

# Multivariate Data Retrieval Modified by Random Noise using Lattice Autoassociative Memories with Eroded or Dilated Input Residuals

Gonzalo Urcid<sup>1,\*</sup>, Rocío Morales-Salgado<sup>2,\*\*</sup>, and José-Angel Nieves-Vázquez<sup>3</sup>

<sup>1</sup>Optics Department, INAOE, Tonantzintla 72840, Puebla, Mexico

<sup>2</sup>Information Technology and Data Science Department, UPAEP, Puebla 72410, Puebla, Mexico

<sup>3</sup>Engineering Division, ITSSAT, Maticapan 95804, San Andres Tuxtla, Veracruz, Mexico

**Abstract.** Lattice associative memories were proposed as an alternative approach to work with a set of associated vector pairs for which the storage and retrieval stages are based in the theory of algebraic lattices. Several techniques have been established to deal with the problem of binary or real valued vector recall from corrupted inputs. This paper presents a thresholding technique coupled with statistical correlation pattern index search to enhance the recall performance of lattice auto-associative memories for multivariate data inputs degraded by random noise. By thresholding a given noisy input, a lower bound is generated to produce an eroded noisy version used to boost the min-lattice auto-associative memory inherent retrieval capability. Similarly, an upper bound is generated to obtain a dilated noisy version used to enhance the max-lattice auto-associative memory response. A self contained theoretical foundation is provided including a visual example of a multivariate data set composed of grayscale images that show the increased retrieval capability of this type of associative memories.

## 1 Introduction

Morphological associative memories (MAMs) were introduced more than two decades ago as a new paradigm for the storage and recall of pattern associations [3]. MAMs are feedforward, fully connected neural networks, in which the interconnection weights between input and output neurons follow a properly defined hebbian type learning rule. In most correlation type associative memory models, the storage and recall stages use conventional algebra whereas computation in MAMs are based on the mathematical theory of minimax algebra [4–6]; since the standard minimax matrix algebra is a specific instance of a lattice algebraic structure, MAMs are also named *lattice associative memories* (LAMs). Once a set of exemplar pattern pairs is imprinted in the neural network, then it is anticipated that such a memory will manifest a certain capability to recall the correct association when presented with an exemplar pattern (perfect input) or with a distorted version (non-perfect input). It is important to reflect that both input cases pose a difficult robustness problem for any associative memory model. For perfect input, it was shown that the canonical auto-associative morphological memories (AMMs) have unlimited storage capacity, give perfect recall for all exemplar patterns, and are robust for *exclusively* erosive or dilative noise; based on stronger assumptions, similar results were corroborated for hetero-associative morphological memories (HMMs). To dispense with inputs corrupted by mixed random noise, i. e. by a combination of both erosive and dilative noise,

the kernel method that binds the max and min AMMs in a two-stage memory scheme was designed for binary [3, 7, 8] and real valued patterns [9]; in addition, distinct computational perspectives or model extensions were developed to enlarge its recall capability or diversify their applicability in pattern recognition problems. In addition to the kernel method, an algorithm based on induced morphological strong independence for real valued exemplar patterns was proposed in [10] to deal with inputs corrupted by mixed random noise. Parallel to these technical developments, a deeper understanding has been achieved that afforded useful results on the fixed point set of LAAMs as well as a complete mathematical characterization of their response to arbitrary inputs [6, 11]. Within this frame of reference, numerical examples and a practical discussion can be found in [6], and a comparison of AMMs with other enhanced associative memory neural networks for grayscale pattern retrieval appears in [11], though *only* for erosive or dilative random noise.

We point out the relevant fact that the strengths and weaknesses of the min and max auto-associative memories have conducted the different developments of previous work done on MAMs as briefly described above. The recall failure of the min AMM for inputs degraded by dilative noise and similarly, the poor performance of the max AMM for inputs corrupted with erosive noise, and consequently their inapplicability to deal with mixed random noise has been exposed repeatedly since their introduction [3, 7] and revisited across the most recent theoretical results discovered about them [6, 11]. By capitalizing the robustness to either erosive or dilative noise, respectively, of the min- and max-LAAMs, a practical solution

\*e-mail: [gurcid@inaoep.mx](mailto:gurcid@inaoep.mx) (corresponding author)

\*\*e-mail: [mariadelrocio.morales@upaep.mx](mailto:mariadelrocio.morales@upaep.mx)

to pattern recall from inputs corrupted by mixed random noise using only one of the canonical LAAMs was initiated by masking the noise contained in the corrupted input pattern [12–14]. Based on selective pattern indexing, further improvement of the noise masking strategy and its extension to fuzzy type associative morphological memories have been recently reported in [15, 16] and [17]. To regain the LAAMs inherent computational functionality to cope with different amounts of impulsive random noise, in this paper, we introduce a novel technique using thresholded noisy inputs followed by statistical correlation pattern indexing, to generate an erosive or dilative noisy version of the input pattern and show the LAAMs enhanced recall performance by means of an illustrative example.

Our work is organized as follows: Section 2 gives some mathematical background on minimax algebra covering the basic lattice matrix operations used to define and apply the canonical LAAMs, including a brief description of the main theoretical results regarding their recall performance; Section 3 explains the idea of noise thresholding and correlation pattern search that will be used to produce lower and upper noise bounds useful for pattern retrieval. In Section 4, we give an algorithm that employs the proposed technique to boost a single LAAM recall capability for real valued input patterns contaminated with mixed random impulse noise. We close our paper by presenting in Section 5 our conclusions and comments to the research presented here.

## 2 Mathematical Background

### 2.1 Basic Lattice Matrix Algebra

The numerical operations of taking the maximum or minimum of two numbers, usually denoted as functions  $\max(x, y)$  and  $\min(x, y)$ , will be written as binary operators using the “join” and “meet” symbols employed in lattice theory, i.e.,  $x \vee y = \max(x, y)$  and  $x \wedge y = \min(x, y)$  [1, 2]. It was shown in [5, 6] that the algebraic systems  $(\mathbb{R}_{-\infty}, \vee, +)$  and  $(\mathbb{R}_{+\infty}, \wedge, +')$  are *semirings* for the corresponding  $\max$ ,  $\vee$ ,  $\min$ ,  $\wedge$  operations and their respective additions,  $+$  and  $+$ ' over the sets  $\mathbb{R}_{-\infty} = \mathbb{R} \cup \{-\infty\}$  and  $\mathbb{R}_{+\infty} = \mathbb{R} \cup \{+\infty\}$ . Since we will consider only *finite* values of  $x$  and  $y$ , then the sum operation in the aforementioned semirings coincide, that is,  $x + ' y = x + y$ . In what follows, if  $n$  is any positive integer, the index set  $I_n$  is an abbreviation for the list of numbers  $\{1, 2, \dots, n\}$ .

We use lattice matrix operations that are defined component-wise using the underlying structure of  $\mathbb{R}_{-\infty}$  or  $\mathbb{R}_{+\infty}$  as semirings. For example, the maximum of two matrices  $X, Y$  of size  $m \times n$  is defined as shown in (1) for all  $i \in I_m$  and  $j \in I_n$ :

$$(X \vee Y)_{ij} = x_{ij} \vee y_{ij}. \quad (1)$$

Inequalities between matrices are also verified element-wise, e. g.,  $X \leq Y$  if and only if  $x_{ij} \leq y_{ij}$  for all  $i, j$ . On the other hand, the *conjugate matrix*  $X^*$  is defined as  $-X^t$  where  $X^t$  denotes usual matrix transposition.

