Quaternionic Recurrent Correlation Neural Networks

Marcos Eduardo Valle

Department of Applied Mathematics
Institute of Mathematics, Statistics, and Scientific Computing
University of Campinas - Brazil
www.ime.unicamp.br/~valle

July 10, 2018

Introduction

Hopfield Network

Introduction

Hopfield Network
Low storage capacity!

Introduction

Hopfield Network
Low storage capacity!

(Chiueh and Goodman, 1991)

Recurrent Correlation Networks
High storage capacity!

Introduction

Quaternionic Hopfield Networks
Low storage capacity!

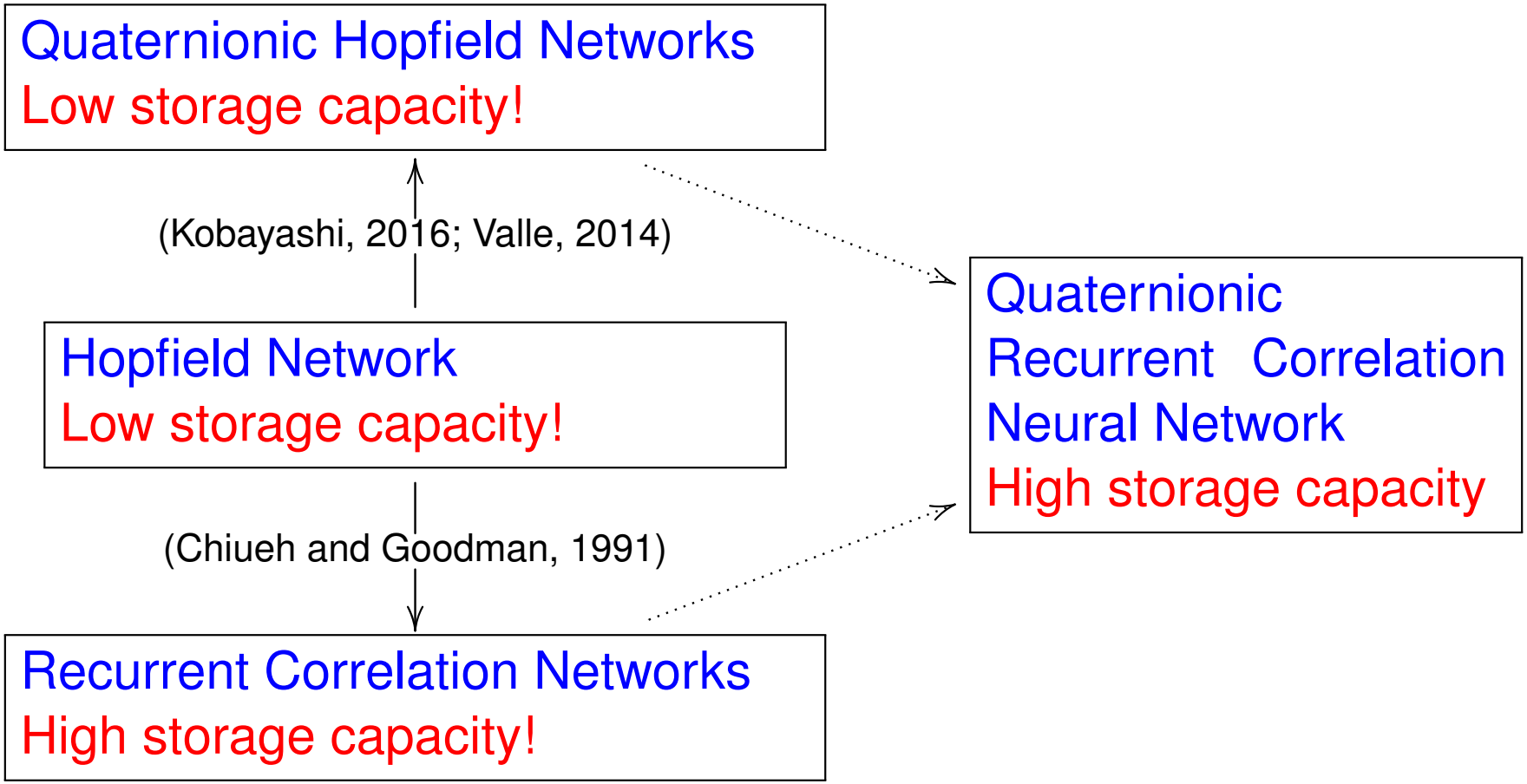
(Kobayashi, 2016; Valle, 2014)

Hopfield Network
Low storage capacity!

