2014 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)
July 6-11, 2014, Beijing, China

An Introduction to the Max-plus Projection Autoassociative Morphological Memory and Some of Its Variations

Marcos Eduardo Valle

***Abstract*—In this paper, we present a novel lattice-based memory model called max-plus projection autoassociative morphological memory (max-plus PAMM). The max-plus PAMM yields the largest max-plus combination of the stored patterns which is less than or equal to the input. Such as the original autoassociative morphological memories (AMMs), it is idempotent and it gives perfect recall of undistorted patterns. Furthermore, the max-plus PAMM is very robust to dilative noise and it has less spurious memories than its corresponding AMM. This paper also presents two variations of the max-plus PAMM. The first yields the max-plus combination that is the Chebyshev-best approximation of the input while the second uses a noise masking strategy.**

I. INTRODUCTION

Lattice computing refers to a broad collection of tools and methodologies that apply or use lattice theory [1], [2], [3]. It includes minimax algebra [4], mathematical morphology [5], fuzzy set theory [6], and a broad class of artificial neural networks [7], [8]. *Autoassociative morphological memories* (AMMs) also belong to the lattice computing framework.

The AMMs, also known as lattice autoassociative memories, have been developed initially by Ritter and Sussner in the middle 1990s [9], [10], [11]. As an associative memory, they are designed for the storage and recall of patterns by association or by their contents – in a manner similar to the human brain [12]. We would like to recall that the AMMs are called *morphological* because they perform the basic operations from mathematical morphology [10], [13].

The AMMs have been formerly conceived for the storage of real-valued pattern [10]. However, since they are built on lattice theory, they can possibly be extended to process lattice-ordered data such as sets, symbols, intervals, fuzzy sets, etc. For instance, the classes of fuzzy and interval-valued fuzzy associative memories are used to process fuzzy sets as well as interval-valued fuzzy sets [14], [15], [16]. Applications of AMMs and some of their variations include restoration of corrupted images [17], [18], [19], [20], identification of structures in resting state fMRI data [21], autonomous determination of endmembers in hyperspectral images [22], [23], vision-based self-localization in mobile robots [24], [25], times-series prediction [26], and classification [17].

In contrast to many traditional models such as the Hopfield network [12], the original AMMs are able to store an unlimited number of patterns and give perfect recall of any undistorted item [10], [17]. Also, they converge in one single step when employed with feedback and are robust in the presence of either dilative or erosive noise. On the downside, the original AMMs have a large amount of spurious memories. In view of this fact, we introduce the max-plus projection AMM (max-plus PAMM). The novel memory has a reduced number of spurious memories and, consequently, is more robust to dilative noise than its corresponding original AMM. In this paper, we also present two variations of the max-plus PAMM.

Marcos Eduardo Valle is with the Department of Applied Mathematics, University of Campinas, Campinas, Brazil (email: valle@ime.unicamp.br).

This work was supported in part by by CNPq under grant no. 304240/2011-7, FAPESP under grant no. 2013/12310-4, and FAEPEX/Unicamp under grant no. 519.292.

The paper is organized as follows. Next section presents the mathematical background. A brief review on the original AMMs are provided subsequently. Section III also presents some variations of the AMMs. The novel max-plus PAMM as well as their variations are introduced in Section IV. Illustrative examples are given in both Sections III and IV. The paper finishes with some concluding remarks and the appendix containing the proofs of lemmas and theorems.

II. SOME MATHEMATICAL BACKGROUND

The memory models considered in this paper are described by lattice-based operations from minimax algebra, a mathematical structure motivated by problems from scheduling theory, graph theory, and dynamic programming [4]. Roughly speaking, the minimax algebra is developed in a mathematical structure obtained by enriching a complete lattice with two group operations [9], [10], [17], [13]. For the purposes of this paper, however, we consider the totally ordered field of real numbers (which is not a complete lattice) as the mathematical background. The supremum and the infimum of a bounded set $X \subseteq \mathbb{R}$ are denoted respectively by the symbols $\bigvee X$ and $\bigwedge X$. In case $X = \{x_1, \ldots, x_n\} \subseteq \mathbb{R}$ is a finite set, the operations of computing the maximum and the minimum are written as $\bigvee_{j=1}^{n} x_j$ and $\bigwedge_{j=1}^{n} x_j$, respectively.

Given two matrices $A \in \mathbb{R}^{n\times k}$ and $B \in \mathbb{R}^{k\times m}$, the *max-product* and the *min-product* of A by B, denoted respectively by $C = A \boxdot B \in \mathbb{R}^{n\times m}$ and $D = A \boxtimes B \in \mathbb{R}^{n\times m}$, are given by the following equations for all $i = 1, \ldots, n$ and $j = 1, \ldots, m$:

$$c_{ij} = \bigvee_{\xi=1}^{k} (a_{i\xi} + b_{\xi j}) \quad \text{and} \quad d_{ij} = \bigwedge_{\xi=1}^{k} (a_{i\xi} + b_{\xi j}). \tag{1}$$

Note that the max-product satisfy

$$A \boxdot (B + \alpha) = (A \boxdot B) + \alpha, \quad \forall \alpha \in \mathbb{R}. \tag{2}$$

Here, $B+\alpha$ is the matrix obtained by adding α to each entry of B. Similarly, we have

$$A \boxtimes (B + \alpha) = (A \boxtimes B) + \alpha, \quad \forall \alpha \in \mathbb{R}. \tag{3}$$

In other words, both lattice-based products are invariant under vertical translations [27].

The conjugate of $A \in \mathbb{R}^{n\times k}$ is the matrix $A^* \in \mathbb{R}^{k\times n}$ whose entries satisfy $a^*_{ij} = -a_{ji}$ for all indexes i, j. Note that $(A^*)^* = A$ for any matrix A. The conjugate can be used to establish the following identities concerning the min-product and the max-product:

$$(A \boxwedge B)^* = B^* \boxvee A^* \quad \text{and} \quad (A \boxvee B)^* = B^* \boxwedge A^*. \tag{4}$$

In addition, the lattice-based matrix operations are related by means of the following adjuction relationship for matrices $A \in \mathbb{R}^{n\times k}, B \in \mathbb{R}^{k\times m}$, and $C \in \mathbb{R}^{n\times m}$:

$$A \boxvee B \leq C \Leftrightarrow B \leq A^* \boxwedge C \Leftrightarrow A \leq C \boxwedge B^*. \tag{5}$$

In analogy to the notion of linear combination, a *max-plus combination* of vectors from a set $\mathcal{X} = \{\mathbf{x}^1, \ldots, \mathbf{x}^k\} \subseteq \mathbb{R}^n$ is any vector $\mathbf{y} \in \mathbb{R}^n$ of the form

$$\mathbf{y} = \bigvee_{\xi=1}^{k} (\alpha_\xi + \mathbf{x}^\xi), \quad \alpha_\xi \in \mathbb{R}. \tag{6}$$

In words, $\mathbf{y}$ is the maximum of vertical translations of $\mathbf{x}^1, \ldots, \mathbf{x}^k$. The set of all max-plus combinations of vectors from $\mathcal{X}$ is denoted by $\mathfrak{V}(\mathcal{X})$, i.e.,

$$\mathfrak{V}(\mathcal{X}) = \left\{ \mathbf{y} \in \mathbb{R}^n : \mathbf{y} = \bigvee_{\xi=1}^{k} (\alpha_\xi + \mathbf{x}^\xi), \alpha^\xi_j \in \mathbb{R} \right\}. \tag{7}$$

Note that $\mathbf{y} \in \mathfrak{V}(\mathcal{X})$ if and only if $\mathbf{y} = X \boxvee \boldsymbol{\alpha}$ for some $\boldsymbol{\alpha} = [\alpha_1, \ldots, \alpha_k]^T \in \mathbb{R}^n$, where $X = [\mathbf{x}^1, \ldots, \mathbf{x}^k] \in \mathbb{R}^{n\times k}$ is the matrix whose columns corresponds to the vectors of $\mathcal{X}$. Similarly, any vector $\mathbf{z} \in \mathbb{R}^n$ of the form

$$\mathbf{z} = \bigwedge_{j=1}^{n} \bigvee_{\xi=1}^{k} (a^\xi_j + \mathbf{x}^\xi), \quad a^\xi_j \in \mathbb{R}. \tag{8}$$

is called a *minimax combination* of vectors from $\mathcal{X}$. The set of all minimax combinations is denoted by $\mathfrak{S}(\mathcal{X})$, i.e.,

$$\mathfrak{S}(\mathcal{X}) = \left\{ \mathbf{z} \in \mathbb{R}^n : \mathbf{z} = \bigwedge_{j=1}^{n} \bigvee_{\xi=1}^{k} (a^\xi_j + \mathbf{x}^\xi), a^\xi_j \in \mathbb{R} \right\}. \tag{9}$$

It is not hard to show that $\mathcal{X} \subseteq \mathfrak{V}(\mathcal{X}) \subseteq \mathfrak{S}(\mathcal{X})$.

