

# Brief Papers

## Complex-Valued Multistate Neural Associative Memory

Stanislaw Jankowski, Andrzej Lozowski, and Jacek M. Zurada

**Abstract**—A model of a multivalued associative memory is presented. This memory has the form of a fully connected attractor neural network composed of multistate complex-valued neurons. Such a network is able to perform the task of storing and recalling gray-scale images. It is also shown that the complex-valued fully connected neural network may be considered as a generalization of a Hopfield network containing real-valued neurons. A computational energy function is introduced and evaluated in order to prove network stability for asynchronous dynamics. Storage capacity as related to the number of accessible neuron states is also estimated.

### I. INTRODUCTION

CLASSIC neural networks are usually based on the McCulloch–Pitts neuron model. This model uses two-state neurons and is well suited to the processing of binary-valued vectors. In many situations, however, discrete but nonbinary state values would offer an advantageous representation of information. This includes processing of gray-scale images or mapping of input data into multiclass partitions using a nonbinary class encoding scheme. The concept presented below addresses the theoretical foundations of multivalued neurons and their use in associative memories.

There are several extensions to the basic Hopfield associative memory [1] leading to multivalued processing. Apparently the most simple is grouping a certain number of neurons into one cluster that functionally represents a single multivalued state. A network composed of such clusters can represent gray-scale simply by adjusting the density of activated neurons. In reality, however, the size of such a network would become enormously large even for small images unless information compression techniques are employed. Another choice is to assign each bit of a multivalued state representation to different neurons in a cluster as proposed in [2]. The corresponding bits can then be split off into separate noninteracting Hopfield networks and processed independently from each other. Once the separate networks have reached a fixed point, the final multivalued state is formed based on the bits recombination. Multivalued patterns storage is very efficient in this approach, however, it suffers from the lack of interactions between the bits describing multivalued states. For instance in an image recognition task a small change in the pattern brightness (e.g.,

each pixel intensity increased by one may result in totally different binary representation than expected. Such pattern cannot be recognized correctly if the bits are stored separately.

A standard continuous bilevel neuron activation function can be expressed as  $f_s(h) = (1 + e^{-\lambda h})^{-1}$  [3]. This function has two saturation levels,  $-1$  and  $1$ , and one inflection point at  $h = 0$ . A Hopfield network can be generalized by replacing standard bilevel activation functions of neurons with their multilevel counterparts [4]. The  $(n + 1)$ -level neuron characteristics  $f(h, n) = 2/n \sum_{i=1}^n f_s(h - \Theta_i) - 1$  with  $n$  inflection points  $\Theta_i$  have been considered. It has been shown that the computational energy function has  $n$  saddle points and  $(n + 1)$  minima when two  $(n + 1)$ -level neurons are used in a conventional bistable connection. This approach demonstrates that the generalized Hopfield network is able to process multilevel signals, and in particular can smooth gray-scale images.

Several examples of a gray-scale restoration application by such a Hopfield network with multilevel neurons have been presented [5]. According to dynamical properties formulated in terms of the energy function, the multilevel network with two cells behaves as a multistable memory cell and has the ability to remove distortions from noisy gray-scale images.

A systematic approach to investigate the multivalued pattern storage performance in such networks has been presented in [6]. With the conventional Hebb rule the storage capacity for  $K$ -valued uncorrelated patterns has been shown to be proportional to  $K^{-2}$ . The maximum number of stored patterns can be improved by using the pseudoinverse learning rule, however, the basin of attractions for nonerroneous retrieval has been found to shrink with the number  $K$  in that case.

Another approach to multivalued information representation in neural networks is based on amplitude and phase coding. An oscillatory dynamic system with periodic attractors and an associated learning algorithm for storing oscillatory patterns of the form  $r\mathbf{x}^s e^{i\Phi^s} e^{i\omega^s t}$  have been investigated in [7]. In such memory, the initial pattern evolves in time in interactions with all other patterns stored in the network and the final state emerges as a result of the phase locking process.

An oscillatory phasor associative memory with phase only coded binary patterns has also been considered [8]. The proposed model has a form of sparsely connected two-vectors (phasors) interacting through complex-valued couplings. The retrieving ability of the memory for noisy patterns has been

Manuscript received April 27, 1995; revised December 18, 1995.

S. Jankowski is with the Institute of Electronics Fundamentals, Warsaw University of Technology, Nowowiejska 15/19, 00-665 Warsaw, Poland.

A. Lozowski and J. M. Zurada are with the Department of Electrical Engineering, University of Louisville, Louisville, KY 40292 USA.