(Chiueh and Goodman, 1991)

Recurrent Correlation Networks
High storage capacity!

Introduction



Outline

From Hopfield Network to RCNNs

Given a set of vectors $\mathcal{U} = \{\mathbf{u}^1, \ldots, \mathbf{u}^K\} \subseteq \{-1, +1\}^N$, compute

$$m_{ij} = \frac{1}{N} \sum_{\xi=1}^{K} u_i^\xi u_j^\xi.$$

Given an input $\mathbf{x}(0) = [x_1(0), x_2(0), \ldots, x_N(0)]^T \in \{-1, +1\}^N$, define

$$x_i(t+1) = \begin{cases} \dfrac{v_i(t)}{|v_i(t)|}, & v_i(t) \neq 0, \\ x_i(t), & v_i(t) = 0, \end{cases}$$

where

$$v_i(t) = \sum_{j=1}^{N} m_{ij} x_j(t), \quad \forall i = 1, \ldots, N,$$

From Hopfield Network to RCNNs

Given a set of vectors $\mathcal{U} = \{\mathbf{u}^1, \ldots, \mathbf{u}^K\} \subseteq \{-1, +1\}^N$, compute

$$m_{ij} = \frac{1}{N} \sum_{\xi=1}^{K} u_i^\xi u_j^\xi.$$

Given an input $\mathbf{x}(0) = [x_1(0), x_2(0), \ldots, x_N(0)]^T \in \{-1, +1\}^N$, define

$$x_i(t+1) = \begin{cases} \dfrac{v_i(t)}{|v_i(t)|}, & v_i(t) \neq 0, \\ x_i(t), & v_i(t) = 0, \end{cases}$$

where

$$v_i(t) = \sum_{j=1}^{N} \left(\frac{1}{N} \sum_{\xi=1}^{K} u_i^\xi u_j^\xi \right) x_j(t), \quad \forall i = 1, \ldots, N,$$

From Hopfield Network to RCNNs

Given a set of vectors $\mathcal{U} = \{\mathbf{u}^1, \ldots, \mathbf{u}^K\} \subseteq \{-1, +1\}^N$, compute

$$m_{ij} = \frac{1}{N} \sum_{\xi=1}^{K} u_i^\xi u_j^\xi.$$

Given an input $\mathbf{x}(0) = [x_1(0), x_2(0), \ldots, x_N(0)]^T \in \{-1, +1\}^N$, define

$$x_i(t+1) = \begin{cases} \dfrac{v_i(t)}{|v_i(t)|}, & v_i(t) \neq 0, \\ x_i(t), & v_i(t) = 0, \end{cases}$$

where

$$v_i(t) = \sum_{j=1}^{N} \sum_{\xi=1}^{K} \left(\frac{1}{N} \sum_{j=1}^{N} x_j(t) u_j^\xi \right) u_i^\xi, \quad \forall i = 1, \ldots, N.$$

From Hopfield Network to RCNNs

Given a set of vectors $\mathcal{U} = \{\mathbf{u}^1, \ldots, \mathbf{u}^K\} \subseteq \{-1, +1\}^N$, compute

$$m_{ij} = \frac{1}{N} \sum_{\xi=1}^{K} u_i^\xi u_j^\xi.$$

Given an input $\mathbf{x}(0) = [x_1(0), x_2(0), \ldots, x_N(0)]^T \in \{-1, +1\}^N$, define

$$x_i(t+1) = \begin{cases} \dfrac{v_i(t)}{|v_i(t)|}, & v_i(t) \neq 0, \\ x_i(t), & v_i(t) = 0, \end{cases}$$

where

$$v_i(t) = \sum_{\xi=1}^{K} \left(\frac{1}{N} \left\langle \mathbf{u}^\xi, \mathbf{x}(t) \right\rangle \right) u_i^\xi, \quad \forall i = 1, \ldots, N.$$

