Finite-time Convergent Complex-Valued Neural Networks for Computing Square Root of Complex Matrices

Zhaonian Pu, and Xuezhong Wang

***Abstract*—In this paper, we propose two complex-valued neural networks for finding complex matrix square root by constructing two new types of nonlinear activation functions. Theoretically, we prove that the complex-valued neural networks are globally stable in the sense of Lyapunov stability theory. The state matrix of the complex-valued neural networks converge to the theoretical complex matrix square root in finite time. Numerical simulations are presented to show the effectiveness of the complex-valued neural networks.**

***Index Terms*—Complex matrix; Square root; Finite time convergence; Nonlinear activation function; Complex-valued neural network.**

I. Introduction

THE, problem of matrix square root is widely encountered in many scientific areas [1], [2], [3], [4], [5], [6], [7], [8]. The square root and the matrix sign functions of complex matrices have several applications in systems and control theory. The square root of a positive definite matrix is required in many algorithms in signal processing and control [2]. In mathematics, in order to solve the matrix square root problem, almost all algorithms/schemes are based on the following defining equation [1], [2], [3], [4], [7], [9], [10], [11], [12], [13], [14], [15]:

$$X^2(t) = A, \tag{I.1}$$

where matrix $A \in \mathbb{C}^{n\times n}$ is assumed to be known. It is universally known that if $A \in \mathbb{C}^{n\times n}$ has no nonpositive real eigenvalues, then there is a unique solution X, which is denoted by $A^{1/2}$ and called the principal square root of A [16]. In fact, for given matrix $A \in \mathbb{C}^{n\times n}$, if $\frac{A+A^{\mathrm{H}}}{2}$, the real part of A, is positive definite, then A has a unique square root whose real part is positive definite. This result has been proved by Kato [17] in the more general setting of an infinite dimensional space using operator theory techniques. In this paper, we assume that the real part of $A \in \mathbb{C}^{n\times n}$ is positive definite and we are interested in the computation of the principal square root.

A number of computational methods have been proposed for finding the square root of a matrix [1], [2], [3], [4], [7], [9], [10], [11], [12], [13], [14], [15]. Being one of the most useful methods, Newton iteration [3], [4], [10] has been investigated for matrix square root finding, owing to its good properties of convergence and stability. After that, many improved algorithms from Newton iteration have further been developed and analyzed for solving matrix square root problems [1], [7], [10], such as simple form of Newton iteration [10], the Meini iteration [7], and the Denman and Beavers iteration [1]. However, these numerical algorithms may encounter serious speed bottleneck due to the serial nature of the digital computer, and may not be efficient enough for large-scale online applications [11]. As another important class of solution approach, many parallel-processing computational schemes have been developed, investigated, and implemented on specific architectures [11], [15] because of the parallel distributed nature. Especially, as a software and hardware implementable approach, recurrent neural network (RNN) has some potential advantages in real-time processing applications as compared to conventional numerical algorithms, such as adaptive ability, hardware realizability, and distributed-storage feature, and thus RNN has become a very active research topic in many fields [15], [18], [11], [12], [13], [14].

Recently, many authors have shown great interest for solving linear matrix equations, time-varying matrix equations and matrix square root on the basis of gradient-based neural networks (GNNs) [12], [13], [19], [20], [21], [22] or Zhang neural networks (ZNNs) [9], [11], [14], [15], [23]. The GNN approach uses the Frobenius norm of the error matrix as the performance criterion and defines a neural network evolving along the negative gradient-descent direction. In the time-varying case, the Frobenius norm of the error matrix cannot decrease to zero even after infinite time due to the lack of velocity compensation of time-varying coefficients [11], [15]. ZNNs are developed for solving online time-varying problems. Their dynamic is designed based on an indefinite error-monitoring function instead of a usually norm-based energy function. Compared with GNNs, a prominent advantage of the ZNNs solution lies in that the lagging error diminishes to zero exponentially as time t goes on [9], [11], [14], [15], [21]. It is well known that the design of ZNN is based on a matrix or vector-valued indefinite error function and an exponent-type formula, which makes every entry/element of the error function exponentially converge to zero. By defining different Zhang functions, a series of ZNNs models can be proposed for solving the same time-varying problem [9], [11], [15].