We would like to recall that reversing the arguments of a partial ordering yields the so called dual partial ordering [28], [5]. For instance, the relation "greater than or equal" is the dual partial ordering of "less than or equal". The dual ordering establishes the duality principle: to every statement or notion there corresponds a dual one. For instance, the max-product is the dual of the min-product. Furthermore, the duality principle can used to define concepts similar to the ones expressed by (6) and (8). For example, a min-plus combination is any vector given by the min-product $X \boxwedge \boldsymbol{\alpha}$, where $\boldsymbol{\alpha} \in \mathbb{R}^n$.

III. A Brief Review on Autoassociative Morphological Memories

Autoassociative memories (AMs) are mathematical constructs inspired by the human brain ability to store and recall information. They are systems designed for the storage of a set $\mathcal{X} = \{\mathbf{x}^1, \ldots, \mathbf{x}^k\}$, called the *fundamental memory set*. For the purposes of this paper, let $\mathbf{x}^\xi$ be a column-vector in $\mathbb{R}^n$ for all $\xi \in \mathcal{K} = \{1, \ldots, k\}$. Hence, given a fundamental memory set $\mathcal{X} \subseteq \mathbb{R}^n$, an AM corresponds to a mapping $\mathcal{M} : \mathbb{R}^n \to \mathbb{R}^n$ such that $\mathcal{M}(\mathbf{x}^\xi) = \mathbf{x}^\xi$ for all $\xi \in \mathcal{K}$. Furthermore, the mapping $\mathcal{M}$ must exhibit some noise tolerance in the sense that $\mathcal{M}(\tilde{\mathbf{x}}^\xi) = \mathbf{x}^\xi$ for a noise version $\tilde{\mathbf{x}}^\xi$ of the fundamental memory $\mathbf{x}^\xi$. We say that a vector $\mathbf{x} \in \mathcal{X}$ is a *fixed point* of the AM $\mathcal{M}$ if $\mathcal{M}(\mathbf{x}) = \mathbf{x}$. A fixed point that does not belong to the fundamental memory set is called a *spurious memory*.

A. Original Autoassociative Morphological Memories

The original AMMs [9], [29], also referred to as *lattice autoassociative memories* [11], [30], are mappings defined in terms of the lattice-based matrix products. Specifically, for any input vector $\mathbf{x} \in \mathbb{R}^n$, the AMM $\mathcal{M}_{XX} : \mathbb{R}^n \to \mathbb{R}^n$ is given by

$$\mathcal{M}_{XX}(\mathbf{x}) = M_{XX} \boxwedge \mathbf{x} \tag{10}$$

for a certain matrix $M_{XX} \in \mathbb{R}^{n\times n}$, called the *synaptic weight matrix*. Similarly, we define the AMM $\mathcal{W}_{XX} : \mathbb{R}^n \to \mathbb{R}^n$ by means of the equation $\mathcal{W}_{XX}(\mathbf{x}) = W_{XX} \boxvee \mathbf{x}$ for all $\mathbf{x} \in \mathbb{R}^n$. It follows from (4) that $\mathcal{M}_{XX}$ and $\mathcal{W}_{XX}$ are dual models, i.e., one AMM can be derived from the other by replacing the max-product by the min-product, and vice-versa. Thus, we shall focus only on the AMM $\mathcal{M}_{XX}$.

The synaptic weight matrix M_{XX} is the solution of the following problem: Given a fundamental memory set $\mathcal{X} = \{\mathbf{x}^1, \ldots, \mathbf{x}^p\}$, the matrix M_{XX} is given by

$$M_{XX} = \bigvee \{A \in \mathbb{R}^{n\times n} : A \boxwedge \mathbf{x}^\xi \leq \mathbf{x}^\xi, \forall \xi \in \mathcal{K}\}, \tag{11}$$

In words, M_{XX} is the greatest matrix A such that the inequality $A \boxwedge \mathbf{x}^\xi \leq \mathbf{x}^\xi$ holds for all $\xi \in \mathcal{K}$. In practice, the solution of (11) can be easily computed using the product

$$M_{XX} = X \boxvee X^*, \tag{12}$$

where $X = [\mathbf{x}^1, \ldots, \mathbf{x}^p] \in \mathbb{R}^{n\times p}$ is the matrix whose columns correspond to the fundamental memories.

The vector recalled by the AMM $\mathcal{M}_{XX}$ upon presentation of an input $\mathbf{x} \in \mathbb{R}^n$ can also be expressed in terms of a minimax combination of the fundamental memories [17], [30]. Precisely, for any input pattern $\mathbf{x} = [x_1, \ldots, x_n]^T \in \mathbb{R}^n$, the output $\mathbf{z} = \mathcal{M}_{XX}(\mathbf{x})$ satisfies

$$\mathbf{z} = \bigwedge_{j=1}^{n} \bigvee_{\xi=1}^{k} \left((x_j - x^\xi_j) + \mathbf{x}^\xi \right), \tag{13}$$

where x^ξ_j denotes the jth entry of the fundamental memory $\mathbf{x}^\xi$. In addition, we have that $\mathbf{z}$ given by (13) is the largest minimax combination of $\mathcal{X}$ which is less than or equal to the input $\mathbf{x}$. Alternatively, $\mathbf{z}$ is the infimum of $\mathbf{x}$ in the set of fixed points of $\mathcal{M}_{XX}$ [17]. Mathematically, the mapping $\mathcal{M}_{XX} : \mathbb{R}^n \to \mathbb{R}^n$ also satisfies the equation

$$\mathcal{M}_{XX}(\mathbf{x}) = \bigvee \{\mathbf{z} \in \mathfrak{S}(\mathcal{X}) : \mathbf{z} \leq \mathbf{x}\}, \tag{14}$$

where $\mathfrak{S}(\mathcal{X})$ is the set given by (8).

From (14), we conclude that the AMM $\mathcal{M}_{XX}$ is idempotent. In other words, the output remain stable under repeated applications of $\mathcal{M}_{XX}$. Moreover, any finite minimax combination of the fundamental memories is a fixed point of $\mathcal{M}_{XX}$. Hence, $\mathcal{M}_{XX}$ projects the input $\mathbf{x}$ onto the set of all minimax combinations of the fundamental memories. Since any $\mathbf{x}^\xi \in \mathcal{X}$ also belongs to $\mathfrak{S}(\mathcal{X})$, we infer that $\mathcal{M}_{XX}$ exhibits perfect recall of any undistorted fundamental memory. Nevertheless, the AMM has many spurious memories. In fact, any finite vector in the set difference $\mathfrak{S}(\mathcal{X}) \setminus \mathcal{X}$ is a spurious memory of $\mathcal{M}_{XX}$.

The characterization of the output of $\mathcal{M}_{XX}$ given by (14) also reveals that $\mathcal{M}_{XX}(\mathbf{x}) \leq \mathbf{x}$ for all $\mathbf{x} \in \mathbb{R}^n$. Hence, this AMM is able to recall an original pattern $\mathbf{x}^\xi$ only if the input $\mathbf{x}$ is greater than $\mathbf{x}^\xi$. In other words, $\mathcal{M}_{XX}$ is suited for the reconstruction of patterns corrupted by dilative noise, but it is incapable of handling erosive or mixed (dilative and erosive) noise. Recall that a distorted version $\mathbf{x}$ of the original pattern $\mathbf{x}^\xi$ has undergone a *erosive change* if $\mathbf{x} \leq \mathbf{x}^\xi$ and a *dilative change* if $\mathbf{x} \geq \mathbf{x}^\xi$ [10].

Example 1. Consider the fundamental memory set

$$\mathcal{X} = \left\{ \mathbf{x}^1 = \begin{bmatrix} 2 \\ 0 \\ 7 \\ 4 \\ 3 \end{bmatrix}, \mathbf{x}^2 = \begin{bmatrix} 3 \\ 7 \\ 2 \\ 3 \\ 4 \end{bmatrix}, \mathbf{x}^3 = \begin{bmatrix} 6 \\ 1 \\ 4 \\ 8 \\ 4 \end{bmatrix} \right\}. \tag{15}$$

Also, consider the vector

$$\mathbf{x} = \begin{bmatrix} 10 & 0 & 7 & 4 & 2 \end{bmatrix}^T, \tag{16}$$

which have been obtained by corrupting $\mathbf{x}^1$ by mixed noise. In fact, we have $\mathbf{x} = \mathbf{x}^1 + [8, 0, 0, 0, -1]^T$.

