Neurocomputing 313 (2018) 25–38

Contents lists available at ScienceDirect

Neurocomputing

journal homepage: www.elsevier.com/locate/neucom

Geometric measures of entanglement in multipartite pure states via complex-valued neural networks

Maolin Che[a,1], Liqun Qi[b,2], Yimin Wei[c,3,*], Guofeng Zhang[d,4]

[a] School of Economic Mathematics, Southwest University of Finance and Economics, Chengdu 611130, P R China
[b] Department of Applied Mathematics, the Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong
[c] School of Mathematical Sciences and Key Laboratory of Mathematics for Nonlinear Sciences, Fudan University, Shanghai, 200433, P R China
[d] Department of Applied Mathematics, the Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong

ARTICLE INFO

Article history:
Received 18 January 2018
Revised 11 May 2018
Accepted 27 May 2018
Available online 19 June 2018

Communicated by Prof. Qiankun Song

AMS subject classications:
15A18
15A69
65F15
65F10

Keywords:
Quantum entanglement
Multipartite state
Geometric measure of entanglement
Complex-valued neural network
Quantum eigenvalue problem
Local optimal complex rank-one approximation
Lyapunov stability theory
Complex symmetric tensors
Complex tensors

ABSTRACT

The geometric measure of entanglement of a multipartite pure state is defined it terms of its geometric distance from the set of separable pure states. The quantum eigenvalue problem is derived to compute the separable pure state nearest to the given multipartite pure state. Computing the modulus largest quantum eigenvalue for a multipartite pure state is equivalent to finding the best complex rank-one approximation of the complex unit tensors, associated with the multipartite pure states. This paper is devoted to present a complex-valued neural networks approach for the computation of the quantum eigenvalue problem for multipartite pure states. We design the neural networks for computing the best rank-one tensor approximation of complex tensors, and prove that the solution of the networks is locally asymptotically stable in the sense of Lyapunov stability theory. This solution also converges to the local optimal solutions of the best complex rank-one tensor approximation. We illustrate our theoretical results via numerical simulations.

1. Introduction

A tensor is an N-dimensional array of numbers denoted by script notation $\mathcal{A} \in \mathbb{C}^{I_1 \times I_2 \times \ldots \times I_N}$ with entries given by

$$a_{i_1 i_2 \ldots i_N} \in \mathbb{C}, \quad \text{for} \quad i_n = 1, 2, \ldots, I_n, \ \text{with} \ n = 1, 2, \ldots, N.$$

We use $CT_{N,I}$ to denote the set of order N dimension I complex tensors in general. That is, when $\mathcal{A} \in CT_{N,I}$, we have $a_{i_1 i_2 \ldots i_N} \in \mathbb{C}$ where $i_n = 1, 2, \ldots, I$ and $n = 1, 2, \ldots, N$.

The problem of best rank-one approximation of $\mathcal{A} \in \mathbb{C}^{I_1 \times I_2 \times \ldots \times I_N}$ is to find a real scalar σ and N unit vectors $\mathbf{x}_n \in \mathbb{C}^{I_n}$ ($\|\mathbf{x}_n\|_2 = 1$) that minimize

$$\sum_{i_1=1}^{I_1} \sum_{i_2=1}^{I_2} \ldots \sum_{i_N=1}^{I_N} \left| a_{i_1 i_2 \ldots i_N} - \sigma \cdot (x_{1,i_1} x_{2,i_2} \ldots x_{N,i_N}) \right|^2,$$

* Corresponding author.
E-mail addresses: chncml@outlook.com, cheml@swufe.edu.cn (M. Che), maqilq@polyu.edu.hk (L. Qi), ymwei@fudan.edu.cn (Y. Wei), magzhang@polyu.edu.hk (G. Zhang).

[1] This author is supported by the Fundamental Research Funds for the Central Universities under grant JBK1801058.
[2] This author's work was supported by the Hong Kong Research Grant Council (Grant No. PolyU 15302114, 15300715, 15301716 and 15300717).
[3] This author is supported by the National Natural Science Foundation of China under grant 11771099.
[4] This author is supported by the National Natural Science Foundation of China and the Hong Kong Research Grant Council (Grant No. 61374057 and 15206915).

https://doi.org/10.1016/j.neucom.2018.05.094
0925-2312/

where x_{n,i_n} is the i_nth element of $\mathbf{x}_n \in \mathbb{C}^{I_n}$ for $i_n = 1, 2, \ldots, I_n$ with $n = 1, 2, \ldots, N$, and $\sigma \in \mathbb{R}$ is a scaling factor. The relationship between the best rank-one approximation of complex tensors and geometric measures of entanglement in multipartite pure states will be discussed in Section 2.1.

There exist numerical methods to compute the best rank-one approximation of real tensors, e.g., the alternating least squares (ALS) method, truncated higher-order singular value decomposition, higher-order power method and semi-definite relaxations. We refer to Zhang and Golub [47], De Lathauwer, De Moor and Vandewalle [13,14], Kofidis and Regalia [27], Qi et al. [38], Ni and Wang [36], Nie and Wang [37] and the references therein.

Ni et al. [35] considered two eigenvalue problems of complex tensors: the U-eigenvalue problem of a complex tensor and the US-eigenvalue problem of a complex symmetric tensor, which are related to the best rank-one approximation of complex tensors. Recently, Ni and Bai [34] proposed an algorithm for computing the US-eigenpairs of complex symmetric tensors based on a spherical optimization problem of real-valued functions with complex variables. This algorithm was used to compute the upper bound of entanglement in an arbitrary multi-partite system [39]. Che et al. [9] presented iterative algorithms for computing US- (or U-) eigenpairs of complex tensors based on the Takagi factorization of complex matrices.

Wang et al. [44] proposed complex-valued neural network models for the computation of the Takagi vector of a complex symmetric matrix that corresponds to the largest Takagi values. The readers can refer to [2,3,24], which studied a complex nonlinear convex programming problem by means of complex-valued neural network models. Generally speaking, complex-valued neural networks have different and more complicated properties than real-valued ones. Thus, it is important to study the dynamical behaviors of complex-valued neural networks.

One important aspects of the dynamics of neural networks is their stability. To analyze the stability of neural networks, various approaches, such as Lyapunov function method and synthesis method, have been proposed [12,30,41]. Che et al. [8] presented a neural dynamical network to compute a local optimal rank-one approximation of a real tensor and proved that the state of the proposed neural network is locally asymptotically stable in the sense of Lyapunov stability theory. The main purpose of this paper is to design complex-valued neural network models for computing the local optimal rank-one approximation of complex tensors. We also derive that the solution of the complex-valued ODEs is locally asymptotically stable in the sense of Lyapunov stability theory. As shown in Section 7, the method of complex-valued neural network models is a strong tool for calculating geometric measure of entanglement.

Throughout this paper, we assume that *I*, *J*, and *N* will be reserved to denote the index upper bounds, unless stated otherwise. Scalars are denoted by lower Greek letters and lower Roman letters, e.g., α and a. Vectors are denoted by boldface letters and are lower case, e.g., $\mathbf{z}$. Matrices are denoted by block capital letters, e.g., $\mathbf{A}$. Tensors are denoted by calligraphic letters, e.g., $\mathcal{A}$. The superscripts $\cdot^\top$, $\bar{\cdot}$ and $\cdot^*$ are used for the transpose, the complex conjugate and conjugate transpose, respectively.

The two-norm and Frobenius norm are denoted by $\|\cdot\|_2$ and $\|\cdot\|_F$, respectively. The entry with row index i and column index j in a matrix $\mathbf{A}$, i.e., $(\mathbf{A})_{ij}$, is symbolized by a_{ij} (also $(\mathbf{z})_i = z_i$ and $(\mathcal{A})_{i_1 i_2 \ldots i_N} = a_{i_1 i_2 \ldots i_N}$). We use parentheses to denote the concatenation of two or more vectors, e.g., $(\mathbf{a}, \mathbf{b})$ is equivalent to $(\mathbf{a}^\top, \mathbf{b}^\top)^\top$. We use $\Re(\mathbf{z})$ and $\Im(\mathbf{z})$ to denote the real and imaginary parts of a vector $\mathbf{z} \in \mathbb{C}^I$.

The rest of this paper is organized as follows. In Section 2, we introduce basic notations about quantum states, convert the problem for measuring entanglement of a multipartite pure state to the complex best rank-one tensor approximation, and present the expressions for the complex gradient of real functions in complex variables. In Section 3, we define the generalized Rayleigh quotient of the complex tensors and establish the relationship between the local optimal complex rank-one tensor approximation and the nonlinear quantum eigenvalue problem (US-eigenvalue problems or U-eigenvalue problems [35]) based on the generalized Rayleigh quotient of any complex tensor. We present neural networks and consider the properties of these neural networks in Section 4. In Section 5, we establish the complex-valued neural networks to find the local optimal complex rank-one tensor approximation and analyze its local asymptotic stability in the sense of Lyapunov stability theory. We illustrate our theory via numerical simulations in Section 6 and conclude our paper in Section 7.

2. Preliminaries

The mode-n product [28] of a complex tensor $\mathcal{A} \in \mathbb{C}^{I_1 \times I_2 \times \ldots \times I_N}$ by a matrix $\mathbf{B} \in \mathbb{C}^{J_n \times I_n}$, denoted by $\mathcal{A} \times_n \mathbf{B}$, is a tensor $\mathcal{C} \in \mathbb{C}^{I_1 \times \ldots \times I_{n-1} \times J_n \times I_{n+1} \times \ldots \times I_N}$, whose entries are given by

$$c_{i_1 \ldots i_{n-1} j i_{n+1} \ldots i_N} = \sum_{i_n=1}^{I_n} a_{i_1 i_2 \ldots i_N} b_{j i_n}, \quad n = 1, 2, \ldots, N.$$

In particular, the mode-n multiplication of a complex tensor $\mathcal{A} \in \mathbb{C}^{I_1 \times I_2 \times \ldots \times I_N}$ by a vector $\mathbf{z} \in \mathbb{C}^{I_n}$ is denoted by $\mathcal{A} \times_n \mathbf{z}^\top$. If we set $\mathcal{C} = \mathcal{A} \times_n \mathbf{z}^\top \in \mathbb{C}^{I_1 \times \ldots \times I_{n-1} \times I_{n+1} \times \ldots \times I_N}$, then we have element-wise [28],

$$c_{i_1 \ldots i_{n-1} i_{n+1} \ldots i_N} = \sum_{i_n=1}^{I_n} a_{i_1 \ldots i_{n-1} i_n i_{n+1} \ldots i_N} x_{i_n}.$$

Given N vectors $\mathbf{z}_n \in \mathbb{C}^{I_n}$ $(n = 1, 2, \ldots, N)$, the notation $\mathcal{A} \times_1 \mathbf{z}_1^\top \times_2 \mathbf{z}_2^\top \ldots \times_N \mathbf{z}_N^\top$ is easy to define. For any given tensor $\mathcal{A} \in \mathbb{C}^{I_1 \times I_2 \times \ldots \times I_N}$ and the matrices $\mathbf{F} \in \mathbb{C}^{J_n \times I_n}$ and $\mathbf{G} \in \mathbb{C}^{J_m \times I_m}$, one has [28]

$$\begin{cases} (\mathcal{A} \times_n \mathbf{F}) \times_m \mathbf{G} = (\mathcal{A} \times_m \mathbf{G}) \times_n \mathbf{F} = \mathcal{A} \times_n \mathbf{F} \times_m \mathbf{G}; \\ (\mathcal{A} \times_n \mathbf{F}) \times_n \mathbf{G} = \mathcal{A} \times_n (\mathbf{G} \cdot \mathbf{F}), \quad \text{with} \quad \mathrm{J_n = I_m}, \end{cases}$$

with $m \neq n \in \{1, 2, \ldots, N\}$, where '$\cdot$' represents the multiplication of two matrices.

If the entries of $\mathcal{A} \in \mathbb{C}^{I_1 \times I_2 \times \ldots \times I_N}$ are given by $a_{i_1 i_2 \ldots i_N} = x_{1,i_1} x_{2,i_2} \ldots x_{N,i_N}$, where x_{n,i_n} is the i_nth element of $\mathbf{x}_n \in \mathbb{C}^{I_n}$ for $n = 1, 2, \ldots, N$, then we call $\mathcal{A}$ a complex rank-one tensor [14,47].

2.1. Geometric measure of entanglement

Entanglement has been identified as a resource central to quantum information processing. As a result, the task of characterizing and quantifying entanglement is vitally important in quantum information theory. The geometric measure of entanglement is one of most natural and important measures for pure states in bipartite and multipartite systems. We refer to [42,45] and their inferences therein. Mathematically speaking, the geometric measure of entanglement is nothing but the injective tensor norm [21], which appears in the theory of operator algebra [15]. The geometric measure of entanglement also has found wild applications in various different topics, such as many-body physics [31,33], entanglement witnesses [17,20] and the study of quantum channel capacities [7,16,46].

Wei and Goldbart [45] extended the geometric measure of entanglement from a bipartite pure state [42] to a multipartite pure state via the entanglement eigenvalue of a nonlinear quantum eigenvalue problem. Ni et al. [35] studied the nonlinear quantum eigenvalue problem in two forms: the U-eigenvalue problem of a complex tensor and the US-eigenvalue problem of a complex symmetric tensor. For a symmetric pure state with nonnegative amplitudes, its geometric measure of entanglement problem can be regarded as a tensor decomposition problem or a rank-one approximation to higher-order tensors problem [21,25]. As shown in [10], there exist two ways for computing geometric measure: the analytic method and the numerical method. More methods for the computation of geometric measure of entanglement can be referred to [18]. For the calculation of geometric measure, the interested readers can refer to [10,11,19,32,48] and the references therein. In Section 6, we will compare our proposed complex-valued neural network approach with the numerical method in [10] for computing geometric measure of some multipartite pure states in [10].

Let us first develop a general formation, which is appropriate for multipartite systems comprising N parts, in which each part has a distinct Hilbert space as state space. An N-partite *pure state* $|\psi\rangle$ of a composite quantum system can be regarded as a normalized element in a tensor product Hilbert space $\mathbb{H}=\mathbb{H}_1\otimes\mathbb{H}_2\otimes\ldots\otimes\mathbb{H}_N$, where $\mathbb{H}_n:=\mathbb{C}^{I_n}$ for $n=1,2,\ldots,N$ and '$\otimes$' is the Kronecker product [23].

Assume that the set $\{|\mathbf{e}_{n,i_n}\rangle : i_n=1,2,\ldots,I_n\}$ is an orthonormal basis of $\mathbb{H}_n$ with $n=1,2,\ldots,N$. Then the set $\{|\mathbf{e}_{1,i_1}\mathbf{e}_{2,i_2}\ldots\mathbf{e}_{N,i_N}\rangle : i_n=1,2,\ldots,I_n; n=1,2,\ldots,N\}$ is an orthonormal basis of $\mathbb{H}$. Any N-partite pure state $|\psi\rangle\in\mathbb{H}$ can be written as

$$|\psi\rangle=\sum_{i_1=1}^{I_1}\sum_{i_2=1}^{I_2}\ldots\sum_{i_N=1}^{I_N}a_{i_1i_2\ldots i_N}|\mathbf{e}_{1,i_1}\mathbf{e}_{2,i_2}\ldots\mathbf{e}_{N,i_N}\rangle \quad \text{with} \quad a_{i_1i_2\ldots i_N}\in\mathbb{C}.$$

Assume that

$$|\varphi\rangle=\sum_{i_1=1}^{I_1}\sum_{i_2=1}^{I_2}\ldots\sum_{i_N=1}^{I_N}b_{i_1i_2\ldots i_N}|\mathbf{e}_{1,i_1}\mathbf{e}_{2,i_2}\ldots\mathbf{e}_{N,i_N}\rangle, \quad \text{with} \quad b_{i_1i_2\ldots i_N}\in\mathbb{C},$$

we define

$$\langle\psi|\varphi\rangle=\sum_{i_1=1}^{I_1}\sum_{i_2=1}^{I_2}\ldots\sum_{i_N=1}^{I_N}\overline{b_{i_1i_2\ldots i_N}}a_{i_1i_2\ldots i_N},$$

and

$$\||\varphi\rangle\|=\sqrt{\langle\varphi|\varphi\rangle}.$$

A *separable* N-partite pure state is expressed by $|\phi\rangle=|\phi_1\rangle\otimes|\phi_2\rangle\otimes\ldots\otimes|\phi_N\rangle$, where each part $|\phi_n\rangle$ is defined by

$$|\phi_n\rangle=\sum_{i_n=1}^{I_n}x_{n,i_n}|\mathbf{e}_{n,i_n}\rangle.$$

One can envisage a geometric definition of entanglement for the state $|\psi\rangle$ via the distance

$$d=\min_{|\phi\rangle}\||\psi\rangle-|\phi\rangle\| \tag{2.1}$$

between $|\psi\rangle$ and the nearest separable state $|\phi\rangle$ [45].