Publisher Item Identifier S 1045-9227(96)07444-9.

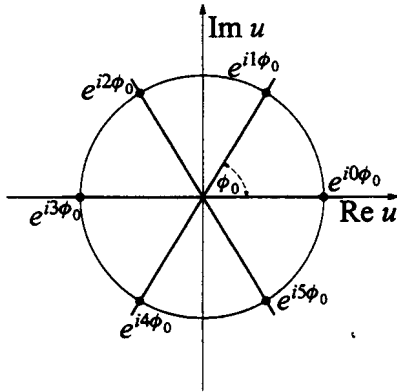


Fig. 1. Complex signum function CSIGN (case shown for  $K = 6$ ).

investigated. By generalization it is expected that attractor neural networks with complex-valued vector states reach the final state in a phase dependent process of excitatory and inhibitory interaction between units.

Multivalued pattern storage can be also obtained by using a continuous-state classifier that performs matching of an input to some "library" patterns. As has been shown in [9] projecting in a signal space between constrain sets corresponding to the patterns can serve as a content-addressable multivalued memory with linear capacity.

The objective of this paper is to introduce a multivalued associative memory model having the architecture of fully connected Hopfield network with neurons possessing a complex-valued discrete activation function. Such memory facilitates natural processing of gray-scale images and has the advantage of mathematical simplicity. The model presented in this paper is based on the concept of multivalued logic [10] and adopts the complex neuron model [11]. The elementary operation used in the multivalued threshold logic is the *complex-signum* operation [10] (see Fig. 1). For a specified number of values  $K$ , called the *resolution factor*, and an arbitrary complex number  $u$  the complex-signum function is defined as follows:

$$\text{CSIGN}(u) \stackrel{\text{df}}{=} \begin{cases} z^0, & 0 \leq \arg(u) < \varphi_0 \\ z^1, & \varphi_0 \leq \arg(u) < 2\varphi_0 \\ \vdots \\ z^{K-1}, & (K-1)\varphi_0 \leq \arg(u) < K\varphi_0 \end{cases} \quad (1)$$

where  $\varphi_0$  is a phase quantum delimited by  $K$ :  $\varphi_0 = 2\pi/K$ , and  $z$  is the corresponding  $K$ th root of one:  $z = e^{i\varphi_0}$ . The result of  $\text{CSIGN}(u)$  belongs to the  $K$ -valued discrete set of complex numbers placed uniformly on the unit circle on the complex plane. Thus, CSIGN generates  $K$  representatives of  $K$  separate sectors of the complex plane and the value of  $\text{CSIGN}(u)$  is one of the complex plane sectors representatives  $e^{ik\varphi_0}$ , such that both  $u$  and  $\text{CSIGN}(u)$  belong to the same sector denoted by  $k\varphi_0 \leq \arg(u) < (k+1)\varphi_0$ . Note that for  $K = 2$  (two-valued threshold logic) and real-valued  $u$  CSIGN returns the real-valued signum since  $\varphi_0 = \pi$ , and  $z = e^{i\pi} = -1$ ; hence  $z^0 = 1, z^1 = -1$ .

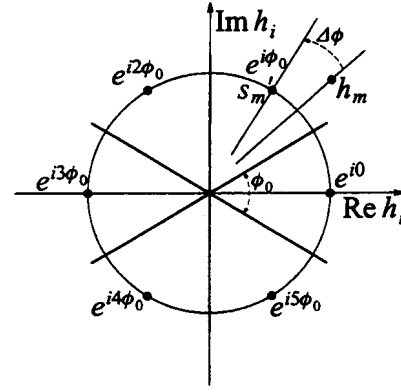


Fig. 2. Complex-valued neuron activation function  $g(h_i) = \text{CSIGN}(h_i e^{i(\varphi_0/2)})$  based on the complex-signum function (case shown for  $K = 6$ ).

## II. COMPLEX-VALUED NEURAL NETWORK

On the basis of the complex-signum function one may postulate a multivalued associative memory model in the following way. Assume a neural network is composed of  $N$  fully connected neurons in a Hopfield-like structure. The output state of each neuron is represented by a complex number from the set  $\{z^0, z^1 \dots z^{K-1}\}$ . Thus the network state  $\vec{s}$  is a complex-valued vector of  $N$  components  $\vec{s} = [s_1, s_2 \dots s_N]$ .