The synaptic weight matrix M_{XX} given by (12) satisfies

$$M_{XX} = \begin{bmatrix} 0 & 5 & 2 & 0 & 2 \\ 4 & 0 & 5 & 4 & 3 \\ 5 & 7 & 0 & 3 & 4 \\ 2 & 7 & 4 & 0 & 4 \\ 1 & 3 & 2 & 1 & 0 \end{bmatrix}.$$

Moreover, upon presentation of $\mathbf{x}$ as input, we obtain from $\mathcal{M}_{XX}$ the pattern

$$\mathcal{M}_{XX}(\mathbf{x}) = \begin{bmatrix} 4 & 0 & 6 & 4 & 2 \end{bmatrix}^T, \tag{17}$$

which differs from the original pattern $\mathbf{x}^1$ except in the second and forth components. Note that $\mathcal{M}_{XX}(\mathbf{x})$ does not belong to the fundamental memory set $\mathcal{X}$. Thus, it is an spurious memory of $\mathcal{M}_{XX}$.

A quantitative measure of the performance of $\mathcal{M}_{XX}$ as well as the other memory models considered in this paper is presented in Table I. The first row of this table contains the *normalized mean squared errors* (NMSEs) between the fundamental memories $\mathbf{x}^\xi$ and the input $\mathbf{x}$, for $\xi = 1, 2, 3$. Recall that the NMSE is defined by

$$\text{NMSE}(\mathbf{x}, \mathbf{y}) = \frac{\sum_{j=1}^n (x_j - y_j)^2}{\sum_{j=1}^n y_j^2} = \frac{\|\mathbf{x} - \mathbf{y}\|_2^2}{\|\mathbf{y}\|_2^2}, \tag{18}$$

where $\|\cdot\|_2$ denotes the usual Euclidean norm. The second row of Table I shows the NMSE produced by $\mathcal{M}_{XX}$. Note that, although $\mathbf{x}^1$ has more components equal to the input $\mathbf{x}$ than to the recalled pattern $\mathcal{M}_{XX}(\mathbf{x})$, $\text{NMSE}\big(\mathcal{M}_{XX}(\mathbf{x}), \mathbf{x}^1\big)$ is significantly smaller than $\text{NMSE}(\mathbf{x}, \mathbf{x}^1)$.

TABLE I
NORMALIZED MEANS SQUARED ERRORS.

1^o argument	$\text{NMSE}(\cdot, \mathbf{x}^1)$	$\text{NMSE}(\cdot, \mathbf{x}^2)$	$\text{NMSE}(\cdot, \mathbf{x}^3)$
$\mathbf{x}$	0.833333	1.471264	0.345865
$\mathcal{M}_{XX}(\mathbf{x})$	0.076923	0.816092	0.218045
$\mathcal{Z}(\mathbf{x})$	**0.025641**	0.816092	0.308271
$\mathcal{M}^{\#}_{XX}(\mathbf{x})$	0.653846	1.126437	0.240602
$\mathcal{M}^{M}_{XX}(\mathbf{x})$	0.435897	1.091954	0.067669
$\mathcal{V}_{XX}(\mathbf{x})$	**0.025641**	0.816092	0.308271
$\mathcal{V}^{\#}_{XX}(\mathbf{x})$	0.846154	1.275862	0.368421
$\mathcal{V}^{M}_{XX}(\mathbf{x})$	0.435897	1.091954	0.067669

B. Generalized Kernel Method for the Original AMMs

The original AMM $\mathcal{M}_{XX}$, which uses the min-product in the retrieval phase, is suitable for patterns degraded by dilative noise. Dually, the AMM $\mathcal{M}_{XX}$, whose retrieval is based on the max-product, is suitable for patterns corrupted by erosive noise. The idea of the kernel method and its generalization is to combine the max-product and the min-product in the retrieval phase of an AMM which, hopefully, will be able to deal with both dilative and erosive noise [10], [31], [32].

The notion of a kernel has primarily been defined in [10], [31]. The following presents the generalized kernel introduced by Sussner in [32]: A matrix $Z \in \mathbb{R}^{p \times k}$, $p \geq k$, is a *generalized kernel* for X if the equation

$$W_{ZX} \boxed{\vee} (M_Z^X \boxed{\wedge} X) = X \tag{19}$$

holds true where the matrices W_{ZX} and M_Z^X are given by

$$W_{ZX} = X \boxed{\wedge} Z^* \quad \text{and} \quad M_Z^X = (Z \boxed{\vee} X^*) \boxed{\wedge} (X \boxed{\vee} X^*). \tag{20}$$

Note that the dimension p of a generalized kernel Z may differ from the dimension n of X. Furthermore, given a matrix $X = [\mathbf{x}^1, \ldots, \mathbf{x}^k]$ whose columns correspond to the fundamental memories and a generalized kernel Z for X, the mapping $\mathcal{Z} : \mathbb{R}^n \to \mathbb{R}^n$ given by

$$\mathcal{Z}(\mathbf{x}) = W_{ZX} \boxed{\vee} (M_Z^X \boxed{\wedge} \mathbf{x}), \quad \forall \mathbf{x} \in \mathbb{R}^n, \tag{21}$$

is called *generalized kernel AMM* (GK-AMM). Theoretical results as well as computational experiments confirmed the excellent noise tolerance of the GK-AMM for the storage and recall of binary patterns [32].

Example 2. Consider the matrix $X = [\mathbf{x}^1, \mathbf{x}^2, \mathbf{x}^3] \in \mathbb{R}^{5 \times 3}$ whose columns correspond to the fundamental memories in (15). It is rather straightforward to verify that the matrix $Z =$

$10I_{3\times 3}$, where $I_{3\times 3}$ denotes the identity matrix of size 3×3, is a generalized kernel for X. Hence, the GK-AMM designed for the storage of $\mathcal{X}$ in (15) is the mapping $\mathcal{Z}$ given by (21) where

$$W_{ZX}=\begin{bmatrix}-8 & -7 & -4\\ -10 & -3 & -9\\ -3 & -8 & -6\\ -6 & -7 & -2\\ -7 & -6 & -6\end{bmatrix} \quad\text{and}\quad M_Z^X=\begin{bmatrix}8 & 10 & 3 & 6 & 7\\ 7 & 3 & 8 & 7 & 6\\ 4 & 9 & 6 & 2 & 6\end{bmatrix}.$$

Upon presentation of the input $\mathbf{x}$ given by (16), we obtain as output of the GK-AMM the pattern

$$\mathcal{Z}(\mathbf{x}) = \begin{bmatrix}2 & 0 & 6 & 4 & 2\end{bmatrix}^T,$$

which coincides with the original pattern $\mathbf{x}^1$ except in the third and fifth components, where the difference is one.

We would like to point out that feeding $\mathcal{Z}$ with the output of $\mathcal{M}_{XX}$ and conversely yields

$$\mathcal{Z}\big(\mathcal{M}_{XX}(\mathbf{x})\big) \neq \mathcal{M}_{XX}(\mathbf{x}) \quad\text{and}\quad \mathcal{M}_{XX}\big(\mathcal{Z}(\mathbf{x})\big) = \mathcal{Z}(\mathbf{x}).$$

Therefore, $\mathcal{Z}(\mathbf{x})$ is a fixed point of $\mathcal{M}_{XX}$ but $\mathcal{M}_{XX}(\mathbf{x})$ is not a fixed point of $\mathcal{Z}$. Indeed, the memory $\mathcal{Z}$ has less spurious memories than the original AMM $\mathcal{M}_{XX}$. Furthermore, as shown in the first column of Table I, the GK-AMM $\mathcal{Z}$ (third row) produced an NMSE smaller than that yielded by $\mathcal{M}_{XX}$ in this example.

C. Chebyshev-Best Approximation AMMs

Recall from (14) that the original AMM $\mathcal{M}_{XX}(\mathbf{x})$ yields the largest minimax combination of $\mathcal{X}$ which is less than or equal to the input $\mathbf{x}$. The retrieval phase of this AMM can be slightly adapted to yield the minimax combination of $\mathcal{X}$ that is the closest approximation $\mathbf{x}$ [17].