Let  $X$  be a matrix of size  $m \times p$  and  $Y$  be a matrix of size  $p \times n$ . The *max-sum* of  $X$  and  $Y$ , denoted by  $X \nabla Y$  is the

matrix operation defined by (2) for all  $i \in I_m$  and  $j \in I_n$ . In similar fashion, the *min-sum* of  $X$  and  $Y$ , denoted by  $X \Delta Y$  is the matrix operation given in (3).

$$(X \nabla Y)_{ij} = \bigvee_{k=1}^p (x_{ik} + y_{kj}), \quad (2)$$

$$(X \Delta Y)_{ij} = \bigwedge_{k=1}^p (x_{ik} + y_{kj}). \quad (3)$$

The *minimax outer sum* of vectors  $\mathbf{x} = (x_1, \dots, x_n)^t \in \mathbb{R}^n$  and  $\mathbf{y} = (y_1, \dots, y_m)^t \in \mathbb{R}^m$ , is given by the  $m \times n$  matrix

$$\mathbf{y} \boxplus \mathbf{x}^t = \begin{pmatrix} y_1 + x_1 & \cdots & y_1 + x_n \\ \vdots & \ddots & \vdots \\ y_m + x_1 & \cdots & y_m + x_n \end{pmatrix}. \quad (4)$$

It is important to observe that the minimax outer sum displayed in (4) is a particular case of the max-sum and min-sum matrix operations since,  $\mathbf{y} \boxplus \mathbf{x}^t = \mathbf{y} \nabla \mathbf{x}^t = \mathbf{y} \Delta \mathbf{x}^t$ .

If  $X = \{\mathbf{x}^1, \dots, \mathbf{x}^k\} \subset \mathbb{R}^n$  represents a multivariate data set, then a *linear minimax combination* of vectors from  $X$  is any vector  $\mathbf{x} \in \mathbb{R}^n$  of the form

$$\mathbf{x} = \bigvee_{i \in I} \bigwedge_{\xi=1}^k (\mathbf{x}^\xi + \alpha_i^\xi), \quad (5)$$

where  $I$  is a finite index set and  $\alpha_i^\xi \in \mathbb{R}$  for all  $i \in I$  and all  $\xi \in I_k$ . The set of all such minimax combinations is called the *real linear minimax span* of  $X$  and is denoted by  $LMS_{\mathbb{R}}(X)$ . Also, given a real  $n$ -dimensional self-transformation,  $\mathcal{T} : \mathbb{R}^n \rightarrow \mathbb{R}^n$ , then  $\mathbf{x} \in \mathbb{R}^n$  is called a *fixed point* of  $\mathcal{T}$  if and only if  $\mathcal{T}(\mathbf{x}) = \mathbf{x}$  (see, e. g. [6]). The previous definitions and equations are the necessary tools to discuss the theoretical background of lattice associative memories as explained in the next subsection.

### 2.2 Lattice Associative Memories

Let  $(\mathbf{x}^1, \mathbf{y}^1), \dots, (\mathbf{x}^k, \mathbf{y}^k)$  be  $k$  vector pairs with  $\mathbf{x}^\xi = (x_1^\xi, \dots, x_n^\xi)^t \in \mathbb{R}^n$  and  $\mathbf{y}^\xi = (y_1^\xi, \dots, y_m^\xi)^t \in \mathbb{R}^m$  for  $\xi \in I_k$ . For a given set of pattern associations  $\{(\mathbf{x}^\xi, \mathbf{y}^\xi) \mid \xi \in I_k\}$  we define a pair of associated pattern matrices  $(X, Y)$ , where  $X = (\mathbf{x}^1, \dots, \mathbf{x}^k)$  and  $Y = (\mathbf{y}^1, \dots, \mathbf{y}^k)$ . Thus,  $X$  is of dimension  $n \times k$  with  $i, j$ th entry  $x_i^j$  and  $Y$  is of dimension  $m \times k$  with  $i, j$ th entry  $y_i^j$ . To store  $k$  vector pairs  $(\mathbf{x}^1, \mathbf{y}^1), \dots, (\mathbf{x}^k, \mathbf{y}^k)$  in an  $m \times n$  *lattice associative memory* we encode associations as in linear or correlation memories but instead of using the linear outer product, the minimax outer sum is used as follows [3]. The  $(X, Y)$ -min lattice associative memory, represented by  $W_{XY}$ , and the  $(X, Y)$ -max lattice associative memory, denoted by  $M_{XY}$ , both of size  $m \times n$ , are given, respectively, by the expressions,

$$W_{XY} = \bigwedge_{\xi=1}^k (\mathbf{y}^\xi \boxplus \mathbf{x}^{\xi*}); \quad w_{ij} = \bigwedge_{\xi=1}^k (y_i^\xi - x_j^\xi), \quad (6)$$

$$M_{XY} = \bigvee_{\xi=1}^k (\mathbf{y}^\xi \boxplus \mathbf{x}^{\xi*}); \quad m_{ij} = \bigvee_{\xi=1}^k (y_i^\xi - x_j^\xi). \quad (7)$$

The expressions to the left of (6) and (7) are in matrix form and the right expression is the  $ij$ th entry of the corresponding memory. We remark that based on (4), for each  $\xi$ ,  $\mathbf{y}^\xi \boxplus (-\mathbf{x}^\xi)^t$  is a matrix  $A^\xi$  of size  $m \times n$  that *memo-rizes* or stores the *association pair*  $(\mathbf{x}^\xi, \mathbf{y}^\xi)$ . Therefore,  $W_{XY} = \bigwedge_{\xi=1}^k A^\xi$  and  $M_{XY} = \bigvee_{\xi=1}^k A^\xi$ , expressions that clarify their names.

The retrieval of pattern  $\mathbf{y}^\xi$  from pattern  $\mathbf{x}^\xi$  can be expressed using either the min-memory  $W_{XY}$  or the max-memory  $M_{XY}$  as computed with the corresponding equation in (8),

$$\mathbf{y}^\xi = W_{XY} \nabla \mathbf{x}^\xi \quad \text{or} \quad \mathbf{y}^\xi = M_{XY} \Delta \mathbf{x}^\xi, \quad (8)$$

We point out the important fact that  $W_{XY}$  operates on  $\mathbf{x}^\xi$  using the max-sum and, dually, that  $M_{XY}$  operates on  $\mathbf{x}^\xi$  using the min-sum. On the other hand, if  $X \neq Y$ , then we speak of a lattice *hetero-associative* memory (LHAM) and, if  $X = Y$  we obtain a lattice *auto-associative* memory (LAAM). In this paper we restrict our discussion to the canonical LAAMs,  $W_{XX}$  and  $M_{XX}$  whose size is  $n \times n$ . Also, in the remaining sections, without loss of generality, we will consider pattern entries to be *non-negative* numbers, i. e.,  $\mathbf{x}^\xi \in \mathbb{R}_0^+$  meaning  $x_i^\xi \geq 0$  for all  $i, \xi$ .

### 2.3 Main Theoretical Results

The theorems listed here give a quick overview of the theoretical foundations of LAAMs that are of practical significance in applications; detailed mathematical proofs can be found in [3, 6] or [7]. In most cases, statements given in the theorems refer only to the min-memory  $W_{XX}$  since analogous properties for the max-memory  $M_{XX}$  are readily obtained using the *principle of duality* in algebraic lattices. LAAMs have *unlimited* storage capacity, the pattern domain can be binary or real valued, give *perfect recall* for perfect input, and computation is performed in one step, free of any convergence problems [3]. In mathematical form:

**Theorem 1**  $W_{XX} \nabla X = X$ , for any matrix  $X$  of size  $n \times k$  whose columns are the exemplar pattern vectors  $\mathbf{x}^1, \dots, \mathbf{x}^k \in \mathbb{R}^n$ . Furthermore, if  $W_{XX} \nabla \mathbf{x} = \mathbf{x}'$ , then  $W_{XX} \nabla \mathbf{x}' = \mathbf{x}'$ , where  $\mathbf{x} \in \mathbb{R}^n$  is any input vector and  $\mathbf{x}'$  is the recalled pattern.