Every two neurons  $i, j$  are coupled via their synaptic connection  $w_{ij}$  represented by a complex number. Similarly to a classic Hopfield network, each input  $h_i$  of the  $i$ th neuron is dependent upon the network state  $\vec{s}$  through synaptic weight  $w_{ij}$

$$h_i = \sum_j w_{ij} s_j. \quad (2)$$

Updating the neuron states as well as the dynamics of transition  $\vec{s} \rightarrow \vec{s}'$  of the network is determined by the following activation function which makes use of the multivalued threshold logic technique:

$$s_i(t+1) = \text{CSIGN}(h_i(t) \cdot z^{1/2}) \quad (3)$$

where  $z^{1/2} = e^{i(\varphi_0/2)}$ . As shown in Fig. 2, the multivalued neuron defined in (3) simultaneously performs phase quantization and magnitude normalization of its input yielding a  $K$ -valued complex output state. As it can be seen from (1), the activation function (3) quantizes the phase of the network input  $h_i$  uniformly since its angle is shifted half a sector in phase. The factor  $z^{1/2}$  places the resulting states in angular centers of each sector. Note that assuming  $K = 2$  the activation function (3) becomes the real-valued signum since the sectors for  $z^0$  and  $z^1$  now become right half-plane and left half-plane, respectively. Although there is no difference in outcome of (1) and (3) on the real axis, (3) produces the phase margin of value  $\varphi_0/2$  for phase noise at the network input.

Consider that a multivalued natural pattern is represented as a vector with entries being  $N$  integers  $\xi = [\xi_1, \xi_2 \dots \xi_N], \xi_i \in \{0 \dots K-1\}$ . To apply the concept of a multivalued associative memory it is necessary to encode the natural pattern  $\xi$  to

its complex representation  $\varepsilon$

$$\xi \rightarrow \varepsilon: \quad \varepsilon_i = z^{\xi_i}, \quad i = 1 \cdots N. \quad (4)$$

Inversely, each state  $\vec{s}$  of the network may be represented by a corresponding natural pattern. This allows for decoding memory states into natural patterns by performing the inverse transformation of (4). The set of complex-valued patterns  $\varepsilon^\mu$ ,  $\mu = 1 \cdots p$ , can be stored in the synaptic matrix according to the modified Hebb rule [1], [11]

$$w_{ij} = \frac{1}{N} \sum_{\mu=1}^p \varepsilon_i^\mu \bar{\varepsilon}_j^\mu \quad i, j = 1 \cdots N. \quad (5)$$

Note that the synaptic matrix created by rule (5) is a Hermitian matrix, i.e.,  $w_{ij} = \bar{w}_{ji}$ .

### III. NETWORK STABILITY

Hopfield's proof of stability of an associative memory has been based on the existence of a Lyapunov-like computational energy function [1]. This function, which possesses lower bounds, does not increase during the network's evolution in time. Thus the dynamics of the network may only converge to an attractor state. Let us postulate the energy function for the complex-valued network to be

$$E(\vec{s}) \stackrel{\text{df}}{=} -\frac{1}{2} \sum_i \sum_j w_{ij} \bar{s}_i s_j. \quad (6)$$

The functional  $E$  is real-valued as long as the synaptic matrix  $[w_{ij}]$  is Hermitian.

Assume that the network has asynchronous dynamics. Each neuron updates its state randomly and independently of the history of activations of other neurons in the network. At a certain instant only one neuron is allowed to change its state by utilizing the dynamic rule (3). Let  $m$  be an index of the neuron to change the state from  $s_m$  to  $s'_m$ . As shown in Fig. 2, both  $s_m$  and  $s'_m$  are discrete complex values of unity magnitude and differing phases. Thus, updating the state of the  $m$ th neuron for asynchronous dynamics is equivalent to a phase-shift with  $k$  quanta  $\varphi_0$  ( $0 \leq k < K$ )

$$s'_i = \begin{cases} z^k s_i, & i = m \\ s_i, & i \neq m. \end{cases} \quad (7)$$

As seen from (3), the output state  $s'_m$  of the  $m$ th neuron at time  $(t+1)$  is determined by the net input  $h_m$  at time  $t$  since  $s'_m = \text{CSIGN}(h_m z^{1/2}) = (h_m/|h_m|)e^{i\Delta\varphi}$ . Here  $h_m/|h_m|$  represents the magnitude normalization while the phase shift  $\Delta\varphi$  represents the phase quantization of the net input. Since both  $h_m$  and  $s'_m$  are located in the same sector, the phase distance  $|\Delta\varphi|$  does not exceed  $\varphi_0/2$ . The relation between  $h_m$  and  $s'_m$  can then be expressed as follows:

$$h_m = |h_m| s'_m e^{-i\Delta\varphi}, \quad |\Delta\varphi| \leq \frac{\varphi_0}{2}. \quad (8)$$

In order to prove the convergence of the multivalued memory dynamics it is sufficient to show that the energy function does not increase for a single update  $\vec{s} \rightarrow \vec{s}'$  since the number of states  $\vec{s}$  that can be reached is finite and limited to  $K^N$ .