First, recall that the *Chebyshev distance* between two vectors $\mathbf{x},\mathbf{z} \in \mathbb{R}^n$, denoted by $\|\mathbf{x}-\mathbf{z}\|_\infty$, is the greatest component-wise absolute difference between $\mathbf{x}$ and $\mathbf{y}$, i.e.,

$$\|\mathbf{x}-\mathbf{z}\|_\infty = \bigvee_{j=1}^{n} |x_j - z_j|. \tag{22}$$

Now, the solution of the constrained optimization problem

$$\text{Minimize } \|\mathbf{x}-\mathbf{z}\|_\infty \text{ subject to } \mathbf{z} \in \mathfrak{S}(\mathcal{X}), \tag{23}$$

is the vector $\mathbf{z}^\# = \mu + M_{XX} \boxtimes \mathbf{x}$, where M_{XX} is the matrix given by (12) and $\mu = \frac{1}{2}(M_{XX}\boxtimes \mathbf{x})^* \boxtimes \mathbf{x}$ [4], [17]. Remembering that $\mathcal{M}_{XX}(\mathbf{x}) = M_{XX}\boxtimes\mathbf{x}$, we can express the Chebyshev-best approximation of $\mathbf{x}$ as $\mu+\mathcal{M}_{XX}(\mathbf{x})\boxtimes\mathbf{x}$ where $\mu = (\mathcal{M}^*_{XX}(\mathbf{x})\boxtimes\mathbf{x})/2$.

The remarks in the previous paragraph provide support to the following memory model [17]: The *Chebyshev-best approximation AMM* (CBA-AMM) corresponds to the mapping $\mathcal{M}^\#_{XX}:\mathbb{R}^n\to\mathbb{R}^n$ defined by

$$\mathcal{M}^\#_{XX}(\mathbf{x}) = \frac{1}{2}\mathcal{M}^*_{XX}(\mathbf{x})\boxtimes\mathbf{x} + \mathcal{M}_{XX}(\mathbf{x}), \quad \forall \mathbf{x}\in\mathbb{R}^n. \tag{24}$$

Such as the original AMM $\mathcal{M}_{XX}$, the CBA-AMM $\mathcal{M}^\#_{XX}$ projects the input $\mathbf{x}$ onto the set $\mathfrak{S}(\mathcal{X})$ of all minimax combinations of the fundamental memories. As a consequence, the CBA-AMM also has optimal absolute storage capacity and the output remain stable under repeated applications of $\mathcal{M}^\#_{XX}$. In addition, since the inequality $\mathcal{M}^\#_{XX}(\mathbf{x}) \leq \mathbf{x}$ does not necessarily hold true, the CBA-AMM exhibits some tolerance to mixed noise [17]. On the downside, because $\mathcal{M}_{XX}$ and $\mathcal{M}^\#_{XX}$ have the same set of fixed points, the CBA-AMM also has a large number of spurious memories.

Example 3. Consider the fundamental memory set $\mathcal{X}$ and the input pattern $\mathbf{x}$ given respectively by (15) and (16). In this case, $\mu = \big(\mathcal{M}^*_{XX}(\mathbf{x})\boxtimes\mathbf{x}\big)/2 = 3$ and the output of the CBA-AMM $\mathcal{M}^\#_{XX}$ is

$$\mathcal{M}^\#_{XX}(\mathbf{x}) = \begin{bmatrix}7 & 3 & 9 & 7 & 5\end{bmatrix}^T.$$

Note that $\zeta\big(\mathcal{M}^\#_{XX}(\mathbf{x}),\mathbf{x}\big) = 3$ while $\zeta\big(\mathcal{M}_{XX}(\mathbf{x}),\mathbf{x}\big) = 6$. Thus, the output of $\mathcal{M}^\#_{XX}$ is closer to the input – using the Chebyshev distance – than $\mathcal{M}_{XX}(\mathbf{x})$. Notwithstanding, Table I reveals that the NMSEs produced by $\mathcal{M}^\#_{XX}$ are larger than those yielded by $\mathcal{M}_{XX}$ in this example.

D. Noise Masking Strategy

Urcid et al. observed that the noise tolerance of the original AMMs can be significantly improved by masking the noise contained in a corrupted input pattern [33], [34]. In few words, noise masking converts an input degraded by mixed noise into a pattern corrupted by either dilative or erosive noise. Let us clarify this idea.

Suppose that $\mathcal{X} = \{\mathbf{x}^1,\ldots,\mathbf{x}^k\}$ have been stored in the AMM $\mathcal{M}_{XX}$. Also, assume that $\mathbf{x}$ is a version of the fundamental memory $\mathbf{x}^\eta$ corrupted by mixed noise. Then, $\mathbf{x}^\eta_p = \mathbf{x}\vee\mathbf{x}^\eta$ is the masked input pattern which contains only dilative noise, i.e., the inequality $\mathbf{x}^\eta_p \geq \mathbf{x}^\eta$ holds true. Since $\mathcal{M}_{XX}$ is robust to dilative noise, we expect the AMM to recall perfectly the original pattern $\mathbf{x}^\eta$ under presentation of the masked vector $\mathbf{x}^\eta_p$.

The noise masking idea has a practical shortcoming: we do not known a priori which fundamental memory have been corrupted. Hence, Urcid and Ritter suggested to compare, for all $\xi = 1,\ldots,k$, the masked pattern $\mathbf{x}^\xi_p = \mathbf{x}\vee\mathbf{x}^\xi$ with the input $\mathbf{x}$ as well as with the fundamental memory $\mathbf{x}^\xi$ [33]. The comparison is based on some meaningful measure such as the NMSE defined by (18). Subsequently, the masked pattern $\mathbf{x}^\eta_p$ that minimizes the average errors

$$D_\xi = \frac{1}{2}\big(\text{NMSE}(\mathbf{x}^\xi,\mathbf{x}^\xi_p) + \text{NMSE}(\mathbf{x},\mathbf{x}^\xi_p)\big), \forall \xi \in \mathcal{K}, \tag{25}$$

is fed into $\mathcal{M}_{XX}$. The following lemma provides an alternative formula for computing the index η that minimizes (25).

Lemma 1. *Given $\mathcal{X} = \{\mathbf{x}^1,\ldots,\mathbf{x}^k\}\subseteq\mathbb{R}^n$ and $\mathbf{x}\in\mathbb{R}^n$, an index η satisfies $D_\eta = \wedge_{\xi=1}^k D_\xi$ if and only if*

$$\frac{\|\mathbf{x}-\mathbf{x}^\eta\|_2^2}{\|\mathbf{x}\vee\mathbf{x}^\eta\|_2^2} = \bigwedge_{\xi=1}^{k}\left\{\frac{\|\mathbf{x}-\mathbf{x}^\xi\|_2^2}{\|\mathbf{x}\vee\mathbf{x}^\xi\|_2^2}\right\}. \tag{26}$$

Concluding, the technique of noise masking for recall of real-valued patterns using $\mathcal{M}_{XX}$ yields the memory $\mathcal{M}^M_{XX}:\mathbb{R}^n\to\mathbb{R}^n$ given by $\mathcal{M}^M_{XX}(\mathbf{x}) = \mathcal{M}_{XX}(\mathbf{x}\vee\mathbf{x}^\eta)$, where η is an index that satisfies (26).

Example 4. Consider the fundamental memory set and the input pattern given respectively by (15) and (16). In this case,

$$\frac{\|\mathbf{x}-\mathbf{x}^1\|_2^2}{\|\mathbf{x}\vee\mathbf{x}^1\|_2^2}=0.37, \frac{\|\mathbf{x}-\mathbf{x}^2\|_2^2}{\|\mathbf{x}\vee\mathbf{x}^2\|_2^2}=0.56, \frac{\|\mathbf{x}-\mathbf{x}^3\|_2^2}{\|\mathbf{x}\vee\mathbf{x}^3\|_2^2}=0.20.$$

Note that $\eta = 3$ satisfies (26). Thus, the output of $\mathcal{M}_{XX}^M$ is

$$\mathcal{M}_{XX}^M(\mathbf{x}) = \mathcal{M}_{XX}(\mathbf{x}\vee\mathbf{x}^3) = \begin{bmatrix}6 & 1 & 7 & 8 & 4\end{bmatrix}^T,$$

which coincides with the fundamental memory $\mathbf{x}^1$ only in the third component. Indeed, the pattern recalled by $\mathcal{M}_{XX}^M$ is more similar to $\mathbf{x}^3$ than to $\mathbf{x}^1$. The NMSEs produced by $\mathcal{M}_{XX}^M$ is also shown in Table I.

IV. Max-plus Projection Autoassociative Morphological Memory

We can design an AMM model with a reduced number of spurious memories by replacing $\mathfrak{S}(\mathcal{X})$ by the set $\mathfrak{V}(\mathcal{X})$ of all max-plus combination of vectors of $\mathcal{X}$ in (14). Precisely, given a fundamental memory set $\mathcal{X} = \{\mathbf{x}^1,\ldots,\mathbf{x}^k\} \subseteq \mathbb{R}^n$, we define an AMM $\mathcal{V}_{XX} : \mathbb{R}^n \to \mathbb{R}^n$ as follows for any input pattern $\mathbf{x}\in\mathbb{R}^n$:

$$\mathcal{V}_{XX}(\mathbf{x}) = \bigvee\{\mathbf{y}\in\mathfrak{V}(\mathcal{X}) : \mathbf{y}\leq\mathbf{x}\}. \tag{27}$$

Such as the original AMM $\mathcal{M}_{XX}$, the novel memory $\mathcal{V}_{XX}$ is idempotent. In fact, the mapping $\mathcal{V}_{XX}$ projects the input pattern $\mathbf{x}$ onto the set of all max-plus combinations of the fundamental memories $\mathbf{x}^1,\ldots,\mathbf{x}^k$. Thus, $\mathcal{V}_{XX}$ is referred to as a *max-plus projection autoassociative morphological memory* or *max-plus PAMM* for short.

