosset, C. Morel, Digital pulse shape discrimination methods for phoswich detectors, *Proceedings of the IEEE Nuclear Science Symposium and Medical imaging*, Puerto Rico, October 23th -29th, 2005, vol. 5, pp. 2979-2983, 2005
- [25] L. Bertalot, B. Esposito, Y. Kascuck, D. Marocco, M. Riva, A. Rizzo, D. Skopintsev, Fast digitizing techniques applied to scintillation detectors, *Nuclear Physics B – Proceedings supplement of the Proceedings of the 9<sup>th</sup> Topical Seminar on innovative Particle and Radiation Detectors*, May 23rd-26<sup>th</sup>, 2004, Siena, Italy, 2004, vol. 150, pp. 78-81, 2006
- [26] J. Gill, T. Persson, K. Sjogren, K. Sols, N. Sundstrom, S. Wranne, Identification of ions by pulse-shape analysis and evaluation of Lyso scintillator crystal, *Technical Report*, Chalmers University of Technology, Goteborg, Sweden, 2008
- [27] G. Liu, M. D. Aspinall, X. Ma, M. J. Joyce, An investigation of the digital discrimination of neutron and  $\gamma$  rays with organic scintillation detectors using a artificial neural network, *Nuclear Instruments and Methods in Physics Research A*, vol. 607, pp. 620-628, 2009
- [28] E. Ronchi, P. A. Soderstrom, J. Nyberg, E. Andersson Sunden, S. Conroy, G. Ericsson, C. Hellen, M. Gatu Johnson, M. Weiszflog, An artificial neural network based neutron-gamma discrimination and pile-up rejection framework for the BC-501 scintillation detector, *Nuclear Instruments and Methods in Physics Research A*, vol. 610, pp. 534-539, 2009
- [29] R. Jimenez, M. Sanchez-Raya, J. A. Gomez-Galan, J. L. Flores, J. A. Duenas, I. Martel, Implementation of a neural network for digital pulse shape analysis on FPGA for on-line identification of heavy ions, *Nuclear Instruments and Methods in Physics Research A*, vol. 674, pp. 99-104, 2012
- [30] N. Yildiz, S. Akkoyun, Neural network consistent empirical physical formula construction for neutron-gamma discrimination in gamma ray tracking, *Annals of Nuclear Energy*, vol. 51, pp. 10-17, 2013
- [31] T. Kohonen, Learning vector quantization for pattern recognition, Report TKK-F-A601, Helsinki University of Technology, Espoo, Finland, 1986
- [32] T. Kohonen, Self-organized formation of topologically correct feature maps, *Biological Cybernetics*, vol. 43, no. 1, pp. 59-69, 1982
- [33] MATLAB R2009a, Mathworks, February 2009
- [34] D. Chernikova, K. Axell, I. Pzsit, A. Nordlund, R. Sarwar, A direct method for evaluating the concentration of boric acid in a fuel pool using scintillation detectors for joint-multiplicity measurements, *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 714, pp. 90-97, 2013
- [35] T. Tambouratzis, D. Chernikova, I. Pzsit, A comparison of artificial neural network performance: the case of neutron/gamma pulse shape discrimination, *2013 IEEE Symposium Series on Computational Intelligence (CISDA)*, Singapore, April 16<sup>th</sup>-19<sup>th</sup>, 2013, pp. 88-95, 2013
- [36] T. Kohonen, Learning vector quantization, *The handbook of brain theory and neural networks* (M.A. Arbib, editor), MIT Press, Cambridge, MA, pp. 537-540, 1995
- [37] T. Kohonen, Improved versions of learning vector quantization, *Proceedings of the International Joint Conference on Neural Networks*, San Diego, California, June 17<sup>th</sup>, 1990, vol. 1, pp. 545-550, 1990
- [38] T. Kohonen, J. Hynninen, J. Kangas, J. Laaksonen, K. Torkkola, LVQ\_PAK: the learning vector quantization programming package, Report A30, Helsinki University of Technology, Laboratory of Computer and Information Science, Espoo, Finland, 1996
- [39] T. Kohonen, The self-organizing map. *Proceedings of the IEEE*, vol. 78, no. 9, pp. 1464-1480, 1990
- [40] T. Kohonen, Statistical pattern recognition revisited, *Advanced Neural Computers*, pp. 137-144, 1990

- [41] C. Zhu, J. Wang, T. Wang, Analysis of learning vector quantization algorithms for pattern classification, International Conference on Acoustics, Speech, and Signal Processing, Detroit, MI, May 9th-12<sup>th</sup>, 1995, vol. 5, pp. 3471-3474, 1995
- [42] M. T. Hagan, H. B. Demuth, M. Beale, Neural Networks Design, PWS Publishing, CO, U.S.A., 1996
- [43] T. Kohonen, Self-Organizing Maps. Springer, Berlin, 1995
- [44] S. Haykin, Self-organizing maps (chapter 9), in Neural Networks - A Comprehensive Foundation (2nd ed.), Prentice-Hall, 1999
- [45] A. Ultsch, Emergence in self-organizing feature maps, in H. Ritter, R. Haschke (eds), Proceedings of the 6th International Workshop on Self-Organizing Maps (WSOM '07). Bielefeld, Germany, 2007
- [46] S. Kaski, Data Exploration Using Self-Organizing Maps. PhD thesis, Helsinki University of Technology, Espoo, Finland, 1997
- [47] A. Hmlinen, Self-organizing map and reduced kernel density estimation, PhD thesis, University of Jyväskylä, Jyväskylä, Finland, 1995
- [48] H. Shah-Hosseini, R. Reza, TASOM: a new time adaptive self-organizing map, IEEE Transactions on Systems, Man, and Cybernetics - Part B: Cybernetics, vol. 33, pp. 271-282, 2003
- [49] D. Alahakoon, S. K. Halgamuge, B. Sirinivasan, Dynamic self organizing maps with controlled growth for knowledge discovery, IEEE Transactions on Neural Networks, Special Issue on Knowledge Discovery and Data Mining, vol. 11, pp. 601-614, 2000
- [50] N. V. Chawla, Data Mining for Imbalanced Datasets: An Overview, pp. pages 875-886, in O. Maimon, L. Rokach (eds), Data Mining and Knowledge Discovery Handbook, Springer, 2010