*Theorem 1:* Given a complex-valued attractor neural network possessing a Hermitian synaptic weight matrix  $[w_{ij}]$  with nonnegative diagonal entries ( $w_{ii} \geq 0$ ) and asynchronous dynamics  $\vec{s} \rightarrow \vec{s}'$  expressed by (7), the change of the energy function  $\Delta E = E(\vec{s}') - E(\vec{s}) \leq 0$  with equality holding only when  $s' = s''$  if the activation function of the network's units performs uniform phase quantization represented by (3).

A detailed proof of this theorem is provided in the Appendix.

Note that the energy  $E$  is properly defined when only  $w_{ij} = \bar{w}_{ji}$ . Also note that for the case  $K = 2$  this condition implies  $w_{ij} = w_{ji}$ . Therefore, the case of Hermitian interconnections in the complex-valued associative memory generalizes the symmetry condition for the Hopfield model. Also, the multivalued associative memory presented here constitutes a generalization of the classical Hopfield binary-valued network.

### IV. STORAGE CAPACITY ESTIMATION

Similarly to the analysis of a Hopfield network storage capacity [12], storage capacity of a complex-valued associative memory can be based on estimation of probability of an error in the network response. First assume that the network state corresponds to one of the stored patterns:  $\vec{s} = \varepsilon^\lambda$ . Due to (2) and (5) the input  $h_i^\lambda$  of  $i$ th neuron is a sum of two terms

$$h_i^\lambda = \varepsilon_i^\lambda + \frac{1}{N} \sum_j \sum_{\mu \neq \lambda} \varepsilon_i^\mu \bar{\varepsilon}_j^\mu \varepsilon_j^\lambda \quad (9)$$

where the first term  $\varepsilon_i^\lambda$  is the proper response of the network and the second can be regarded as cross-correlation noise. Note that  $h_i^\lambda$  can belong to any of  $K$  sectors. Rotation of  $h_i^\lambda$  to the first sector would be more convenient for analysis. The rotated representation of (9) becomes

$$h_i^\lambda \bar{\varepsilon}_i^\lambda = 1 + \frac{1}{N} \sum_j \sum_{\mu \neq \lambda} \varepsilon_i^\mu \bar{\varepsilon}_j^\mu \bar{\varepsilon}_i^\lambda \varepsilon_j^\lambda \quad (10)$$

where the second term of (10) is a *crosstalk* term  $c_i^\lambda$ , which represents the distance between  $h_i^\lambda \bar{\varepsilon}_i^\lambda$  and one on the complex plane, as shown in Fig. 3. According to (3), if the sum  $1 + c_i^\lambda$  does not belong to the first sector of the complex plane, the state of the  $i$ th neuron after its update changes to an erroneous state. In order to estimate the storage capacity of the network it is necessary to find the probability of such an event by calculating the probability distribution of the crosstalk values  $c_i^\lambda$ .

The crosstalk  $c_i^\lambda$  is a sum of  $N(p-1)$  components  $\varepsilon_i^\mu \bar{\varepsilon}_j^\mu \bar{\varepsilon}_i^\lambda \varepsilon_j^\lambda$ . Assuming that the  $\varepsilon^\mu$  are randomly chosen patterns, we consider the components  $\varepsilon_i^\mu \bar{\varepsilon}_j^\mu \bar{\varepsilon}_i^\lambda \varepsilon_j^\lambda$  as independent random complex  $K$ -valued numbers from the distribution

$$\text{Pr}(\varepsilon_i^\mu \bar{\varepsilon}_j^\mu \bar{\varepsilon}_i^\lambda \varepsilon_j^\lambda = z^k) = \frac{1}{K} \quad k = 0 \cdots K-1 \quad (11)$$

with a mean value of zero and a variance of one. According to the Lindenberg-Levy central limit theorem [15], as  $Np$  tends to infinity, the distribution for the crosstalk  $c_i^\lambda$  becomes a

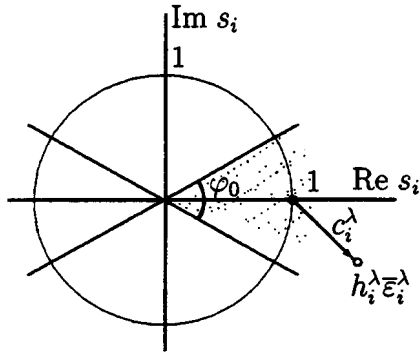


Fig. 3. Crosstalk  $c_i^\lambda$  in a complex-valued associative memory in state  $\bar{s}$  corresponding to one of the stored patterns  $\epsilon^\lambda$ .