Obviously, $\mathcal{V}_{XX}$ also exhibit perfect recall of undistorted patterns because $\mathbf{x}^\xi\in\mathfrak{V}(\mathcal{X})$ for all $\xi\in\mathcal{K}$. Notwithstanding, the inequality $\mathcal{V}_{XX}(\mathbf{x})\leq\mathbf{x}$ holds for all $\mathbf{x}$. Therefore, such as the original AMM $\mathcal{M}_{XX}$, the max-plus PAMM is suited for the reconstruction of patterns corrupted by dilative noise, but it is incapable of handling erosive or mixed noise.

The advantage of $\mathcal{V}_{XX}$ is that it has less spurious memories than $\mathcal{M}_{XX}$. In fact, any finite vector in the set difference $\mathfrak{V}(\mathcal{X})\setminus\mathcal{X}$ is a spurious memory of $\mathcal{V}_{XX}$. It follows from the inclusion $\mathfrak{V}(\mathcal{X})\subseteq\mathfrak{S}(\mathcal{X})$ that any spurious memory of $\mathcal{V}_{XX}$ is also a spurious memory of $\mathcal{M}_{XX}$. However, the converse $\mathfrak{S}(\mathcal{X})\subseteq\mathfrak{V}(\mathcal{X})$ does not hold true. Let us conclude by remarking that the inequalities $\mathcal{V}_{XX}(\mathbf{x})\leq\mathcal{M}_{XX}(\mathbf{x})\leq\mathbf{x}$ hold for all $\mathbf{x}\in\mathbb{R}^n$.

The following theorem provides a formula to compute the output of $\mathcal{V}_{XX}$ given by (27). Furthermore, analogous to (13), the pattern recalled by the novel max-plus PAMM is characterized in terms of extrema operations and the fundamental memories.

Theorem 1. *Let $X = [\mathbf{x}^1,\ldots,\mathbf{x}^k]\in\mathbb{R}^{n\times k}$ be the matrix whose columns correspond to the fundamental memories stored in a max-plus PAMM $\mathcal{V}_{XX}$. For any input pattern $\mathbf{x}\in\mathbb{R}^n$, the output of $\mathcal{V}_{XX}$ satisfies*

$$\mathcal{V}_{XX}(\mathbf{x}) = X \boxdot \boldsymbol{\alpha}, \quad \textit{where} \quad \boldsymbol{\alpha} = X^* \boxslash \mathbf{x}. \tag{28}$$

Alternatively, the output of $\mathcal{V}_{XX}(\mathbf{x})$ can be expressed as

$$\mathcal{V}_{XX}(\mathbf{x}) = \bigvee_{\xi=1}^{k}\bigwedge_{j=1}^{n}\big((x_j - x_j^\xi)+\mathbf{x}^\xi\big). \tag{29}$$

At this point, we invite the reader to compare (29) with (13), the equations that characterize the patterns recalled by $\mathcal{V}_{XX}$ and $\mathcal{M}_{XX}$, respectively. Note that the maximum and minimum operations have exchanged their places. However, the following example shows that the two memories may produce different output upon presentation of the same input.

Example 5. Consider the fundamental memory set $\mathcal{X}$ and the input pattern $\mathbf{x}$ given respectively by (15) and (16). The vector $\boldsymbol{\alpha}$ given by (28) satisfies

$$\boldsymbol{\alpha} = \begin{bmatrix}-1 & -7 & -4\end{bmatrix}^T,$$

and the output of the max-plus PAMM is

$$\begin{aligned}\mathcal{V}_{XX}(\mathbf{x}) &= \big(\alpha_1+\mathbf{x}^1\big)\vee\big(\alpha_2+\mathbf{x}^2\big)\vee\big(\alpha_3+\mathbf{x}^3\big)\\ &= \left((-1)+\begin{bmatrix}2\\0\\7\\4\\3\end{bmatrix}\right)\vee\left((-7)+\begin{bmatrix}3\\7\\2\\3\\4\end{bmatrix}\right)\vee\left((-4)+\begin{bmatrix}6\\1\\4\\8\\4\end{bmatrix}\right)\\ &= \begin{bmatrix}1\\-1\\6\\3\\2\end{bmatrix}\vee\begin{bmatrix}-4\\0\\-5\\-4\\-3\end{bmatrix}\vee\begin{bmatrix}2\\-3\\0\\4\\0\end{bmatrix} = \begin{bmatrix}2\\0\\6\\4\\2\end{bmatrix}\end{aligned}$$

Note that $\mathcal{V}_{XX}(\mathbf{x})$ differs from $\mathbf{x}^1$ only in the third and fifth components. As expected, the inequalities $\mathcal{V}_{XX}(\mathbf{x})\leq\mathcal{M}_{XX}\leq\mathbf{x}$ hold true. Also, the pattern recalled by $\mathcal{V}_{XX}$ is equal to the pattern recalled by the GK-AMM of Example 2. Hence, the pattern recalled by $\mathcal{V}_{XX}$ is a fixed point of $\mathcal{M}_{XX}$ but the converse does not hold true. In addition, as shown in the first column of Table I, the memory $\mathcal{V}_{XX}$ outperformed the original AMM $\mathcal{M}_{XX}$ as well as its variations $\mathcal{M}_{XX}^{\#}$ and $\mathcal{M}_{XX}^M$. In the following subsection, we investigate the relationship between $\mathcal{Z}$ and $\mathcal{V}_{XX}$.

Let us now briefly examine the coefficients α_ξ of the vector $\boldsymbol{\alpha}$ given by (28). It is not hard to show that $\alpha_\xi\geq 0$ if and only if $\mathbf{x}\geq\mathbf{x}^\xi$. Conversely, $\alpha_\xi<0$ if and only if there is at least one index $j\in\mathcal{N}$ such that $x_j\leq x_j^\xi$. Hence, in some sense, α_ξ measures the truth of the inequality $\mathbf{x}^\xi\leq\mathbf{x}$. In view of this remark, let us define the mapping $\mathcal{A}:\mathbb{R}^n\times\mathbb{R}^n\to\mathbb{R}$ as follows for any vectors $\mathbf{x}$ and $\mathbf{y}$:

$$\mathcal{A}(\mathbf{y},\mathbf{x}) = \bigwedge_{j=1}^{n}(x_j-y_j) = \mathbf{y}^*\boxslash\mathbf{x}, \tag{30}$$

Therefore, $\alpha_\xi = \mathcal{A}(\mathbf{x}^\xi,\mathbf{x})$ for all $\xi\in\mathcal{K}$ and the max-plus PAMM satisfies

$$\mathcal{V}_{XX}(\mathbf{x}) = \bigvee_{\xi=1}^{k}\big(\mathcal{A}(\mathbf{x}^\xi,\mathbf{x})+\mathbf{x}^\xi\big), \tag{31}$$

for any $\mathbf{x} \in \mathbb{R}^n$. Alternatively, the pattern recalled by $\mathcal{V}_{XX}$ can be determined as follows:

Theorem 2 (Dual Representation of $\mathcal{V}_{XX}$)**.** *Given a fundamental memory set $\mathcal{X} = \{\mathbf{x}^1, \ldots, \mathbf{x}^k\}$, the memory $\mathcal{V}_{XX}$ given by* (27) *satisfies*

$$\mathcal{V}_{XX}(\mathbf{x}) = \bigwedge \{\mathbf{y} \in \mathbb{R}^n : \mathcal{A}(\mathbf{x}^\xi, \mathbf{x}) \leq \mathcal{A}(\mathbf{x}^\xi, \mathbf{y}), \forall \xi \in \mathcal{K}\}, \tag{32}$$

for all input $\mathbf{x} \in \mathbb{R}^n$.

Example 6. In Example 5, for any $\xi \in \{1, 2, 3\}$, we have $\alpha_\xi = \mathcal{A}(\mathbf{x}^\xi, \mathbf{x}) < 0$. Thus, there exists at least one index j such that $x_j \leq x_j^\xi$. Indeed, $x_5 = 2 \leq 3 = x_5^1$, $x_2 = 0 \leq 7 = x_2^2$, and $x_4 = 4 \leq 8 = x_4^3$. Note that the memory $\mathcal{V}_{XX}$ failed to recall the fundamental memory $\mathbf{x}^1$ because $\mathcal{A}(\mathbf{x}^1, \mathbf{x}) < 0$.