In the literature, the *geometric measure* for pure states is also taken as

$$\frac{1}{2}\min\||\psi\rangle-|\phi\rangle\|^2=1-G(\psi),$$

for separable pure states $|\phi\rangle$, where $G(\psi)$ is the maximal overlap:

$$G(\psi):=\max|\langle\psi|\phi\rangle|$$

for separable pure states $|\phi\rangle$ with $\langle\phi|\phi\rangle=1$. In the bipartite case, we can solve the minimization problem (2.1) via the Schmidt decomposition (a restatement of the singular value decomposition) [42].

According to the above description, to solve the minimization problem (2.1), we focus on the following best complex rank-one tensor approximation problem: find N nonzero vectors $\mathbf{z}_n$ to minimize

$$\|\mathcal{A}-\mathbf{z}_1\circ\mathbf{z}_2\circ\ldots\circ\mathbf{z}_N\|_F^2, \tag{2.2}$$

where the tensor $\mathcal{A}\in\mathbb{C}^{I_1\times I_2\times\cdots\times I_N}$ is associated with the multipartite pure state $|\phi\rangle$ and the Frobenius norm of $\mathcal{A}$ is the square root of the sum of the squares of the moduli of all its elements, i.e.,

$$\|\mathcal{A}\|_F=\sqrt{\sum_{i_1=1}^{I_1}\sum_{i_2=1}^{I_2}\ldots\sum_{i_N=1}^{I_N}|a_{i_1i_2\ldots i_N}|^2}.$$

Meanwhile, we can impose the constraints $\|\mathbf{z}_n\|_2=1$ and rewrite the best complex rank-one tensor approximation (2.2) as

$$\min\|\mathcal{A}-\sigma\cdot(\mathbf{z}_1\circ\mathbf{z}_2\circ\ldots\circ\mathbf{z}_N)\|_F^2. \tag{2.3}$$

If the scalar σ is complex, then we can define $\mathbf{y}_n:=\exp(\iota\theta/N)\mathbf{z}_n$, $(\iota=\sqrt{-1})$ to ensure that σ is an associated real number, where $\theta\in(-\pi,\pi]$ is the argument of the scalar σ. Hence, in this paper, we always assume that the scalar σ is real.

2.2. Real functions with complex variables

Consider a function $f:\mathbb{C}\to\mathbb{C}$. The complex derivative of f at $z\in\mathbb{C}$ is defined as the limit, if it exists,

$$f'(z):=\lim_{\delta z\to 0}\frac{f(z+\delta z)-f(z)}{\delta z}.$$

It is well known that f is differentiable in the complex sense, if and only if the Cauchy–Riemann conditions hold. However, in many practical applications, many functions are not differentiable in the complex sense [4]. In order to deal with these problems, we develop an alternative formulation that is based on the real derivatives, and it is similar to that of the complex derivative.

The purpose of this section is thus to gather some results concerning Wirtinger derivatives of real valued functions over complex variables. Suppose that $f:\mathbb{C}^I\to\mathbb{R}$ is a real function with complex variables. Let $\mathbf{z}=\mathbf{x}+\iota\mathbf{y}\in\mathbb{C}^I$ with $\mathbf{x},\mathbf{y}\in\mathbb{R}^I$. We introduce a calculus of the differential operators, which was developed principally by Wirtinger, often called the *Wirtinger calculus*. We refer to [5,29,40,43] for the underlying framework of the complex derivatives.

Definition 2.1. Let $\mathbf{z}=\mathbf{x}+\iota\mathbf{y}\in\mathbb{C}^I$ with $\mathbf{x},\mathbf{y}\in\mathbb{R}^I$. The cogradient operator $\frac{\partial}{\partial\mathbf{z}}$ and conjugate cogradient operator $\frac{\partial}{\partial\bar{\mathbf{z}}}$ are defined as

$$\frac{\partial}{\partial\mathbf{z}}=\frac{1}{2}\begin{bmatrix}\frac{\partial}{\partial x_1}-\iota\frac{\partial}{\partial y_1}\\ \vdots\\ \frac{\partial}{\partial x_I}-\iota\frac{\partial}{\partial y_I}\end{bmatrix},\quad \frac{\partial}{\partial\bar{\mathbf{z}}}=\frac{1}{2}\begin{bmatrix}\frac{\partial}{\partial x_1}+\iota\frac{\partial}{\partial y_1}\\ \vdots\\ \frac{\partial}{\partial x_I}+\iota\frac{\partial}{\partial y_I}\end{bmatrix},$$

where $\iota=\sqrt{-1}$ is the imaginary unit.

Note that the Cauchy–Riemann conditions for f to be analytic in $\mathbf{z}$ can be expressed compactly, using the cogradient as $\frac{\partial f}{\partial\bar{\mathbf{z}}}=0$, i.e., f is a function only of $\mathbf{z}$. Analogously, f is analytic in $\bar{\mathbf{z}}$ if and only if $\frac{\partial f}{\partial\mathbf{z}}=0$.

Let $\frac{\partial}{\partial\mathbf{x}}=(\frac{\partial}{\partial x_1},\ldots,\frac{\partial}{\partial x_I})^\top$ and $\frac{\partial}{\partial\mathbf{y}}=(\frac{\partial}{\partial y_1},\ldots,\frac{\partial}{\partial y_I})^\top$, it is clear that

$$\frac{\partial}{\partial\mathbf{x}}=\frac{\partial}{\partial\mathbf{z}}+\frac{\partial}{\partial\bar{\mathbf{z}}},\quad \frac{\partial}{\partial\mathbf{y}}=\iota\left(\frac{\partial}{\partial\mathbf{z}}-\frac{\partial}{\partial\bar{\mathbf{z}}}\right).$$

The complex Taylor series as presented in this paper is due to van den Bos [43], who found a compact way of transforming the real gradient and Hessian into their complex counterparts. Kreutz-Delgado points out that there is actually more than one way to define the *complex Hessian* [29].

Lemma 2.1. ([29, 34]) *Suppose that a real valued function $f(\mathbf{z})$ with complex variables $\mathbf{z} \in \mathbb{C}^I$ is first-order differentiable. Then, the first-order Taylor series of $f(\mathbf{z})$ is*

$$\begin{aligned} f(\mathbf{z}+\delta\mathbf{z}) &= f(\mathbf{z}) + 2\Re\left[\left(\frac{\partial f(\mathbf{z})}{\partial \mathbf{z}}\right)^\top \delta\mathbf{z}\right] + \mathcal{O}(\|\delta\mathbf{z}\|_2^2) \\ &= f(\mathbf{z}) + \left(\left(\frac{\partial f(\mathbf{z})}{\partial \mathbf{x}}\right)^\top \delta\mathbf{x} + \left(\frac{\partial f(\mathbf{z})}{\partial \mathbf{y}}\right)^\top \delta\mathbf{y}\right) \\ &\quad + \mathcal{O}(\|\delta\mathbf{x}\|_2^2 + \|\delta\mathbf{y}\|_2^2), \end{aligned}$$

where $\delta\mathbf{z} = \delta\mathbf{x} + \iota\delta\mathbf{y}$ with $\delta\mathbf{x}, \delta\mathbf{y} \in \mathbb{R}^I$.

3. Equivalent rank-one formulations

Similar to the generalized Rayleigh quotient of any real tensor [47], we define the the generalized Rayleigh quotient of any complex tensor in $\mathbb{C}^{I_1\times I_2\times\cdots\times I_N}$ as follows.

Definition 3.1. Suppose that $\mathcal{A} \in \mathbb{C}^{I_1\times I_2\times\cdots\times I_N}$. For N given nonzero vectors $\mathbf{z}_n \in \mathbb{C}^{I_n}$, the generalized Rayleigh quotient of the tensor $\mathcal{A}$ is defined as

$$\mathrm{GRQ}(\mathbf{z}_1,\mathbf{z}_2,\ldots,\mathbf{z}_N) = \frac{1}{2}\cdot\frac{\mathcal{A}\times_1\mathbf{z}_1^*\times_2\mathbf{z}_2^*\cdots\times_N\mathbf{z}_N^* + \overline{\mathcal{A}}\times_1\mathbf{z}_1^\top\times_2\mathbf{z}_2^\top\cdots\times_N\mathbf{z}_N^\top}{\|\mathbf{z}_1\|_2\|\mathbf{z}_2\|_2\ldots\|\mathbf{z}_N\|_2}.$$

Similar to the standard Rayleigh quotient with real symmetric matrices, we define the generalized Rayleigh quotient of the tensor $\mathcal{A}$ in a way that is invariant under a positive scaling of one of the vectors $\mathbf{z}_1, \mathbf{z}_2, \ldots, \mathbf{z}_N$. We have the following lemma to establish the relationship between the generalized Rayleigh quotient of a tensor $\mathcal{A} \in \mathbb{C}^{I_1\times I_2\times\cdots\times I_N}$ and the minimization problem (2.3).

Lemma 3.1. *Suppose that $\mathcal{A} \in \mathbb{C}^{I_1\times I_2\times\ldots\times I_N}$. If the vectors $\mathbf{z}_n$ satisfy $\|\mathbf{z}_n\|_2 = 1$ with $n = 1, 2, \ldots, N$, then $\sigma_* = \mathrm{GRQ}(\mathbf{z}_1,\mathbf{z}_2,\ldots,\mathbf{z}_N)$ minimizes (2.3), and the minimum value is*

$$\|\mathcal{A}\|_F^2 - \mathrm{GRQ}(\mathbf{z}_1,\mathbf{z}_2,\ldots,\mathbf{z}_N)^2.$$

Proof. For two given scalars $z_1, z_2 \in \mathbb{C}$, it is easy to see that $|z_1 - z_2|^2 = |z_1|^2 + |z_2|^2 - (\bar{z}_1 z_2 + z_1\bar{z}_2)$. According to the Frobenius norm of $\mathcal{A}$, the formula $\|\mathcal{A} - \sigma\cdot(\mathbf{z}_1\circ\mathbf{z}_2\circ\ldots\circ\mathbf{z}_N)\|_F^2$ can be represented as

$$\|\mathcal{A}\|_F^2 - \sigma(\mathcal{A}\times_1\mathbf{z}_1^*\times_2\mathbf{z}_2^*\cdots\times_N\mathbf{z}_N^* + \overline{\mathcal{A}}\times_1\mathbf{z}_1^\top\times_2\mathbf{z}_2^\top\cdots\times_N\mathbf{z}_N^\top) + \sigma^2(\|\mathbf{z}_1\|_2\|\mathbf{z}_2\|_2\ldots\|\mathbf{z}_N\|_2)^2.$$

Then, the minimizer of (2.3) is $\sigma = \mathrm{GRQ}(\mathbf{z}_1,\mathbf{z}_2,\ldots,\mathbf{z}_N)$ and the minimum value is

$$\|\mathcal{A}\|_F^2 - \mathrm{GRQ}(\mathbf{z}_1,\mathbf{z}_2,\ldots,\mathbf{z}_N)^2.$$

Hence, this proof is completed. □

It follows that the minimization problem (2.3) is equivalent to the problem of maximizing the absolute value of the GRQ

$$\max\left|\frac{1}{2}\cdot\left(\mathcal{A}\times_1\mathbf{z}_1^*\times_2\mathbf{z}_2^*\cdots\times_N\mathbf{z}_N^* + \overline{\mathcal{A}}\times_1\mathbf{z}_1^\top\times_2\mathbf{z}_2^\top\ldots\times_N\mathbf{z}_N^\top\right)\right| \tag{3.1}$$

under the constraints that $\|\mathbf{z}_n\|_2 = 1$ with $n = 1, 2, \ldots, N$. As we know, the above optimization problem is equivalent to

$$\max\frac{1}{2}\cdot\left(\mathcal{A}\times_1\mathbf{z}_1^*\times_2\mathbf{z}_2^*\cdots\times_N\mathbf{z}_N^* + \overline{\mathcal{A}}\times_1\mathbf{z}_1^\top\times_2\mathbf{z}_2^\top\cdots\times_N\mathbf{z}_N^\top\right) \tag{3.2}$$

under the constraints that $\|\mathbf{z}_n\|_2 = 1$ with $n = 1, 2, \ldots, N$; or

$$\min\frac{-1}{2}\cdot\left(\mathcal{A}\times_1\mathbf{z}_1^*\times_2\mathbf{z}_2^*\cdots\times_N\mathbf{z}_N^* + \overline{\mathcal{A}}\times_1\mathbf{z}_1^\top\times_2\mathbf{z}_2^\top\cdots\times_N\mathbf{z}_N^\top\right) \tag{3.3}$$

under the constraints that $\|\mathbf{z}_n\|_2 = 1$ with $n = 1, 2, \ldots, N$.

We write the Lagrangian function of the maximization problem (3.2) as

$$\begin{aligned}\mathfrak{L} &= \frac{1}{2}\cdot\left(\mathcal{A}\times_1\mathbf{z}_1^*\times_2\mathbf{z}_2^*\cdots\times_N\mathbf{z}_N^* + \overline{\mathcal{A}}\times_1\mathbf{z}_1^\top\times_2\mathbf{z}_2^\top\cdots\times_N\mathbf{z}_N^\top\right) \\ &\quad - \sum_{n=1}^N\frac{\mu_n}{2}(\|\mathbf{z}_n\|_2^2 - 1).\end{aligned}$$

By differentiating $\mathfrak{L}$ with respect to $\mathbf{z}_n$ and $\overline{\mathbf{z}}_n$ separately, we obtain the following system at a critical point $\mathbf{z}_n$:

$$F(\mathbf{z}_1,\mathbf{z}_2,\ldots,\mathbf{z}_N)_{-n} = \mu_n\mathbf{z}_n,\quad G(\mathbf{z}_1,\mathbf{z}_2,\ldots,\mathbf{z}_N)_{-n} = \mu_n\overline{\mathbf{z}}_n, \tag{3.4}$$

where

$$\begin{cases}F(\mathbf{z}_1,\mathbf{z}_2,\ldots,\mathbf{z}_N)_{-n} = \mathcal{A}\times_1\mathbf{z}_1^*\ldots\times_{n-1}\mathbf{z}_{n-1}^*\times_{n+1}\mathbf{z}_{n+1}^*\ldots\times_N\mathbf{z}_N^*,\\ G(\mathbf{z}_1,\mathbf{z}_2,\ldots,\mathbf{z}_N)_{-n} = \overline{\mathcal{A}}\times_1\mathbf{z}_1^\top\ldots\times_{n-1}\mathbf{z}_{n-1}^\top\times_{n+1}\mathbf{z}_{n+1}^\top\ldots\times_N\mathbf{z}_N^\top.\end{cases}$$

When we multiply $\mathbf{z}_n^*$ and $\mathbf{z}_n^\top$ to each part of the equations (3.4) with $n = 1, 2, \ldots, N$, respectively, we can derive that $\mu_1 = \mu_2 = \ldots = \mu_N = \mathrm{GRQ}(\mathbf{z}_1,\mathbf{z}_2,\ldots,\mathbf{z}_N)$.