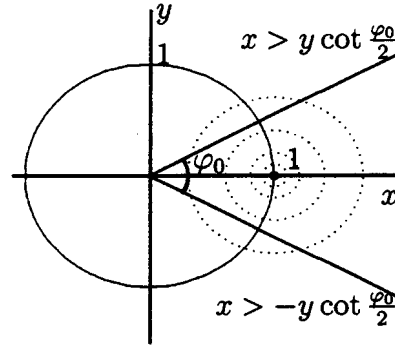


Fig. 4. Probability distribution  $f_{h_i^\lambda \bar{\epsilon}_i^\lambda}(x, y)$ . The distribution is Gaussian with mean  $(1, 0)$  and variance  $p - 1/N$ .

complex Gaussian distribution with mean zero and variance  $\sigma^2 = (p - 1)/N$

$$\lim_{Np \rightarrow \infty} \Pr(\text{Re } c_i^\lambda < c_x, \text{Im } c_i^\lambda < c_y) = \frac{1}{2\pi\sigma^2} \int_{-\infty}^{c_x} \int_{-\infty}^{c_y} e^{-u^2/2\sigma^2} e^{-v^2/2\sigma^2} du dv \quad (12)$$

where  $c_x$  and  $c_y$  are the real and imaginary parts of the variable  $c_i^\lambda$ .

Since  $h_i^\lambda \bar{\epsilon}_i^\lambda = 1 + c_i^\lambda$ , we are able to characterize the distribution for  $h_i^\lambda \bar{\epsilon}_i^\lambda$  by the following two-dimensional (2-D) Gaussian distribution:

$$f_{h_i^\lambda \bar{\epsilon}_i^\lambda}(x, y) = \frac{1}{2\pi\sigma^2} e^{-(x-m_x)^2/2\sigma^2} e^{-(y-m_y)^2/2\sigma^2} \quad (13)$$

where  $x$  and  $y$  are the real and imaginary parts of the variable  $h_i^\lambda \bar{\epsilon}_i^\lambda$  with means  $m_x = 1$  and  $m_y = 0$  (see Fig. 4).

The probability  $P_e$  that the state of the  $i$ th neuron is erroneous is equal to the probability that  $h_i^\lambda \bar{\epsilon}_i^\lambda$  does not fall into the first sector of the plane  $x$ - $y$ . This is equivalent to the fact that the coordinates of  $h_i^\lambda \bar{\epsilon}_i^\lambda$  do not satisfy the condition  $x > |y| \cot(\varphi_0/2)$ . Thus

$$P_e = 1 - \Pr\left(\text{Re } h_i^\lambda \bar{\epsilon}_i^\lambda > |\text{Im } h_i^\lambda \bar{\epsilon}_i^\lambda| \cot \frac{\varphi_0}{2}\right). \quad (14)$$

The probability  $P_e$  may be estimated by integrating the distribution  $f_{h_i^\lambda \bar{\epsilon}_i^\lambda}(x, y)$  in the region corresponding to the first sector of the  $x$ - $y$  plane and using the axial symmetry of the distribution. Thus we obtain

$$P_e = 1 - 2 \int_0^\infty \int_{y \cot \varphi_0/2}^\infty f_{h_i^\lambda \bar{\epsilon}_i^\lambda}(x, y) dx dy. \quad (15)$$

By changing the variables  $\omega = (x - m_x)/\sqrt{2}\sigma$  and  $\tau = (y - m_y)/\sqrt{2}\sigma$ , an explicit formula can be obtained for calculating  $P_e$  for case when  $\varphi_0$  and  $\sigma$  are specified

$$P_e = \frac{1}{2} \left[ 1 + \int_0^\infty \frac{2}{\sqrt{\pi}} e^{-\tau^2} \text{erf} \left( \tau \cot \frac{\varphi_0}{2} - \frac{1}{\sqrt{2}\sigma} \right) d\tau \right]. \quad (16)$$

It should be noted that for the case  $K = 2$ , formula (16) is identical with the one obtained for the classic Hopfield model