We would like to point out that the coefficients $\alpha_\xi = \mathcal{A}(\mathbf{x}^\xi, \mathbf{x})$ can be computed in parallel. Thus, they do not pose any computational burden in applications of $\mathcal{V}_{XX}$ for the storage of large-scale patterns. Indeed, in contrast to the original AMM and the Hopfield network [12], the novel memory does not require the storage of a synaptic weight matrix of size $n \times n$. On the downside, the information on the fundamental memories is not distributed over the weights of the max-plus PAMM.

A. Max-Plus PAMM as a GK-AMM

In accordance with the idea used in the kernel method, (28) reveals that the max-product and the min-product are combined in the retrieval phase of the max-plus PAMM. In this subsection, we explore further the relationship between the novel memory and the GK-AMM given by (21). Precisely, we provide a generalized kernel Z such that the mappings $\mathcal{V}_{XX}$ and $\mathcal{Z}$ coincide in an hyperbox.

In many practical situations, the entries of the fundamental memories as well as the components of the input pattern are confined into an interval $\mathbb{I} = [a, b]$. In this case, we write $\mathcal{X} = \{\mathbf{x}^1, \ldots, \mathbf{x}^k\} \subseteq \mathbb{I}^n$ and $\mathbf{x} \in \mathbb{I}^n$. Also, we are able to state the following theorem:

Theorem 3. *Given a set $\mathcal{X} = \{\mathbf{x}^1, \ldots, \mathbf{x}^k\}$ of fundamental memories with entries in $\mathbb{I} = [a, b]$, define the diameter $\omega = b - a$ and the matrices $X = [\mathbf{x}^1, \ldots, \mathbf{x}^k]$ and $Z = \omega I_{k\times k}$, where $I_{k\times k}$ is the identity matrix of size k. The matrix Z is a generalized kernel for X. Furthermore, this generalized kernel yields a GK-AMM $\mathcal{Z}$ such that $\mathcal{Z}(\mathbf{x}) = \mathcal{V}_{XX}(\mathbf{x})$ for all input pattern $\mathbf{x} \in \mathbb{I}^n$.*

Example 7. Note that the entries of the fundamental memories $\mathbf{x}^1, \mathbf{x}^2, \mathbf{x}^3$ in (15) belong to the interval $\mathbb{I} = [0, 10]$, whose diameter is $\omega = 10$. From Theorem 3, we conclude that the max-plus PAMM of Example 5 coincides with the GK-AMM of Example 2 for all input in $\mathbb{I}^5$, including the pattern $\mathbf{x}$ given by (16).

B. Two Variations of the Max-plus PAMM

The max-plus memory $\mathcal{V}_{XX}$ shares many properties with the original memory $\mathcal{M}_{XX}$. As a consequence, the recall performance of $\mathcal{V}_{XX}$ can be improved in a straightforward way by noise masking the input pattern. Alternatively, the recall phase of $\mathcal{V}_{XX}$ can be adapted to yield the Chebyshev best-approximation max-plus combination of the input pattern. The following presents these two variations of $\mathcal{V}_{XX}$.

First, analogous to the CBA-AMM $\mathcal{M}_{XX}^{\#}$, the retrieval phase of the max-plus PAMM $\mathcal{V}_{XX}$ can be modified to yield the following memory model: The *Chebyshev-best approximation PAMM* (CBA-PAMM) is the mapping $\mathcal{V}_{XX}^{\#} : \mathbb{R}^n \to \mathbb{R}^n$ given by the following equation for any input $\mathbf{x} \in \mathbb{R}^n$:

$$\mathcal{V}_{XX}^{\#}(\mathbf{x}) = \frac{1}{2}\mathcal{V}_{XX}^{*}(\mathbf{x}) \boxtimes \mathbf{x} + \mathcal{V}_{XX}(\mathbf{x}). \tag{33}$$

Using previous results (e.g. Theorem 7 of [17]), it is rather straightforward to show that $\mathcal{V}_{XX}^{\#}$ yields the Chebyshev-best approximation of $\mathbf{x}$ by max-plus combinations of the fundamental memories $\mathbf{x}^1, \ldots, \mathbf{x}^k$. In other words, $\mathcal{V}_{XX}^{\#}$ is the solution of the constrained optimization problem:

$$\text{Minimize } \|\mathbf{x} - \mathbf{y}\|_\infty \text{ subject to } \mathbf{y} \in \mathfrak{V}(\mathcal{X}), \tag{34}$$

where $\mathfrak{V}(\mathcal{X})$ is the set defined in (7). Needless to say that $\mathcal{V}_{XX}^{\#}$ inherits from $\mathcal{V}_{XX}$ all desired properties except the anti-extensiveness expressed by the inequality $\mathcal{V}_{XX}(\mathbf{x}) \leq \mathbf{x}$, for all $\mathbf{x}$.

Also, a robust memory model $\mathcal{V}_{XX}^{M}$ is obtained by replacing $\mathcal{M}_{XX}$ by $\mathcal{V}_{XX}$ in the original technique of noise masking proposed by Urcid et at. Precisely, the mapping $\mathcal{V}_{XX}^{M} : \mathbb{R}^n \to \mathbb{R}^n$ is defined by means of the equation

$$\mathcal{V}_{XX}^{M}(\mathbf{x}) = \mathcal{V}_{XX}(\mathbf{x} \vee \mathbf{x}^\eta), \quad \forall \mathbf{x} \in \mathbb{R}^n, \tag{35}$$

where η is an index that satisfies (26).

Example 8. Consider the fundamental memory set $\mathcal{X}$ given by (15). Presenting the pattern $\mathbf{x}$ in (16) as input to $\mathcal{V}_{XX}^{\#}$ and $\mathcal{V}_{XX}^{M}$, we obtain as output the patterns

$$\mathcal{V}_{XX}^{\#}(\mathbf{x}) = \begin{bmatrix} 6 \\ 4 \\ 10 \\ 8 \\ 6 \end{bmatrix} \quad \text{and} \quad \mathcal{V}_{XX}^{M}(\mathbf{x}) = \begin{bmatrix} 6 \\ 1 \\ 7 \\ 8 \\ 4 \end{bmatrix}.$$

Note that $\mathcal{V}_{XX}^{*}(\mathbf{x}) \boxtimes \mathbf{x} = 8$ and $\mathcal{V}_{XX}^{\#}(\mathbf{x}) = 4 + \mathcal{V}_{XX}(\mathbf{x})$. Also, such as in Example 4, the index $\eta = 3$ have been used in (35). As a consequence, the NMSE between the pattern recalled by $\mathcal{V}_{XX}^{M}$ and $\mathbf{x}^3$ is less than $\text{NMSE}\big(\mathcal{V}_{XX}^{M}(\mathbf{x}), \mathbf{x}^1\big)$. The NMSEs produced by $\mathcal{V}_{XX}^{\#}$ and $\mathcal{V}_{XX}^{M}$ are also shown in Table I.

V. Concluding Remarks

In this paper, we briefly reviewed the original autoassociative morphological memories (AMMs) introduced by Ritter and Sussner in the middle 1990s [9], [10]. In particular, we recalled that the output of the AMM $\mathcal{M}_{XX}$ corresponds to the largest minimax combination of the fundamental memories which is less than or equal to the input (cf. Equation 14). We also reviewed some of their variations including the generalized kernel method [32], the Chebyshev-best approximation AMM [17], and the noise masking technique proposed by Urcid and Ritter [33].

Subsequently, we introduced the max-plus projection AMM (max-plus PAMM) $\mathcal{V}_{XX}$ by replacing the set of all minimax combinations $\mathfrak{S}(\mathcal{X})$ by $\mathfrak{V}(\mathcal{X})$, the set of all max-plus combinations of the fundamental memories (cf. Equation 27). We showed that the output of the novel memory can be expressed as a maximum of minima in vertical translations of the fundamental memories (cf. Equation 29). In addition, we showed that the novel memory corresponds to a certain generalized-kernel AMM. We also presented two variations of the max-plus AMM: one based on the Chebyshev-best approximation and the other using the noise masking technique.