Notice that the max-sum of  $W_{XX}$  with vector  $\mathbf{x}$  is an example of a *lattice matrix transform* defined by  $\mathcal{T}_W(\mathbf{x}) = W_{XX} \nabla \mathbf{x}$ ; therefore, Theorem 1 means that any exemplar pattern  $\mathbf{x}^\xi$  for  $\xi \in I_k$  is a fixed point of  $\mathcal{T}_W$  (complete perfect recall of all exemplar data). In addition, the recalled pattern  $\mathbf{x}'$  obtained in the first step is a fixed point for the second application of  $W_{XX}$  (one step convergence). The *fixed point set* of the min-memory  $W$  is defined as

$$F(X) = \{\mathbf{x} \in \mathbb{R}^n : W_{XX} \nabla \mathbf{x} = \mathbf{x}\}. \quad (9)$$

The next theorem gives a complete algebraic characterization of the recall response of the LAAMs in terms of linear minimax combinations.

**Theorem 2** The fixed point set  $F(X)$  of  $W_{XX}$  and  $M_{XX}$  is the same and is equal to  $LMS_{\mathbb{R}}(X)$ .

Consequently, from (5) and (9), the result settled in Theorem 1 means that, for any input vector  $\mathbf{x} \in \mathbb{R}^n$  the recalled pattern  $\mathbf{x}'$  belongs to  $F(X)$ . Hence, if  $\alpha_i^\xi \in \mathbb{R}$  for all  $i \in I_n$  and all  $\xi \in I_k$ , the following equation is valid

$$W_{XX} \nabla \mathbf{x} = \mathbf{x}' = \bigvee_{i \in I} \bigwedge_{\xi=1}^k (\mathbf{x}^\xi + \alpha_i^\xi). \quad (10)$$

Observe that the input vector  $\mathbf{x}$  appearing in (10) can be an exemplar pattern  $\mathbf{x}^\xi$ , a noisy version of it, denoted by  $\tilde{\mathbf{x}}^\xi$ , or even a non-corrupted non-exemplar pattern. In addition, the right-hand side of (10) means also that  $W_{XX}$  has many spurious states that are stable. Although equation (10) conveys a deep and nice result, in general *does not* provide a practical criterion to enhance the recall performance of the canonical LAAMs for inputs corrupted by mixed noise.

Missing parts, oclusions or corruption of exemplar patterns can be considered as “noise”. Particularly, if alterations in pattern entries follow a probability law, then we refer to *random noise*. In addition, to specify the recall capabilities of LAAMs when presented with non-perfect inputs, noise can be classified plainly in three types with respect to numerical ordering. Thus, let  $I = \{1, \dots, n\}$ , then, a distorted version  $\tilde{\mathbf{x}}$  of pattern  $\mathbf{x}$  has undergone an *erosive change* whenever  $\tilde{\mathbf{x}} \leq \mathbf{x}$  or equivalently if  $\forall i \in I, \tilde{x}_i \leq x_i$ . A *dilative change* occurs whenever  $\tilde{\mathbf{x}} \geq \mathbf{x}$  or equivalently if  $\forall i \in I, \tilde{x}_i \geq x_i$ . Let  $L, G \subset I$  be two non-empty disjoint sets of indices. If  $\forall i \in L, \tilde{x}_i < x_i$  and  $\forall i \in G, \tilde{x}_i > x_i$ , then the distorted pattern  $\tilde{\mathbf{x}}$  is said to contain *mixed noise* (random or non-random). Due to the preceding inequalities, eroded and dilated versions of a modified vector correspond respectively to lower and upper bounds of  $\tilde{\mathbf{x}}$ .

**Theorem 3** Let  $\tilde{\mathbf{x}}^\gamma$  be a distorted version of pattern  $\mathbf{x}^\gamma$ , then  $W_{XX} \nabla \tilde{\mathbf{x}}^\gamma = \mathbf{x}^\gamma$  holds if and only if the following conditions are satisfied for all  $i \in I_n$ :

- 1)  $w_{ij} \leq x_i^\gamma - \tilde{x}_j^\gamma$ , for each  $j \in I_n$ , and
- 2)  $w_{ij_i} = x_i^\gamma - \tilde{x}_{j_i}^\gamma$ , for at least an index  $j_i \in I_n$ .

Similarly,  $M_{XX} \Delta \tilde{\mathbf{x}}^\gamma = \mathbf{x}^\gamma$  holds if and only if the following conditions are satisfied for all  $i \in I_n$ :

- 1)  $m_{ij} \geq x_i^\gamma - \tilde{x}_j^\gamma$ , for each  $j \in I_n$ , and
- 2)  $m_{ij_i} = x_i^\gamma - \tilde{x}_{j_i}^\gamma$ , for at least an index  $j_i \in I_n$ .

**Corollary 4** If  $W_{XX} \nabla \tilde{\mathbf{x}}^\gamma = \mathbf{x}^\gamma$  then  $\tilde{\mathbf{x}}^\gamma$  must be an eroded version of  $\mathbf{x}^\gamma$ , i. e.,  $\tilde{\mathbf{x}}^\gamma \leq \mathbf{x}^\gamma$ . Dually, if  $M_{XX} \Delta \tilde{\mathbf{x}}^\gamma = \mathbf{x}^\gamma$  then  $\tilde{\mathbf{x}}^\gamma$  must be a dilated version of  $\mathbf{x}^\gamma$ , i. e.,  $\tilde{\mathbf{x}}^\gamma \geq \mathbf{x}^\gamma$ .

The fundamental results stated in Theorem 3 and Corollary 4 give the conditions to guarantee *perfect recall* from a distorted exemplar corrupted by erosive or dilative noise applying the min-memory  $W_{XX}$  or the max-memory  $M_{XX}$ . Additional theoretical discussions [7, 11],

useful commentaries and explanations using numerical examples [3, 6], and a wide spectrum of computational experiments with different sets of binary or gray-scale images [3, 7, 9, 11], have demonstrated the surprising fact that  $W_{XX}$  and  $M_{XX}$  are quite sturdy to deal, respectively, with erosive or dilative noise. However, both LAAMs fail for inputs corrupted with mixed noise, a limitation that has been emphasized several times in all enhanced models developed so far. Nevertheless, from a pragmatic point of view a simple reasonable question comes to our minds. If the canonical LAAMs have a firm mathematical foundation, give results of far reaching applications, and are quite robust to erosive or dilative distortions, is it possible to implement a mechanism that can take advantage of all their known properties and therefore be capable of dealing with mixed noise? The answer is affirmative and we explain our proposal in detail in the following two sections.

To attain our objective we need to relax the concept of “perfect recall” based on an adequate error measure between a LAAM’s recalled vector from a given noisy input and its corresponding associated exemplar. Here we use the *normalized mean squared error* (NMSE), denoted by  $\rho$ , to measure the proximity between two vectors [18]. In common practice,  $\rho$  returns values that agree more objectively with the perceived differences when comparing, for example, two one-dimensional signal graphs or two-dimensional grayscale images. Thus, the idea of “almost perfect recall” can be formalized in the following sense [12, 13]: the min-W or max-M memories give *almost perfect recall* of an exemplar vector  $\mathbf{x}^\xi$  from a noisy input version  $\tilde{\mathbf{x}}^\xi$  for  $\xi \in I_k$  if and only if each LAAM satisfies the corresponding inequality given in (11),

$$\rho(W_{XX} \nabla \tilde{\mathbf{x}}^\xi, \mathbf{x}^\xi) \leq \epsilon \quad \text{and} \quad \rho(M_{XX} \Delta \tilde{\mathbf{x}}^\xi, \mathbf{x}^\xi) \leq \epsilon, \quad (11)$$

where  $\epsilon > 0$  is a small positive number close to zero.