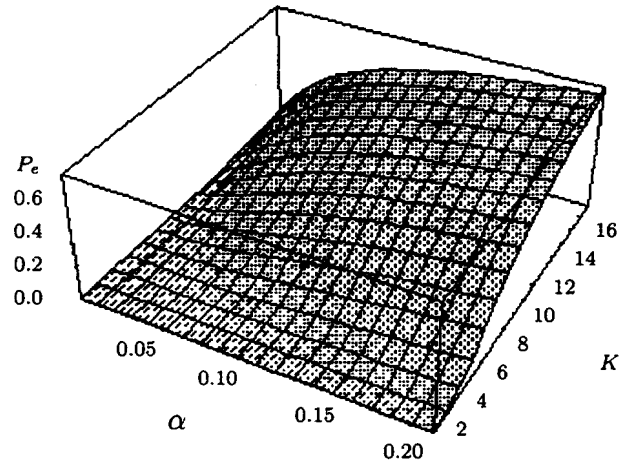


Fig. 5. The probability  $P_e$  of an erroneous recall versus load parameters  $\alpha$  and resolution factors  $K$ .

[12] and takes the form

$$P_e|_{K=2} = \frac{1}{2} \left[ 1 - \text{erf} \left( \frac{1}{\sqrt{2}\sigma} \right) \right]. \quad (17)$$

The results of the numerical calculation of probability  $P_e$  for various resolution factors  $K$  and a range of load parameters  $\alpha$  are shown in Fig. 5. The load parameter  $\alpha$  is related to the variance  $\sigma$  and is defined as follows:

$$\alpha = \frac{p}{N}. \quad (18)$$

Small values for the probability of an erroneous recall are expected for low values of  $\alpha K$ . It is evident that the probability increases either when storing new patterns (increase of  $\alpha$ ) or when shrinking the sectors of the complex plane (increase of  $K$ ).

We may define the storage capacity by choosing a criterion  $P_{\max}$  for an acceptable level of error probability. Thus the storage capacity is the largest load parameter  $\alpha_{\max}$  for which the corresponding error probability still satisfies  $P_e < P_{\max}$ . Fig. 6 shows several  $\alpha - K$  contour plots satisfying the condition  $P_e = P_{\max}$  for various criteria  $P_{\max}$ . It is apparent from the figure that the storage capacity  $\alpha_{\max}$  is dependent upon the number of values  $K$  of the neuron model. As the resolution  $K$  is increased the capacity of the memory

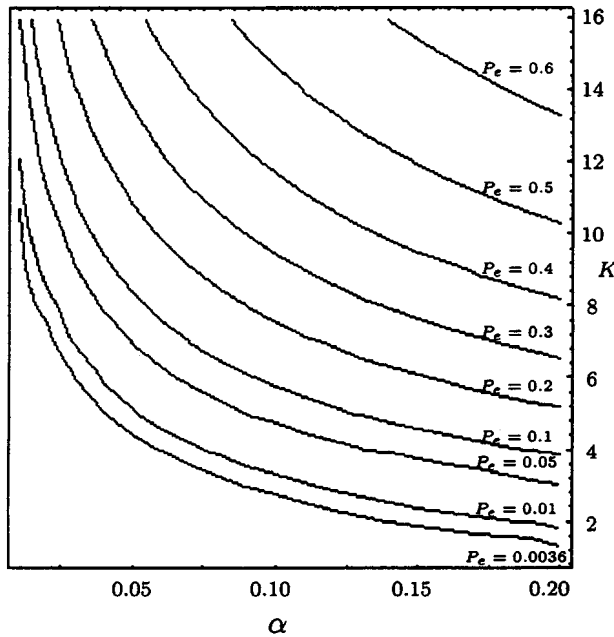


Fig. 6. The  $\alpha - K$  contour plots satisfying  $P_e = P_{max}$  for various  $P_{max}$ .

decreases, as shown in Table I. The storage capacity estimation for  $K = 2$  agrees with the one known for the classic Hopfield model [12].

Note that the used definition of storage capacity is related to the number of stored patterns regardless of the resolution factor  $K$ . Although memory capacity for higher  $K$  reduces, each of the patterns contains more information. Thus, one is capable to distribute same amount of information either in various patterns of limited resolution  $K$  or in a fewer high- $K$ -valued patterns. For a class of information processing tasks, the multivalued encoding of information can be recognized as an advantage of the present approach.