From Hopfield Network to RCNNs

Given a set of vectors $\mathcal{U} = \{\mathbf{u}^1, \ldots, \mathbf{u}^K\} \subseteq \{-1, +1\}^N$, compute

$$m_{ij} = \frac{1}{N} \sum_{\xi=1}^{K} u_i^\xi u_j^\xi.$$

Given an input $\mathbf{x}(0) = [x_1(0), x_2(0), \ldots, x_N(0)]^T \in \{-1, +1\}^N$, define

$$x_i(t+1) = \begin{cases} \dfrac{v_i(t)}{|v_i(t)|}, & v_i(t) \neq 0, \\ x_i(t), & v_i(t) = 0, \end{cases}$$

where

$$v_i(t) = \sum_{\xi=1}^{K} f\left(\frac{1}{N}\left\langle \mathbf{u}^\xi, \mathbf{x}(t)\right\rangle\right) u_i^\xi, \quad \forall i = 1, \ldots, N,$$

and $f : [-1, 1] \to \mathbb{R}$ is a continuous nondecreasing function.

From Hopfield Network to RCNNs

Given a set of vectors $\mathcal{U} = \{\mathbf{u}^1, \ldots, \mathbf{u}^K\} \subseteq \{-1, +1\}^N$, compute

$$m_{ij} = \frac{1}{N} \sum_{\xi=1}^{K} u_i^\xi u_j^\xi.$$

Given an input $\mathbf{x}(0) = [x_1(0), x_2(0), \ldots, x_N(0)]^T \in \{-1, +1\}^N$, define

$$x_i(t+1) = \begin{cases} \dfrac{v_i(t)}{|v_i(t)|}, & v_i(t) \neq 0, \\ x_i(t), & v_i(t) = 0, \end{cases}$$

where

$$v_i(t) = \sum_{\xi=1}^{K} w_\xi u_i^\xi \quad \text{and} \quad w_\xi = f\left(\frac{1}{N}\left\langle \mathbf{u}^\xi, \mathbf{x}(t)\right\rangle\right).$$

Here, $f : [-1, 1] \to \mathbb{R}$ is a continuous nondecreasing function.

Quaternions

A quaternion is a four-dimensional hyper-complex number

$$q = q_0 + q_1\mathbf{i} + q_2\mathbf{j} + q_3\mathbf{k},$$

where $\mathbf{i}, \mathbf{j}$, and $\mathbf{k}$ are imaginary numbers such that

$$\mathbf{i}^2 = \mathbf{j}^2 = \mathbf{k}^2 = \mathbf{ijk} = -1.$$

The real part and the vector part of q are

$$\mathrm{Re}\{q\} \equiv q_0 \quad \text{and} \quad \mathrm{Ve}\{q\} \equiv \vec{q} = q_1\mathbf{i} + q_2\mathbf{j} + q_3\mathbf{k}.$$

The sum and the product of quaternions are

$$p + q = (p_0 + q_0) + (p_1 + q_1)\mathbf{i} + (p_2 + q_2)\mathbf{j} + (p_3 + q_3)\mathbf{k}.$$

and

$$pq = p_0 q_0 - \vec{p} \cdot \vec{q} + p_0\vec{q} + q_0\vec{p} + \vec{p} \times \vec{q}.$$

The conjugate and the norm are defined by

$$\bar{q} \equiv q_0 - \vec{q} \quad \text{and} \quad |q| = \sqrt{\bar{q}q} = \sqrt{q_0^2 + q_1^2 + q_2^2 + q_3^2}.$$

We say that q is a *unit* quaternion if $|q| = 1$.

We denote by $\mathbb{S}$ the set of all unit quaternions, i.e.,

$$\mathbb{S} = \{q \in \mathbb{H} : |q| = 1|\}.$$

Quaternionic Recurrent Correlation Networks

Let $\mathcal{U} = \{\mathbf{u}^1, \ldots, \mathbf{u}^K\} \subseteq \mathbb{S}^N$ be a set of vectors $\mathbf{u}^\xi = [u_1^\xi, \ldots, u_N^\xi]^T$ such that $|u_j^\xi| = 1$ for all $j = 1, \ldots, N$ and $\xi = 1, \ldots, K$.