Such as the original AMM $\mathcal{M}_{XX}$, the novel memory is idempotent and anti-extensive, i.e., $\mathcal{V}_{XX}\big(\mathcal{V}_{XX}(\mathbf{x})\big) = \mathcal{V}_{XX}(\mathbf{x})$ and $\mathcal{V}_{XX}(\mathbf{x}) \leq \mathbf{x}$ for all $\mathbf{x} \in \mathbb{R}^n$. Furthermore, any fundamental memory is a fixed point of $\mathcal{V}_{XX}$, i.e., $\mathcal{V}_{XX}(\mathbf{x}^\xi) = \mathbf{x}^\xi$. The advantage of the max-plus PAMM $\mathcal{V}_{XX}$ is that it has less spurious memories than $\mathcal{M}_{XX}$. Consequently, it is more robust to dilative noise than the latter. Also, the recall phase of $\mathcal{V}_{XX}$ can be optimized by parallelizing the computation of the vector $\boldsymbol{\alpha}$ and do not require the storage of a synaptic weight of size $n \times n$.

In the future, we plan to further investigate the noise tolerance of the max-plus PAMM. In particular, we intent to evaluate the performance of the novel memories for the reconstruction of images corrupted by different types of noise.

Appendix
Proofs of Theorems and Lemmas

Proof of Lemma 1: The components of the masked pattern $\mathbf{x}_p^\xi$ satisfy $(x_p^\xi)_j = x_j^\xi$ for all $j \in J_\xi = \{j : x_j^\xi \geq x_j\}$. Similarly, we have $(x_p^\xi)_j = x_j$ for $j \notin J_\xi$. Therefore, the average distance given by (25) satisfies

$$\begin{aligned} D_\xi &= \frac{1}{2\|\mathbf{x}_p^\xi\|_2^2} \sum_{j=1}^n \Big(\big(x_j^\xi - (x_p^\xi)_j\big)^2 + \big(x_j - (x_p^\xi)_j\big)^2 \Big) \\ &= \frac{1}{2\|\mathbf{x}_p^\xi\|_2^2} \Bigg[\sum_{j \in J_\xi} \Big(\big(x_j^\xi - (x_p^\xi)_j\big)^2 + \big(x_j - (x_p^\xi)_j\big)^2 \Big) \\ &\qquad + \sum_{j \notin J_\xi} \Big(\big(x_j^\xi - (x_p^\xi)_j\big)^2 + \big(x_j - (x_p^\xi)_j\big)^2 \Big) \Bigg] \\ &= \frac{1}{2\|\mathbf{x}_p^\xi\|_2^2} \Bigg[\sum_{j \in J_\xi} (x_j - x_j^\xi)^2 + \sum_{j \notin J_\xi} (x_j^\xi - x_j)^2 \Bigg] \\ &= \frac{1}{2\|\mathbf{x}_p^\xi\|_2^2} \sum_{j=1}^n (x_j - x_j^\xi)^2 = \frac{1}{2} \frac{\|\mathbf{x} - \mathbf{x}^\xi\|_2^2}{\|\mathbf{x} \vee \mathbf{x}^\xi\|_2^2}. \end{aligned}$$

Since the masked input $\mathbf{x}_p^\eta$ is computed using an index η that minimizes D_ξ, we may discard the factor $1/2$. ■

Proof of Theorem 1: First of all, recall that $\mathbf{y} \in \mathfrak{V}(\mathcal{X})$ if and only if it satisfies (6). Furthermore, the following equivalences hold true for $\mathbf{y} \in \mathfrak{V}(\mathcal{X})$, $\mathcal{N} = \{1, \ldots, n\}$, and $\mathcal{K} = \{1, \ldots, k\}$:

$$\begin{aligned} \mathbf{y} \leq \mathbf{x} \quad &\Leftrightarrow \quad y_j \leq x_j, \forall j \in \mathcal{N} && (36) \\ &\Leftrightarrow \quad \bigvee_{\xi=1}^k (\alpha_\xi + x_j^\xi) \leq x_j, \forall j \in \mathcal{N} && (37) \\ &\Leftrightarrow \quad \alpha_\xi + x_j^\xi \leq x_j, \forall j \in \mathcal{N}, \forall \xi \in \mathcal{K} && (38) \\ &\Leftrightarrow \quad \alpha_\xi \leq x_j - x_j^\xi, \forall j \in \mathcal{N}, \forall \xi \in \mathcal{K} && (39) \\ &\Leftrightarrow \quad \alpha_\xi \leq \bigwedge_{j=1}^n (x_j - x_j^\xi), \forall \xi \in \mathcal{K}. && (40) \end{aligned}$$

Therefore, the largest $\mathbf{y} \in \mathfrak{V}(\mathcal{X})$ such that $\mathbf{y} \leq \mathbf{x}$ is given by (6) with

$$\alpha_\xi = \bigwedge_{j=1}^n (x_j - x_j^\xi), \quad \forall \xi = 1, \ldots, k. \tag{41}$$

Writing $\boldsymbol{\alpha} = [\alpha_1, \ldots, \alpha_k]^T \in \mathbb{R}^k$, we obtain (28). Finally, we derive (29) by substituting the expression for α_ξ given by (41) in (6). Here, recall that the addition is monotonic and, thus, it commutes with the minimum operation. ■

Proof of Theorem 2: From (28) and (5), we conclude that

$$\begin{aligned} \mathcal{V}_{XX}(\mathbf{x}) &= \bigwedge \{\mathbf{y} \in \mathbb{R}^n : \mathcal{V}_{XX}(\mathbf{x}) \leq \mathbf{y}\} \\ &= \bigwedge \{\mathbf{y} \in \mathbb{R}^n : X \boxed{\vee} (X^* \boxed{\wedge} \mathbf{x}) \leq \mathbf{y}\} \\ &= \bigwedge \{\mathbf{y} \in \mathbb{R}^n : X^* \boxed{\wedge} \mathbf{x} \leq X^* \boxed{\wedge} \mathbf{y}\} \\ &= \bigwedge \{\mathbf{y} \in \mathbb{R}^n : \mathcal{A}(\mathbf{x}^\xi, \mathbf{x}) \leq \mathcal{A}(\mathbf{x}^\xi, \mathbf{y}), \forall \xi \in \mathcal{K}\}, \end{aligned}$$

for any input $\mathbf{x} \in \mathbb{R}^n$. ■

Proof of Theorem 3: First of all, it is rather straightforward to show that if $\mathcal{X} \subseteq \mathbb{I}^n$, $\mathbf{x} \in \mathbb{I}^n$, and $\omega = b - a$ denotes the diameter of the interval $\mathbb{I}$ then the inequalities

$$(x_j - \omega) \leq \left(\bigwedge_{\xi=1}^k \bigwedge_{j=1}^n x_j^\xi \right) \leq \left(\bigvee_{\xi=1}^k \bigvee_{j=1}^n x_j^\xi \right) \leq (x_i + w), \tag{42}$$

hold true for all $j \in \mathcal{N}$.

Now, we only need to show that the mapping $\mathcal{Z}$ given by (21) with $Z = \omega I_{k \times k}$ coincide with the max-plus PAMM $\mathcal{V}_{XX}$. Indeed, since $\mathcal{V}_{XX}$ has optimal absolute storage capacity, we conclude that

$$\mathcal{V}_{XX}(\mathbf{x}^\xi) = \mathcal{Z}(\mathbf{x}^\xi) = W_{ZX} \boxed{\vee} (M_Z^X \boxed{\wedge} \mathbf{x}^\xi) = \mathbf{x}^\xi.$$

In a matrix form, we have $W_{ZX} \boxed{\vee} (M_Z^X \boxed{\wedge} X) = X$. Thus, the matrix Z is a generalized kernel for X.

Let us now establish a relationship between X and the matrices W_{ZX} and M_Z^X. From (42), we conclude that the entries of $W_{ZX} = X \boxed{\wedge} Z^* \in \mathbb{R}^{n \times k}$ satisfies the following equalities for all $i \in \mathcal{N}$ and $\eta \in \mathcal{K}$:

$$(W_{ZX})_{i\eta} = \bigwedge_{\xi=1}^k (x_i^\xi - z_\eta^\xi) = (x_i^\xi - \omega) \wedge \Big(\bigwedge_{\xi \neq \eta} x_i^\xi \Big) = x_i^\xi - \omega.$$

In other words, we have $W_{ZX} = X - \omega$ if $Z = \omega I_{k\times k}$. From the dually relationship given by (4), we obtain

$$Z \boxed{\vee} X^* = (X \boxed{\wedge} Z^*)^* = (X - \omega)^* = X^* + \omega.$$

Recalling that the original AMM $\mathcal{W}_{XX}$ also exhibit optimal absolute storage capacity, we conclude that the identities $X = W_{XX} \boxed{\vee} X = (X \boxed{\wedge} X^*) \boxed{\vee} X$ hold true. Therefore, by (3) and (4), we deduce the identities

$$\begin{aligned} M_Z^X &= (Z \boxed{\vee} X^*) \boxed{\wedge} (X \boxed{\vee} X^*) = (X^* + \omega) \boxed{\wedge} (X \boxed{\vee} X^*) \\ &= X^* \boxed{\wedge} (X \boxed{\vee} X^*) + \omega = \big((X \boxed{\wedge} X^*) \boxed{\vee} X\big)^* + \omega \\ &= X^* + \omega. \end{aligned}$$

Finally, the memories $\mathcal{Z}$ and $\mathcal{V}_{XX}$ satisfy the following equations for all $\mathbf{x} \in \mathbb{I}^n$:

$$\begin{aligned} \mathcal{Z}(\mathbf{x}) &= W_{ZX} \boxed{\vee} (M_Z^X \boxed{\wedge} \mathbf{x}) = (X - \omega) \boxed{\vee} \big((X^* + \omega) \boxed{\wedge} \mathbf{x}\big) \\ &= X \boxed{\vee} (X^* \boxed{\wedge} \mathbf{x}) - \omega + \omega = \mathcal{V}_{XX}(\mathbf{x}), \end{aligned}$$

Here, we used the identities in (2), (3), and (28). ■

REFERENCES

[1] V. Kaburlasos and G. Ritter, Eds., *Computational Intelligence Based on Lattice Theory*. Heidelberg, Germany: Springer-Verlag, 2007.