For given matrix $A \in \mathbb{R}^{n\times n}$ with no nonpositive real eigenvalues, Xiao [9], [11] has proposed real-valued finite time convergence ZNNs model with the sign-bi-power function to find the real matrix square root. In addition, the upper bound of convergence time for the proposed model

Z. Pu is withthe School of Mathematics and Statistics, Hexi University, Zhangye, 734000, P. R. of China. E-mail:puzhaonian@163.com.

X. Wang is with the School of Mathematics and Statistics, Hexi University, Zhangye, 734000, P. R. of China. E-mail:xuezhongwang77@126.com. This author is supported by the National Natural Science Foundation of China under grant 11771099. Natural Science Foundation of Gansu Province and Innovative Ability Promotion Project in Colleges and Universities of Gansu Province 2019B-146.

is theoretically derived and estimated by solving differential inequalities. However, if matrix $A \in \mathbb{C}^{n\times n}$ with the real part of A is positive definite, then A has a unique square root. In the general case, we need convert the complex matrix equation (I.1) into a real one for finding the square root of complex matrix $A \in \mathbb{C}^{n\times n}$. Thus, we have to solve matrix equation (I.1) in a double real-valued space. Theorefore, the main motivation and the novelty of us, in this paper, is to propose a complex-valued neural network for solving the matrix equation (I.1), where this neural network can avoid redundant computation in a double real-valued space and reduce a low model complexity and storage capacity.

Throughout this paper, we use $\|A\|_F$, $A^\top$, A^{H}, $\Re(A)$ and $\Im(A)$ to denote the Frobenius norm, the transpose, the complex conjugated transpose, the real part and the imaginary part of a given matrix $A \in \mathbb{C}^{m\times n}$, respectively. This notation is consistently used for lower-order parts of a given structure. For example, the entry with row index i and column index j in a matrix A, i.e., A_{ij}, is symbolized by a_{ij} (also $(\mathbf{x})_i = x_i$). Hence, we use $|A| = (|a_{kj}|)$, $\Theta(A) = (\Theta(a_{kj}))$ and $\exp(A) = (\exp(a_{kj}))$ denote the element-wise modulus, the element-wise argument and the element-wise exponential of the matrix $a \in \mathbb{C}^{m\times n}$, respectively. For two given matrices $A, B \in \mathbb{C}^{m\times n}$, $A \circ B$ denotes the Hadamard product of matrices A and B, i.e., $(A\circ B)_{ij} = a_{ij}b_{ij}$.

This paper is organized as follows. In Section II, we recall some preliminary results. Complex-valued neural network models with the weighted sign-bi-power activation functions for online solution of the time-varying complex matrix square root are presented in Section III. Convergence properties of the complex-valued neural network models will be discussed in Section IV. Illustrative numerical examples are presented in Section V.

Before ending this section, the main contributions of this paper are summarized and listed as follows:

1) This paper focuses on finding complex matrix square root in complex domain rather than conventionally investigated real matrix square root in real domain.
2) Two types of activation functions are constructed and two finite-time convergent complex-valued neural networks are proposed and investigated for online solution of the complex matrix square root finding in complex domain.
3) The paper carries out an in-depth theoretical analysis for our proposed ZNN models. It is theoretically proved that our models can converge to the theoretical solution of the complex matrix square root finding with in finite time. In addition, the upper bound of the convergence time are derived analytically via Lyapunov theory.

II. Preliminary

By Euler's formula, a complex number $\alpha \in \mathbb{C}$ can be represented as $\alpha = |\alpha|\exp(\iota\theta)$, where $\iota = \sqrt{-1}$ is imaginary unit and $\theta \in (-\pi, \pi]$ is the augment of the number α. Meanwhile, we can also rewrite a complex matrix $A \in \mathbb{C}^{m\times n}$ as $|A| \circ \exp(\iota\Theta(A))$.

The following two lemmas are needed to analyze the convergence and stability of the complex-valued neural networks.

Lemma II.1. [24] *The following identity holds for arbitrary time-varying complex matrix $Z(t) \in \mathbb{C}^{m\times p}$:*

$$\frac{\mathrm{d}Z^H(t)}{\mathrm{d}t} = \left(\frac{\mathrm{d}Z(t)}{\mathrm{d}t}\right)^H.$$

Lemma II.2. [24] *For any two time-varying complex matrices $Y(t) \in \mathbb{C}^{m\times n}, Z(t) \in \mathbb{C}^{n\times p}$, the next identity is satisfied:*

$$\frac{\mathrm{d}(Y(t)Z(t))}{\mathrm{d}t} = \frac{\mathrm{d}Y(t)}{\mathrm{d}t}Z(t) + Y(t)\frac{\mathrm{d}Z(t)}{\mathrm{d}t}.$$

