Hypercomplex-valued Neural Networks

Part 5 – Hypercomplex-Valued Hopfield Neural Networks

POLISH NATIONAL AGENCY
FOR ACADEMIC EXCHANGE

Marcos Eduardo Valle
Universidade Estadual de Campinas (Unicamp)
Campinas, Brazil.

UNICAMP

Introduction

The Hopfield network is a recurrent neural network model that played an essential role in the 1980s and the subsequent developments in neural networks.

Applications of the Hopfield network include implementing associative memories and solving optimization problems, among others (Hassoun, 1995).

Despite the advancement of deep learning, Hopfield networks remain an active field of research (Krotov and Hopfield, 2020; Liang et al., 2022; Ramsauer et al., 2020).

Research on Hopfield networks with hypercomplex values also remains an important research topic with significant contributions in the last years (Kobayashi, 2020a,b).

In the following, we briefly review the mathematical aspects of the Hopfield network.

Hypercomplex-valued versions of the Hopfield networks are presented subsequently de Castro and Valle (2020).

Discrete-Time Hopfield Neural Network

The Hopfield network is a recurrent neural network with McCulloch and Pitts neurons (Hopfield, 1982).

Consider a network with N fully connected neurons. Let w_{ij} and θ_i be the jth synaptic connection and the threshold of the i-th neuron.

Let $x_i(t) \in \{-1, +1\}$ be the state of the i-th neuron in discrete time t.

The dynamics of the Hopfield network is described by the equation

$$x_i(t+1) = \begin{cases} \operatorname{sgn}(h_i(t)), & h_i(t) \neq 0, \\ x_i(t), & \text{otherwise}, \end{cases} \tag{1}$$

where

$$h_i(t) = \sum_{j=1}^{N} w_{ij} x_j(t) - \theta_i, \quad \forall i = 1, \ldots, N. \tag{2}$$

Update and Convergence

Hopfield network neurons can be updated synchronously (parallel) or asynchronously (sequential).

Theorem 1 (Convergence (Hopfield, 1982))

The Hopfield network, operating asynchronously, produces a convergent sequence if the synaptic weights satisfy $w_{ii} = 0$ and $w_{ji} = w_{ij}$ for all $i, j = 1, \ldots, N$.

The theorem is proved by showing that $E : \{-1, +1\}^N \to \mathbb{R}$ given by

$$E(\mathbf{x}) = -\frac{1}{2}\mathbf{x}^T W \mathbf{x} + \theta^T \mathbf{x}. \tag{3}$$

is an energy or Lyapunov function, that is, E is bounded from below and $\mathbf{x}(t+1) \neq \mathbf{x}(t)$ implies $E(\mathbf{x}(t+1)) < E(\mathbf{x}(t))$.

The Hopfield network can implement associative memories or solve optimization problems!

Complex-valued Hopfield Network

Jankowski et al. (1996) presented a complex version of the Hopfield network using the complex-valued signum function.

The complex-valued signum function is defined as follows: Given a positive integer $K > 1$, called *resolving factor*, set $\Delta\theta = \pi/K$ and

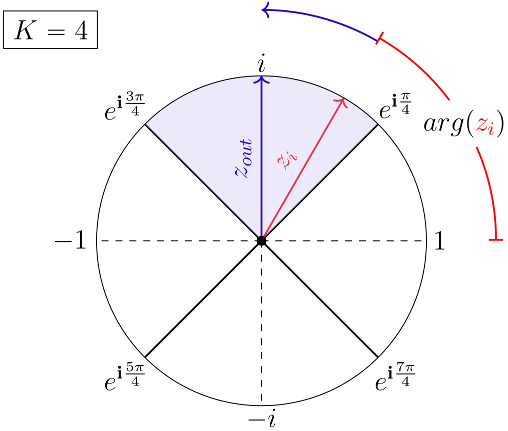
$$\mathcal{D} = \{z \in \mathbb{C} \setminus \{0\} : \arg(z) \neq (2k-1)\Delta\theta, \forall k = 1,\ ldots, K\}.$$

