Research Article

Storage Capacities of Twin-Multistate Quaternion Hopfield Neural Networks

Masaki Kobayashi [ID]

Mathematical Science Center, University of Yamanashi, Takeda 4-3-11, Kofu, Yamanashi 400-8511, Japan

Correspondence should be addressed to Masaki Kobayashi; k-masaki@yamanashi.ac.jp

A twin-multistate quaternion Hopfield neural network (TMQHNN) is a multistate Hopfield model and can store multilevel information, such as image data. Storage capacity is an important problem of Hopfield neural networks. Jankowski et al. approximated the crosstalk terms of complex-valued Hopfield neural networks (CHNNs) by the 2-dimensional normal distributions and evaluated their storage capacities. In this work, we evaluate the storage capacities of TMQHNNs based on their idea.

1. Introduction

A complex-valued Hopfield neural network (CHNN) is a multistate model of Hopfield neural network. CHNNs can deal with multilevel information and have been applied to the storage of image data [1–7]. They have also been extended using Clifford algebra, which includes a complex field, hyperbolic algebra, and quaternion field. Several models of hyperbolic Hopfield neural network have been proposed [8–11]. Isokawa et al. [12] proposed quaternion Hopfield neural networks (QHNNs) employing the split activation function. Minemoto et al. [13] studied QHNNs using the polar-represented activation function. Several other models of QHNNs have also been proposed [14–17]. In this work, we study the twin-multistate quaternion Hopfield neural networks (TMQHNNs) [18]. The neuron of a TMQHNN consists of a pair of complex-valued multistate neurons. The TMQHNN requires only half the connection weight parameters of the CHNN.

Storage capacity is an important issue in Hopfield neural networks. When a Hopfield neural network is given a training pattern, the weighted sum input is decomposed into main and crosstalk terms. The main term enables the Hopfield neural network to memorize the training patterns. The crosstalk term interferes with the storage of training patterns. The storage capacity of conventional Hopfield neural network has been investigated by evaluating the crosstalk term [19]. Jankowski et al. [1] and Kobayashi [20] applied this technique to CHNNs and rotor Hopfield neural networks (RHNNs), respectively. The RHNN is an extension of the CHNN using vectors and matrices [21]. In this work, we provide the Hebbian learning rule for TMQHNNs and evaluate the storage capacity based on Jankowski's concept. In the case of TMQHNNs, the cross talk term is decomposed into two complex parts. By evaluating both parts, we determine the storage capacity of TMQHNNs. The theory suggests that TMQHNNs have half the storage capacities of CHNNs. In addition, we compared the storage capacities of CHNNs and TMQHNNs by computer simulation.

The rest of this paper is organized as follows: Sections 2 and 3 introduce CHNNs and TMQHNNs, respectively. Section 4 provides the Hebbian learning rule for the TMQHNN and evaluates the storage capacity. It also contains descriptions of the computer simulations conducted to verify our analysis. Section 5 concludes this paper.

2. Complex-Valued Hopfield Neural Networks

The CHNNs are briefly described [1]. Let z_a and w_{ab} be the state of neuron a and the connection weight from neuron b to neuron a, respectively. The weighted sum input to neuron a is given by

$$I_a = \sum_{b=1}^{N_C} w_{ab} z_b, \tag{1}$$

where N_C is the number of neurons. For the resolution factor K, we define $\theta_K = \pi/K$. For the weighted sum, input $I = r \exp(i\theta)$, where $r \geq 0$ and $0 \leq \theta < 2\pi$; the complex-valued multistate activation function is defined by

$$f_C(I) = \begin{cases} 1 & 0 \leq \theta < \theta_K, \\ \exp(2i\theta_K) & \theta_K \leq \theta < 3\theta_K, \\ \vdots & \vdots \\ \exp(2(K-1)i\theta_K) & (2K-3)\theta_K \leq \theta < (2K-1)\theta_K, \\ 1 & (2K-1)\theta_K \leq \theta < 2\pi. \end{cases} \tag{2}$$

We define the set of neuron states and denote it as follows: $S = \{1, \exp(2i\theta_K), \ldots, \exp(2(K-1)i\theta_K)\}$. The connection weights must satisfy the following conditions:

$$\overline{w_{ab}} = w_{ba}, \tag{3}$$

$$w_{aa} \geq 0. \tag{4}$$

Then, the CHNN converges to a fixed point.