Let $\sigma = \mathrm{GRQ}(\mathbf{z}_1,\mathbf{z}_2,\ldots,\mathbf{z}_N)$, then we can rewrite the above system as [22,35,45]

$$\begin{aligned}F(\mathbf{z}_1,\mathbf{z}_2,\ldots,\mathbf{z}_N)_{-n} = \sigma\mathbf{z}_n,\quad G(\mathbf{z}_1,\mathbf{z}_2,\ldots,\mathbf{z}_N)_{-n} = \sigma\overline{\mathbf{z}}_n,\\ F(\mathbf{z}_1,\mathbf{z}_2,\ldots,\mathbf{z}_N) + G(\mathbf{z}_1,\mathbf{z}_2,\ldots,\mathbf{z}_N) = 2\sigma,\end{aligned} \tag{3.5}$$

where

$$\begin{cases}F(\mathbf{z}_1,\mathbf{z}_2,\ldots,\mathbf{z}_N) = \mathcal{A}\times_1\mathbf{z}_1^*\times_2\mathbf{z}_2^*\ldots\times_N\mathbf{z}_N^*,\\ G(\mathbf{z}_1,\mathbf{z}_2,\ldots,\mathbf{z}_N) = \overline{\mathcal{A}}\times_1\mathbf{z}_1^\top\times_2\mathbf{z}_2^\top\ldots\times_N\mathbf{z}_N^\top.\end{cases}$$

When $\mathcal{A}$ is the associated tensor corresponding to a multipartite pure state $|\psi\rangle$, we call the vector $\mathbf{z}_n$ as a mode-n *quantum eigenvector* of the complex tensor $\mathcal{A}$ with $\|\mathbf{z}_n\|_2 = 1$, corresponding to the *quantum eigenvalue* $\sigma \in [-1, 1]$. Meanwhile, we can also derive the first-order optimal condition for the dual problem (3.3), as shown in the nonlinear quantum eigenvalue problem (3.5). For the modulus largest quantum eigenvalue of a complex tensor $\mathcal{A} \in \mathbb{C}^{I_1\times I_2\times\cdots\times I_N}$, we have the following remark.

Remark 3.1. The modulus largest quantum eigenvalue ϱ of a tensor $\mathcal{A} \in \mathbb{C}^{I_1\times I_2\times\cdots\times I_N}$ is called the entanglement eigenvalue and we have

$$|\varrho| = \max|\langle\mathcal{A},\mathcal{B}\rangle|,$$

for all complex rank-one tensors $\mathcal{B} \in \mathbb{C}^{I_1\times I_2\times\cdots\times I_N}$.

Let an $(N+1)$-tuple $(\sigma, \mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_N)$ be a quantum eigenpair of a tensor $\mathcal{A} \in \mathbb{C}^{I_1\times I_2\times\cdots\times I_N}$ with $\|\mathbf{v}_n\|_2 = 1$ for $n = 1, 2, \ldots, N$, then we have the following results.

- The vector $\exp(\iota\theta_n)\mathbf{v}_n$ is the mode-n quantum eigenvector of the tensor $\mathcal{A}$, corresponding to the quantum eigenvalue σ with $\exp(\iota(\theta_1+\theta_2+\cdots+\theta_N)) = 1$ and $\theta_n \in (-\pi, \pi]$; the vector $\exp(\iota\theta_n)\mathbf{v}_n$ is the mode-n quantum eigenvector of the tensor $\mathcal{A}$, corresponding to the quantum eigenvalue $-\sigma$ with $\exp(\iota(\theta_1+\theta_2+\cdots+\theta_N)) = -1$ and $\theta_n \in (-\pi, \pi]$.
- If the 2-norm of the $\mathbf{v}_n$ is not one, then, the nonlinear quantum eigenvalue problem (3.5) should be rewritten as

$$\begin{cases}F(\mathbf{v}_1,\mathbf{v}_2,\ldots,\mathbf{v}_N)_{-n} = \sigma\left(\prod_{m=1, m\neq n}^N\|\mathbf{v}_m\|_2\right)\mathbf{v}_n/\|\mathbf{v}_n\|_2,\\ G(\mathbf{v}_1,\mathbf{v}_2,\ldots,\mathbf{v}_N)_{-n} = \sigma\left(\prod_{m=1, m\neq n}^N\|\mathbf{v}_m\|_2\right)\overline{\mathbf{v}}_n/\|\mathbf{v}_n\|_2,\end{cases}$$

  and

$$F(\mathbf{v}_1,\mathbf{v}_2,\ldots,\mathbf{v}_N) + G(\mathbf{v}_1,\mathbf{v}_2,\ldots,\mathbf{v}_N) = 2\sigma\left(\prod_{n=1}^N\|\mathbf{v}_m\|_2\right),$$

We establish an important property for the generalized Rayleigh quotient of the complex tensor $\mathcal{A} \in \mathbb{C}^{I_1 \times I_2 \times \ldots \times I_N}$.

Theorem 3.1. *Assume that the* $(N+1)$*-tuple* $(\sigma_*, \mathbf{u}_1, \ldots, \mathbf{u}_N)$ *is a nonzero solution to the nonlinear quantum eigenvalue problem (3.5). Then for N small perturbations* $\delta\mathbf{z}_n$*, we have*

$$\mathrm{GRQ}(\mathbf{u}_1 + \delta\mathbf{z}_1, \mathbf{u}_2 + \delta\mathbf{z}_1, \ldots, \mathbf{u}_N + \delta\mathbf{z}_N) = \sigma_* + \mathcal{O}(\|\delta\mathbf{z}_1\|_2^2 + \|\delta\mathbf{z}_2\|_2^2 + \ldots + \|\delta\mathbf{z}_N\|_2^2).$$

Proof. For clarity, let $\mathbf{w} = (\mathbf{z}_1, \mathbf{z}_2, \ldots, \mathbf{z}_N)$ and define

$$f(\mathbf{w}) = \mathrm{GRQ}(\mathbf{z}_1, \mathbf{z}_2, \ldots, \mathbf{z}_N) = \frac{1}{2} \cdot \frac{\mathcal{A} \times_1 \mathbf{z}_1^* \times_2 \mathbf{z}_2^* \cdots \times_N \mathbf{z}_N^* + \overline{\mathcal{A}} \times_1 \mathbf{z}_1^\top \times_2 \mathbf{z}_2^\top \cdots \times_N \mathbf{z}_N^\top}{\|\mathbf{z}_1\|_2 \|\mathbf{z}_2\|_2 \ldots \|\mathbf{z}_N\|_2}.$$

In order to prove the result, we need to introduce the first-order (complex multivariate) Taylor series of f at $\mathbf{w} = \mathbf{w}_*$ with $\mathbf{w}_* = (\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_N)$:

$$f(\mathbf{w}_* + \delta\mathbf{w}) = f(\mathbf{w}_*) + 2\Re\left[\left(\left.\frac{\partial f(\mathbf{w})}{\partial \mathbf{w}}\right|_{\mathbf{w}=\mathbf{w}_*}\right)^\top \delta\mathbf{w}\right] + \mathcal{O}(\|\delta\mathbf{w}\|_2^2). \tag{3.6}$$

According to the assumption, we have

$$f(\mathbf{w}_*) = \sigma_*, \quad \text{and} \quad \|\delta\mathbf{w}\|_2^2 = \|\delta\mathbf{z}_1\|_2^2 + \|\delta\mathbf{z}_2\|_2^2 + \ldots + \|\delta\mathbf{z}_N\|_2^2,$$

where $\delta\mathbf{w} = (\delta\mathbf{z}_1, \delta\mathbf{z}_2, \ldots, \delta\mathbf{z}_N)$.

Next, we prove that the second part on the right-hand side of the equation (3.6) is equal to zero. Since the pair $(\sigma_*, \mathbf{u}_1, \ldots, \mathbf{u}_N)$ is a nonzero solution to the quantum eigenvalue problem (3.5), then we have

$$F(\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_N)_{-n} = \sigma_* \mathbf{u}_n,$$
$$G(\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_N)_{-n} = \sigma_* \overline{\mathbf{u}}_n,$$
$$F(\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_N) + G(\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_N) = 2\sigma_*,$$

Note that the nonzero solution $(\sigma_*, \mathbf{u}_1, \ldots, \mathbf{u}_N)$ to the quantum eigenvalue problem (3.5) automatically guarantees that $\|\mathbf{u}_n\|_2 = 1$ for $n = 1, 2, \ldots, N$. Some tedious manipulation yields that the value of $\frac{\partial f(\mathbf{w})}{\partial \mathbf{w}}$ at $\mathbf{w}_*$ vainshes. Hence, this theorem is proved. □

4. Complex-valued neural networks for (3.1)

We propose a neural network model for computing the local optimal rank-one approximation of the tensors $\mathcal{A} \in \mathbb{C}^{I_1 \times I_2 \times \ldots \times I_N}$. Some properties for the states of the complex-valued neural networks are also presented.

4.1. Case I: complex symmetric tensors

A tensor $\mathcal{A} \in CT_{N,I}$ is called *symmetric*, if its entries $a_{i_1 i_2 \ldots i_N}$ are invariant under any permutation of their indices. Denote by $CST_{N,\,I}$ all complex symmetric N-order I-dimensional tensors.

For a given $\mathcal{A} \in CST_{N,I}$, the nonlinear quantum eigenvalue problem (3.5) can be reduced to the following nonlinear quantum eigenvalue problem

$$\mathcal{A}\overline{\mathbf{v}}^{N-1} = \sigma \mathbf{v}, \quad \overline{\mathcal{A}}\mathbf{v}^{N-1} = \sigma \overline{\mathbf{v}}, \tag{4.1}$$

where the vector $\mathbf{v} \in \mathbb{C}^I$ satisfies $\|\mathbf{v}\|_2 = 1$ with $\sigma \in \mathbb{R}$.

Let the pair $(\sigma, \mathbf{v})$ be any quantum eigenpair of the complex symmetric tensor $\mathcal{A}$. Then, it is easy to derive the following results.

- For even N, the vectors $-\mathbf{v}$, $\exp(2\iota\pi/N + 2k\pi\iota)\mathbf{v}$ and $\exp(-2\iota\pi/N + 2k\pi\iota)\mathbf{v}$ are the quantum eigenvectors of the tensor $\mathcal{A}$, corresponding to the quantum eigenvalue σ, the vectors $\exp(\iota\pi/N + 2k\pi\iota)\mathbf{v}$ and $\exp(-\iota\pi/N + 2k\pi\iota)\mathbf{v}$ are the quantum eigenvectors of the tensor $\mathcal{A}$, corresponding to the quantum eigenvalue $-\sigma$, with $k = 0, \pm 1, \pm 2, \ldots$.
- For odd N, the vectors $\exp(2\iota\pi/N + 2k\pi\iota)\mathbf{v}$ and $\exp(-2\iota\pi/N + 2k\pi\iota)\mathbf{v}$ are the quantum eigenvectors of the tensor $\mathcal{A}$, corresponding to the quantum eigenvalue σ, the vectors $-\mathbf{v}$, $\exp(\iota\pi/N + 2k\pi\iota)\mathbf{v}$ and $\exp(-\iota\pi/N + 2k\pi\iota)\mathbf{v}$ are the quantum eigenvectors of the tensor $\mathcal{A}$, corresponding to the quantum eigenvalue $-\sigma$, with $k = 0, \pm 1, \pm 2, \ldots$.
- If $\|\mathbf{v}\|_2 \neq 1$, then the nonlinear Eq. (4.1) should be rewritten as
$$\mathcal{A}\overline{\mathbf{v}}^{N-1} = \sigma \|\mathbf{v}\|_2^{N-2}\mathbf{v}, \quad \overline{\mathcal{A}}\mathbf{v}^{N-1} = \sigma \|\mathbf{v}\|_2^{N-2}\overline{\mathbf{v}}.$$
where the nonzero vector $\mathbf{v} \in \mathbb{C}^I$ with $\sigma \in \mathbb{R}$.

We shall use the fact that $\exp(\pm 2k\pi\iota) = 1$ and $\exp(\pm(2k-1)\pi\iota) = -1$ with $k = 0, \pm 1, \pm 2, \ldots$. The dynamics of the complex-valued neural network model for computing the quantum eigenvectors of a complex symmetric tensor $\mathcal{A} \in CST_{N,I}$ is described by

$$\frac{d\mathbf{z}(t)}{dt} = \mathcal{A}\overline{\mathbf{z}}(t)^{N-1} - \frac{\mathcal{A}\overline{\mathbf{z}}(t)^N + \overline{\mathcal{A}}\mathbf{z}(t)^N}{2\|\mathbf{z}(t)\|_2^2}\mathbf{z}(t), \tag{4.2}$$

or,

$$\frac{d\overline{\mathbf{z}}(t)}{dt} = \overline{\mathcal{A}}\mathbf{z}(t)^{N-1} - \frac{\mathcal{A}\overline{\mathbf{z}}(t)^N + \overline{\mathcal{A}}\mathbf{z}(t)^N}{2\|\mathbf{z}(t)\|_2^2}\overline{\mathbf{z}}(t),$$

for $t \geq 0$, where $\mathbf{z} = (z_1, z_2, \ldots, z_I)^\top \in \mathbb{C}^I$ represents the state of the network.

We have the following lemma to state the property for the solution of the neural network described by (4.2).

Lemma 4.1. *If the vector* $\mathbf{z}(t)$ *is a solution of the neural network described by (4.2) for all* $t \geq 0$*, then we have* $\|\mathbf{z}(t)\|_2^2 = \|\mathbf{z}(0)\|_2^2$ *for all* $t \geq 0$*, where the nonzero vector* $\mathbf{z}(0) \in \mathbb{C}^I$ *is any nonzero initial value.*

Proof. Suppose that $\mathbf{z}(t) = \mathbf{x}(t) + \iota\mathbf{y}(t)$ for all $t \geq 0$. Then, we have

$$\frac{d\mathbf{z}(t)}{dt} = \frac{\mathbf{x}(t)}{dt} + \iota\frac{\mathbf{y}(t)}{dt}, \quad \frac{d\overline{\mathbf{z}}(t)}{dt} = \frac{\mathbf{x}(t)}{dt} - \iota\frac{\mathbf{y}(t)}{dt},$$

that is,

$$\frac{d\mathbf{x}(t)}{dt} = \frac{1}{2}\left(\frac{d\mathbf{z}(t)}{dt} + \frac{d\overline{\mathbf{z}}(t)}{dt}\right), \quad \frac{d\mathbf{y}(t)}{dt} = \frac{\iota}{2}\left(\frac{d\overline{\mathbf{z}}(t)}{dt} - \frac{d\mathbf{z}(t)}{dt}\right).$$

According to the neural network described by (4.2), we have

$$\begin{aligned}\frac{d\mathbf{x}(t)}{dt} &= \frac{\mathcal{A}\overline{\mathbf{z}}(t)^{N-1} + \overline{\mathcal{A}}\mathbf{z}(t)^{N-1}}{2} - \frac{\mathcal{A}\overline{\mathbf{z}}(t)^N + \overline{\mathcal{A}}\mathbf{z}(t)^N}{4\|\mathbf{z}(t)\|_2^2}(\mathbf{z}(t) + \overline{\mathbf{z}}(t))\\ &= \Re(\overline{\mathcal{A}}\mathbf{z}(t)^{N-1}) - \frac{\Re(\overline{\mathcal{A}}\mathbf{z}(t)^N)}{\|\mathbf{z}(t)\|_2^2}\mathbf{x}(t),\\ \frac{d\mathbf{y}(t)}{dt} &= \iota\left(\frac{\overline{\mathcal{A}}\mathbf{z}(t)^{N-1} - \mathcal{A}\overline{\mathbf{z}}(t)^{N-1}}{2} - \frac{\mathcal{A}\overline{\mathbf{z}}(t)^N + \overline{\mathcal{A}}\mathbf{z}(t)^N}{4\|\mathbf{z}(t)\|_2^2}(\overline{\mathbf{z}}(t) - \mathbf{z}(t))\right)\\ &= -\Im(\overline{\mathcal{A}}\mathbf{z}(t)^{N-1}) - \frac{\Re(\overline{\mathcal{A}}\mathbf{z}(t)^N)}{\|\mathbf{z}(t)\|_2^2}\mathbf{y}(t).\end{aligned}$$

We have $\Re(\overline{\mathcal{A}}\mathbf{z}(t)^N) = \mathbf{x}(t)^\top\Re(\overline{\mathcal{A}}\mathbf{z}(t)^{N-1}) - \mathbf{y}(t)^\top\Im(\overline{\mathcal{A}}\mathbf{z}(t)^{N-1})$ and $\|\mathbf{z}(t)\|_2^2 = \|\mathbf{x}(t)\|_2^2 + \|\mathbf{y}(t)\|_2^2$. Hence, we have

$$\frac{d\|\mathbf{z}(t)\|_2^2}{dt} = 2\left(\mathbf{x}(t)^\top\frac{\mathbf{x}(t)}{dt} + \mathbf{y}(t)^\top\frac{\mathbf{y}(t)}{dt}\right) = 0.$$

The analytical solutions of the network described by (4.2) are given by $\|\mathbf{z}(t)\|_2^2 = \|\mathbf{z}(0)\|_2^2$. This completes the proof. □

4.2. Case II: general complex tensors

More generally, we now consider to design the complex-valued neural networks for solving the best complex rank-one approximation problem with a general complex tensor, which is not necessarily symmetric. Consider the complex-valued neural network model for computing the mode-n quantum eigenvectors of the tensor $\mathcal{A}$ is described by

$$\frac{d\mathbf{z}_n(t)}{dt} = F(\mathbf{z}_1, \mathbf{z}_2, \ldots, \mathbf{z}_N)_{-n} - \frac{F(\mathbf{z}_1, \mathbf{z}_2, \ldots, \mathbf{z}_N) + G(\mathbf{z}_1, \mathbf{z}_2, \ldots, \mathbf{z}_N)}{2\|\mathbf{z}_n\|_2^2}\mathbf{z}_n, \tag{4.3}$$

or,

$$\frac{d\bar{\mathbf{z}}_n(t)}{dt} = G(\mathbf{z}_1, \mathbf{z}_2, \ldots, \mathbf{z}_N)_{-n} - \frac{F(\mathbf{z}_1, \mathbf{z}_2, \ldots, \mathbf{z}_N) + G(\mathbf{z}_1, \mathbf{z}_2, \ldots, \mathbf{z}_N)}{2\|\mathbf{z}_n\|_2^2}\bar{\mathbf{z}}_n,$$

for all $t \geq 0$, where $\mathbf{z}_n = (z_{n,1}, z_{n,2}, \ldots, z_{n,I_n})^\top \in \mathbb{C}^{I_n}$ represents the state of the network with $n = 1, 2, \ldots, N$.