The complex-valued signum function $\text{csgn} : \mathcal{D} \to \mathbb{S} = \{1, 2^{2\boldsymbol{i}\Delta\theta}, \ldots, e^{2(K-1)\boldsymbol{i}\Delta\theta}\}$ is

$$\text{csgn}(z) = \begin{cases} 1, & 0 \leq \arg(z) < \Delta\theta, \\ e^{2\boldsymbol{i}\Delta\theta}, & \Delta\theta < \arg(z) < 3\Delta\theta, \\ \vdots & \vdots \\ 1, & (2K-1)\Delta\theta < \arg(z) < 2\pi. \end{cases}$$

Complex-valued signum function

Geometrical interpretation of the complex-valued signum function:



Source: Rodolfo Lobo (2021).

Dynamics of Complex-Valued Hopfield Network

The dynamics of the complex-valued Hopfield network is given by

$$z_i(t+1) = \begin{cases} \text{csgn}(h_i(t)), & h_i(t) \in \mathcal{D}, \\ z_i(t), & \text{otherwise}, \end{cases} \tag{4}$$

where

$$h_i(t) = \sum_{j \neq i}^{N} w_{ij} z_j(t), \quad \forall i = 1, \ldots, N. \tag{5}$$

Theorem 2 (Convergence)

The complex-valued Hopfield network, operating asynchronously, produces a convergent sequence if the synaptic weights satisfy $w_{ii} \geq 0$ *and* $w_{ji} = \bar{w}_{ij}$ *for all* $i, j = 1, \ldots, N$.

Hyperbolic-valued Hopfield Network

Inspired by the complex-valued network, Kobayashi (2018b, 2019, 2020b) presented a version of the Hopfield network with hyperbolic numbers.

The dynamics of the hyperbolic-valued Hopfield network is given by

$$z_i(t+1) = \begin{cases} \text{csgn}(h_i(t)), & h_i(t) \in \mathcal{D}, \\ z_i(t), & \text{otherwise}, \end{cases} \tag{6}$$

where

$$h_i(t) = \sum_{j=1}^{N} w_{ij} z_j(t).$$

Theorem 3 (Convergence)

The hyperbolic-valued Hopfield network, operating asynchronously, produces a convergent sequence if the synaptic weights satisfy $Re(w_{ii}) \geq |Im(w_{ii})|$ *and* $w_{ji} = w_{ij}$.

The energy function of the hyperbolic-valued Hopfield network is

$$E(\mathbf{z}) = -\frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N}\text{Re}(z_i w_{ij} z_j).$$

Computational experiments showed that the hyperbolic-valued Hopfield network has better noise tolerance than the Hopfield network with complex values (Kobayashi, 2020b).

Quaternion-Valued Hopfield Networks

Quaternion-valued versions of the Hopfield network have also been developed in recent decades (Isokawa et al., 2013; Kobayashi, 2017; Valle, 2014).

Like complex and hyperbolic networks, the quaternion-valued Hopfield network is given by

$$q_i(t+1) = \begin{cases} \varphi(h_i(t)), & h_i(t) \in \mathcal{D}, \\ q_i(t), & \text{otherwise}, \end{cases} \tag{7}$$

where $h_i(t) = \sum_{j \neq i}^{N} w_{ij} q_j(t)$ and $\varphi : \mathcal{D} \to \mathcal{S}$ is an appropriate activation function.

Continuous-valued Activation Function

The continuous-valued activation function $\sigma : \mathcal{D} \rightarrow \mathcal{S}$ is given by

$$\sigma(q) = \frac{q}{|q|},$$

on what

$$\mathcal{D} = \mathbb{Q} \setminus \{0\} \quad \text{and} \quad \mathcal{S} = \{z \in \mathbb{Q} : |q| = 1\}.$$

This activation function appears as an alternative to the multi-state quaternionic activation function developed by Isokawa et al. (2013).