For a given matrix $A \in \mathbb{R}^{m\times n}$, the function $\mathcal{F}(A)$ is defined to be element-wise applicable, odd and monotonically increasing, i.e., $\mathcal{F}(A) = (f(a_{kj}))$, with an odd and monotonically increasing sign-bi-power function [25] $f(\cdot)$, where

$$f(a_{kj}) = \mathrm{Lip}^{\sigma}(a_{kj}) + \mathrm{Lip}^{\frac{1}{\sigma}}(a_{kj}), \quad \sigma \in (0,1),$$

with

$$\mathrm{Lip}^{\sigma}(a_{kj}) = \begin{cases} a_{kj}^{\sigma}, & a_{kj} > 0, \\ 0, & a_{kj} = 0, \\ -a_{kj}^{\sigma}, & a_{kj} < 0. \end{cases}$$

Now, we construct two new activation functions to analyze the complex-valued neural networks for solving the equation (I.1). For a given complex matrix $A = \Re(A) + \iota\Im(A) \in \mathbb{C}^{m\times n}$, two types of the activation functions $\Psi_k(A) = (\psi_k(e_{ij}))$ $(k = 1, 2)$ are as follows:

(a) Type I activation function array is defined by

$$\Psi_1(\Re(A) + \iota\Im(A)) = \mathcal{F}(\Re(A)) + \iota\mathcal{F}(\Im(A)). \quad \text{(II.1)}$$

(b) Type II activation function array is defined by the expression

$$\Psi_2(\Re(A) + \iota\Im(A)) = \mathcal{F}(\Gamma) \circ \exp(\iota\Theta), \quad \text{(II.2)}$$

where $\Gamma = |A| \in \mathbb{R}^{m\times n}$ (resp. $\Theta = \Theta(A) \in \mathbb{R}^{m\times n}$) denotes element-wise modulus (resp. element-wise arguments) of the complex matrix A.

III. Neural network models based on ZNNs

Here, the nonlinear methods of ZNNs design for finite-time convergent complex-valued ZNNs model are presented. Then, by exploiting this method, two finite time convergent ZNNs model are first proposed for complex matrix square root finding based on two basic ZFs. For presentation convenience, such two ZNNs models are termed ZNN-I model and ZNN-II model.

As usual, the time derivative of the complex function $E(t)$ is denoted by $\dot{E}(t)$. The complex-valued neural network model is developed by employing three basic steps from [9], [11], [15]. An application of these steps in our case is described in the following.

Step 1. (Choose a convenient Zhang Function). The first step assumes definition of a proper fundamental matrix-valued error-monitoring function (Zhang Function, or ZF1 shortly) is defined as follows

$$E(t) = X^2(t) - A. \quad \text{(III.1)}$$

Step 2. (Define Zhang design formula). In the second step, with the aim to achieve global convergence of $E(t)$ to zero, it is necessary to use the general design pattern

$$\dot{E}(t) := \frac{\mathrm{d}E(t)}{\mathrm{d}t} = -\gamma\Psi_k(E(t)),\ k = 1, 2, \qquad \text{(III.2)}$$

where the design parameter $\gamma > 0$ corresponds to the inductance parameter or reciprocal of a capacitance parameter, and $\Psi_k(\cdot)$, $(k = 1, 2)$ denotes an especially constructed activation-function matrix mapping of neural networks.

In this paper, we apply the sign-bi-power activation function [25] to accelerate the ZNN to finite-time convergence to the complex matrix square root of the complex matrix.

Step 3. (Generate a ZNN model). In the last step, the dynamic equation of a complex neural network model for finding complex matrix square root of the complex matrix can be established by expanding (III.2). The complex matrix-valued error-monitoring function $E(t)$ defined in (III.1) possesses the following time derivative:

$$\dot{E}(t) = \dot{X}(t)X(t) + X(t)\dot{X}(t). \qquad \text{(III.3)}$$

Combining (III.2) and (III.3), we can obtain the following implicit dynamic equation of ZNN model for $k = 1, 2$:

$$\dot{X}(t)X(t) + X(t)\dot{X}(t) = -\gamma\Psi_k\left(X^2(t) - A\right), \qquad \text{(III.4)}$$

where $X(t)$, starting from initial state matrix $X(0) \in \mathbb{C}^{n\times n}$, denotes the state matrix corresponding to the theoretical matrix square root of equation (I.1).