### 3 Noisy Input Bounds for LAAMs

Based on the theoretical results given by Theorem 3 and Corollary 4 in Section 2 as well as the qualitative comments made in the last paragraph of the same section, the following two subsections present an alternative approach to the noise masking technique developed in earlier papers [12–14] and its improvements as described in [15–17].

#### 3.1 Noise Lower-Upper Thresholding

Once a dataset of exemplar vectors are stored in a LAAM we can use either the min-memory  $W_{XX}$  or the max-memory  $M_{XX}$  as data recognizers by association. However, before the recalling stage, the data used as input can be altered due to different causes, such as, missing, saturated or changed values. Although data alteration may occur in a “structured” way because of periodic or deterministic variations, we will focus our attention to the more general case of additive random noise with varying degrees of impulsiveness. Thus, let  $\mathbf{x} \in \mathbb{R}_0^+{}^n$  be a vector with entries  $x_i$  belonging to a bounded interval  $[a, b]$  of  $\mathbb{R}_0^+$

for all  $i \in I_n$ . Let  $\epsilon \in (a, b)$ , then a noisy version  $\tilde{\mathbf{x}}$  will have entries given by,

$$\tilde{x}_i = \begin{cases} x_i & \Leftrightarrow i \in J \subseteq I_n, \\ x_i \pm \epsilon & \Leftrightarrow i \in K \subseteq I_n, \end{cases} \quad (12)$$

where,  $J \cup K = I_n$ ,  $J \cap K = \emptyset$ , and the constant  $\epsilon > 0$  is added or subtracted following a uniform probability density function on the unit interval  $(0, 1)$ . The index subsets  $J$  and  $K$  correspond, respectively, to the *unmodified* and *modified* vector entries. Since  $n = |J| + |K|$ , the cardinalities of  $J$  and  $K$  determine the amount of noise present in  $\tilde{\mathbf{x}}$ , and if  $\mathbf{X} = (\mathbf{x}^1, \dots, \mathbf{x}^k)$  is the matrix of all exemplar vectors, then the data domain bounds,  $a = \min(\mathbf{X})$  and  $b = \max(\mathbf{X})$ , are given by,

$$a = \bigwedge_{\xi=1}^k \bigwedge_{i=1}^n x_i^\xi \quad \text{and} \quad b = \bigvee_{\xi=1}^k \bigvee_{i=1}^n x_i^\xi. \quad (13)$$

For certain vector operations we will make use of the constant bound vectors,  $\mathbf{a}$  and  $\mathbf{b}$ , whose components are defined respectively by,  $a_i = a$  and  $b_i = b$  for all  $i \in I_n$ . Under the assumption that the underlying dataset domain specified by  $[a, b]$  can not be extended in a meaningful manner, the modified vector entries  $\tilde{x}_i$  given in (12) are clipped to fall within  $[a, b]$ . Hence, if  $x_i + \epsilon > b$  then  $\tilde{x}_i = b$ ; similarly, if  $x_i - \epsilon < a$  then  $\tilde{x}_i = a$ . Algorithm I lists the steps that generate additive impulse random noise with varying offsets ( $\epsilon$  parameter) and variable amounts of noise ( $\%p$  parameter) where,  $\epsilon \in (a, b)$  and  $p \in (0, 100)$ . Note that, the arrow ‘ $\leftarrow$ ’ designates an *assignment* operation.

#### Algorithm I-Impulsive Noise Generation (ING)

```
[Select an exemplar vector to simulate a noisy version]
input exemplar  $\mathbf{x}^\xi \in X$  where  $\xi \in I_k$ , dimension  $n$ , noise level  $p$ , and domain bounds  $a, b$ .
[Compute fraction  $q$  of unmodified vector entries in  $\mathbf{x}^\xi$ ]
 $q \leftarrow 1 - p/100$ 
for  $i = 1$  to  $n$ 
     $f \leftarrow \text{rnd}(1)$ 
    if  $f \leq q$ 
         $\tilde{\mathbf{x}}_i^\xi \leftarrow \mathbf{x}_i^\xi$ 
    else [here,  $f > q$ ]
        [Alter vector entries by  $\pm\epsilon$  with equal probability]
         $g \leftarrow \text{rnd}(1)$ 
        if  $g > 0.5$ 
             $c \leftarrow \mathbf{x}_i^\xi + \epsilon, \mathbf{x}_i^\xi \leftarrow \text{if}(c > b, b, c)$ 
        else [here,  $g \leq 0.5$ ]
             $c \leftarrow \mathbf{x}_i^\xi - \epsilon, \mathbf{x}_i^\xi \leftarrow \text{if}(c < a, a, c)$ 
        end if
    end if
next  $i$ 
output exemplar noisy version  $\tilde{\mathbf{x}}^\xi$ 
```

In Algorithm I, the scalar function  $\text{rnd}(1)$  returns a random real number with uniform distribution on the unit interval  $(0, 1)$  and the compound function  $\text{if}(\text{cond.}, \text{true exp.}, \text{false exp.})$  evaluates the true (clipped value) or false

expression (non-clipped value) upon verification of a logical condition. Assuming that  $\tilde{\mathbf{x}}$  is a noisy vector obtained with Algorithm I, then the proposed *thresholding functions* (14) or (15) defined component-wise, will produce an eroded  $\tilde{\mathbf{e}}$  or dilated version  $\tilde{\mathbf{o}}$  of  $\tilde{\mathbf{x}}$ .

$$\tilde{e}_i = \text{if}(\tilde{x}_i \geq \tau_e, a, \tilde{x}_i) \quad \forall i \in I_n, \quad (14)$$

$$\tilde{o}_i = \text{if}(\tilde{x}_i \leq \tau_d, b, \tilde{x}_i) \quad \forall i \in I_n, \quad (15)$$

where,  $\tau_e = \varepsilon - 1$ ,  $\tau_d = b - \varepsilon$ , and  $\varepsilon = (a + b)/2$ . In addition, thresholds  $\tau_e$  and  $\tau_d$  should satisfy the complementary inequalities with respect to the data domain  $[a, b]$ , i. e.,  $a < \tau_e \leq \varepsilon - 1$  and  $\varepsilon - 1 \leq \tau_d < b$ . A functional representation of (14) and (15) is  $\tilde{\mathbf{e}} = T_e(\tilde{\mathbf{x}}, \tau_e, a)$  and  $\tilde{\mathbf{o}} = T_d(\tilde{\mathbf{x}}, \tau_d, b)$ , where the threshold and domain bound parameters appear explicitly. By construction, it is not difficult to see that  $\tilde{\mathbf{e}} \leq \tilde{\mathbf{x}}$  and  $\tilde{\mathbf{o}} \geq \tilde{\mathbf{x}}$ .