A quaternionic RCNN (QRCNN) defines the sequence

$$x_j(t+1) = \begin{cases} v_j(t)/|v_j(t)|, & 0 < |v_j(t)| < +\infty, \\ x_j(t), & \text{otherwise}, \end{cases}$$

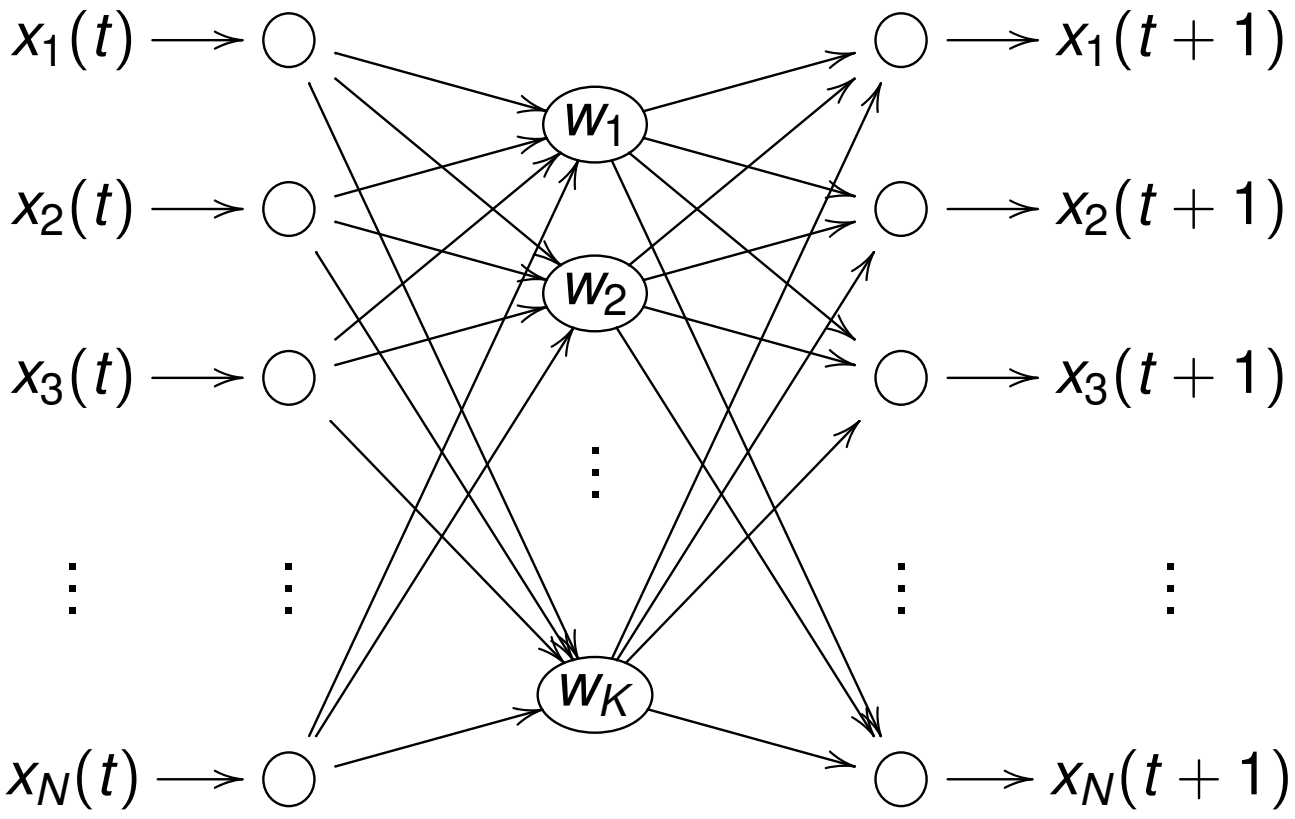
where

$$v_j(t) = \sum_{\xi=1}^{K} w_\xi(t) u_j^\xi \quad \text{and} \quad w_\xi(t) = f\left(\frac{1}{N}\mathrm{Re}\left\{\left\langle \mathbf{u}^\xi, \mathbf{x}(t)\right\rangle\right\}\right),$$

and $f : [-1, 1] \to \mathbb{R}$ is a continuous nondecreasing function.

Recall that $\left\langle \mathbf{u}^\xi, \mathbf{x}(t)\right\rangle = \sum_{j=1}^{N} \bar{x}_j(t) u_j^\xi$.