Let $\mathbf{z}^p = (z_1^p, z_2^p, \ldots, z_{N_C}^p)$ $(1 \leq p \leq P_C)$ be the pth training pattern, where P_C is the number of training patterns. The Hebbian learning rule is defined as

$$w_{ab} = \frac{1}{N_C} \sum_{p=1}^{P_C} z_a^p \overline{z_b^p}. \tag{5}$$

Then, the connection weights satisfy $\overline{w_{ab}} = w_{ba}$. Giving the qth training pattern to the CHNN, the weighted sum input to neuron a is

$$I_a^q = \sum_{b=1}^{N_C} w_{ab} z_b^q \tag{6}$$

$$= z_a^q + \frac{1}{N_C} \sum_{p \neq q} \sum_{b=1}^{N_C} z_a^p \overline{z_b^p} z_b^q. \tag{7}$$

The second term of (7) is referred to as the crosstalk term. The crosstalk term interferes with the storage of training patterns. We define

$$A_a^q = \frac{1}{N_C} \sum_{p \neq q} \sum_{b=1}^{N_C} z_a^p \overline{z_b^p} z_a^q z_b^q. \tag{8}$$

Then, we have

$$I_a^q = z_a^q (1 + A_a^q). \tag{9}$$

If $|\arg(1 + A_a^q)| < \theta_K$, then we have $f_C(I_a^q) = z_a^q$. Therefore, if $|\arg(1 + A_a^q)| < \theta_K$ for all a, the qth training pattern is a fixed point. We regard $N_C A_a^q$ as the summation of $P_C N_C$ random variables of V for simplicity, although the summation consists of exactly $(P-1)N_C$ terms. The real and imaginary parts of each random variable have the equal

variance σ and do not have correlations. Setting $P_C = \alpha N_C$, $N_C A_a^q$ is regarded as the summation of αN_C^2 random variables. Then, we have

$$\frac{1}{\sqrt{\alpha}} A_a^q = \frac{1}{\sqrt{\alpha} N_C} \sum_{p \neq q} \sum_{b=1}^{N_C} z_a^p \overline{z_b^p} z_a^q z_b^q. \tag{10}$$

Let X_A and Y_A be the real and imaginary parts of $(1/\sqrt{\alpha}) A_a^q$, respectively. From the central limit theorem, we have

$$(X_A, Y_A) \sim \frac{1}{2\pi\sigma^2} \exp\left(-\frac{X_A^2 + Y_A^2}{2\sigma^2}\right). \tag{11}$$

For $A_a^q = X_A' + Y_A' i$, we have

$$(X_A', Y_A') \sim \frac{1}{2\alpha\pi\sigma^2} \exp\left(-\frac{X_A'^2 + Y_A'^2}{2\alpha\sigma^2}\right). \tag{12}$$

3. Twin-Multistate Quaternion Hopfield Neural Networks

A quaternion is expressed by $q = q_0 + q_1 i + q_2 j + q_3 k$ using real numbers $q_0, q_1, q_2,$ and q_3. The imaginary units i, j, and k satisfy the following properties:

$$\begin{aligned} i^2 &= j^2 = k^2 = -1, \\ ij &= -ji = k, \\ jk &= -kj = i, \\ ki &= -ik = j. \end{aligned} \tag{13}$$

The quaternions satisfy the associative and distributive laws. For a complex number c, we have the important equality:

$$jc = \overline{c} j. \tag{14}$$

Putting $x = q_0 + q_1 i$ and $y = q_2 + q_3 i$, the quaternion q is described as $q = x + yj$. For quaternions $x + yj$ and $x' + y'j$, the addition and multiplication are described as

$$\begin{aligned} (x + yj) + (x' + y'j) &= (x + x') + (y + y')j, \\ (x + yj)(x' + y'j) &= (xx' - y\overline{y'}) + (xy' + \overline{x'}y)j. \end{aligned} \tag{15}$$