### 3.2 Eroded and Dilated Input Residuals

Before a meaningful association using a LAAM can be established from a noisy exemplar input, a search must be performed to ensure that a stored memory pattern resembles most the given input due to the fact that for any significant amount of noise no prediction can be made in advance of which exemplar was altered. Thus, if  $\tilde{\mathbf{x}}$  is the given corrupted input, a similarity measure must be used to find  $\xi \in I_k$  such that  $\tilde{\mathbf{x}} \approx \tilde{\mathbf{x}}^\xi$ . Since the model we consider for impulsive noise depends on signed equiprobable values of  $\varepsilon$ , index search uses the statistical linear *correlation coefficient* (Pearson's) calculated between the vectors  $\tilde{\mathbf{x}}$  and  $\tilde{\mathbf{x}}^\xi$ , or  $\text{cc}(\tilde{\mathbf{x}}, \tilde{\mathbf{x}}^\xi)$ . Therefore, an index  $\gamma \in I_k$  is selected such that,

$$\text{cc}(\tilde{\mathbf{x}}, \tilde{\mathbf{x}}^\gamma) = \bigvee_{\xi=1}^k \text{cc}(\tilde{\mathbf{x}}, \tilde{\mathbf{x}}^\xi), \quad (16)$$

Once the index value  $\gamma$  is found, the eroded and dilated *residuals* of the noisy exemplar used as candidate inputs to the corresponding LAAM, are computed as,

$$\tilde{\mathbf{x}}_e = (\tilde{\mathbf{x}}^\gamma - \tilde{\mathbf{e}}) + \mathbf{a} \quad \text{and} \quad \tilde{\mathbf{x}}_d = \mathbf{b} - (\tilde{\mathbf{o}} - \tilde{\mathbf{x}}^\gamma). \quad (17)$$

The vector operations (17) are devised in order to guarantee that  $\tilde{\mathbf{x}}_e \approx \tilde{\mathbf{x}}_e^\gamma$  and  $\tilde{\mathbf{x}}_d \approx \tilde{\mathbf{x}}_d^\gamma$ . Also, using (13) to (16), it follows that,  $\tilde{\mathbf{x}}_e \leq \tilde{\mathbf{x}}_e^\gamma$  and  $\tilde{\mathbf{x}}_d \geq \tilde{\mathbf{x}}_d^\gamma$ . Consequently, in the recall stage,  $\tilde{\mathbf{x}}_e$  can be fed into the min-memory  $W_{XX}$  or dually  $\tilde{\mathbf{x}}_d$  can be used as an adequate input for the max-memory  $M_{XX}$ . The corresponding output vectors are then given by,  $\mathbf{y}_e = W_{XX} \nabla \tilde{\mathbf{x}}_e$  or  $\mathbf{y}_d = M_{XX} \Delta \tilde{\mathbf{x}}_d$  and because of the robustness of each LAAM it is no surprise to expect that  $\mathbf{y}_e \approx \mathbf{x}^\gamma \approx \mathbf{y}_d$  with low error (NMSE) as shown in Section 4. Another useful similarity measure is the cosine of the angle between vectors. However, if both vectors are displaced with respect to their means the *centered cosine angle*, as it is known, turns to be equivalent to the linear correlation coefficient.

## 4 LAAMs Enhanced Retrieval Capability

Integrating the ‘‘almost perfect recall’’ criterion with noise thresholding and the generation of eroded or dilated

residuals, a simple procedure applying equations (11) through (17) together with Algorithm I, can be used to test the recall performance of either the min- $W_{XX}$  or max- $M_{XX}$  memories over a run of trials for different amounts of impulsive random noise with varying offset values. Hence, let  $m > 0$  represent the number of trials realized during the recall stage, let  $h$  be the percentage increment for the amount of noise corrupting a vector, and let  $\bar{\rho}$  stand for the average NMSE error computed over a run of trials. The mathematical pseudo-code shown in Algorithm II lists the steps required to test either LAAM by carrying on the following computational subtasks: 1) input noise generation, 2) correlation coefficient based index search, 3) thresholding low or up a noisy input, 4) erosion or dilation residuation, 5)  $W_{XX}$  max-sum with eroded residual calculation (resp.,  $M_{XX}$  min-sum with dilated residual), and 6) average NMSE determination over a trial run for any amount of noise determined by  $h$ . In the code phrase ‘*expr1 | expr2*’, the symbol ‘|’ means ‘xor’ (exclusive or).

### Algorithm II-LAAMs Enhanced Recall (LAAM-ER)

```
[Test recall improvement over  $m$  trials of noisy inputs]
input matrix  $X$  of exemplars, test vector index  $\gamma \in I_k$ ,
memory  $W_{XX} | M_{XX}$ , dimension  $n$ , noise offset  $\varepsilon$ , domain
bounds  $a, b$ , number of trials  $m$ , noise increment  $h$ .
[Initialize threshold, noise level, results table index]
 $\varepsilon \leftarrow (a + b)/2$ ,  $\tau_e \leftarrow \varepsilon - 1$  |  $\tau_d \leftarrow b - \varepsilon$ ,  $i \leftarrow 1$ ,  $p \leftarrow h$ 
[External cycle increments noise level by  $h$ ]
while  $p \leq 100$ 
  [Initialize error sum and inequality hit count]
   $s \leftarrow 0$ ,  $c \leftarrow 0$ 
  [Internal cycle is a run of  $m$  trials]
  for  $j = 1$  to  $m$ 
     $\tilde{\mathbf{x}} \leftarrow \text{call ING}(\tilde{\mathbf{x}}^\gamma, p, \varepsilon)$ 
     $\tilde{\mathbf{e}} \leftarrow T_e(\tilde{\mathbf{x}}, \tau_e, a)$  |  $\tilde{\mathbf{o}} \leftarrow T_d(\tilde{\mathbf{x}}, \tau_d, b)$ 
     $\gamma = \text{argmax}\{\text{cc}(\tilde{\mathbf{x}}, \tilde{\mathbf{x}}^\xi), \xi \in I_k\}$ 
     $\tilde{\mathbf{x}}_e \leftarrow (\tilde{\mathbf{x}}^\gamma - \tilde{\mathbf{e}}) + \mathbf{a}$  |  $\tilde{\mathbf{x}}_d \leftarrow \mathbf{b} - (\tilde{\mathbf{o}} - \tilde{\mathbf{x}}^\gamma)$ 
     $\mathbf{y}_e \leftarrow W_{XX} \nabla \tilde{\mathbf{x}}_e$  |  $\mathbf{y}_d \leftarrow M_{XX} \Delta \tilde{\mathbf{x}}_d$ 
     $s \leftarrow s + \rho(\tilde{\mathbf{x}}^\gamma, \mathbf{y}_e)$  |  $\rho(\tilde{\mathbf{x}}^\gamma, \mathbf{y}_d)$ 
     $c \leftarrow c + (\tilde{\mathbf{e}} \leq \tilde{\mathbf{x}}^\gamma)$  |  $(\tilde{\mathbf{o}} \geq \tilde{\mathbf{x}}^\gamma)$ 
  next  $j$ 
  [Save noise level, average NMSE, ( $\leq$  |  $\geq$ ) hit count]
   $T_{i1} \leftarrow p$ ,  $T_{i2} \leftarrow s/m$ ,  $T_{i3} \leftarrow c$ 
  [Increment noise level, advance table row]
   $p \leftarrow p + h$ ,  $i \leftarrow i + 1$ 
end while
output results table  $T$ 
```

A dataset of 30 images will serve our aim to illustrate the retrieval performance of LAAMs. More specifically, each element in our dataset is a grayscale digital image of size  $64 \times 64$  pixels, whose domain of values are integer numbers belonging to the interval  $[0, 255]$ . Following a row-order scan of the underlying image data matrix, an equivalent column vector is obtained for each image. Thus, the exemplar image dataset can be represented by  $X = (\mathbf{x}^1 \dots \mathbf{x}^k)$  where  $k = 30$  and vector dimensionality  $n$  equals 4096.

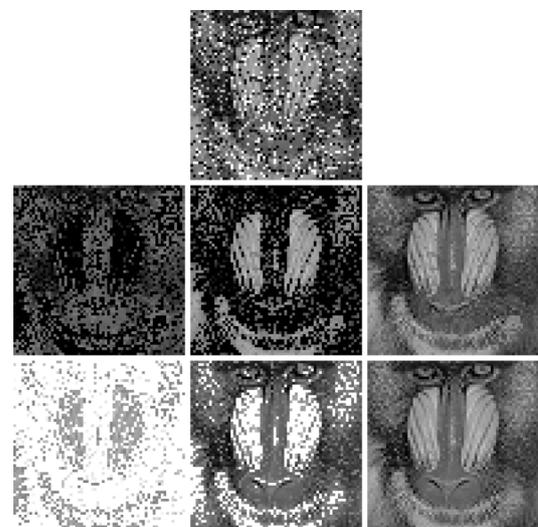


**Figure 1.** Set  $X$  of 30 exemplar grayscale images of size  $64 \times 64$  pixels. 1st row (left to right),  $\mathbf{x}^1, \dots, \mathbf{x}^6$ ; 2nd row,  $\mathbf{x}^7, \dots, \mathbf{x}^{12}$ ; 3rd row,  $\mathbf{x}^{13}, \dots, \mathbf{x}^{18}$ ; 4th row,  $\mathbf{x}^{19}, \dots, \mathbf{x}^{24}$ ; 5th row,  $\mathbf{x}^{25}, \dots, \mathbf{x}^{30}$ . Each column vector  $\mathbf{x}^\xi$  is of dimension 4096 with integer values in  $[0, 255]$ .