We present the following lemma that gives the property for the solution of the neural network described by (4.3).

Lemma 4.2. *If the vector $\mathbf{z}_n(t)$ is a solution of the neural network described by (4.3) for all $t \geq 0$, then we have $\|\mathbf{z}_n(t)\|_2^2 = \|\mathbf{z}_n(0)\|_2^2$ for all $t \geq 0$, where the nonzero vector $\mathbf{z}(0) \in \mathbb{C}^{I_n}$ is any nonzero initial value with $n = 1, 2, \ldots, N$.*

Proof. The process for proving this lemma is similar to that for Lemma 4.1. Hence, we omit its proof. □

For clarity, we suppose that $\|\mathbf{z}_n(0)\|_2 = 1$. By Lemma 4.2, we have $\|\mathbf{z}_n(t)\|_2 = 1$ for all $t \geq 0$ with $n = 1, 2, \ldots, N$. When the vector $\mathbf{z}_n(t)$ converges to a nonzero vector $\mathbf{u}_n \in \mathbb{C}^{I_n}$ as $t \to +\infty$, we have

$$\begin{cases} F(\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_N)_{-n} = \frac{F(\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_N) + G(\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_N)}{2}\mathbf{u}_n, \\ G(\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_N)_{-n} = \frac{F(\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_N) + G(\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_N)}{2}\bar{\mathbf{u}}_n, \end{cases}$$

that is, the vector $\mathbf{u}_n$ is a mode-n quantum eigenvector of the tensor $\mathcal{A}$, corresponding to the value $(F(\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_N) + G(\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_N))/2$.

5. Asymptotic stability analysis

In this section, we prove that the solutions of (4.2) and (4.3) are locally asymptotically stable in the sense of Lyapunov stability theory.

5.1. Case I: general complex tensors

For any given locally maximizer $(\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_N)$ of the maximization problem (3.1), we define a neighbourhood of $\mathbf{u}$ with $\mathbf{u} = (\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_N)$ as

$$\mathbb{B}(\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_N, \epsilon) := \mathbb{B}(\mathbf{u}, \epsilon) = \{\mathbf{z} = (\mathbf{z}_1, \mathbf{z}_2, \ldots, \mathbf{z}_N) := \|\mathbf{z} - \mathbf{u}\|_2 \leq \hat{\epsilon}\}$$

where $\mathbf{z}_n \in \mathbb{C}^{I_n}$ and $0 \leq \hat{\epsilon} \leq \epsilon_0$, with $\epsilon_0 = \min \|\mathbf{v} - \mathbf{u}\|_2$ for any other local maximizer $\mathbf{v} := (\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_N)$ of the maximization problem (3.1).

Let $\mathbf{z}_n = \mathbf{x}_n + \iota\mathbf{y}_n$ with $n = 1, 2, \ldots, N$. Since $F(\mathbf{z}_1, \mathbf{z}_2, \ldots, \mathbf{z}_N) + G(\mathbf{z}_1, \mathbf{z}_2, \ldots, \mathbf{z}_N)$ is continuous and differentiable with respect to its real and imaginary parts, then, there exists $0 \leq \epsilon \leq \hat{\epsilon}$ such that the sign of $F(\mathbf{z}_1, \mathbf{z}_2, \ldots, \mathbf{z}_N) + G(\mathbf{z}_1, \mathbf{z}_2, \ldots, \mathbf{z}_N)$ is the same as the sign of $F(\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_N) + G(\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_N)$ for all $\mathbf{z} = (\mathbf{z}_1, \mathbf{z}_2, \ldots, \mathbf{z}_N) \in \mathbb{B}(\mathbf{u}, \epsilon)$.

We obtain the following theorem to show that the solution of the network described by (4.3) is locally asymptotically stable in the sense of Lyapunov stability theory at any locally maximizer of the maximization problem (3.1).

Theorem 5.1. *Suppose that the N vectors $\mathbf{u}_n \in \mathbb{C}^{I_n}$ forms a locally maximizer of the optimal problem (3.1), with $\|\mathbf{u}_n\|_2 = 1$. If the initial value $\mathbf{z}_0(0)$ of the neural network described by (4.3) belongs to the set $\mathbb{B}(\exp(\iota\theta_1)\mathbf{u}_1, \exp(\iota\theta_2)\mathbf{u}_2, \ldots, \exp(\iota\theta_N)\mathbf{u}_N, \epsilon)$ with $\|\mathbf{z}_n(0)\|_2 = 1$, then the solution of the neural network described by (4.3) is locally asymptotically stable in the sense of Lyapunov stability theory at $(\exp(\iota\theta_1)\mathbf{u}_1, \exp(\iota\theta_2)\mathbf{u}_2, \ldots, \exp(\iota\theta_N)\mathbf{u}_N)$, where the N scalars $\theta_n \in (-\pi, \pi]$ satisfy $\exp(\iota(\theta_1 + \theta_2 + \ldots + \theta_N)) = 1$ with $F(\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_N) + G(\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_N) > 0$ and $\exp(\iota(\theta_1 + \theta_2 + \ldots + \theta_N)) = -1$ with $F(\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_N) + G(\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_N) < 0$.*

Proof. For simplicity, we assume that $N = 3$ and $\sigma_* = (F(\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3) + G(\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3))/2 > 0$. We can see that $F(\mathbf{z}_1, \mathbf{z}_2, \mathbf{z}_3) + G(\mathbf{z}_1, \mathbf{z}_2, \mathbf{z}_3) > 0$ for all $\mathbf{z} \in \mathbb{B}(\mathbf{u}, \epsilon)$ with $\mathbf{z} = (\mathbf{z}_1, \mathbf{z}_2, \mathbf{z}_3)$. For the network described by (4.3), we define the associated Lyapunov function as

$$V(\mathbf{z}_1, \mathbf{z}_2, \mathbf{z}_3) = \sigma_* - \frac{F(\mathbf{z}_1, \mathbf{z}_2, \mathbf{z}_3) + G(\mathbf{z}_1, \mathbf{z}_2, \mathbf{z}_3)}{2\|\mathbf{z}_1\|_2\|\mathbf{z}_2\|_2\|\mathbf{z}_3\|_2},$$

for all nonzero vectors $\mathbf{z}_n \in \mathbb{C}^{I_n}$ and $n = 1, 2, 3$.

Clearly, $V(\mathbf{z}_1, \mathbf{z}_2, \mathbf{z}_3) > 0$, where $\mathbf{z} \in \mathbb{B}(\mathbf{u}, \epsilon)$ and $\mathbf{z}_n = \exp(\iota\theta_n)\mathbf{u}_n$ with $\exp(\iota(\theta_1 + \theta_2 + \theta_3)) = 1$. Meanwhile, the function $V(\mathbf{z}_1, \mathbf{z}_2, \mathbf{z}_3) := V(\mathbf{x}_1, \mathbf{y}_1, \mathbf{x}_2, \mathbf{y}_2, \mathbf{x}_3, \mathbf{y}_3)$ can be viewed as a real continuous and differentiable function with respective to its real and imaginary parts.

We shall prove the local asymptotic stability of the point $(\mathbf{x}_1, \mathbf{y}_1, \mathbf{x}_2, \mathbf{y}_2, \mathbf{x}_3, \mathbf{y}_3)$ by the Lyapunov function $V(\mathbf{z}_1, \mathbf{z}_2, \mathbf{z}_3)$ for the network described by (4.3), that is, we deduce that the inequality

$$\sum_{n=1}^{3}\left(\left[\frac{d\mathbf{x}_n}{dt}\right]^\top \frac{\partial V(\mathbf{z}_1, \mathbf{z}_2, \mathbf{z}_3)}{\partial \mathbf{x}_n} + \left[\frac{d\mathbf{y}_n}{dt}\right]^\top \frac{\partial V(\mathbf{z}_1, \mathbf{z}_2, \mathbf{z}_3)}{\partial \mathbf{y}_n}\right) \leq 0$$

holds for all $(\mathbf{z}_1, \mathbf{z}_2, \mathbf{z}_3) \in \mathbb{B}(\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3, \epsilon)$, and the equality holds if and only if $(\mathbf{z}_1, \mathbf{z}_2, \mathbf{z}_3) = (\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3)$.

Without of loss generality, we derive the explicit expressions for $\frac{\partial V(\mathbf{z}_1, \mathbf{z}_2, \mathbf{z}_3)}{\partial \mathbf{x}_1}$ and $\frac{\partial V(\mathbf{z}_1, \mathbf{z}_2, \mathbf{z}_3)}{\partial \mathbf{y}_1}$. According to the expression of $V(\mathbf{z}_1, \mathbf{z}_2, \mathbf{z}_3)$, we obtain

$$\begin{cases} \frac{\partial V(\mathbf{z}_1, \mathbf{z}_2, \mathbf{z}_3)}{\partial \mathbf{z}_1} = -\left[\frac{G(\mathbf{z}_1, \mathbf{z}_2, \mathbf{z}_3)_{-1}}{\|\mathbf{z}_1\|_2\|\mathbf{z}_2\|_2\|\mathbf{z}_3\|_2} - \frac{F(\mathbf{z}_1, \mathbf{z}_2, \mathbf{z}_3) + G(\mathbf{z}_1, \mathbf{z}_2, \mathbf{z}_3)}{4\|\mathbf{z}_1\|_2^3\|\mathbf{z}_2\|_2\|\mathbf{z}_3\|_2}\bar{\mathbf{z}}_1\right], \\ \frac{\partial V(\mathbf{z}_1, \mathbf{z}_2, \mathbf{z}_3)}{\partial \bar{\mathbf{z}}_1} = -\left[\frac{F(\mathbf{z}_1, \mathbf{z}_2, \mathbf{z}_3)_{-1}}{\|\mathbf{z}_1\|_2\|\mathbf{z}_2\|_2\|\mathbf{z}_3\|_2} - \frac{F(\mathbf{z}_1, \mathbf{z}_2, \mathbf{z}_3) + G(\mathbf{z}_1, \mathbf{z}_2, \mathbf{z}_3)}{4\|\mathbf{z}_1\|_2^3\|\mathbf{z}_2\|_2\|\mathbf{z}_3\|_2}\mathbf{z}_1\right], \end{cases}$$

By using the fact $\mathbf{z}_1 = \mathbf{x}_1 + \iota\mathbf{y}_1$ and $\bar{\mathbf{z}}_1 = \mathbf{x}_1 - \iota\mathbf{y}_1$, we have

$$\begin{cases} \frac{\partial V(\mathbf{z}_1, \mathbf{z}_2, \mathbf{z}_3)}{\partial \mathbf{x}_1} = -\left[\frac{F(\mathbf{z}_1, \mathbf{z}_2, \mathbf{z}_3)_{-1} + G(\mathbf{z}_1, \mathbf{z}_2, \mathbf{z}_3)_{-1}}{\|\mathbf{z}_1\|_2\|\mathbf{z}_2\|_2\|\mathbf{z}_3\|_2} \right. \\ \left. - \frac{F(\mathbf{z}_1, \mathbf{z}_2, \mathbf{z}_3) + G(\mathbf{z}_1, \mathbf{z}_2, \mathbf{z}_3)}{4\|\mathbf{z}_1\|_2^3\|\mathbf{z}_2\|_2\|\mathbf{z}_3\|_2}(\bar{\mathbf{z}}_1 + \mathbf{z}_1)\right], \\ \frac{\partial V(\mathbf{z}_1, \mathbf{z}_2, \mathbf{z}_3)}{\partial \mathbf{y}_1} = -\iota\left[\frac{F(\mathbf{z}_1, \mathbf{z}_2, \mathbf{z}_3)_{-1} - G(\mathbf{z}_1, \mathbf{z}_2, \mathbf{z}_3)_{-1}}{\|\mathbf{z}_1\|_2\|\mathbf{z}_2\|_2\|\mathbf{z}_3\|_2} \right. \\ \left. - \frac{F(\mathbf{z}_1, \mathbf{z}_2, \mathbf{z}_3) + G(\mathbf{z}_1, \mathbf{z}_2, \mathbf{z}_3)}{4\|\mathbf{z}_1\|_2^3\|\mathbf{z}_2\|_2\|\mathbf{z}_3\|_2}(\overline{\mathbf{z}_1} - \mathbf{z}_1)\right]. \end{cases}$$

According to the network described by (4.3) with $N = 3$, we present the expressions for $d\mathbf{x}_1/dt$ and $d\mathbf{y}_1/dt$ as follows:

$$\begin{cases} \frac{d\mathbf{x}_1}{dt} = \frac{F(\mathbf{z}_1, \mathbf{z}_2, \mathbf{z}_3)_{-1} + G(\mathbf{z}_1, \mathbf{z}_2, \mathbf{z}_3)_{-1}}{2} \\ \quad - \frac{F(\mathbf{z}_1, \mathbf{z}_2, \mathbf{z}_3) + G(\mathbf{z}_1, \mathbf{z}_2, \mathbf{z}_3)}{4\mathbf{z}_1^*\mathbf{z}_1}(\mathbf{z}_1 + \bar{\mathbf{z}}_1), \\ \frac{d\mathbf{y}_1}{dt} = \iota\left[\frac{G(\mathbf{z}_1, \mathbf{z}_2, \mathbf{z}_3)_{-1} - F(\mathbf{z}_1, \mathbf{z}_2, \mathbf{z}_3)_{-1}}{2} \right. \\ \left. - \frac{F(\mathbf{z}_1, \mathbf{z}_2, \mathbf{z}_3) + G(\mathbf{z}_1, \mathbf{z}_2, \mathbf{z}_3)}{4\mathbf{z}_1^*\mathbf{z}_1}(\bar{\mathbf{z}}_1 - \mathbf{z}_1)\right]. \end{cases}$$