Twin-Multistate Activation Function

The twin-multistate activation function, introduced by Kobayashi (Kobayashi, 2017), is defined using the complex-valued signum function and the relation $\boldsymbol{i}\boldsymbol{j} = \boldsymbol{k}$.

Specifically, a quaternion can be written as:

$$q = q_0 + q_1\boldsymbol{i} + q_2\boldsymbol{j} + q_3\boldsymbol{k} = (q_0 + q_1\boldsymbol{i}) + (q_2 + q_3\boldsymbol{i})\boldsymbol{j} = z_0 + z_1\boldsymbol{j},$$

where z_0 and z_1 are complex numbers.

Using the complex-valued signum function, the twin-multistate activation function is defined by

$$\mathrm{tsgn}(q) = \mathrm{csgn}(z_0) + \mathrm{csgn}(z_1)\boldsymbol{j}.$$

Convergence of Quaternion-valued Hopfield Nets

Like the complex-valued Hopfield network, quaternionic Hopfield networks satisfy the following theorem:

Theorem 4 (Convergence)

The quaternion-valued Hopfield network, with the continuous-valued function or the twin complex-valued signum function, operating asynchronously, produces a convergent sequence if the synaptic weights satisfy $w_{ii} \geq 0$ *and* $w_{ji} = \bar{w}_{ij}$ *for all* $i, j = 1, \ldots, N$.

Besides the quaternion-valued models, Hopfield networks based on commutative quaternions (Isokawa et al., 2010) and the Klein four-group (Kobayashi, 2020a) have also been proposed and investigated.

Hypercomplex-valued Hopfield Networks

Motivated by Hopfield networks' developments, we proposed a broad framework for hypercomplex-valued models (de Castro and Valle, 2020)

We begin by generalizing the conjugate operation:

Definition 5 (Reverse Involution)

Let $\mathbb{H}$ be a hypercomplex algebra. An operator $\eta : \mathbb{H} \to \mathbb{H}$ is a reverse involution if the following holds for all $x, y \in \mathbb{H}$ and $\alpha \in \mathbb{R}$:

$$\eta\big(\eta(x)\big) = x, \qquad \text{(involution)}$$

$$\eta(xy) = \eta(y)\eta(x), \qquad \text{(anti-homomorphism)}$$

$$\eta(\alpha x + y) = \alpha\eta(x) + \eta(y). \qquad \text{(linear)}$$

Example 6 (Natural Conjugation)

The natural conjugation of a hypercomplex number $x = x_0 + x_1 \boldsymbol{i}_1 + \cdots + x_n \boldsymbol{i}_n$, denoted by $\bar{x}$ and defined by

$$\bar{x} = x_0 - x_1 \boldsymbol{i}_1 - \ldots - x_n \boldsymbol{i}_n,$$

is an example of reverse involution in hypercomplex algebras, such as complex numbers and quaternions.

Example 7 (Trivial Reverse Involution)

The identity $\eta(x) = x$ is a reverse involution if the multiplication is commutative. In this case, it is called trivial reverse involution.

Symmetric Bilinear Form and $\mathcal{B}$-Function

A reverse involution is used for defining the symmetric bilinear form $\mathcal{B} : \mathbb{H} \times \mathbb{H} \to \mathbb{R}$ given by the equation

$$\mathcal{B}(x, y) = \text{Re}\left\{\eta(x)y\right\}, \quad \forall x, y \in \mathbb{H}.$$

Note that $\mathcal{B}$ establishes a relationship between x and y, taking into account the algebraic properties of multiplication and the reverse involution η.