For presentation convenience, if $k = 1$, we call above dynamic equation as ZNN-I model. If $k = 2$, we call above dynamic equation as ZNN-II model.

IV. Convergence analysis

In this section, we prove that both ZNN-I and ZNN-II can be globally convergent to the time-varying theoretical solution of equation (I.1).

A. Convergence of the model ZNN-I

The convergence performances in finite time as the sign-bi-power activation function of the ZNN-I model, defined on the basis of (II.1), is investigated in this subsection. The following result is valid for the ZNN-I model with the type I activation function.

Theorem IV.1. *Given matrix $A \in \mathbb{C}^{n\times n}$ with the real part of A is positive definite. If the Type I activation function is used, then the state matrix $X(t) \in \mathbb{C}^{n\times n}$ of the neural network* (III.4)*, starting from an arbitrary initial state $X(0) \in \mathbb{C}^{n\times n}$, converges to the theoretical matrix square root $X^*(t) \in \mathbb{C}^{n\times n}$ of the equation* (I.1) *in finite time:*

$$t_f < \frac{|e^+(0)|^{1-\sigma}}{\gamma(1-\sigma)},$$

where $e^+(0)$ is the largest element in

$$E(0) = X^2(0) - A.$$

Proof. According to the definition of $\Psi_1(\cdot)$, the following two equivalent formulae in the real numbers domain appear:

$$\Re(\dot{E}(t)) = -\gamma\mathcal{F}(\Re(E(t)))$$

and

$$\Im(\dot{E}(t)) = -\gamma\mathcal{F}(\Im(E(t))).$$

We construct the following Lyapunov function:

$$L(t) = \frac{\|E(t)\|_F^2}{2} = \frac{\mathrm{Tr}\left(E(t)^H E(t)\right)}{2},$$

where $\mathrm{Tr}(P) = \sum_{i=1}^n p_{ii}$ for any matrix $P \in \mathbb{C}^{n\times n}$. Since $E(t) = \Re(E(t)) + \iota\Im(E(t))$, the time derivative of $L(t)$ satisfies the following identities:

$$\begin{aligned}
\frac{\mathrm{d}L(t)}{\mathrm{d}t} &= \frac{\mathrm{Tr}\left(\dot{E}(t)^H E(t) + E(t)^H \dot{E}(t)\right)}{2}\\
&= -\frac{1}{2}\gamma\mathrm{Tr}\left\{\left(\mathcal{F}\left(\Re\left(E(t)\right)\right)^\top - \iota\mathcal{F}\left(\Im\left(E(t)\right)\right)^\top\right)\right.\\
&\quad\times\left(\Re\left(E(t)\right) + \iota\Im\left(E(t)\right)\right) + \left(\Re\left(E(t)\right)^\top - \iota\right.\\
&\quad\left.\left.\times\ \Im\left(E(t)\right)^\top\right)\left(\mathcal{F}(\Re\left(E(t)\right))\right)^\top + \iota\mathcal{F}(\Im\left(E(t)\right))\right\}\\
&= -\gamma\mathrm{Tr}\left\{\Re\left(E(t)\right)^\top\mathcal{F}\left(\Re\left(E(t)\right)\right)\right.\\
&\quad\left.+\Im\left(E(t)\right)^\top\mathcal{F}\left(\Im\left(E(t)\right)\right)\right\}.
\end{aligned}$$

Since $\mathcal{F}(\cdot)$ is odd and monotonically increasing, we conclude

$$\Re\left(E(t)\right)^\top\mathcal{F}\left(\Re\left(E(t)\right)\right) + \Im\left(E(t)\right)^\top\mathcal{F}\left(\Im\left(E(t)\right)\right) \geq 0,$$

and then $\frac{\mathrm{d}L(t)}{\mathrm{d}t} \leq 0$. According to the Lyapunov stability theory,

$$E(t) = X^2(t) - A$$

is globally convergent to zero matrix, regardless of the initial value. That is to say, as $t \to \infty$, we have $X^2(t) \to A$. In view of $\lambda \to 0$, the state matrix $X(t)$ globally converges to the time-varying theoretical solution of (I.1) starting from arbitrary initial state $X(0)$.

Next, it is necessary to prove the finite-time convergent performance of the ZNN-I model.