Some computation gets that

$$\left[\frac{d\mathbf{x}_1}{dt}\right]^\top \frac{\partial V(\mathbf{z}_1, \mathbf{z}_2, \mathbf{z}_3)}{\partial \mathbf{x}_1} = -\left[\Re(F(\mathbf{z}_1, \mathbf{z}_2, \mathbf{z}_3)_{-1}) - \frac{\Re(F(\mathbf{z}_1, \mathbf{z}_2, \mathbf{z}_3))}{\mathbf{z}_1^*\mathbf{z}_1}\mathbf{x}_1\right]^\top$$

$$\left[\Re(F(\mathbf{z}_1,\mathbf{z}_2,\mathbf{z}_3)_{-1})-\frac{\Re(F(\mathbf{z}_1,\mathbf{z}_2,\mathbf{z}_3))}{\mathbf{z}_1^*\mathbf{z}_1}\mathbf{x}_1\right]\leq 0,$$

or,

$$\left[\frac{d\mathbf{x}_1}{dt}\right]^\top\frac{\partial V(\mathbf{z}_1,\mathbf{z}_2,\mathbf{z}_3)}{\partial\mathbf{x}_1}=-\left[\Re(G(\mathbf{z}_1,\mathbf{z}_2,\mathbf{z}_3)_{-1})-\frac{\Re(F(\mathbf{z}_1,\mathbf{z}_2,\mathbf{z}_3))}{\mathbf{z}_1^*\mathbf{z}_1}\mathbf{x}_1\right]^\top\left[\Re(G(\mathbf{z}_1,\mathbf{z}_2,\mathbf{z}_3)_{-1})-\frac{\Re(F(\mathbf{z}_1,\mathbf{z}_2,\mathbf{z}_3))}{\mathbf{z}_1^*\mathbf{z}_1}\mathbf{x}_1\right]\leq 0;$$

and

$$\left[\frac{d\mathbf{y}_1}{dt}\right]^\top\frac{\partial V(\mathbf{z}_1,\mathbf{z}_2,\mathbf{z}_3)}{\partial\mathbf{y}_1}=-\left[\Im(F(\mathbf{z}_1,\mathbf{z}_2,\mathbf{z}_3)_{-1})-\frac{\Re(F(\mathbf{z}_1,\mathbf{z}_2,\mathbf{z}_3))}{\mathbf{z}_1^*\mathbf{z}_1}\mathbf{y}_1\right]^\top\left[\Im(F(\mathbf{z}_1,\mathbf{z}_2,\mathbf{z}_3)_{-1})-\frac{\Re(F(\mathbf{z}_1,\mathbf{z}_2,\mathbf{z}_3))}{\mathbf{z}_1^*\mathbf{z}_1}\mathbf{y}_1\right]\leq 0,$$

or,

$$\left[\frac{d\mathbf{y}_1}{dt}\right]^\top\frac{\partial V(\mathbf{z}_1,\mathbf{z}_2,\mathbf{z}_3)}{\partial\mathbf{y}_1}=-\left[\Im(G(\mathbf{z}_1,\mathbf{z}_2,\mathbf{z}_3)_{-1})+\frac{\Re(F(\mathbf{z}_1,\mathbf{z}_2,\mathbf{z}_3))}{\mathbf{z}_1^*\mathbf{z}_1}\mathbf{y}_1\right]^\top\left[\Im(G(\mathbf{z}_1,\mathbf{z}_2,\mathbf{z}_3)_{-1})+\frac{\Re(F(\mathbf{z}_1,\mathbf{z}_2,\mathbf{z}_3))}{\mathbf{z}_1^*\mathbf{z}_1}\mathbf{y}_1\right]\leq 0.$$

In the above two equations, we note that the equality holds if and only if $\mathbf{z}_n=\exp(\iota\tau_n)\mathbf{u}_n$, where the scalars $\tau_n\in(-\pi,\pi]$ satisfy $\exp(\iota(\tau_1+\tau_2+\tau_3))=1$, with $n=1,2,3$. Similarly, for the pair ($\mathbf{x}_2$, $\mathbf{y}_2$), some algebra yields that

$$\left[\frac{d\mathbf{x}_2}{dt}\right]^\top\frac{\partial V(\mathbf{z}_1,\mathbf{z}_2,\mathbf{z}_3)}{\partial\mathbf{x}_2}=-\left[\Re(F(\mathbf{z}_1,\mathbf{z}_2,\mathbf{z}_3)_{-2})-\frac{\Re(F(\mathbf{z}_1,\mathbf{z}_2,\mathbf{z}_3))}{\mathbf{z}_2^*\mathbf{z}_2}\mathbf{x}_2\right]^\top\left[\Re(F(\mathbf{z}_1,\mathbf{z}_2,\mathbf{z}_3)_{-2})-\frac{\Re(F(\mathbf{z}_1,\mathbf{z}_2,\mathbf{z}_3))}{\mathbf{z}_2^*\mathbf{z}_2}\mathbf{x}_2\right]\leq 0,$$

or,

$$\left[\frac{d\mathbf{x}_2}{dt}\right]^\top\frac{\partial V(\mathbf{z}_1,\mathbf{z}_2,\mathbf{z}_3)}{\partial\mathbf{x}_2}=-\left[\Re(G(\mathbf{z}_1,\mathbf{z}_2,\mathbf{z}_3)_{-2})-\frac{\Re(F(\mathbf{z}_1,\mathbf{z}_2,\mathbf{z}_3))}{\mathbf{z}_2^*\mathbf{z}_2}\mathbf{x}_2\right]^\top\left[\Re(G(\mathbf{z}_1,\mathbf{z}_2,\mathbf{z}_3)_{-2})-\frac{\Re(F(\mathbf{z}_1,\mathbf{z}_2,\mathbf{z}_3))}{\mathbf{z}_2^*\mathbf{z}_2}\mathbf{x}_2\right]\leq 0;$$

and

$$\left[\frac{d\mathbf{y}_2}{dt}\right]^\top\frac{\partial V(\mathbf{z}_1,\mathbf{z}_2,\mathbf{z}_3)}{\partial\mathbf{y}_2}=-\left[\Im(F(\mathbf{z}_1,\mathbf{z}_2,\mathbf{z}_3)_{-2})-\frac{\Re(F(\mathbf{z}_1,\mathbf{z}_2,\mathbf{z}_3))}{\mathbf{z}_2^*\mathbf{z}_2}\mathbf{y}_2\right]^\top\left[\Im(F(\mathbf{z}_1,\mathbf{z}_2,\mathbf{z}_3)_{-2})-\frac{\Re(F(\mathbf{z}_1,\mathbf{z}_2,\mathbf{z}_3))}{\mathbf{z}_2^*\mathbf{z}_2}\mathbf{y}_2\right]\leq 0,$$

or,

$$\left[\frac{d\mathbf{y}_2}{dt}\right]^\top\frac{\partial V(\mathbf{z}_1,\mathbf{z}_2,\mathbf{z}_3)}{\partial\mathbf{y}_2}=-\left[\Im(G(\mathbf{z}_1,\mathbf{z}_2,\mathbf{z}_3)_{-2})+\frac{\Re(F(\mathbf{z}_1,\mathbf{z}_2,\mathbf{z}_3))}{\mathbf{z}_2^*\mathbf{z}_2}\mathbf{y}_2\right]^\top\left[\Im(G(\mathbf{z}_1,\mathbf{z}_2,\mathbf{z}_3)_{-2})+\frac{\Re(F(\mathbf{z}_1,\mathbf{z}_2,\mathbf{z}_3))}{\mathbf{z}_2^*\mathbf{z}_2}\mathbf{y}_2\right]\leq 0;$$

For the pair ($\mathbf{x}_3$, $\mathbf{y}_3$), some tedious manipulation also leads to

$$\left[\frac{d\mathbf{x}_3}{dt}\right]^\top\frac{\partial V(\mathbf{z}_1,\mathbf{z}_2,\mathbf{z}_3)}{\partial\mathbf{x}_3}=-\left[\Re(F(\mathbf{z}_1,\mathbf{z}_2,\mathbf{z}_3)_{-3})-\frac{\Re(F(\mathbf{z}_1,\mathbf{z}_2,\mathbf{z}_3))}{\mathbf{z}_1^*\mathbf{z}_1}\mathbf{x}_3\right]^\top\left[\Re(F(\mathbf{z}_1,\mathbf{z}_2,\mathbf{z}_3)_{-3})-\frac{\Re(F(\mathbf{z}_1,\mathbf{z}_2,\mathbf{z}_3))}{\mathbf{z}_1^*\mathbf{z}_1}\mathbf{x}_3\right]\leq 0,$$

or,

$$\left[\frac{d\mathbf{x}_3}{dt}\right]^\top\frac{\partial V(\mathbf{z}_1,\mathbf{z}_2,\mathbf{z}_3)}{\partial\mathbf{x}_3}=-\left[\Re(G(\mathbf{z}_1,\mathbf{z}_2,\mathbf{z}_3)_{-3})-\frac{\Re(F(\mathbf{z}_1,\mathbf{z}_2,\mathbf{z}_3))}{\mathbf{z}_1^*\mathbf{z}_1}\mathbf{x}_3\right]^\top\left[\Re(G(\mathbf{z}_1,\mathbf{z}_2,\mathbf{z}_3)_{-3})-\frac{\Re(F(\mathbf{z}_1,\mathbf{z}_2,\mathbf{z}_3))}{\mathbf{z}_1^*\mathbf{z}_1}\mathbf{x}_3\right]\leq 0;$$

and

$$\left[\frac{d\mathbf{y}_3}{dt}\right]^\top\frac{\partial V(\mathbf{z}_1,\mathbf{z}_2,\mathbf{z}_3)}{\partial\mathbf{y}_3}=-\left[\Im(F(\mathbf{z}_1,\mathbf{z}_2,\mathbf{z}_3)_{-3})-\frac{\Re(F(\mathbf{z}_1,\mathbf{z}_2,\mathbf{z}_3))}{\mathbf{z}_1^*\mathbf{z}_1}\mathbf{y}_3\right]^\top\left[\Im(F(\mathbf{z}_1,\mathbf{z}_2,\mathbf{z}_3)_{-3})-\frac{\Re(F(\mathbf{z}_1,\mathbf{z}_2,\mathbf{z}_3))}{\mathbf{z}_1^*\mathbf{z}_1}\mathbf{y}_3\right]\leq 0,$$

or,

$$\left[\frac{d\mathbf{y}_3}{dt}\right]^\top\frac{\partial V(\mathbf{z}_1,\mathbf{z}_2,\mathbf{z}_3)}{\partial\mathbf{y}_3}=-\left[\Im(G(\mathbf{z}_1,\mathbf{z}_2,\mathbf{z}_3)_{-3})+\frac{\Re(F(\mathbf{z}_1,\mathbf{z}_2,\mathbf{z}_3))}{\mathbf{z}_1^*\mathbf{z}_1}\mathbf{y}_3\right]^\top\left[\Im(G(\mathbf{z}_1,\mathbf{z}_2,\mathbf{z}_3)_{-3})+\frac{\Re(F(\mathbf{z}_1,\mathbf{z}_2,\mathbf{z}_3))}{\mathbf{z}_1^*\mathbf{z}_1}\mathbf{y}_3\right]\leq 0.$$

We note that $\Re(F(\mathbf{z}_1,\mathbf{z}_2,\mathbf{z}_3))=\Re(G(\mathbf{z}_1,\mathbf{z}_2,\mathbf{z}_3))$ and $\Im(F(\mathbf{z}_1,\mathbf{z}_2,\mathbf{z}_3))=-\Im(G(\mathbf{z}_1,\mathbf{z}_2,\mathbf{z}_3))$. Then, the solution of the network described by (4.3) is locally asymptotically stable in the sense of Lyapunov stability theory at ($\mathbf{u}_1$, $\mathbf{u}_2$, $\mathbf{u}_3$) with $F(\mathbf{u}_1,\mathbf{u}_2,\mathbf{u}_3)+G(\mathbf{u}_1,\mathbf{u}_2,\mathbf{u}_3)>0$.

When the scalars $\theta_n\in(-\pi,\pi]$ satisfy $\exp(\iota(\theta_1+\theta_2+\theta_3))=1$, we have

$$F(\exp(\iota\theta_1)\mathbf{u}_1,\exp(\iota\theta_2)\mathbf{u}_2,\exp(\iota\theta_3)\mathbf{u}_3)+G(\exp(\iota\theta_1)\mathbf{u}_1,\exp(\iota\theta_2)\mathbf{u}_2,\exp(\iota\theta_3)\mathbf{u}_3)>0.$$

According to the above procedure, we can prove that the solution of the network described by (4.3) is locally asymptotically stable in the sense of Lyapunov stability theory at

$$(\exp(\iota\theta_1)\mathbf{u}_1,\exp(\iota\theta_2)\mathbf{u}_2),\exp(\iota\theta_3)\mathbf{u}_3).$$

Furthermore, we show that the solution of the network described by (4.3) is locally asymptotically stable in the sense of Lyapunov stability theory at $(\exp(\iota\theta_1)\mathbf{u}_1,\exp(\iota\theta_2)\mathbf{u}_2,\dots,\exp(\iota\theta_N)\mathbf{u}_N)$, where the N scalars $\theta_n \in (-\pi, \pi]$ satisfy $\exp(\iota(\theta_1 + \theta_2 + \ldots + \theta_N)) = 1$ with $F(\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_N) + G(\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_N) > 0$.

Now, we suppose that $\sigma_* := F(\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_N) + G(\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_N) < 0$. According to the analysis in Section 3, there exist N vectors $\mathbf{v}_n = \exp(\iota\tau_n)\mathbf{u}_n$ such that $F(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_N) + G(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_N) = -\sigma_* > 0$, where $\exp(\iota(\tau_1 + \tau_2 + \cdots + \tau_N)) = -1$ with $\tau_n \in (-\pi, \pi]$ for $n = 1, 2, \ldots, N$. For the network described by (4.3), the associated Lyapunov function is defined as

$$V(\mathbf{z}_1, \mathbf{z}_2, \ldots, \mathbf{z}_N) = -\sigma_* - \frac{F(\mathbf{z}_1, \mathbf{z}_2, \ldots, \mathbf{z}_N) + G(\mathbf{z}_1, \mathbf{z}_2, \ldots, \mathbf{z}_N)}{2\|\mathbf{z}_1\|_2\|\mathbf{z}_2\|_2 \ldots \|\mathbf{z}_N\|_2}.$$

For the case of $F(\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_N) + G(\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_N) > 0$, we can prove that the solution of the network described by (4.3) is locally asymptotically stable in the sense of Lyapunov stability theory at $(\exp(\iota\tau_1)\mathbf{u}_1, \exp(\iota\tau_2)\mathbf{u}_2, \ldots, \exp(\iota\tau_N)\mathbf{u}_N)$. Furthermore, we prove that the solution of the neural network described by (4.3) is locally asymptotically stable in the sense of Lyapunov stability theory at $(\exp(\iota\theta_1)\mathbf{u}_1, \exp(\iota\theta_2)\mathbf{u}_2, \ldots, \exp(\iota\theta_N)\mathbf{u}_N)$, where the N scalars $\theta_n \in (-\pi, \pi]$ satisfy $\exp(\iota(\theta_1 + \theta_2 + \cdots + \theta_N)) = -1$ with $F(\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_N) + G(\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_N) < 0$. □

5.2. Case II: complex symmetric tensors

In Section 5.1, we analyzed that the solution of the network described by (4.3) is locally asymptotically stable in the sense of Lyapunov stability theory. The main result is summarized in Theorem 5.1. In this subsection, we explore a special case of general complex tensors, that is, complex symmetric tensors. The goal is to analyze the local asymptotic stability in the sense of Lyapunov stability theory for the neural network described by (4.2).

For a given $\mathcal{A} \in CST_{N,I}$, the maximization problem (3.1) can be reduced by

$$\max |f(\mathbf{z})|, \quad \text{with} \quad f(\mathbf{z}) = \frac{\mathcal{A}\overline{\mathbf{z}}^N + \overline{\mathcal{A}}\mathbf{z}^N}{2} \tag{5.1}$$

under the constraint that $\|\mathbf{z}\|_2 = 1$ with $\mathbf{z} \in \mathbb{C}^I$. Suppose that $\mathbf{u} \in \mathbb{C}^I$ is a local maximizer of the maximization problem (5.1), we define a neighborhood of the vector $\mathbf{u}$ as

$$\mathbb{B}(\mathbf{u}, \widehat{\epsilon}) = \{\mathbf{z} \in \mathbb{C}^I : \|\mathbf{z} - \mathbf{u}\|_2 \leq \widehat{\epsilon}\},$$

for all $0 < \widehat{\epsilon} \leq \epsilon_0 := \min \|\mathbf{u} - \mathbf{v}\|_2$, where the vector $\mathbf{v}$ is also a local maximizer of (5.1) with $\mathbf{v} \neq \mathbf{u}$. Then, there exists $0 \leq \epsilon \leq \widehat{\epsilon}$ such that the sign of $f(\mathbf{z})$ is the same as the sign of $f(\mathbf{u})$ for all $\mathbf{z} \in \mathbb{B}(\mathbf{u}, \epsilon)$. All local maximizers of the maximization problem (5.1) are the quantum eigenvectors of the tensor $\mathcal{A}$.