Definition 8 ($\mathcal{B}$-Function)

A hypercomplex-valued function $\phi : \mathcal{D} \subset \mathbb{H} \to \mathcal{S} \subset \mathbb{H}$ is a $\mathcal{B}$-function if

$$\mathcal{B}(\phi(x), x) > \mathcal{B}(s, x), \quad \forall x \in D, \forall s \in \mathcal{S} \setminus \left\{\phi(x)\right\}.$$

Real-part Associative Hypercomplex Algebras

Definition 9 (Real-part Associative Hypercomplex Algebras)

A hypercomplex algebra $\mathbb{H}$ equipped with a reverse involution η is a *real-part associative hypercomplex algebra* (Re-AHA) if the following identity holds for all x, y, z:

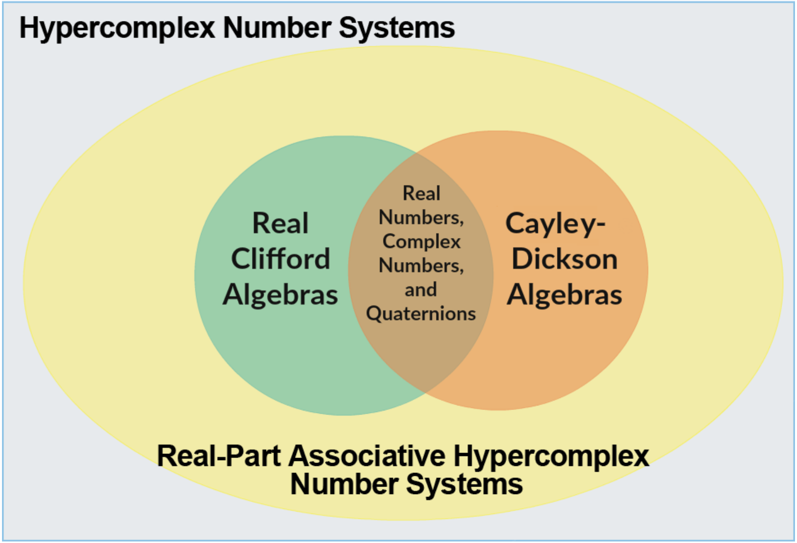
$$\text{Re}\left\{(xy)z - x(yz)\right\} = 0.$$

In particular, we say that a Re-AHA is positive semi-definite (or non-negative definite) if the symmetric bilinear form $\mathcal{B}$ satisfies

$$\mathcal{B}(x, x) \geq 0, \quad \forall x \in \mathbb{H}.$$

The Cayley-Dickson algebras, with natural conjugation, are examples of Re-AHAs.

Visual interpretation of hypercomplex algebras:



Source: Castro and Valle (2020).

Hypercomplex-Valued Hopfield Neural Networks

Let $\mathbb{H}$ be a Re-AHA and $\phi : \mathcal{D} \to \mathcal{S}$ be a $\mathcal{B}$-function with $\mathcal{S}$ a compact subset of $\mathbb{H}$.

Let $w_{ij} \in \mathbb{H}$ be the jth hypercomplex synaptic weight of the ith neuron of a Hopfield network with hypercomplex values with N neurons.

The state of the Hopfield network at time t is represented by a vector

$$\mathbf{x}(t) = [x_1(t), \ldots, x_N(t)]^T \in \mathcal{S}^N,$$

where

$$x_i(t) = x_{i0}(t) + x_{i1}(t)\boldsymbol{i}_1 + \ldots + x_{in}(t)\boldsymbol{i}_n,$$

corresponds to the state of the ith neuron.

Given an initial state (input vector)

$$\mathbf{x}(0) = [x_1(0), \ldots, x_N(0)]^T \in \mathcal{S}^N,$$

a hypercomplex-valued Hopfield network (HvHN) defines the sequence $\{\mathbf{x}(t)\}_{t \geq 0}$ by the equation

$$x_i(t + \Delta t) = \begin{cases} \phi\big(h_i(t)\big), & h_i(t) \in \mathcal{D}, \\ x_i(t), & \text{otherwise}, \end{cases}$$

where

$$h_i(t) = \sum_{j=1}^{N} w_{ij} x_j(t),$$

is the activation potential of the *i*th neuron at time *t*.