Figure 1 displays a panel of all images belonging to  $X$  and selected ones will be used as examples in additional figures, numerical tables or graphs. In this example, data domain bounds are  $a = 0$  and  $b = 255$ . On the other hand, the min- $W_{XX}$  and max- $M_{XX}$  lattice associative memories that store all exemplar images are matrices of size  $4096 \times 4096$ . The computer tests consider three runs where the impulse noise offset value  $\varepsilon$  was set correspondingly, to 128, 96, and 64. Each run goes through  $m = 50$  trials for variable amounts of noise increasing in  $h = 5\%$  steps. Hence, after applying Algorithm II to any exemplar vector in  $X$ , the output table  $T$  has 20 rows and 3 columns. A particular row will give in column-order, the % noise level, the average NMSE value, and the inequality check count overall trials.

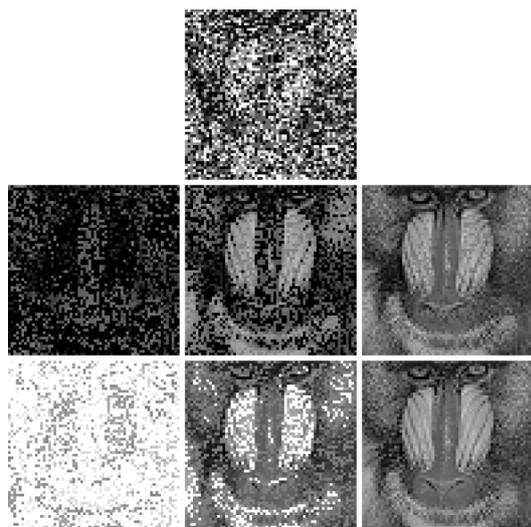
The information depicted in Figs. 2, 3, and 4 has the following structure. The input image with a certain amount of impulsive noise appears on top (first row). The eroded thresholded input, the eroded generated residual, and the recalled image using the min- $W_{XX}$  lattice memory are displayed, respectively, to the left, at the center, and to the right of the second row. Similarly, the dilated thresholded input, the generated dilated residual, and the recalled image using the max- $M_{XX}$  lattice memory are shown to the left, center, and right of the third row. Note that, in this example,  $\tilde{\mathbf{e}} = T_e(\tilde{\mathbf{x}}^2, 127, 0)$  and  $\tilde{\mathbf{o}} = T_d(\tilde{\mathbf{x}}^2, 127, 255)$ . Numerical results relative to Figs. 2 3, and 4 are listed in Table 1.

As expected from the comments given in the last paragraph of Section 3, after (17), the hit counts of the vector inequalities  $\tilde{\mathbf{e}} \leq \tilde{\mathbf{x}}^\gamma$  and  $\tilde{\mathbf{o}} \geq \tilde{\mathbf{x}}^\gamma$ , as well as the vector in-

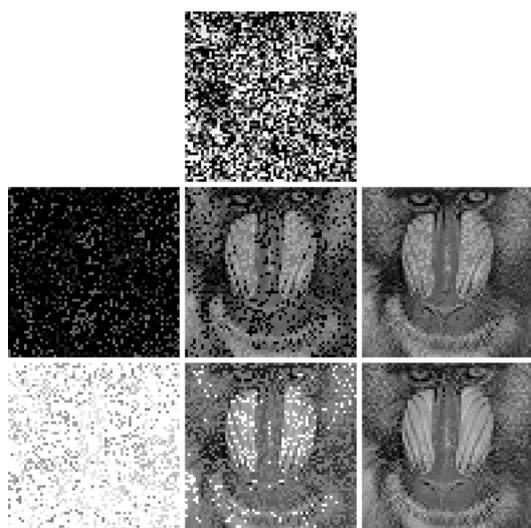


**Figure 2.** Top, noisy input  $\tilde{\mathbf{x}}^2$  ( $\varepsilon = 128, p = 25\%$ ); 2nd row (left to right), eroded thresholding  $\tilde{\mathbf{e}}$ , eroded residual  $\tilde{\mathbf{x}}_e$ , and min-memory recalled output  $\mathbf{y}_e$ ; 3rd row, dilated thresholding  $\tilde{\mathbf{o}}$ , dilated residual  $\tilde{\mathbf{x}}_d$ , and max-memory recalled output  $\mathbf{y}_d$ .

equalities  $\tilde{\mathbf{x}}_e \leq \tilde{\mathbf{x}}^\gamma$  and  $\tilde{\mathbf{x}}_d \geq \tilde{\mathbf{x}}^\gamma$ , all sum up to  $m$  (number of tested trials), i. e., inequalities are satisfied no matter the amount of noise the input undergoes. Observe that the  $\varepsilon$  values equal to 64, 96, and 128, may be interpreted subjectively as “mild”, “moderate”, and “strong” variations of impulsiveness. In Table 1, the numerical results in lines 5 and 6 for data retrieval altered by moderate impulsive noise as well as the values in lines 8 and 9 for recalling



**Figure 3.** Top, noisy input  $\hat{x}^2$  ( $\varepsilon = 128$ ,  $p = 50\%$ ); 2nd row (left to right), eroded thresholding  $\tilde{e}$ , eroded residual  $\hat{x}_e$ , and min-memory recalled output  $y_e$ ; 3rd row, dilated thresholding  $\tilde{o}$ , dilated residual  $\hat{x}_d$ , and max-memory recalled output  $y_d$ .



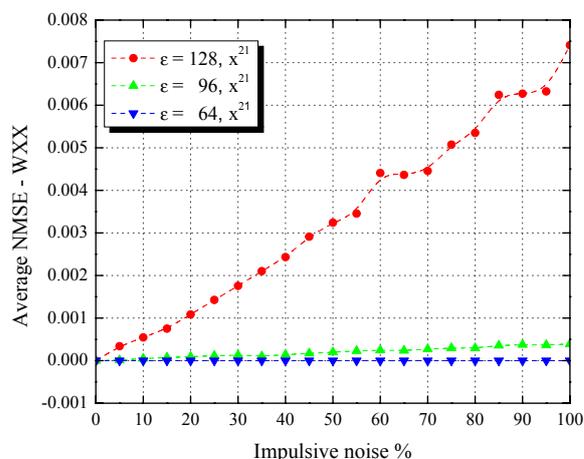
**Figure 4.** Top, noisy input  $\hat{x}^2$  ( $\varepsilon = 128$ ,  $p = 75\%$ ); 2nd row (left to right), eroded thresholding  $\tilde{e}$ , eroded residual  $\hat{x}_e$ , and min-memory recalled output  $y_e$ ; 3rd row, dilated thresholding  $\tilde{o}$ , dilated residual  $\hat{x}_d$ , and max-memory recalled output  $y_d$ .

data modified by mild impulse noise show the level of robustness achieved by the proposed mechanism in using a single LAAM.