Network Topology



Examples of QRCNNs

- The correlation QRCNN is obtained by considering

$$f_c(x) = x.$$

- The exponential QRCNN is obtained considering

$$f_e(x) = e^{\alpha x}, \quad \alpha > 0.$$

- The high-order QRCNN is obtained by considering

$$f_h(x) = (1 + x)^q, \quad q > 1 \text{ is an integer.}$$

- The potential-function QRCNN is obtained by considering

$$f_p(x) = \frac{1}{(1 - x + \varepsilon_{mach})^L}, \quad L \geq 1.$$

Convergence

Theorem (Convergence of the QRCNNs)

The sequence produced by a QRCNN always converges.

The sequence $\{\mathbf{x}(t)\}_{t\geq 0}$ yields a local minima of the energy function

$$E(\mathbf{x}) = -\sum_{\xi=1}^{K} F\left(\frac{\mathrm{Re}\left\{\left\langle \mathbf{u}^{\xi}, \mathbf{x}\right\rangle\right\}}{N}\right), \quad \forall\, \mathbf{x} \in \mathbb{S}^N,$$

where F is a certain primitive of the real-valued function f.

Thus, we can define the mapping $\psi : \mathbb{S}^N \to \mathbb{S}^N$ by setting

$$\psi(\mathbf{x}) = \lim_{t\to\infty} \mathbf{x}(t) \quad \text{with} \;\; \mathbf{x}(0) = \mathbf{x}.$$

QRCNN as Associative Memories

We say that ψ implements a τ-associative memory designed for the storage of $\mathbf{u}^1, \ldots, \mathbf{u}^K$ if

$$\|\psi(\mathbf{u}^\xi) - \mathbf{u}^\xi\| \leq \tau, \quad \forall \xi = 1, \ldots, K,$$

where $\|\cdot\|$ denotes the Euclidean norm and $\tau > 0$ is a tolerance.

Theorem

Consider a fundamental memory set $\mathcal{U} = \{\mathbf{u}^1, \ldots, \mathbf{u}^K\} \subseteq \mathbb{S}^N$ and a sufficiently small tolerance $\tau > 0$. If $f(x; \lambda) = [A(x)]^\lambda$, then there exists a sufficiently large parameter λ such that

$$\|\mathbf{x}(1) - \mathbf{u}^\gamma\|_2 < \tau,$$

where $\mathbf{u}^\gamma$ is the fundamental memory closest $\mathbf{x}(0)$.

Consequences of the Theorem

If the input $\mathbf{x}(0)$ corresponds to a corrupted version of $\mathbf{u}^\gamma$, then we can always fine tune the parameter λ of a such that $\mathbf{x}(1)$ sufficiently close to $\mathbf{u}^\gamma$.