Neurons are usually updated either synchronously or asynchronously.

Dynamics of the HvHNs

Theorem 10 (Dynamics of the HvHNs)

The sequence of states produced by a hypercomplex-valued Hopfield network is convergent in asynchronous mode if the synaptic weights satisfy

$$w_{ij} = \eta(w_{ji})$$

and one of the following cases occurs:

- $w_{ii} = 0$ *for all* $i \in \{1, \ldots, N\}$.
- w_{ii} *is a non-negative real number for all* $i \in \{1, \ldots, N\}$ *and* $\mathbb{H}$ *is a non-negative definite Re-AHA.*

The proof of this theorem is given by de Castro and Valle (2020).

Briefly, the proof follows by showing that

$$E(\mathbf{x}) = -\frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N} \text{Re}\left\{\eta(x_i)(w_{ij}x_j)\right\} = -\frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N} \mathcal{B}(x_i, w_{ij}x_j),$$

is an energy function of the hypercomplex-valued Hopfield network.

In other words, $E : \mathcal{S}^N \to \mathbb{R}$ is bounded and decreasing along any non-stationary trajectory, i.e, the inequality

$$\Delta E = E(\mathbf{x}(t + \Delta t)) - E(\mathbf{x}(t)) < 0,$$

holds whenever $\mathbf{x}(t + \Delta t) \neq \mathbf{x}(t)$.

Example - Complex-Valued Multistate Hopfield Nets

Consider the complex numbers $\mathbb{C}$ with the natural conjugation.

Note that

$$\mathcal{B}(z_1, z_2) = |z_1||z_2| \cos |\theta_1 - \theta_2|,$$

where $z_1 = |z_1|e^{\boldsymbol{i}\theta_1}$ and $z_2 = |z_2|e^{\boldsymbol{i}\theta_2}$ are such that $0 \leq |\theta_1 - \theta_2| \leq \pi$.

In particular, $\mathcal{B}(z, z) = |z|^2 \geq 0$. Thus, $\mathbb{C}$ with the natural conjugation is a positive semi-definite Re-AHN.

Furthermore, the complex-valued signum function $\text{csgn} : \mathcal{D} \to \mathcal{S}$ is a $\mathcal{B}$-function.

The complex-valued multistate Hopfield neural network with $f \equiv \text{csgn}$ yields a convergent sequence if $w_{ij} = \bar{w}_{ji}$ and $w_{ii} \geq 0$.

Tessarine-Valued Hopfield Network

The tessarines $\mathbb{T}$, introduced in 1849 by James Cockle and later by Corrado Segre as the bicomplex numbers (Cerroni, 2017), is a commutative four-dimension hypercomplex algebra characterized by the multiplication table

$\mathbb{T}$	$\boldsymbol{i}$	$\boldsymbol{j}$	$\boldsymbol{k}$
$\boldsymbol{i}$	-1	$\boldsymbol{k}$	$-\boldsymbol{j}$
$\boldsymbol{j}$	$\boldsymbol{k}$	1	$\boldsymbol{i}$
$\boldsymbol{k}$	$-\boldsymbol{j}$	$\boldsymbol{i}$	-1

Consider the tessarines $\mathbb{T}$ with the reverse-involution

$$\eta(x) = x_0 - x_1\boldsymbol{i} + x_2\boldsymbol{j} - x_3\boldsymbol{k}, \quad \forall x = x_0 + x_1\boldsymbol{i} + x_2\boldsymbol{j} + x_3\boldsymbol{k}.$$

In this case, $\mathbb{T}$ is a positive semi-definite Re-AHA and the symmetric bilinear form is

$$\mathcal{B}_{\mathbb{T}}(x, y) = x_0 y_0 + x_1 y_1 + x_2 y_2 + x_3 y_3, \quad \forall x, y \in \mathbb{T}.$$