The conjugate of q is defined as

$$\begin{aligned} \overline{q} &= q_0 - q_1 i - q_2 j - q_3 k \\ &= \overline{x} - yj. \end{aligned} \tag{16}$$

Then, we have the equality

$$\begin{aligned} q\overline{q} &= x\overline{x} + y\overline{y} \\ &= |x|^2 + |y|^2. \end{aligned} \tag{17}$$

In the TMQHNNs, the neuron states and connection weights are represented by quaternions. These neuron states and connection weights are denoted in the same way as those of CHNNs. The number of neurons in a TMQHNN is denoted as N_Q. The weighted sum input to neuron a is given by

$$I_a = \sum_{b=1}^{N_Q} w_{ab} z_b. \tag{18}$$

For the weighted sum, input $I = I_x + I_y j$, the activation function is defined as

$$f_Q(I) = f_C(I_x) + f_C(I_y)j. \tag{19}$$

Therefore, the set of neuron states is $S + Sj$. The connection weights must satisfy conditions (3) and (4). Then, the TMQHNN converges to a fixed point.

4. Storage Capacity of Twin-Multistate Quaternion Hopfield Neural Networks

We provide the Hebbian learning rule for TMQHNNs. Let $\mathbf{z}^p = (z_1^p, z_2^p, \ldots, z_{N_Q}^p)$ $(1 \le p \le P_Q)$ be the pth training pattern, where P_Q is the number of training patterns. The Hebbian learning rule is given by

$$w_{ab} = \frac{1}{2N_Q} \sum_{p=1}^{P_Q} z_a^p \overline{z_b^p}. \tag{20}$$

Then, the connection weights satisfy $\overline{w}_{ab} = w_{ba}$. Giving the qth training pattern to the TMQHNN, the weighted sum input to neuron a is

$$I_a^q = \sum_{b \ne a} w_{ab} z_b^q \tag{21}$$

$$= z_a^q + \frac{1}{2N_Q} \sum_{p \ne q} \sum_{b=1}^{N_Q} z_a^p \overline{z_b^p} z_b^q. \tag{22}$$

The second term of (22) is also referred to as the crosstalk term and interferes the storage of training patterns. We decompose the quaternion z_a^p into a pair of complex numbers by $z_a^p = x_a^p + y_a^p j$ to investigate the storage capacity. Then, we have

$$I_a^q = x_a^q + y_a^q j + \frac{1}{2N_Q} \sum_{p \ne q} \sum_{b=1}^{N_Q} (x_a^p + y_a^p j)(x_b^p - y_b^p j)(x_b^q + y_b^q j)$$

$$= x_a^q + \frac{1}{2N_Q} \sum_{p \ne q} \sum_{b=1}^{N_Q} \left(x_a^p x_b^p x_b^q + x_a^p y_b^p \overline{y_b^q} - y_a^p \overline{x_b^p y_b^q} + y_a^p \overline{y_b^p} x_b^q \right)$$

$$+ y_a^q j + \frac{1}{2N_Q} \sum_{p \ne q} \sum_{b=1}^{N_Q} \left(x_a^p x_b^p y_b^q - x_a^p y_b^p \overline{x_b^q} - y_a^p \overline{x_b^p x_b^q} + y_a^p \overline{y_b^p} y_b^q \right) j. \tag{23}$$

We define

$$B_a^q = \frac{1}{2N_Q} \sum_{p \ne q} \sum_{b=1}^{N_Q} \left(x_a^p x_b^p x_b^q + x_a^p y_b^p \overline{y_b^q} - y_a^p \overline{x_b^p y_b^q} + y_a^p \overline{y_b^p} x_b^q \right) \overline{x_a^q}, \tag{24}$$

$$C_a^q = \frac{1}{2N_Q} \sum_{p \ne q} \sum_{b=1}^{N_Q} \left(x_a^p x_b^p y_b^q - x_a^p y_b^p \overline{x_b^q} - y_a^p \overline{x_b^p x_b^q} + y_a^p \overline{y_b^p} y_b^q \right) \overline{y_a^q}. \tag{24}$$