Any input in the open ball

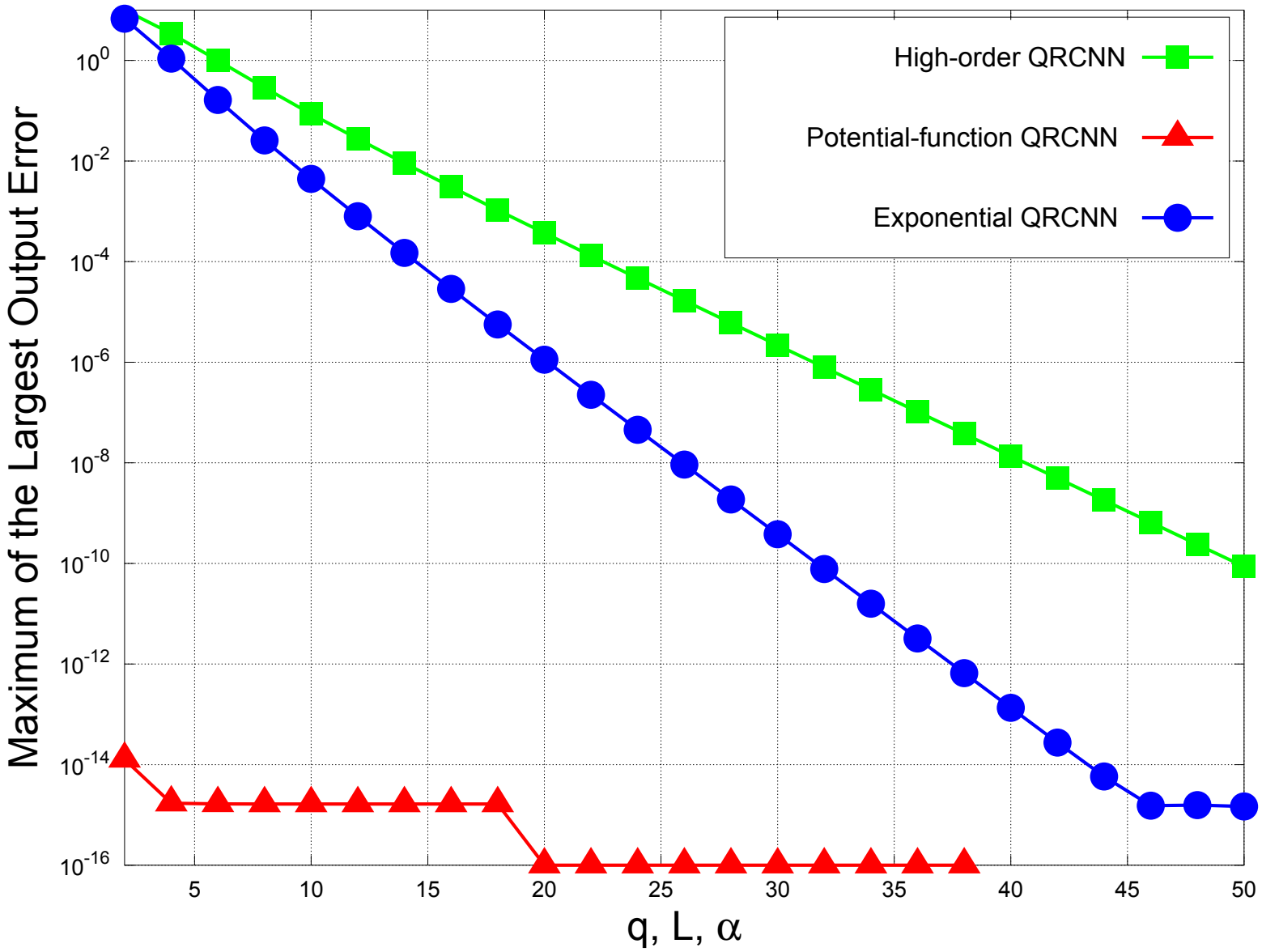
$$\mathcal{B}(\mathbf{u}^\gamma) = \left\{ \mathbf{x} \in \mathbb{S}^N : \|\mathbf{x} - \mathbf{u}^\gamma\|_2 < \min_{\xi \neq \gamma} \{\|\mathbf{x} - \mathbf{u}^\xi\|\} \right\},$$

is attracted to $\mathbf{u}^\gamma$.