## V. CONCLUSIONS

This paper formulates the novel concept of a complex-valued fully coupled neural network which computes associations based on phase-encoded information. Each neuron can be assigned a multivalued state from the finite set of complex numbers of unity magnitudes and various discrete phases. Since the set of available states possesses a circular symmetry on the complex plane, the states are indistinguishable from each other and none of the states is privileged. Moreover, the memory is able to recognize patterns correctly as long as the relative phase shifts (or gray-level distances) between the patterns components are preserved. In other words, recognition of images is invariant under the absolute level of the pattern brightness. The result of multivalued processing in complex-valued neural networks appears to be more general and yet computationally efficient as compared to real-valued multistate networks.

In a sense, complex-valued calculus generalizes the definitions of Hopfield's bilevel associative memory. Usually, preprocessing of gray-scale (multivalued) patterns is required for attractor neural-network processing since the network

TABLE I  
MULTIVALUED MEMORY CAPACITIES

resolution factors $K$	memory capacities $\alpha_{max}$
2	0.15
3	0.10
4	0.07
6	0.04
8	0.025

operates with bivalent neurons. Here, multivalued information representation corresponds to a natural characteristics of complex-valued neural networks and such additional preprocessing step is eliminated.

The complex multivalued associative memory is particularly well suited to interpret images transformed by 2-D Fourier transformation and the 2-D autocorrelation functions. Even for binary images, both of these algorithms result in multivalued transformations whose storage requires a multivalued representation.

## APPENDIX

### PROOF OF ENERGY FUNCTION MINIMIZATION FOR ASYNCHRONOUS DYNAMICS (SEE THEOREM 1)

The change of the energy function  $\Delta E$  for two consecutive states  $\vec{s}$  and  $\vec{s}'$  can be derived by inspection of energy levels of both states

$$\Delta E = E(\vec{s}') - E(\vec{s}). \quad (19)$$

Since the  $m$ th neuron is the only neuron that changes its state, the value of  $E(\vec{s})$  is expressed, from (6) as follows:

$$E(\vec{s}) = -\frac{1}{2} \left( \sum_{i \neq m} \sum_{j \neq m} w_{ij} \bar{s}_i s_j + \sum_{j \neq m} w_{mj} \bar{s}_m s_j + \sum_{i \neq m} w_{im} \bar{s}_i s_m + w_{mm} \bar{s}_m s_m \right). \quad (20)$$

Equation (7) can be thought of as the relationship between the consecutive states  $\vec{s}$  and  $\vec{s}'$ . Consequently, for some  $k \in \{0, 1, \dots, K-1\}$

$$E(\vec{s}') = -\frac{1}{2} \left( \sum_{i \neq m} \sum_{j \neq m} w_{ij} \bar{s}_i s_j + \sum_{j \neq m} w_{mj} \bar{z}^k \bar{s}_m s_j + \sum_{i \neq m} w_{im} \bar{s}_i z^k s_m + w_{mm} \bar{z}^k \bar{s}_m z^k s_m \right). \quad (21)$$

Comparing (20) and (21) one may obtain the energy change  $\Delta E$

$$\Delta E = -\frac{1}{2} \left( \sum_{j \neq m} (\bar{z}^k - 1) w_{mj} \bar{s}_m s_j + \sum_{i \neq m} (z^k - 1) w_{im} \bar{s}_i s_m \right). \quad (22)$$

Hermitian couplings ensure that  $w_{mj}$  and  $w_{im}$  are mutually conjugate for  $i = j$ . Thus both sums in (22) are composed of mutually conjugate terms  $(\bar{z}^k - 1)w_{mj}\bar{s}_m s_j$  and  $(z^k - 1)w_{im}\bar{s}_i s_m$ . This yields that the energy change is real valued and reads

$$\Delta E = -\text{Re} \left[ \sum_{j \neq m} (\bar{z}^k - 1)w_{mj}\bar{s}_m s_j \right]. \quad (23)$$

After completing the sum in the brackets with the  $m$ th component, the change of energy function becomes

$$\Delta E = -\text{Re} \left[ \sum_j (\bar{z}^k - 1)w_{mj}\bar{s}_m s_j - (z^k - 1)w_{mm}\bar{s}_m s_m \right]. \quad (24)$$

Since  $w_{mm}$  is greater or equal to zero and  $\bar{s}_m s_m = 1$ , the term  $-(z^k - 1)w_{mm}\bar{s}_m s_m$  of (24) is located in the right half-plane. Hence, the following holds:

$$\text{Re} \left[ \sum_j (\bar{z}^k - 1)w_{mj}\bar{s}_m s_j \right] \geq 0 \Rightarrow \Delta E \leq 0. \quad (25)$$