Then, we have

$$I_a^q = x_a^q (1 + B_a^q) + y_a^q (1 + C_a^q) j. \tag{25}$$

If $|\arg(1 + B_a^q)| < \theta_K$ and $|\arg(1 + C_a^q)| < \theta_K$, then we have $f_Q(I_a^q) = z_a^q$. We regard $2N_Q B_a^q$ and $2N_Q C_a^q$ as the summations of $4N_Q P_Q$ random variables of S. Then, B_a^q and C_a^q follow the same distributions. Thus, we can discuss only B_a^q. If the TMQHNN is used instead of the CHNN, $N_C = 2N_Q$ is required, since a twin-multistate quaternion neuron consists of two complex-valued multistate neurons. Setting $P_Q = \beta N_C = 2\beta N_Q$, $2N_Q B_a^q$ is regarded as the summation of $8\beta N_Q^2$ random variables, and we have

$$\frac{1}{\sqrt{2\beta}} B_a^q = \frac{1}{2\sqrt{2\beta} N_Q} \sum_{p \ne q} \sum_{b=1}^{N_Q} \left(x_a^p x_b^p x_b^q + x_a^p y_b^p \overline{y_b^q} - y_a^p \overline{x_b^p y_b^q} \right.$$

$$\left. + y_a^p \overline{y_b^p} x_b^q \right) \overline{x_a^q}. \tag{26}$$

Let X_B and Y_B be the real and imaginary parts of $(1/2\sqrt{\beta}) B_a^q$, respectively. From the central limit theorem, we have

$$(X_B, Y_B) \sim \frac{1}{2\pi\sigma^2} \exp\left(-\frac{X_B^2 + Y_B^2}{2\sigma^2} \right). \tag{27}$$

Putting $B_a^q = X_B' + Y_B' i$, we have

$$(X_B', Y_B') \sim \frac{1}{4\beta\pi\sigma^2} \exp\left(-\frac{X_B'^2 + Y_B'^2}{4\beta\sigma^2} \right). \tag{28}$$

We require the same distributions for (12) and (22) and obtained $\alpha = 2\beta$. Thus, the CHNN has double the storage capacity of the TMQHNN.

Computer simulations were conducted to verify our analysis. K was varied from 4–12 in steps of 2, and P was varied from 1–20. For each K and P, 100 sets of training patterns were generated randomly; the number of trials was 100. The CHNN and TMQHNN attempted to store the training patterns by the Hebbian learning rule. If all the training patterns were fixed, the trial was regarded as successful, otherwise, as failed. Figure 1 shows the simulation results. The horizontal and vertical axes indicate the number of training patterns and success rate, respectively. The simulation results showed that the storage capacity of the TMQHNN was a bit larger than half that of the CHNN.