[2] M. Graña, "A brief review of lattice computing," in *Proceedings of the IEEE International Conference on Fuzzy Systems (FUZZ-IEEE 2008)*, Hong Kong, China, June 2008, pp. 1777–1781.

[3] M. Grana, "Lattice computing in hybrid intelligent systems," in *Hybrid Intelligent Systems (HIS), 2012 12th International Conference on*, 2012, pp. 1–5.

[4] R. Cuninghame-Green, *Minimax Algebra: Lecture Notes in Economics and Mathematical Systems 166*. New York: Springer-Verlag, 1979.

[5] H. Heijmans, *Morphological Image Operators*. New York, NY: Academic Press, 1994.

[6] V. Kaburlasos and A. Kehagias, "Fuzzy Inference System (FIS) Extensions Based on Lattice Theory," 2013, accepted for publication.

[7] V. Petridis and V. Kaburlasos, "Fuzzy lattice neural network (FLNN): a hybrid model for learning," *IEEE Transactions on Neural Networks*, vol. 9, no. 5, pp. 877–890, Sep. 1998.

[8] G. X. Ritter and G. Urcid, "Lattice Algebra Approach to Single-Neuron Computation," *IEEE Transactions on Neural Networks*, vol. 14, no. 2, pp. 282–295, 2003.

[9] G. X. Ritter and P. Sussner, "An Introduction to Morphological Neural Networks," in *Proceedings of the 13th International Conference on Pattern Recognition*, Vienna, Austria, 1996, pp. 709–717.

[10] G. X. Ritter, P. Sussner, and J. L. D. de Leon, "Morphological Associative Memories," *IEEE Transactions on Neural Networks*, vol. 9, no. 2, pp. 281–293, 1998.

[11] G. X. Ritter, G. Urcid, and M. S. Schmalz, "Lattice associative memories that are robust in the presence of noise," in *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series: Mathematical Methods in Pattern and Image Analysis*, vol. 5916. SPIE Publications, 2005, pp. 255–260.

[12] M. H. Hassoun, Ed., *Associative Neural Memories: Theory and Implementation*. Oxford, U.K.: Oxford University Press, 1993.

[13] P. Sussner and E. L. Esmi, "Morphological Perceptrons with Competitive Learning: Lattice-Theoretical Framework and Constructive Learning Algorithm," *Information Sciences*, vol. 181, no. 10, pp. 1929–1950, May 2011.

[14] P. Sussner and M. E. Valle, "Fuzzy Associative Memories and Their Relationship to Mathematical Morphology," in *Handbook of Granular Computing*, W. Pedrycz, A. Skowron, and V. Kreinovich, Eds. New York: John Wiley and Sons, Inc., 2008, ch. 33, pp. 733–754.

[15] M. E. Valle and P. Sussner, "A General Framework for Fuzzy Morphological Associative Memories," *Fuzzy Sets and Systems*, vol. 159, no. 7, pp. 747–768, 2008.

[16] P. Sussner and T. Schuster, "Interval-valued fuzzy associative memories based on representable conjunctions with applications in prediction," in *IFSA World Congress and NAFIPS Annual Meeting (IFSA/NAFIPS), 2013 Joint*, 2013, pp. 344–349.

[17] P. Sussner and M. E. Valle, "Grayscale Morphological Associative Memories," *IEEE Transactions on Neural Networks*, vol. 17, no. 3, pp. 559–570, May 2006.

[18] P. Sussner and C. Medeiros, "An introduction to morphological associative memories in complete lattices and inf-semilattices," in *Fuzzy Systems (FUZZ-IEEE), 2012 IEEE International Conference on*, 2012, pp. 1–8.

[19] R. A. Vazquez and H. Sossa, "Behavior of morphological associative memories with true-color image patterns," *Neurocomputing*, vol. 73, no. 1-3, pp. 225–244, December 2009.

[20] ——, "Behavioural study of median associative memory under true-colour image patterns," *Neurocomputing*, vol. 74, no. 17, pp. 2985–2997, 2011.

[21] D. Chyzhyk and M. Graña, "Results on a Lattice Computing Based Group Analysis of Schizophrenic Patients on Resting State fMRI," in *Natural and Artificial Computation in Engineering and Medical Applications*, ser. Lecture Notes in Computer Science, J. Ferrández Vicente, J. Álvarez Sánchez, F. Paz López, and F. Toledo Moreo, Eds. Springer Berlin Heidelberg, 2013, vol. 7931, pp. 131–139.

[22] G. X. Ritter and G. Urcid, "A lattice matrix method for hyperspectral image unmixing," *Information Sciences*, vol. 181, no. 10, pp. 1787–1803, May 2011.

[23] M. Graña and M. A. Veganzones, "An endmember-based distance for content based hyperspectral image retrieval," *Pattern Recognition*, vol. 45, no. 9, pp. 3472–3489, 2012, best Papers of Iberian Conference on Pattern Recognition and Image Analysis (IbPRIA'2011).

[24] M. Graña, I. Villaverde, J. M. López-Guede, and B. Fernández-Gauna, "Lattice independent component analysis for appearance-based mobile robot localization." *Neural Computing and Applications*, vol. 21, no. 5, pp. 1031–1042, 2012.

[25] P. Sussner, E. L. Esmi, I. Villaverde, and M. Graña, "The Kosko Subsethood Fuzzy Associative Memory (KS-FAM): Mathematical Background and Applications in Computer Vision," *Journal of Mathematical Imaging and Vision*, vol. 42, no. 2–3, pp. 134–149, Feb. 2012.

[26] P. Sussner and M. E. Valle, "Recall of Patterns Using Morphological and Certain Fuzzy Morphological Associative Memories," in *Proceedings of the IEEE World Conference on Computational Intelligence 2006*, Vancouver, Canada, 2006, pp. 209–216.

[27] P. Maragos, "Lattice Image Processing: A Unification of Morphological and Fuzzy Algebraic Systems," *Journal of Mathematical Imaging and Vision*, vol. 22, no. 2-3, pp. 333–353, 2005.

[28] G. Birkhoff, *Lattice Theory*, 3rd ed. Providence: American Mathematical Society, 1993.

[29] G. X. Ritter, J. L. D. de Leon, and P. Sussner, "Morphological Bidirectional Associative Memories," *Neural Networks*, vol. 6, no. 12, pp. 851–867, 1999.

[30] G. X. Ritter and P. Gader, "Fixed Points of Lattice Transforms and Lattice Associative Memories," in *Advances in Imaging and Electron Physics*, P. Hawkes, Ed. New York, NY: Academic Press, 2006, vol. 144.

[31] P. Sussner, "Observations on Morphological Associative Memories and the Kernel Method," *Neurocomputing*, vol. 31, pp. 167–183, Mar. 2000.

[32] ——, "Associative morphological memories based on variations of the kernel and dual kernel methods," *Neural Networks*, vol. 16, no. 5, pp. 625–632, Jul. 2003.

[33] G. Urcid and G. Ritter, "Noise Masking for Pattern Recall Using a Single Lattice Matrix Auto-Associative Memory," in *Fuzzy Systems, 2006 IEEE International Conference on*, 2006, pp. 187–194.

[34] G. Urcid and G. X. Ritter, "Noise Masking for Pattern Recall Using a Single Lattice Matrix Associative Memory," in *Computational Intelligence Based on Lattice Theory*, V. Kaburlasos and G. Ritter, Eds. Heidelberg, Germany: Springer-Verlag, 2007, ch. 5, pp. 81–100.