The content of the brackets in (25) is a product of  $(\bar{z}^k - 1)\bar{s}_m$  and  $\sum_j w_{mj}s_j$ . Thus, according to (2) and (7), (25) is equivalent to the following implication:

$$\text{Re}[(1 - z^k)\bar{s}'_m h_m] \geq 0 \Rightarrow \Delta E \leq 0. \quad (26)$$

Taking (8) into account and using  $\bar{s}'_m s'_m = 1$  we obtain

$$\begin{aligned} (1 - z^k)\bar{s}'_m h_m &= |h_m|\bar{s}'_m s'_m (1 - z^k)e^{-i\Delta\varphi} \\ &= |h_m|[e^{-i\Delta\varphi} - e^{i(k\varphi_0 - \Delta\varphi)}]. \end{aligned} \quad (27)$$

Since  $|h_m| \geq 0$ , implication (26) may be rewritten by using Euler's equations for the exponential terms as

$$\cos \Delta\varphi - \cos(k\varphi_0 - \Delta\varphi) \geq 0 \Rightarrow \Delta E \leq 0. \quad (28)$$

For all  $k$  from the set  $\{0, \dots, K - 1\}$  the condition of implication (28) yields that the bound of the phase shift  $\Delta\varphi$

which preserves energy minimization needs to be

$$|\Delta\varphi| \leq \frac{\varphi_0}{2} \Rightarrow \Delta E \leq 0. \quad (29)$$

Note that condition (29) is equivalent to uniform phase-quantization of the neurons activation function ensured in (3). This concludes the proof of Theorem 1.

#### REFERENCES

- [1] J. J. Hopfield, "Neural networks and physical systems with emergent collective computational abilities," in *Proc. Nat. Academy Sci. USA*, vol. 79, 1982, pp. 2554-2558.
- [2] G. A. Kohring, "On the problems of neural networks with multistate neurons," *J. De Physique I*, vol. 2, pp. 1549-1552, Aug. 1992.
- [3] J. J. Hopfield, "Neurons with graded responses have collective computational properties like those of two-state neurons," in *Proc. Nat. Academy Sci. USA*, vol. 81, 1984, pp. 3088-3092.
- [4] J. M. Zurada, I. Cloete, and E. van der Poel, "Generalized Hopfield networks with multiple stable states," *Neurocomputing*, vol. 13, nos. 2-4, Oct. 1997.
- [5] —, "Neural associative memories with multiple stable states," in *Proc. 3rd Int. Conf. Fuzzy Logic, Neural Nets, and Soft Comput.*, Iizuka, Fukuoka, Japan, 1994, pp. 45-51.
- [6] H. Rieger, "Storing an extensive number of gray-toned patterns in a neural network using multistate neurons," *J. Phys. A*, vol. 23, pp. L1273-L1280, 1990.
- [7] B. Baird and F. Eeckman, "A normal form projection algorithm for associative memory," in *Associative Neural Memories—Theory and Implementation*, M. H. Hassoun, Ed. Oxford, U.K.: Oxford Univ. Press, 1993, pp. 135-166.
- [8] A. J. N. Noest, "Associative memory in sparse phasor neural network," *Europhysics Lett.*, vol. 6, pp. 469-474, 1988.
- [9] R. J. Marks II, S. Oh, and L. E. Atlas, "Alternating projection neural networks," *IEEE Trans. Circuit Syst.*, vol. 36, pp. 846-857, 1989.
- [10] N. N. Aizenberg and Y. L. Ivaskiv, *Multivalued Threshold Logic*. Kiev, Ukraine: Naukova Dumka, 1977 (in Russian).
- [11] N. N. Aizenberg and I. N. Aizenberg, "CNN based on multivalued neuron as a model of associative memory for gray-scale images," in *Proc. IEEE 2nd Int. Wkshp. CNN's and Their Applcat. CNNA '92*, Munich, Germany, pp. 36-41.
- [12] J. Hertz, A. Krogh, and R. G. Palmer, *Introduction to the Theory of Neural Computation*. Reading, MA: Addison-Wesley, 1991.
- [13] J. M. Zurada, *Introduction to Artificial Neural Systems*. Boston, MA: PWS, 1992.
- [14] S. Jankowski and A. Lozowski, "Complex-valued neural networks," in *Proc. XVI Nat. Conf., Circuit Theory Electron. Syst.*, Kolobrzeg, Poland, 1993, pp. 582-587.
- [15] W. Feller, *An Introduction to Probability Theory and its Applications*. New York: Wiley, 1960, pp. 228-230.