The following theorem states that the solution of the network described by (4.2) is locally asymptotically stable in the sense of Lyapunov stability theory at any local maximizer of the maximization problem (5.1).

Theorem 5.2. *Suppose that the vector $\mathbf{u} \in \mathbb{C}^I$ is a local maximizer of the maximization problem (5.1) with $\|\mathbf{u}\|_2 = 1$. If the initial value $\mathbf{z}(0)$ of the neural network described by (4.2) belongs to the set $\mathbb{B}(\exp(\iota\theta)\mathbf{u}, \epsilon)$, with $\|\mathbf{z}(0)\|_2 = 1$, then the solution of the neural network described by (4.2) is locally asymptotically stable in the sense of Lyapunov stability theory at $\exp(\iota\theta)\mathbf{u}$, where θ belongs to the set*

$$\{0, \pm\pi/N, \pm 2\pi/N, \ldots, \pm\pi\}.$$

Since the network described by (4.2) is a special case of the network described by (4.3), we omit a rigorous proof of Theorem 5.2. Instead, we give the Lyapunov function, corresponding to the network described by (4.2), as follows:

$$V(\mathbf{z}) = V(\mathbf{x}, \mathbf{y}) = \begin{cases} \sigma_* - \frac{\mathcal{A}\overline{\mathbf{z}}^N + \overline{\mathcal{A}}\mathbf{z}^N}{2\|\mathbf{z}\|_2^N} \geq 0, & f(\mathbf{u}) > 0; \\ -\sigma_* - \frac{\mathcal{A}\overline{\mathbf{z}}^N + \overline{\mathcal{A}}\mathbf{z}^N}{2\|\mathbf{z}\|_2^N} \geq 0, & f(\mathbf{u}) \leq 0. \end{cases}$$

6. Numerical examples

In this section, the computations are implemented in Matlab Version 2013a and the Matlab Tensor Toolbox [6] on a laptop with Intel Core i5-4200M CPU (2.50GHz) and 8.00GB RAM. All floating point numbers in each example have four decimal digits. We suppose that $\epsilon = 1e-10$. We use functions “ttv” and “ttm” in [6] to implement the tensor-vector operation and the tensor-matrix operation.

In order to compute the mode-n quantum eigenvectors of a given tensor $\mathcal{A} \in \mathbb{C}^{I_1 \times I_2 \times \ldots \times I_N}$, we utilize the following first-order difference equations to approximate the network described by (4.3):

$$\mathbf{z}_n(k+1) = \mathbf{z}_n(k) + \alpha_k\Big(F(\mathbf{z}_1(k), \mathbf{z}_2(k), \ldots, \mathbf{z}_N(k))_{-n} - \frac{F(\mathbf{z}_1(k), \mathbf{z}_2(k), \ldots, \mathbf{z}_N(k)) + G(\mathbf{z}_1(k), \mathbf{z}_2(k), \ldots, \mathbf{z}_N(k))}{2\|\mathbf{z}_n(k)\|_2^2}\mathbf{z}_n(k)\Big), \tag{6.1}$$

where α_k is a learning rate, for any nonzero vectors $\mathbf{z}_n(0) \in \mathbb{C}^{I_n}$.

According to Lemma 4.2, we have that $\|\mathbf{z}_n(0)\|_2 = 1$ implies $\|\mathbf{z}_n(k)\|_2 = 1$. Then, for the pair $(\sigma(k), \mathbf{z}_1(k), \ldots, \mathbf{z}_N(k))$ with $\sigma(k) = (F(\mathbf{z}_1(k), \mathbf{z}_2(k), \ldots, \mathbf{z}_N(k)) + G(\mathbf{z}_1(k), \mathbf{z}_2(k), \ldots, \mathbf{z}_N(k)))/2$, we define

$$\mathrm{ERR}(k) := \max_{n=1,2,\ldots,N} \|F(\mathbf{z}_1(k), \mathbf{z}_2(k), \ldots, \mathbf{z}_N(k))_{-n} - \sigma(k)\mathbf{z}_n(k)\|_2.$$

If there exists a positive integer k_0 such that $\mathrm{ERR}(k) \leq \epsilon$ for all $k \geq k_0$, then we terminate the iteration scheme (6.1) and the pair $(\sigma(k), \mathbf{z}_1(k), \ldots, \mathbf{z}_N(k))$ is an approximate quantum eigenpair of the tensor $\mathcal{A}$. If the number of iterations reaches 10000, we just terminate. Meanwhile, according to Theorem 5.1, we have that $(\mathbf{z}_1(k), \ldots, \mathbf{z}_N(k))$ is a locally approximate maximizer of the maximization problem (3.1). The symbols $\sigma(k)$ and $\mathrm{ERR}(k)$ are also suitable for complex symmetric tensors and complex symmetric matrices.

Note that all initial values are randomly selected. Alteratively, we can also use the truncated HOSVD to generate starting points [14].

Example 6.1. In this example, the entries of the testing complex tensors are random variables with independent identically distributed entries, each distributed as a real Gaussian random variable of zero mean and variance. Meanwhile, the number of the testing tensors is 100.

We suppose that $\mathcal{A} \in \mathbb{C}^{10\times 10\times 10}$. Che et al. [9] designed a higher order power type method, denoted by *Power method* in this paper for comparison, for computing the best complex rank-one approximation of the tensor $\mathcal{A}$, which is similar to Algorithm 3.1 in [14] to derive the best rank-one approximation of any real tensor.

We compare the neural network described by (4.3) with *Power method* for computing the best complex rank-one approximation of the testing tensor on these two sides: iteration steps and CPU times. Simulation results are shown in Fig. 1.

Example 6.2. The following symmetric pure states are considered in [20,21],

$$|S(N,K)\rangle \equiv \sqrt{\frac{K!(N-K)!}{N!}} \sum_{\text{permutations}} |\underbrace{0\ldots0}_{K}\underbrace{1\ldots1}_{N-K}\rangle.$$

As the amplitudes are all positive, one can assume that the closest separable state is of the form

$$|\phi\rangle = \underbrace{(\sqrt{p}|0\rangle + \sqrt{1-p}|1\rangle) \otimes (\sqrt{p}|0\rangle + \sqrt{1-p}|1\rangle) \otimes \ldots \otimes (\sqrt{p}|0\rangle + \sqrt{1-p}|1\rangle)}_{N}$$

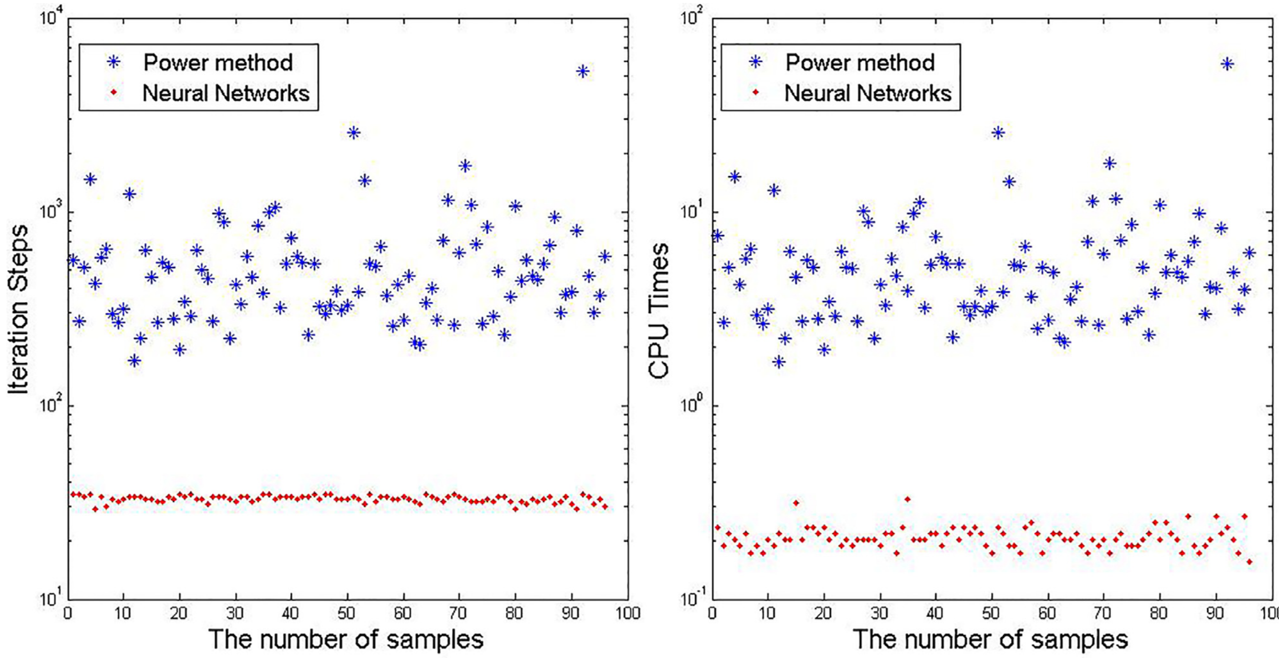

Fig. 1. Comparison for the network described by (4.3) with *Power method* to find the approximate maximizers of the maximization problem (3.1) with 100 random complex tensors.

Table 1
Relative results derived by the network described by (4.2) to find the local maximizers of the maximization problem (5.1) with $K = 0, 1, 2, 3, 4, 5$.

K	EEE	AMV	|EEE − AMV|	AIS	ACT (seconds)	Occurrence
0	1.0000	1.0000	1.7764e-15	240.0000	0.4153	19
1	0.6400	0.6400	5.1494e-12	170.1053	0.3359	18
2	0.5879	0.5879	7.3205e-12	182.1579	0.3109	19
3	0.5879	0.5879	4.1280e-12	180.1111	0.2995	18
4	0.6400	0.6400	7.9898e-11	205.3158	0.3544	18
5	1.0000	1.0000	1.3323e-15	257.0000	0.5383	20

for which the maximal overlap (with respect to p) gives the entanglement eigenvalue for $|S(N, K)\rangle$:

$$|\varrho| = \sqrt{\frac{K!(N-K)!}{N!}\left(\frac{K}{N}\right)^{K/2}\left(\frac{N-K}{N}\right)^{(N-K)/2}}.$$

For the given integers N and K, we use $\mathcal{A} \in CST_{N,2}$ to denote the tensor corresponding to the state $|S(N, K)\rangle$. To compute the entanglement eigenvalue of the state $|S(N, K)\rangle$ is to solve the maximization problem (5.1) with the tensor $\mathcal{A}$. When we fix N, we implement the network described by (4.2) to find the local maximizers of the maximization problem (5.1) 20 times, for each $K \in \{0, 1, \ldots, N\}$.

Set $N = 5$. For each $K \in \{0, 1, \ldots, 5\}$, the approximate maximal values (abbreviated as AMV) of the maximization problem (5.1), the exact entanglement eigenvalue (abbreviated as EEE) of the state $|S(5, K)\rangle$, the average iteration steps (abbreviated as AIS) and the average CPU times (abbreviated as ACT) are given in Table 1.

Example 6.3. The testing state is given as [45]

$$|W\widetilde{W}(s,\phi)\rangle \equiv \sqrt{s}|W\rangle + \sqrt{1-s}\exp(\iota\phi)|\widetilde{W}\rangle,$$

where the states $|W\rangle$ and $|\widetilde{W}\rangle$ are given as follows:

$$|W\rangle = (|001\rangle + |010\rangle + |100\rangle)/\sqrt{3}, \quad |\widetilde{W}\rangle = (|101\rangle + |011\rangle + |110\rangle)/\sqrt{3}.$$

The entanglement is independent of ϕ: the transformation $\{|0\rangle, |1\rangle\} \to \{|0\rangle, \exp(-\iota\phi)|1\rangle\}$ induces $|W\widetilde{W}(s,\phi)\rangle \to \exp(-\iota\phi)|W\widetilde{W}(s,0)\rangle$. To compute the entanglement eigenvalue of the state $|W\widetilde{W}(s,\phi)\rangle$, we assume that the separable state is

$$(\cos(\phi)|0\rangle + \sin(\phi)|1\rangle) \otimes (\cos(\phi)|0\rangle + \sin(\phi)|1\rangle) \otimes (\cos(\phi)|0\rangle + \sin(\phi)|1\rangle)$$

and maximize its overlap with $|W\widetilde{W}(s,0)\rangle$. We have $0 \le s \le 1$. The exact entanglement eigenvalue of the state $|W\widetilde{W}(s,\phi)\rangle$ can be expressed by

$$|\varrho| = \frac{1}{2}[\sqrt{s}\cos(\theta) + \sqrt{1-s}\sin(\theta)]\sin(2\theta),$$

where $t \equiv \tan(\theta) \in [\sqrt{1/2}, \sqrt{2}]$ is the particular root of the cubic polynomial equation

$$\sqrt{1-s}t^3 + 2\sqrt{s}t^2 - 2\sqrt{1-s}t - \sqrt{s} = 0.$$

For a given scalar $s \in [0, 1]$, we denote $\mathcal{A} \in CST_{3,2}$ by the tensor corresponding to the state $|W\widetilde{W}(s,0)\rangle$. In order to compute the entanglement eigenvalue of the state $|W\widetilde{W}(s,0)\rangle$, we need to find the modulus largest quantum eigenvalue of the tensor $\mathcal{A}$. We implement the network described by (4.2) to find the local maximizers of the maximization problem (5.1) 20 times, with $s \in \{0, 0.1, \ldots, 1\}$.

The approximate maximal values (abbreviated as AMV) of the maximization problem (5.1), the exact entanglement eigenvalue (abbreviated as EEE) of the state $|W\widetilde{W}(s,0)\rangle$, the average iteration steps (abbreviated as AIS) and the average CPU times (abbreviated as ACT) are given in Table 2, with $s \in \{0, 0.1, \ldots, 1\}$.

Table 2
Relative results derived by the network described by (4.2) to find the local maximizers of the maximization problem (5.1) with $s \in \{0, 0.1, \dots, 1\}$.

s	EEE	AMV	|EEE − AMV|	AIS	ACT (seconds)	Occurrence
0	0.6667	0.6667	1.7764e−15	162.4211	0.3150	20
0.1	0.7933	0.7933	4.1494e−14	1.1162e+3	2.1039	20
0.2	0.8306	0.8306	7.3215e−15	854.8500	1.7117	20
0.3	0.8514	0.8514	4.1380e−16	782.6500	1.5492	20
0.4	0.8625	0.8625	7.9093e−14	758.1000	1.3961	20
0.5	0.8660	0.8660	1.3001e−15	731.0500	1.4727	20
0.6	0.8625	0.8625	5.4123e−16	762.9500	1.4367	20
0.7	0.8514	0.8514	6.3301e−16	787.7500	1.5508	20
0.8	0.8306	0.8306	5.9822e−15	883.0500	1.6414	20
0.9	0.7933	0.7933	1.4567e−14	1.1737e+03	2.2773	20
1	0.6667	0.6667	2.0984e−14	162.2632	0.4441	20

Similar to the case of the state $|W\widetilde{W}(s,\phi)\rangle$, we can also use the network described by (4.2) to compute the entanglement eigenvalue of the state

$$|SS_{N;K_1K_2}(r,\phi)\rangle = \sqrt{r}|S(N,K_1)\rangle + \sqrt{1-r}\exp(\iota\phi)|S(N,K_2)\rangle,$$

where the states $|S(N,K_1)\rangle$ and $|S(N,K_2)\rangle$ are given in Example 6.2.