A tessarine $x = x_0 + x_1 \boldsymbol{i}_1 + \cdots + x_n \boldsymbol{i}_n$ can also be written as

$$x = (x_0 + x_1 \boldsymbol{i}) + (x_2 + x_3 \boldsymbol{i})\boldsymbol{j} = z_0 + z_1 \boldsymbol{j},$$

where $z_0 = x_0 + x_1 \boldsymbol{i} \in \mathbb{C}$ and $z_1 = x_2 + x_3 \boldsymbol{i} \in \mathbb{C}$.

Using such identity, we can show that the twin-multistate activation function is a $\mathcal{B}$-function in this algebra.

Thus, the tessarine-valued multistate Hopfield neural network yields a convergent sequence in an asynchronous update mode if $w_{ij} = \eta(w_{ji})$ and $w_{ii} \geq 0$ for all $i, j = 1, \ldots, N$.

The theory presented in this paper provides a unified mathematical explanation for the stability analysis of many hypercomplex-valued neural networks such as the one detailed by Isokawa et al. (2010); Kobayashi (2018a).

More examples are given in de Castro and Valle (2020).

Concluding Remarks

In this talk, we presented a general framework for hypercomplex-valued Hopfield networks.

These models, operating in asynchronous mode, produce a convergent sequence of states under certain conditions in the hypercomplex algebra and synaptic weights.

This result is significant for implementing associative memories using hypercomplex-valued Hopfield networks. It is also necessary for solving optimization problems.

Future Research

Besides the discrete-time model, the real-valued continuous-time Hopfield neural network is given by (Hassoun and Watta, 1997):

$$C\frac{du_i}{dt} = \alpha_i u_i + \sum_{j=1}^{n} w_{ij} x_j - \theta_i, \quad \text{with} \quad x_i = \phi(u_i), \forall i = 1, \ldots, N. \quad (8)$$

The continuous-time Hopfield network always settles at equilibrium if $w_{ij} = w_{ji}$ and the function ϕ is smooth and monotonically increasing (Hopfield, 1984).

Can we extend this result for a broad class of hypercomplex algebras?

Thanks for your attention!

Acknowledge

These slides are part of a mini-course given during the workshop on **hypercomplex-valued neural networks**, which took place at the **Institute for Research and Applications of Fuzzy Modeling**, **University of Ostrava**, Ostrava, Checz Republic, *06-10 February 2023*, with the support of

POLISH NATIONAL AGENCY FOR ACADEMIC EXCHANGE

My research on hypercomplex-valued neural networks has been partially supported by:

FAPESP and CNPq Conselho Nacional de Desenvolvimento Científico e Tecnológico

References (1)

C. Cerroni. From the theory of “congeneric surd equations” to “Segre’s bicomplex numbers”. *Historia Mathematica*, 44(3): 232–251, 8 2017. doi: 10.1016/j.hm.2017.03.001.

F. Z. de Castro and M. E. Valle. A broad class of discrete-time hypercomplex-valued Hopfield neural networks. *Neural Networks*, 122:54–67, 2020. doi: https://doi.org/10.1016/j.neunet.2019.09.040.

M. H. Hassoun. *Fundamentals of Artificial Neural Networks*. MIT Press, Cambridge, MA, 1995.

M. H. Hassoun and P. B. Watta. Associative Memory Networks. In E. Fiesler and R. Beale, editors, *Handbook of Neural Computation*, pages C1.3:1–C1.3:14. Oxford University Press, 1997.

References (2)

J. J. Hopfield. Neural networks and physical systems with emergent collective computational abilities. *Proceedings of the National Academy of Sciences of the United States of America*, 79(8): 2554–2558, 1982. ISSN 00278424. doi: 10.1073/pnas.79.8.2554.