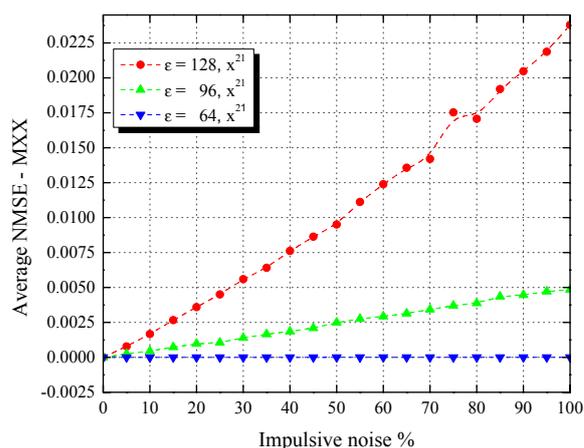
The graphs in Figs. 5 and 6 relative to vector  $x^{21}$  is an example showing an almost linear relationship between the amount of impulsive noise and the average relative error (NMSE) in recalling the corresponding vectors, respectively, with either the min-memory  $W_{XX}$  or the max-memory  $M_{XX}$ . More specifically, the recall capability displayed by the performance lines corresponding to  $\varepsilon = 96$  and  $\varepsilon = 64$  for any  $p \in [0, 100]$  verify the partial results given in Table 1, showing again, the low order error attained during the retrieval stage by either LAAM enhanced

**Table 1.** Average Normalized Mean Square Error

	$\bar{\rho}$	$\varepsilon$	$p = 25\%$	$p = 50\%$	$p = 75\%$
1	$\bar{\rho}(x^2, \hat{x}^2)$	128	2.65e-1	5.49e-1	7.97e-1
2	$\bar{\rho}(x^2, y_e)$	128	8.47e-2	8.42e-2	9.42e-2
3	$\bar{\rho}(x^2, y_d)$	128	4.09e-2	8.23e-2	1.13e-1
4	$\bar{\rho}(x^2, \hat{x}^2)$	96	1.54e-1	3.09e-1	4.56e-1
5	$\bar{\rho}(x^2, y_e)$	96	1.81e-4	1.86e-4	4.62e-4
6	$\bar{\rho}(x^2, y_d)$	96	1.09e-3	2.69e-3	3.65e-3
7	$\bar{\rho}(x^2, \hat{x}^2)$	64	7.12e-2	1.38e-1	2.12e-1
8	$\bar{\rho}(x^2, y_e)$	64	6.45e-6	2.29e-6	1.99e-5
9	$\bar{\rho}(x^2, y_d)$	64	0.00e-6	2.18e-6	2.18e-6

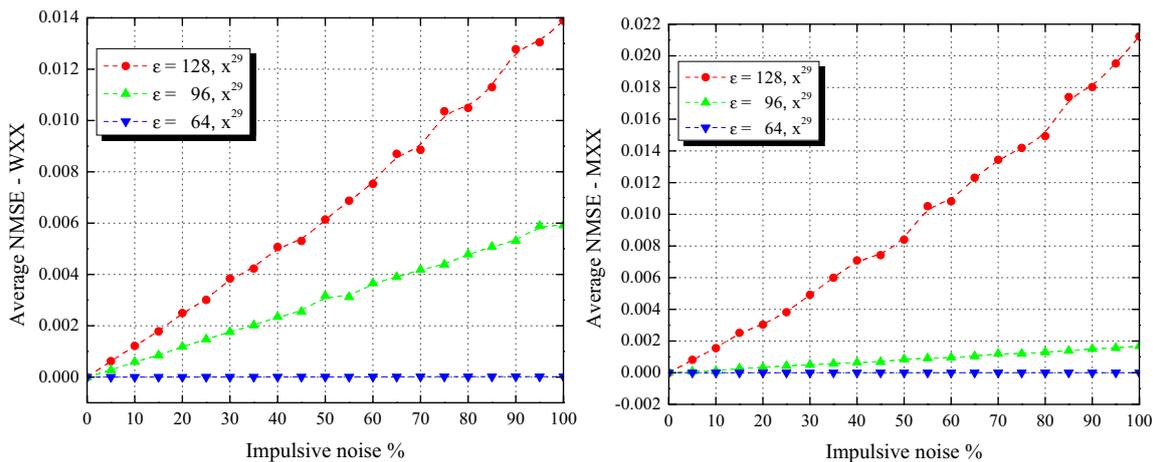


**Figure 5.** Recall performance  $\bar{\rho}$ -curves for noisy versions of exemplar  $x^{21}$  using the min- $W_{XX}$  LAAM.



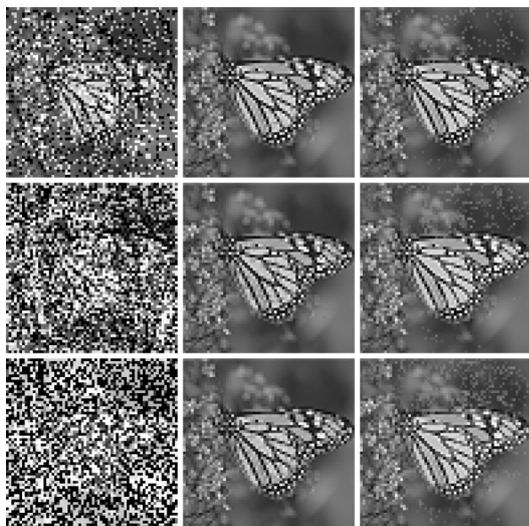
**Figure 6.** Recall performance  $\bar{\rho}$ -curves for noisy versions of exemplar  $x^{21}$  using the max- $M_{XX}$  LAAM.

with the proposed thresholding and residuation technique. Note that  $\bar{\rho} < 10^{-3}$  for  $W_{XX}$  whereas  $\bar{\rho} < 5 \times 10^{-3}$  for  $M_{XX}$ . Also, if  $\varepsilon = 128$ ,  $\bar{\rho} < 7.5 \times 10^{-3}$  for  $W_{XX}$  whereas  $\bar{\rho} < 2.5 \times 10^{-2}$  for  $M_{XX}$ . Thus, in this case, data retrieval from noisy inputs using the min-memory  $W_{XX}$  is slightly better than using the max-memory  $M_{XX}$ .

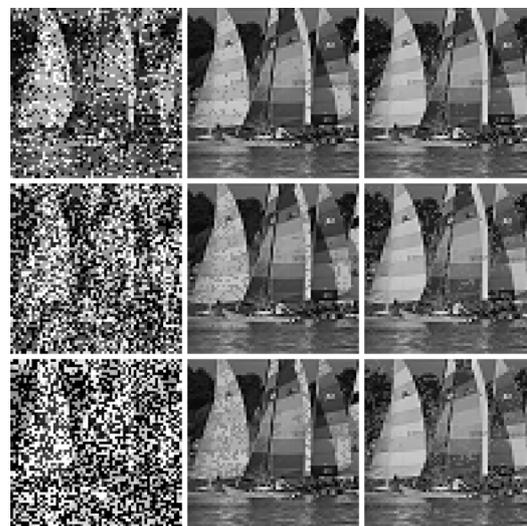


**Figure 7.** Recall performance  $\bar{\rho}$ -curves for noisy versions of exemplar  $\mathbf{x}^{29}$  using the min- $W_{XX}$  and max- $M_{XX}$  LAAMs.

A similar recall performance for vector  $\mathbf{x}^{29}$  is depicted in Fig. 7 except that, for  $\varepsilon = 96$ , better retrieval is obtained with  $M_{XX}$ . Specific instances of strong impulsive noisy versions of exemplar vectors  $\mathbf{x}^{21}$  and  $\mathbf{x}^{29}$  together with the outputs recalled with both LAAMs are shown in Figs. 8 and 9. Also, numerical results relative to these figures are listed in Table 2.



**Figure 8.** 1st column, top to bottom:  $\tilde{\mathbf{x}}^{21}$  where  $\varepsilon = 128$  for  $p = 30\%$ ,  $60\%$ , and  $90\%$ ; 2nd column, outputs  $\mathbf{y}_e$  recalled with  $W_{XX}$ ; 3rd column, outputs  $\mathbf{y}_d$  recalled with  $M_{XX}$ .