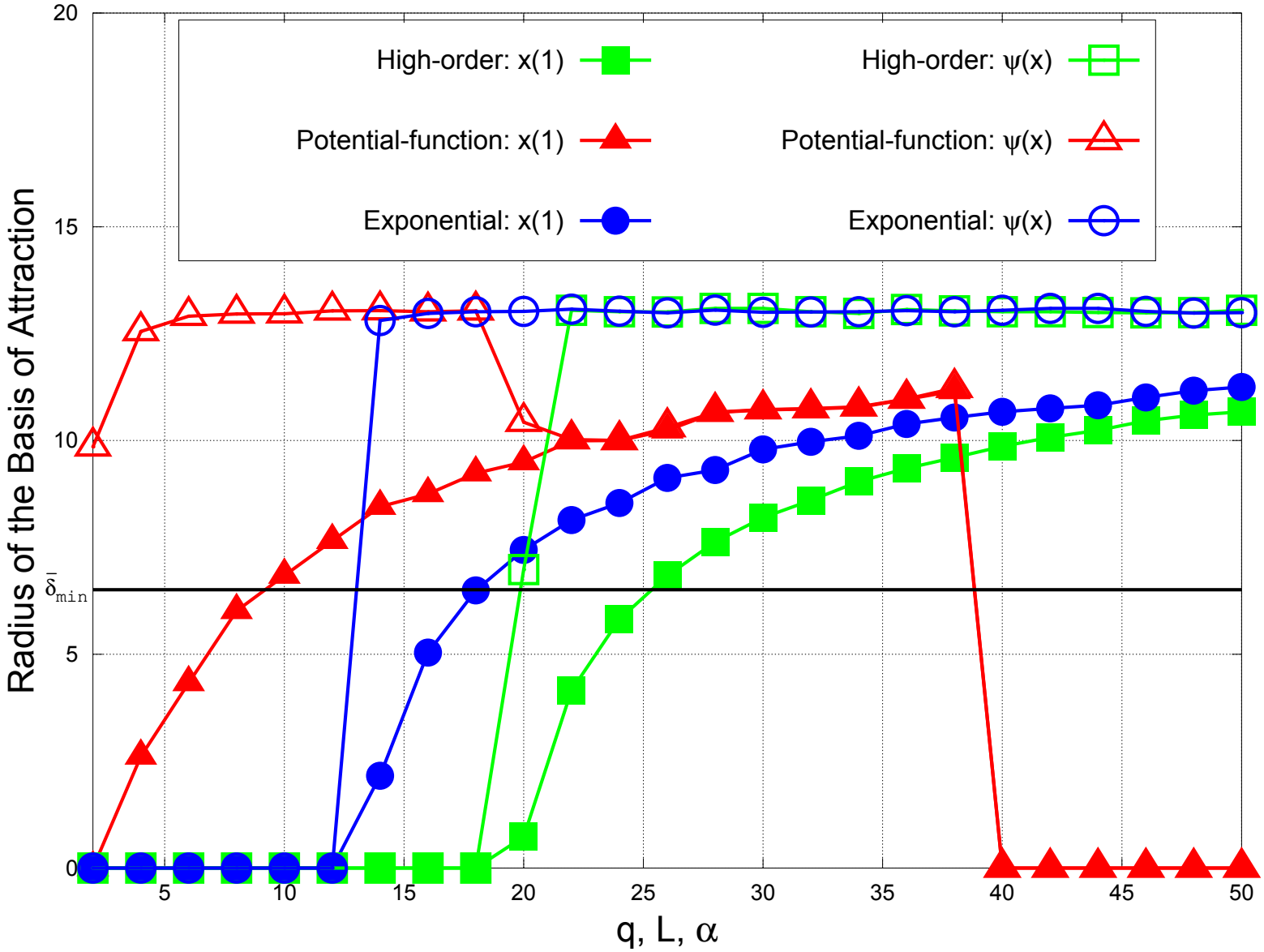
The radius of attraction of $\mathbf{u}^\gamma$ is, at least, half the Euclidean distance between the two closest fundamental memories.

Since $\mathbf{u}^\gamma \in \mathcal{B}(\mathbf{u}^\gamma)$, the inequality $\|\psi(\mathbf{u}^\gamma) - \mathbf{u}^\gamma\|_2 \leq \tau$ holds for all γ.

Capability to Implement Associative Memory



Radius of the Basis of Attraction



Concluding Remarks

We generalized the bipolar RCNN using unit quaternions.

Quaternionic recurrent correlation neural networks (QRCNNs) can implement high-capacity associative memories.

- The sequence produced by an QRCNN always converge.
- But it may converge to an undesired spurious state.

If the activation function is an exponential in its parameter, the QRCNN can retrieve an item if the input is sufficiently close to it.

The radius of the basis of attraction of the resulting associative memory is at least half the distance between the two closest fundamental memories.

Thank you very much!

References (1)

T. Chiueh and R. Goodman. Recurrent Correlation Associative Memories. *IEEE Trans. on Neural Networks*, 2:275–284, Feb. 1991.

M. Kobayashi. Rotational invariance of quaternionic hopfield neural networks. *IEEJ Transactions on Electrical and Electronic Engineering*, 11(4):516–520, 2016. ISSN 1931-4981. doi: 10.1002/tee.22269.

M. E. Valle. A Novel Continuous-Valued Quaternionic Hopfield Neural Network. In *Proceedings of the Brazilian Conference on Intelligent Systems 2014 (BRACIS 2014)*, São Carlos, Brazil, October 2014.