5. Conclusions

A TMQHNN needs only half the connection weight parameters of CHNN and is expected to have a smaller storage capacity. However, the storage capacity had not yet been analyzed. In this work, we defined the Hebbian learning rule for TMQHNNs and analyzed their storage capacity. The analysis demonstrated that a TMQHNN had half of the storage capacity of a CHNN. In addition, a computer simulation was conducted to verify our analysis. The simulation results confirmed our analysis. In future, we intend
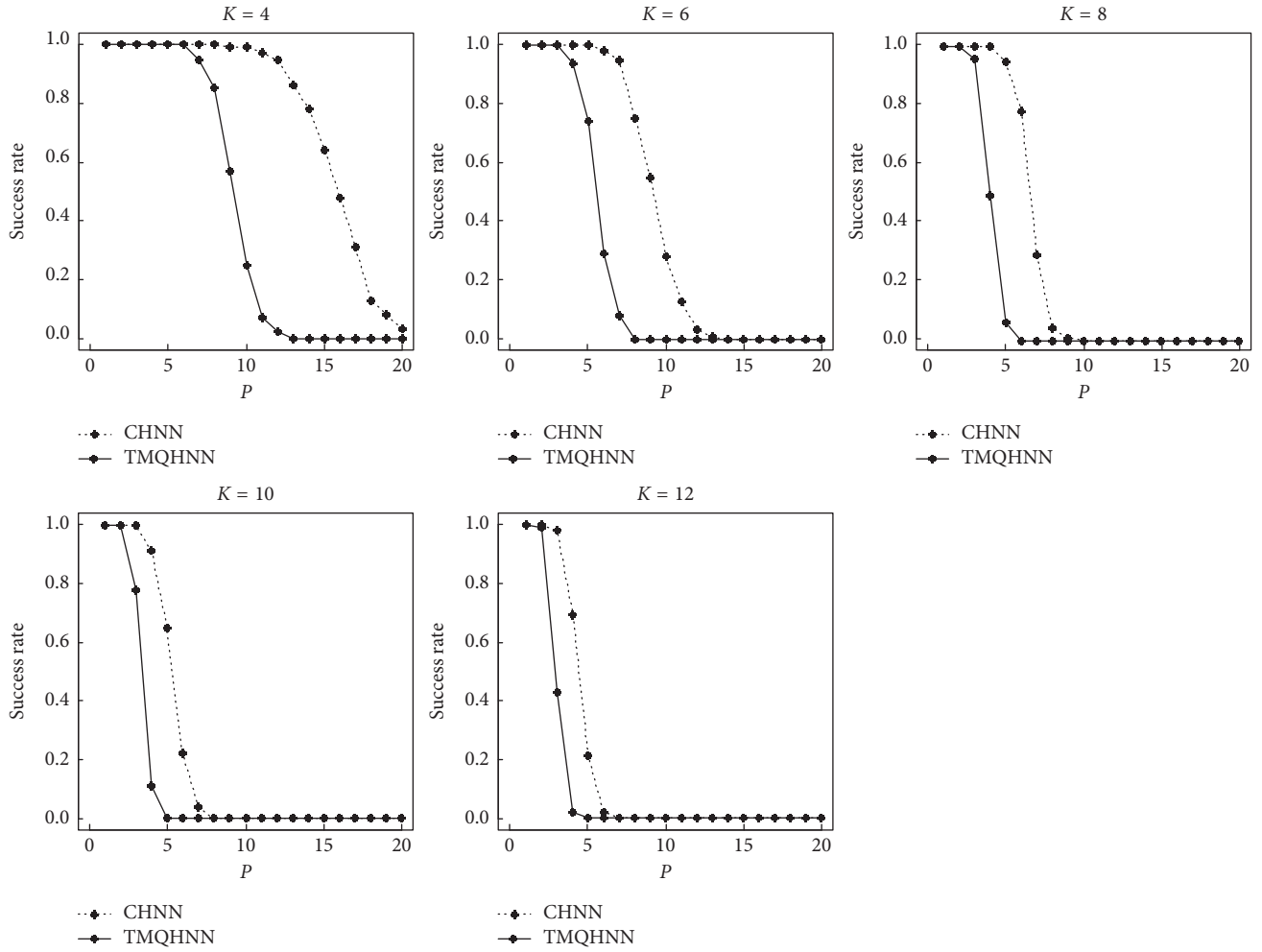
Figure 1: The simulation result for storage capacities.

to study the storage capacity using different methods [19, 22, 23].

Data Availability

No data were used to support this study.

Conflicts of Interest

The author declares that there are no conflicts of interest regarding the publication of this article.

References

[1] S. Jankowski, A. Lozowski, and J. M. Zurada, "Complex-valued multistate neural associative memory," *IEEE Transactions on Neural Networks*, vol. 7, no. 6, pp. 1491–1496, 1996.

[2] H. Aoki and Y. Kosugi, "An image storage system using complex-valued associative memories," in *Proceedings of International Conference on Pattern Recognition*, vol. 2, pp. 626–629, Barcelona, Spain, September 2000.

[3] H. Aoki, "A complex-valued neuron to transform gray level images to phase information," in *Proceedings of the International Conference on Neural Information Processing*, vol. 3, pp. 1084–1088, Singapore, November 2002.