**Figure 9.** 1st column, top to bottom:  $\tilde{\mathbf{x}}^{29}$  where  $\varepsilon = 128$  for  $p = 30\%$ ,  $60\%$ , and  $90\%$ ; 2nd column, outputs  $\mathbf{y}_e$  recalled with  $W_{XX}$ ; 3rd column, outputs  $\mathbf{y}_d$  recalled with  $M_{XX}$ .

We remark that for any exemplar in  $X$  tested with Algorithm II, the retrieval performance of either the min-LAAM or max-LAAM in the presence of noisy inputs is quite similar to those example images used in our exposition.

## 5 Conclusions

Noise masking [12–14] was originally proposed to endow a single LAAM with recall capability for multivariate data degraded by random noise. Furthermore, the noise masking strategy has been recently used to good advantage to increase retrieval performance in max-plus projection autoassociative morphological memories and their fuzzy analogs [15–17]. In this paper, we have introduced a different computational method based on eroded-dilated noise thresholds and eroded-dilated generated residuals in pursuance of increasing a single LAAM recall capability

**Table 2.** Normalized Mean Square Error for  $\varepsilon = 128$

	$\rho$	$p = 30\%$	$p = 60\%$	$p = 90\%$
1	$\rho(\mathbf{x}^{21}, \tilde{\mathbf{x}}^{21})$	2.80e-1	5.60e-1	8.51e-1
2	$\rho(\mathbf{x}^{21}, \mathbf{y}_e)$	1.66e-3	4.83e-3	6.03e-3
3	$\rho(\mathbf{x}^{21}, \mathbf{y}_d)$	6.35e-3	1.11e-2	1.95e-2
4	$\rho(\mathbf{x}^{29}, \tilde{\mathbf{x}}^{29})$	2.60e-1	5.26e-1	7.67e-1
5	$\rho(\mathbf{x}^{29}, \mathbf{y}_e)$	3.50e-3	7.99e-3	1.31e-2
6	$\rho(\mathbf{x}^{29}, \mathbf{y}_d)$	5.19e-3	1.24e-2	1.68e-2

in the presence of random impulsive noise. This technique performs better if compared to similar approaches already discussed elsewhere [9, 10, 19, 20], which in turn outperform other well-known past associative memory models such as the Hopfield continuous recurrent neural network or the Anderson-Kohonen linear correlation model, to name a few. Algorithm II involves simple arithmetical operations and works fast on vectors in high dimensional spaces. As its name implies, impulsive noise alters the most a vector's data content and serves as an extreme model for data variation. Nonetheless, in many application scenarios, less aggressive noise probability densities such as uniform or Gaussian are more adequate for modeling data changes. Although not developed here, we point out that the proposed improved technique for LAAMs pattern recall works the same if other noise models based on symmetric probability densities are used. Continuation of this research will pay attention in doing tests with different multivariate data sets, use of asymmetric probability noise models, and the application of other vector metrics to further verify the robustness of our proposed model.

## Acknowledgments

Gonzalo Urcid thanks the National Council of Science and Technology (CONACYT) in Mexico City for partial support through SNI grant # 22036, Rocío Morales-Salgado thanks the Information Technologies & Data Science Department at Universidad Popular Autónoma del Estado de Puebla (UPAEP) for partial financial support, and José Angel Nieves-Vázquez thanks the Engineering Division of the Instituto Tecnológico Superior de San Andrés Tuxtla (ITSSAT) for partial support and travel permits during the realization of this research project.

## References

- [1] S. Roman, *Lattices and Ordered Sets* (Springer Verlag, New York, NY, 2008).
- [2] G. Grätzer, *Lattice Theory: Foundation* (Birkhäuser-Springer Basel, Berlin, 2011).
- [3] G. X. Ritter, P. Sussner, and J. L. Diaz de Leon, "Morphological associative memories," *IEEE Transactions on Neural Networks*, **9**(2), 281–293 (1998).
- [4] R. Cuninghame-Green, "Minimax Algebra and Applications," in *Advances in Imaging and Electron Physics* **90**, P. Hawkes (Academic Press, New York, NY, 1995), 1–121.
- [5] G. X. Ritter, "Lattice algebra" and "Minimax algebra," in *Image Algebra*, unpublished manuscript available via anonymous ftp from <https://www.cise.ufl.edu/~jnw/CVAIIA/> (University of Florida, CCVV/CISE Department, Gainesville, FL, 1999) 121–135.
- [6] G. X. Ritter and P. Gader, "Fixed Points of Lattice Transforms and Lattice Associative Memories," in *Advances in Imaging and Electron Physics* **144**, P. Hawkes (Elsevier, 2006), 165–242.
- [7] P. Sussner, "Observations on morphological associative memories and the kernel method," *Neurocomputing*, **31**(1)-(4), 167–183 (2000).
- [8] P. Sussner, "Associative morphological memories based on variations of the kernel and dual kernel methods," *Neural Networks*, **16**, 625–632 (2003). Jul. 2003.
- [9] G. X. Ritter, G. Urcid, and L. Iancu, "Reconstruction of noisy patterns using morphological associative memories," *Journal of Mathematical Imaging and Vision*, **19**(5), 95–111 (2003).
- [10] G. Urcid, "Kernel computation in morphological associative memories for grayscale image recollection," *IASTED Proceedings 5th International Conference on Signal and Image Processing*, 450–455 (2003).
- [11] P. Sussner, "Recall of patterns using binary and grayscale autoassociative morphological memories," *Proceedings of SPIE: Mathematical Methods in Pattern and Image Analysis*, **5916**, 59160M:1–10 (2005).
- [12] G. Urcid and G. X. Ritter, "Noise masking for pattern recall using a single lattice matrix auto-associative memory," *Proceedings 2006 IEEE International Conference on Fuzzy Systems*, 187–194 (2006).
- [13] G. Urcid and G. X. Ritter, Noise Masking for Pattern Recall using a Single Lattice Matrix Associative Memory," in *Studies in Computational Intelligence*, **67** (Springer, Berlin, 2007), 81–100.
- [14] G. Urcid, J.A. Nieves-V., A. Garcia-A., and J.C. Valdiviezo-N., "Robust image retrieval from noisy inputs using lattice associative memories," *Proceedings of SPIE-IS&T Electronic Imaging: Image Processing, Algorithms and Systems VII*, **7245**, 724517:1–12 (2009).
- [15] M. E. Valle, "An introduction to the max-plus projection autoassociative morphological memory and some of its variations," *Proceedings 2014 IEEE International Conference on Fuzzy Systems*, 53–60 (2014).
- [16] A. S. Santos and M. E. Valle, "A fast and robust max–C projection fuzzy autoassociative memory with application for face recognition," *IEEE Proceedings 2017 Brazilian Conference on Intelligent Systems*, 306–311 (2017).
- [17] A. S. Santos and M. E. Valle, "Max–C and Min–D projection autoassociative fuzzy morphological memories: theory and applications for face recognition," arXiv:1902.04144v (Preprint submitted to Elsevier), 1–33 (2019).
- [18] Y. Wang, "Model Selection," Ch. 16 in *Handbook of Computational Statistics: Concepts and Methods, 2nd Ed.*, J. E. Gentle, W. K. Härdle, Y. Mori (Springer Berlin, Heidelberg 2012), 469–497.
- [19] M. Wang and S. Chen, "Enhanced fuzzy morphological auto-associative memory based on empirical kernel map," *IEEE Transactions on Neural Networks*, **16**(3), 557–564 (2005).
- [20] P. Sussner and M.E. Valle, "Gray-scale morphological associative memories," *IEEE Transactions on Neural Networks*, **17**(3), 559–570 (2006).