J. J. Hopfield. Neurons with graded response have collective computational properties like those of two-state neurons. *Proceedings of the National Academy of Sciences*, 81(10 I): 3088–3092, 1984. ISSN 00278424. doi: 10.1073/pnas.81.10.3088.

T. Isokawa, H. Nishimura, and N. Matsui. Commutative quaternion and multistate Hopfield neural networks. In *The 2010 International Joint Conference on Neural Networks (IJCNN)*, pages 1–6, 2010. doi: 10.1109/IJCNN.2010.5596736.

References (3)

T. Isokawa, H. Nishimura, and N. Matsui. Quaternionic Neural Networks for Associative Memories. In *Complex-Valued Neural Networks: Advances and Applications*, pages 103–131. John Wiley and Sons, 5 2013. ISBN 9781118344606. doi: 10.1002/9781118590072.ch5.

S. Jankowski, A. Lozowski, and J. M. Zurada. Complex-valued multistate neural associative memory. *IEEE Transactions on Neural Networks*, 7(6):1491–1496, 1996. doi: 10.1109/72.548176.

M. Kobayashi. Quaternionic Hopfield neural networks with twin-multistate activation function. *Neurocomputing*, 267:304–310, 2017. doi: 10.1016/j.neucom.2017.06.013.

M. Kobayashi. Twin-multistate commutative quaternion Hopfield neural networks. *Neurocomputing*, 320:150–156, 2018a. ISSN 0925-2312. doi: https://doi.org/10.1016/j.neucom.2018.09.023.

References (4)

M. Kobayashi. Hyperbolic Hopfield neural networks with directional multistate activation function. *Neurocomputing*, 275:2217–2226, 2018b. ISSN 0925-2312. doi: https://doi.org/10.1016/j.neucom.2017.10.053.

M. Kobayashi. Storage capacity of hyperbolic Hopfield neural networks. *Neurocomputing*, 369:185–190, 2019. ISSN 0925-2312. doi: https://doi.org/10.1016/j.neucom.2019.08.064.

M. Kobayashi. Hopfield neural networks using Klein four-group. *Neurocomputing*, 387:123–128, 4 2020a. ISSN 18728286. doi: 10.1016/j.neucom.2019.12.127.

M. Kobayashi. Noise Robust Projection Rule for Hyperbolic Hopfield Neural Networks. *IEEE Transactions on Neural Networks and Learning Systems*, 31(1):352–356, 1 2020b. ISSN 21622388. doi: 10.1109/TNNLS.2019.2899914.

References (5)

D. Krotov and J. Hopfield. Large Associative Memory Problem in Neurobiology and Machine Learning. 8 2020. URL http://arxiv.org/abs/2008.06996.

Y. Liang, D. Krotov, and M. J. Zaki. Modern Hopfield Networks for graph embedding. *Frontiers in Big Data*, 5, 11 2022. ISSN 2624-909X. doi: 10.3389/fdata.2022.1044709.

H. Ramsauer, B. Schäfl, J. Lehner, P. Seidl, M. Widrich, L. Gruber, M. Holzleitner, M. Pavlović, G. K. Sandve, V. Greiff, D. Kreil, M. Kopp, G. Klambauer, J. Brandstetter, and S. Hochreiter. Hopfield Networks is All You Need. 7 2020. URL http://arxiv.org/abs/2008.02217.

References (6)

Rodolfo Lobo. *Hypercomplex-valued recurrent neural networks and their applications for image reconstruction and pattern classification*. PhD thesis, Universidade Estadual de Campinas, Campinas, 4 2021.

M. E. Valle. A novel continuous-valued quaternionic Hopfield neural network. In *Proceedings - 2014 Brazilian Conference on Intelligent Systems, BRACIS 2014*, pages 97–102. Institute of Electrical and Electronics Engineers Inc., 12 2014. doi: 10.1109/BRACIS.2014.28.