Example 6.4. The testing state is given as [45]:

$$|\psi(x,y,\phi_1,\phi_2,\phi_3)\rangle \equiv \sqrt{x}\exp(\iota\phi_1)|\mathrm{GHZ}\rangle + \sqrt{y}\exp(\iota\phi_2)|W\rangle + \sqrt{1-x-y}\exp(\iota\phi_3)|\widetilde{W}\rangle,$$

for all $\phi_i \in (-\pi,\pi]$ with $i=1,2,3$, where the states $|W\rangle$ and $|\widetilde{W}\rangle$ are given in Example 6.3, and the state $|\mathrm{GHZ}\rangle$ is $(|000\rangle + |111\rangle)/\sqrt{2}$. Similar to the description in Example 6.3, the entanglement eigenvalue of the state $|\psi(x,y,\phi_1,\phi_2,\phi_3)\rangle$ is equal to the entanglement eigenvalue of the state $|\psi(x,y)\rangle$, given as:

$$|\psi(x,y)\rangle \equiv \sqrt{x}|\mathrm{GHZ}\rangle + \sqrt{y}|W\rangle + \sqrt{1-x-y}|\widetilde{W}\rangle.$$

The entanglement eigenvalue of the state $|\psi(x,y)\rangle$ has been calculated, and one obtains

$$|\varrho| = \frac{1}{(1+t^2)^{3/2}}\left\{\sqrt{\frac{x}{z}}(1+t^3) + \sqrt{3y}t + \sqrt{3(1-x-y)}t^2\right\},$$

where t is the (unique) nonnegative real root of the following cubic polynomial equation:

$$3\sqrt{\frac{x}{2}}(-t+t^2) + \sqrt{3y}(-3t^2+1) + \sqrt{3(1-x-y)}(-t^3+2t) = 0.$$

We denote $\mathcal{A} \in CST_{3,2}$ by the tensor corresponding to the state $|\psi(x,y)\rangle$. When we apply the network described by (4.2) with the tensor $\mathcal{A}$, the approximate maximal values of the maximization problem (5.1), the iteration steps and the CPU times are shown in Fig. 2, with $0 \le x, y \le 1$ and $0 \le 1-x-y \le 1$.

The testing tensors in the above three examples are complex symmetric. In the following example we consider the entanglement eigenvalue of a complex nonsymmetric tensor.

Example 6.5. The testing state is given as [1]:

$$|\Psi\rangle = \sqrt{\lambda_1}|000\rangle + \sqrt{\lambda_2}\exp(\iota\phi)|100\rangle + \sqrt{\lambda_3}|101\rangle + \sqrt{\lambda_4}|110\rangle + \sqrt{\lambda_5}|111\rangle,$$

where $\phi \in [0,\pi]$ and $\lambda_1 + \lambda_2 + \cdots + \lambda_5 = 1$ with $\lambda_i \ge 0$. As we know, we can rewrite the state $|\Psi\rangle$ as

$$|\Psi\rangle = \sqrt{\lambda_1}|000\rangle + \sqrt{\lambda_2}\exp(\iota\phi)|100\rangle + \sqrt{\lambda_3}|101\rangle + \sqrt{\lambda_4}|110\rangle + \sqrt{1-\lambda_1-\cdots-\lambda_4}|111\rangle,$$

where $\phi \in [0,\pi]$ and $\lambda_i \ge 0$ with $i=1,2,3,4$.

We denote the tensor corresponding to the state $|\Psi\rangle$ by $\mathcal{A} \in CST_{3,2}$. In order to compute the entanglement eigenvalue of the state $|\Psi\rangle$, we need to find the modulus largest quantum eigenvalue (or the entanglement eigenvalue) of the tensor $\mathcal{A}$. We implement the network described by (4.3) to find the global maximizers of the maximization problem (3.1) for each 6-tuple $(\phi, \lambda_1, \lambda_2, \dots, \lambda_5)$.

On one hand, when the value of the 5-tuple $(\lambda_1, \lambda_2, \dots, \lambda_5)$ is chosen from the following two 5-tuples

$$\left\{\left(\frac{1}{15},\frac{2}{15},\frac{3}{15},\frac{4}{15},\frac{5}{15}\right),\left(\frac{1}{25},\frac{3}{25},\frac{5}{25},\frac{7}{25},\frac{9}{25}\right)\right\},$$

the relationship between the entanglement eigenvalue of the tensor $\mathcal{A}$ and the parameter ϕ is shown in Fig. 3. On the other hand, for any given scalar ϕ, the relationship between the entanglement eigenvalue of the tensor $\mathcal{A}$ and the 3-tuple $(\lambda_2, \lambda_3, \lambda_4)$ is shown in Figs. 4 (for the case of $\phi = \pi/4$) and 5 (for the case of $\phi = \pi/2$), for each $\lambda_1 \in \{0.1, 0.2, \dots, 0.9\}$ with $\lambda_i \ge 0$ and $1-\lambda_1-\lambda_2-\lambda_3-\lambda_4 \ge 0$.

Remark 6.1. For the above numerical examples, the learning rate in the iterative scheme (6.1) is chosen by 0.1. However, this learning rate may not be optimal.

Example 6.6. Consider a generic $I \times I \times I$ symmetric tripartite state [10] as

$$|\psi^{syt}\rangle := \sum_{i=0}^{I-1} a_i|i,i,i\rangle + \sum_{j>i=0}^{I-1} a_{ij}P_3(|i,j,j\rangle) + a_{ijk}\sum_{k>j>i=0}^{I-1} a_{ijk}P_3(|i,j,k\rangle),$$

where P_3's denote the set of all permutations on three parties again. For example, $P_3(|i,j,j\rangle) = |i,j,j\rangle + |j,i,j\rangle + |j,j,i\rangle$. In this example, we compare the network described by (4.3) with the numerical method in [10] for computing geometric measure of entanglement of $|\psi^{syt}\rangle$.

The main idea of the numerical method in [10] is to combine simulated annealing [26] with fmincon (a function in MATLAB attempts to find a constrained minimum of a scalar function of several variables starting from an initial estimate) to construct an algorithm for finding the global minimizer in the whole region. From [10], when considering the case of $\sum_{i=0}^{I-1}|a_i|^2 = 1$, $a_{ij} = a_{ijk} = 0$ for all $i,j,k = 0,1,\dots,I-1$, then the state $|\psi^{syt}\rangle$ is just the definition of GHZ states and the analytical result for geometric measure simply reads $\min\{1-|a_i|^2, i=0,1,\dots,I-1\}$. We denote $\mathcal{A} \in CST_{3,I}$ by the tensor corresponding to the state $|\psi^{syt}\rangle$. For different values of I and a_i, we have the following results.

(a) If $I=3$ and $\{a_0,a_1,a_2\} = \{0.8075-0.1790\iota,\ 0.5427-0.0203\iota, 0.0822-0.1202\iota\}$, then geometric measure of entanglement of $|\psi^{syt}\rangle$, derived from the numerical method in [10], is 0.172898. On the other hand, the largest entanglement eigenvalue, derived from (4.3), is 0.827102 and it costs 0.976574 seconds. So geometric measure of entanglement is 0.172898.

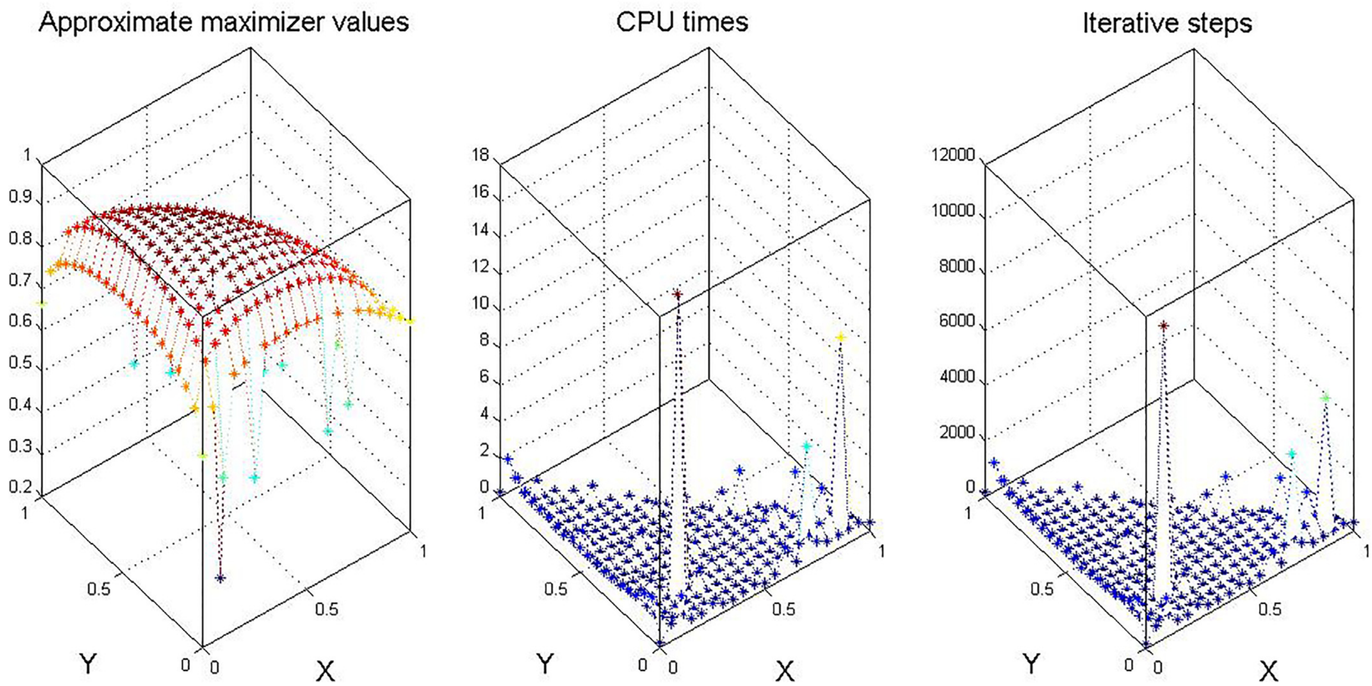

Fig. 2. Illustrations for the network described by (4.2) to find the approximate maximizers of the maximization problem (5.1) with $x, y = 0, 0.05, \ldots, 0.95, 1$ and $x + y \leq 1$.

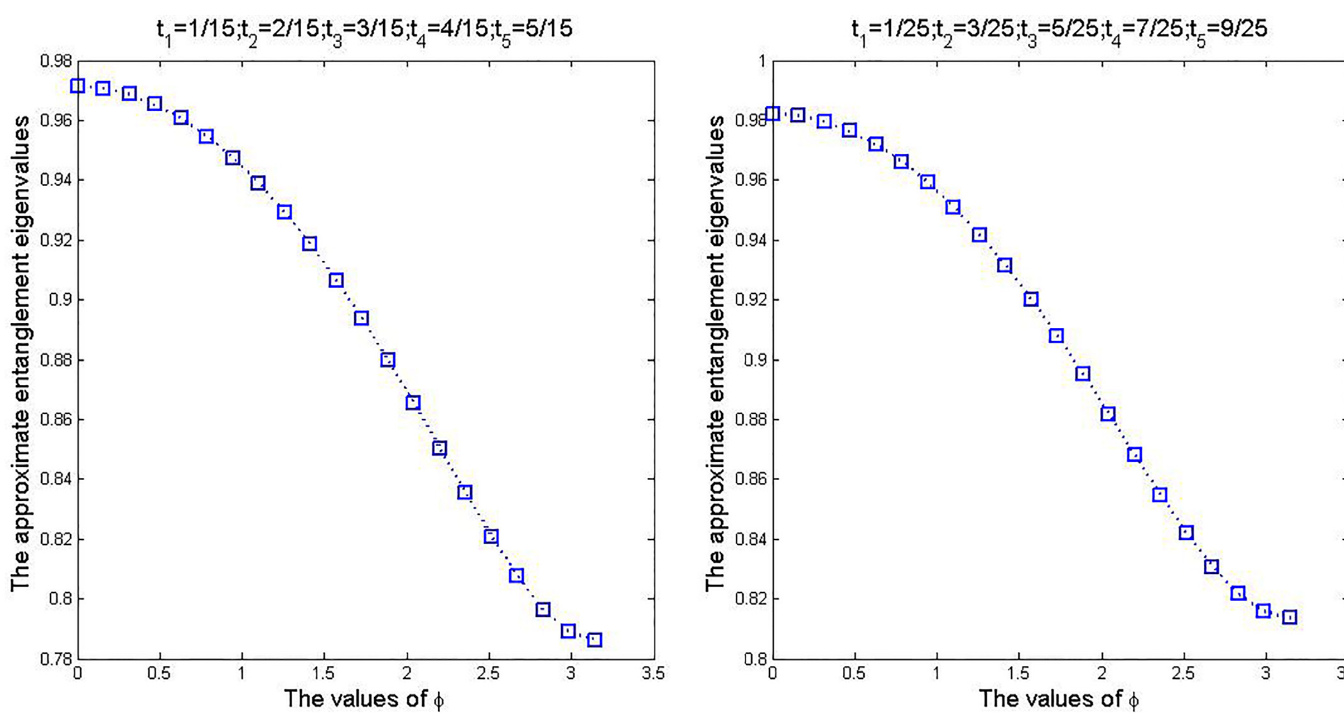

Fig. 3. Illustrations of the network described by (4.3) to find the approximate maximizers of the maximization problem (3.1) with $\phi \in \{0, \pi/20, \ldots, \pi\}$.

(b) If $I = 5$ and $\{a_0, a_1, a_2, a_3, a_4\} = \{0.6378 - 0.4942\iota, -0.5477 - 0.0567\iota, 0.2086 - 0.0457\iota, -0.0092 + 0.0041\iota, -0.0050 - 0.0018\iota\}$, then geometric measure of entanglement of $|\psi^{syt}\rangle$, derived from the numerical method in [10], is 0.193140. On the other hand, the largest entanglement eigenvalue, derived from (4.3), is 0.806860 and it costs 0.994064 seconds. So geometric measure of entanglement is 0.193140.

Next, we suppose that $I = 3$, $a_{210} = 1/\sqrt{6}$ and $a_i = a_{ij} = 0$ for all $i, j = 0, 1, 2$. This gives a 2-qutrit symmetric state whose geometric measure of entanglement has been derived in [45], which is

$$\sqrt{2/9} \approx 0.471405.$$

Geometric measure of entanglement of $|\psi^{syt}\rangle$, derived from the numerical method in [10], is 0.528596. On the other hand, the largest entanglement eigenvalue, derived from (4.3), is 0.471405 and it costs 0.5890 seconds. So geometric measure of entanglement is 0.528596.

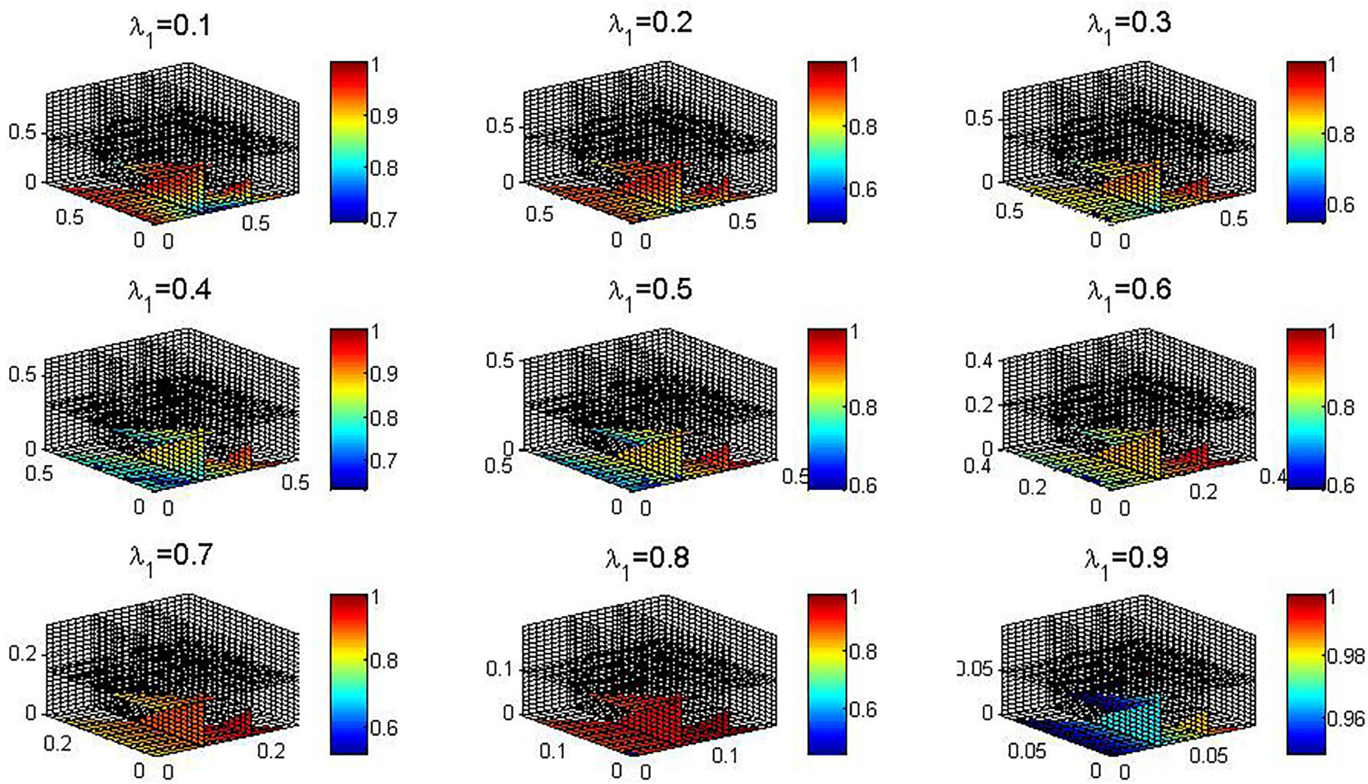

Fig. 4. Illustrations of the network (4.3) to find the approximate maximizers of the maximization problem (3.1) with $\lambda_1 \in \{0.1, 0.2, \ldots, 0.9\}$ and $\phi = \pi/4$.