[4] G. Tanaka and K. Aihara, "Complex-valued multistate associative memory with nonlinear multilevel functions for gray-level image reconstruction," *IEEE Transactions on Neural Networks*, vol. 20, no. 9, pp. 1463–1473, 2009.

[5] M. K. Muezzinoglu, C. Guzelis, and J. M. Zurada, "A new design method for the complex-valued multistate Hopfield associative memory," *IEEE Transactions on Neural Networks*, vol. 14, no. 4, pp. 891–899, 2003.

[6] P. Zheng, "Threshold complex-valued neural associative memory," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 25, no. 9, pp. 1714–1718, 2014.

[7] T. Isokawa, H. Yamamoto, H. Nishimura, T. Yumoto, N. Kamiura, and N. Matsui, "Complex-valued associative memories with projection and iterative learning rules," *Journal of Artificial Intelligence and Soft Computing Research*, vol. 8, no. 3, pp. 237–249, 2018.

[8] Y. Kuroe, S. Tanigawa, and H. Iima, "Models of Hopfield-type Clifford neural networks and their energy functions-hyperbolic and dual valued networks," in *Proceedings of International Conference on Neural Information Processing*, pp. 560–569, Doha, Qatar, November 2011.

[9] M. Kobayashi, "Hyperbolic Hopfield neural networks," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 24, no. 2, pp. 335–341, 2013.

[10] M. Kobayashi, "Global hyperbolic Hopfield neural networks," *IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences*, vol. E99-A, no. 12, pp. 2511–2516, 2016.

[11] M. Kobayashi, "Hyperbolic Hopfield neural networks with directional multistate activation function," *Neurocomputing*, vol. 275, pp. 2217–2226, 2018.

[12] T. Isokawa, H. Nishimura, N. Kamiura, and N. Matsui, "Associative memory in quaternionic Hopfield neural network," *International Journal of Neural Systems*, vol. 18, no. 2, pp. 135–145, 2008.

[13] T. Minemoto, T. Isokawa, H. Nishimura, and N. Matsui, "Quaternionic multistate Hopfield neural network with extended projection rule," *Artificial Life and Robotics*, vol. 21, no. 1, pp. 106–111, 2016.

[14] T. Minemoto, T. Isokawa, H. Nishimura, and N. Matsui, "Pseudo-orthogonalization of memory patterns for complex-valued and quaternionic associative memories," *Journal of Artificial Intelligence and Soft Computing Research*, vol. 7, no. 4, pp. 257–264, 2017.

[15] M. Kobayashi, "Hybrid quaternionic Hopfield neural network," *IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences*, vol. E98-A, no. 7, pp. 1512–1518, 2015.

[16] M. Kobayashi, "Symmetric quaternionic Hopfield neural networks," *Neurocomputing*, vol. 240, pp. 110–114, 2017.

[17] M. Kobayashi, "Three-dimensional quaternionic Hopfield neural networks," *IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences*, vol. E100-A, no. 7, pp. 1575–1577, 2017.

[18] M. Kobayashi, "Quaternionic Hopfield neural networks with twin-multistate activation function," *Neurocomputing*, vol. 267, pp. 304–310, 2017.

[19] J. Hertz, A. Krogh, and R. G. Palmer, *Introduction to the Theory of Neural Computation*, Addison-Weasley, Redwood City, CA, USA, 1991.

[20] M. Kobayashi, "Storage capacity of rotor Hopfield neural networks," *Neurocomputing*, vol. 316, pp. 30–33, 2018.

[21] M. Kitahara and M. Kobayashi, "Projection rule for rotor Hopfield neural networks," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 25, no. 7, pp. 1298–1307, 2014.

[22] J. Cook, "The mean-field theory of a Q-state neural network model," *Journal of Physics A: Mathematical and General*, vol. 22, no. 12, pp. 2057–2067, 1989.

[23] Y. Nakamura, K. Torii, and T. Munakata, "Neural-network model composed of multidimensional spin neurons," *Physical Review E*, vol. 51, no. 2, pp. 1538–1546, 1995.