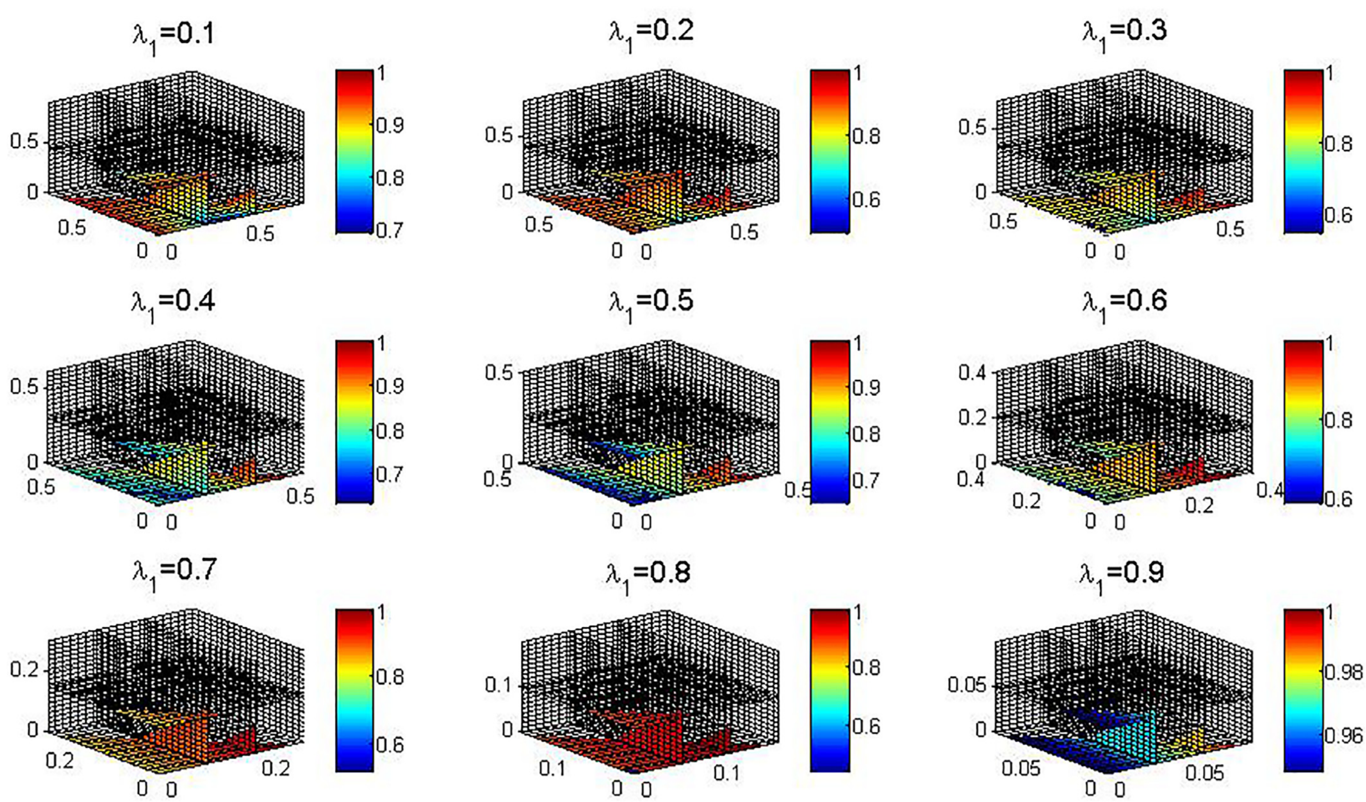

Fig. 5. Illustrations of the network (4.3) to find the approximate maximizers of the maximization problem (3.1) with $\lambda_1 \in \{0.1, 0.2, \ldots, 0.9\}$ and $\phi = \pi/2$.

7. Conclusion

In this paper, we consider how to compute the entanglement eigenvalue of multipartite pure states by the theory of the complex-valued neural networks. Meanwhile, we also prove that the solutions of the neural networks described by (4.3) are locally asymptotically stable in the sense of Lyapunov stability theory. Finally, we can find the global maximizers of the maximization problem (3.1) by the neural network described by (4.3) via the numerical examples with high probability. Unfortunately, we cannot present a rigorous proof for this statement, which will be our future research topic.

Acknowledgment

We would like to thank Editor-in-Chief Prof. Zidong Wang, the associate editor and the anonymous reviewers for their detailed and helpful comments.

References

[1] A. Acin, A. Andrianov, L. Costa, E. Jané, J. Latorre, R. Tarrach, Generalized schmidt decomposition and classification of three-quantum-bit states, Phys. Rev. Lett. 85 (7) (2000) 1560.
[2] I. Aizenberg, Complex-valued neural networks with multi-valued neurons, Studies in Computational Intelligence, 353, Springer-Verlag, Berlin, 2011, doi:10.1007/978-3-642-20353-4.
[3] M. Bohner, V.S.H. Rao, S. Sanyal, Global stability of complex-valued neural networks on time scales, Differ. Equ. Dyn. Syst. 19 (1–2) (2011) 3–11.
[4] P. Bouboulis, Wirtinger's calculus in general Hilbert spaces, arXiv:1005.5170(2010).
[5] D. Brandwood, A complex gradient operator and its application in adaptive array theory, IEE Proc. F: Commun. Radar Signal Process. 130 (1) (1983) 11–16.
[6] W.B. Brett, G.K. Tamara, Matlab tensor toolbox version 2.6, 2015, (Available online) http://www.sandia.gov/~tgkolda/TensorToolbox/
[7] D. Buhr, M. Carrington, T. Fugleberg, R. Kobes, G. Kunstatter, D. McGillis, C. Pugh, D. Ryckman, Geometrical entanglement of highly symmetric multipartite states and the Schmidt decomposition, J. Phys. A: Math. Theor. 44 (36) (2011) 365305.
[8] M. Che, A. Cichocki, Y. Wei, Neural networks for computing best rank-one approximations of real tensor and generalized tensor eigenvalue problems, Neurocomputing 267 (2017a) 114–133.
[9] M. Che, L. Qi, Y. Wei, Iterative algorithms for computing US- and U-eigenpairs of complex tensors, J. Comput. Appl. Math. 317 (2017b) 547–564.
[10] L. Chen, A. Xu, H. Zhu, Computation of the geometric measure of entanglement for pure multiqubit states, Phys. Rev. A 82 (3) (2010) 032301.
[11] L. Chen, H. Zhu, T.C. Wei, Connections of geometric measure of entanglement of pure symmetric states to quantum state estimation, Phys. Rev. A 83 (1) (2011) 012305.
[12] A. Cichocki, R. Unbehauen, Neural Networks for Optimization and Signal Processing, Wiley, New York, 1993.
[13] L. De Lathauwer, B. De Moor, J. Vandewalle, A multilinear singular value decomposition, SIAM J. Matrix Anal. Appl. 21 (4) (2000a) 1253–1278, doi:10.1137/S0895479896305696.
[14] L. De Lathauwer, B. De Moor, J. Vandewalle, On the best rank-1 and rank-$(r_1, r_2, \ldots, r_n)$ approximation of higher-order tensors, SIAM J. Matrix Anal. Appl. 21 (4) (2000b) 1324–1342.
[15] A. Defant, K. Floret, Tensor Norms and Operator Ideals, North-Holland Mathematics Studies, 176, North-Holland Publishing Co., Amsterdam, 1993.
[16] C.E. González-Guillén, Multipartite maximally entangled states in symmetric scenarios, Phys. Rev. A 86 (2) (2012) 022304.
[17] O. Gühne, G. Tóth, Entanglement detection, Phys. Rep. 474 (1) (2009) 1–75.
[18] M. Hayashi, A Group Theoretic Approach to Quantum Information, Springer, New York, 2017.
[19] M. Hayashi, D. Markham, M. Murao, M. Owari, S. Virmani, Bounds on multipartite entangled orthogonal state discrimination using local operations and classical communication, Phys. Rev. Lett. 96 (4) (2006) 040501.
[20] M. Hayashi, D. Markham, M. Murao, M. Owari, S. Virmani, Entanglement of multiparty-stabilizer, symmetric, and antisymmetric states, Phys. Rev. A 77 (1) (2008) 012104.
[21] M. Hayashi, D. Markham, M. Murao, M. Owari, S. Virmani, The geometric measure of entanglement for a symmetric pure state with non-negative amplitudes, J. Math. Phys. 50 (12) (2009) 122104.
[22] J.J. Hilling, A. Sudbery, The geometric measure of multipartite entanglement and the singular values of a hypermatrix, J. Math. Phys. 51 (7) (2010) 072102.
[23] R.A. Horn, C.R. Johnson, Matrix Analysis, 2nd, Cambridge University Press, Cambridge, 2012.
[24] J. Hu, J. Wang, Global stability of complex-valued recurrent neural networks with time-delays, IEEE Trans. Neural Netw. Learn. Syst. 23 (6) (2012) 853–865.
[25] S. Hu, L. Qi, G. Zhang, Computing the geometric measure of entanglement of multipartite pure states by means of non-negative tensors, Phys. Rev. A 93 (1) (2016) 012304.
[26] S. Kirkpatrick, C.D. Gelatt, M.P. Vecchi, Optimization by simulated annealing, Science 220 (4598) (1983) 671–680.
[27] E. Kofidis, P.A. Regalia, On the best rank-1 approximation of higher-order supersymmetric tensors, SIAM J. Matrix Anal. Appl. 23 (3) (2002) 863–884.
[28] T.G. Kolda, B.W. Bader, Tensor decompositions and applications, SIAM Rev. 51 (3) (2009) 455–500, doi:10.1137/07070111X.
[29] K. Kreutz-Delgado, The complex gradient operator and the CR-calculus, arXiv:0906.4835(2009).
[30] L.-Z. Liao, H. Qi, L. Qi, Neurodynamical optimization, J. Global Optim. 28 (2) (2004) 175–195.
[31] D. Markham, J. Anders, V. Vedral, M. Murao, A. Miyake, Survival of entanglement in thermal states, Europhys. Lett. 81 (4) (2008) 40006.
[32] D. Markham, A. Miyake, S. Virmani, Entanglement and local information access for graph states, New J. Phys. 9 (6) (2007) 194–199.
[33] Y. Nakata, D. Markham, M. Murao, Thermal robustness of multipartite entanglement of the one-dimensional spin-1/2 XY model, Phys. Rev. A 79 (4) (2009) 042313.
[34] G. Ni, M. Bai, Spherical optimization with complex variables for computing unitary symmetric eigenpairs, Comput. Optim. Appl. 65 (2016) 799–820.
[35] G. Ni, L. Qi, M. Bai, Geometric measure of entanglement and U-eigenvalues of tensors, SIAM J. Matrix Anal. Appl. 35 (1) (2014) 73–87, doi:10.1137/120892891.
[36] G. Ni, Y. Wang, On the best rank-1 approximation to higher-order symmetric tensors, Math. Comput. Model. 46 (2007) 1345–1352.
[37] J. Nie, L. Wang, Semidefinite relaxations for best rank-1 tensor approximations, SIAM J. Matrix Anal. Appl. 35 (3) (2014) 1155–1179.
[38] L. Qi, F. Wang, Y. Wang, *Z*-eigenvalue methods for a global polynomial optimization problem, Math. Progr. 118 (2, Ser. A) (2009) 301–316.
[39] L. Qi, G. Zhang, G. Ni, How entangled can a multi-party system possibly be? Phys. Lett. A 382 (22) (2018) 1465–1471, doi:10.1016/j.physleta.2018.04.007.
[40] R. Remmert, Theory of Complex Functions, Graduate Texts in Mathematics, 122, Springer-Verlag, New York, 1991, doi:10.1007/978-1-4612-0939-3.
[41] A. Rodriguezvazquez, R. Dominguezcastro, A. Rueda, J.L. Huertas, E. Sanchezsinencio, Nonlinear switched capacitor 'neural' networks for optimization problems, IEEE Trans. Circuits Syst. 37 (3) (1990) 384–398.
[42] A. Shimony, Degree of entanglementa, Ann. NY. Acad. Sci. 755 (1) (1995) 675–679.
[43] A. Van Den Bos, Complex gradient and Hessian, in: Proceedings of the IEEE Proceedings Vision, Image and Signal Processing, 141, 1994, pp. 380–383.
[44] X. Wang, M. Che, Y. Wei, Complex-valued neural networks for the Takagi vector of complex symmetric matrices, Neurocomputing 223 (2017) 77–85.
[45] T.-C. Wei, P.M. Goldbart, Geometric measure of entanglement and applications to bipartite and multipartite quantum states, Phys. Rev. A 68 (4) (2003) 042307.
[46] R.F. Werner, A.S. Holevo, Counterexample to an additivity conjecture for output purity of quantum channels, J. Math. Phys. 43 (9) (2002) 4353–4357.
[47] T. Zhang, G.H. Golub, Rank-one approximation to high order tensors, SIAM J. Matrix Anal. Appl. 23 (2) (2001) 534–550.
[48] H. Zhu, L. Chen, M. Hayashi, Additivity and non-additivity of multipartite entanglement measures, New J. Phys. 12 (8) (2010) 083002.

Maolin Che is a Lecturer of Southwest University of Finance and Economics in Chengdu of China. He obtained Ph.D. degree in computational mathematics from School of Mathematical Sciences at Fudan University in 2017. He received the B.S. degree in applied mathematcs from Neijiang Normal University of Sichuan in China, 2012. His current interests include numerical linear algebra and tensor theory and computation.

Liqun Qi received the B.S. degree at Tsinghua University, Beijing, China, in 1968 and the M.S. and Ph.D. degrees from the University of Wisconsin-Madison in 1981 and 1984, respectively. He has taught at Tsinghua University, the University of Wisconsin-Madison, the University of New South Wales, and the City University of Hong Kong. He is now the Chair Professor of applied mathematics of the Department of Applied Mathematics, The Hong Kong Polytechnic University. He has published more than 200 research papers in international journals and 2 monographs published by SIAM and Springer. His research interests include multilinear algebra, tensor computation, nonlinear numerical optimization and applications.

Yimin Wei received the B.S. degree in computational mathematics from Shanghai Normal University of Shang hai of China, and Ph.D. degree in computational mathematics from Fudan University of Shanghai in China, 1991 and 1997, respectively. He was a Lecturer with the Department of Mathematics, Fudan University, from 1997 to 2000, a Visiting Scholar with the Division of Engineering and Applied Science, Harvard University, Boston, MA, USA, from 2000 to 2001, and an Associate Professor with the School of Mathematical Sciences, Fudan University, from 2001 to 2005. He is currently a Professor with the School of Mathematical Sciences, Fudan University. His current research interests include multi-linear algebra numerical linear algebra and its applications, sensitivity in computational linear control, and perturbation analysis of the matrix and the tensor. He is author of more than 150 technical journal papers and 4 monographs published by Elsevier, Springer and Science Press.

Guofeng Zhang obtained Ph.D. in Applied Mathematics at The University of Alberta in 2005, received the B.S. degree and Msc from Northeastern University, in 1998 and 2000, respectively. He is an Associate Professor of applied mathematics of the Department of Applied Mathematics, The Hong Kong Polytechnic University. He has published more than 40 research papers in international journals.