The initial value of the matrix valued error function $E(t)$ is

$$E(0) = X^2(0) - A.$$

We define

$$|e^+(0)| = \max\left\{|E(0)|\right\},$$

for all possible values of indices $i \in \{1, 2, \ldots, n\}$ and $j \in \{1, 2, \ldots, m\}$. This means that $|e_{ij}(t)|$ converges to zero for all possible $i \in \{1, 2, \ldots, n\}$ and $j \in \{1, 2, \ldots, m\}$ when $|e^+(t)|$ reach zero. In other words, the convergence time t_f of the ZNN-I model (III.4) is bounded by t_f^+ of the dynamics of $|e^+(t)|$, where t_f^+ represent the convergence time of the dynamics of $|e^+(t)|$.

It is clear that

$$\dot{e}^+ = -\gamma\psi\left(e^+(t)\right).$$

Another Lyapunov function candidate is defined as

$$l_+(t) = |e^+(t)|^2.$$

Since $l_+(t) \geq 0$, the time derivative of l_+ is equal to

$$\begin{aligned}
\dot{l}_+(t) &= -2\gamma e^+(t)\psi\left(e^+(t)\right)\\
&= -2\gamma\left(|e^+(t)|^{\sigma+1} + |e^+(t)|^{1/\sigma+1}\right)\\
&\leq -2\gamma|e^+(t)|^{\sigma+1}\\
&= -2\gamma\, l_+^{(\sigma+1)/2}(t).
\end{aligned}$$

The above results mean that, if the Type I activation function is adopted, neural state $X(t)$ of the neural network (III.4) with Ψ_1 converges to the theoretical matrix square root $X^*(t)$ of the equation (I.1) in finite time t_f. The proof is complete. □

B. Convergence of the model ZNN-II

In the following, we investigate the convergence of the complex neural network model ZNN-II, defined by (III.4) for $k = 2$. The following result can be verified about the complex-valued neural network model ZNN-II based on a type II activation function.

Theorem IV.2. *Given matrix $A \in \mathbb{C}^{n\times n}$ with the real part of A is positive definite. If the Type II activation function is used, then the state matrix $X(t) \in \mathbb{C}^{n\times n}$ of the neural network* (III.4)*, starting from an arbitrary initial state $X(0) \in \mathbb{C}^{n\times n}$, converges to the theoretical matrix square root $X^*(t) \in \mathbb{C}^{n\times n}$ of the equation* (I.1) *in finite time:*

$$t_f < \frac{|e^+(0)|^{1-\sigma}}{\gamma(1-\sigma)},$$

where $e^+(0)$ is the largest element in

$$E(0) = X^2(0) - A.$$

Proof. Analogically as in the proof of Theorem IV.1, the error dynamic is given by

$$\dot{E}(t) = -\gamma\Psi_2(E(t)),$$

where $E(t) = (A(t) + \lambda I)\, X(t) - B(t)$. According to the definition of $\Psi_2(\cdot)$, immediately follows

$$\Psi_2(E(t)) = \mathcal{F}(|E(t)|) \circ \exp(\iota\Theta(E(t))).$$

We construct the following Lyapunov function:

$$L(t) = \frac{\|E(t)\|_F^2}{2} = \frac{\mathrm{Tr}\left(E(t)^H E(t)\right)}{2},$$

which further implies

$$\begin{aligned}\frac{\mathrm{d}L(t)}{\mathrm{d}t} &= \frac{\mathrm{Tr}\left(E(t)^H\dot{E}(t) + \dot{E}(t)^H E(t)\right)}{2}\\ &= -\frac{1}{2}\gamma\mathrm{Tr}\left(E(t)^H\mathcal{H}_2\left(E(t)\right) + E(t)\mathcal{H}_2\left(E(t)\right)^H\right)\\ &= -\frac{1}{2}\gamma\mathrm{Tr}\left(E(t)^H\mathcal{H}_2\left(E(t)\right)\right.\\ &\quad\left. + \left(E(t)^H\mathcal{H}_2\left(E(t)\right)\right)^H\right)\\ &= -\gamma\mathrm{Tr}\left(\Re\left(E(t)^H\mathcal{H}_2\left(E(t)\right)\right)\right)\\ &= -\gamma\mathrm{Tr}\left\{\Re\left[E(t)^H\mathcal{F}\left(|E(t)|\right)\circ\exp\left(\iota\Theta(E(t))\right)\right]\right\}.\end{aligned}$$

Since $E(t) = |E(t)| \circ \exp(\iota\Theta(E(t)))$, one can verify

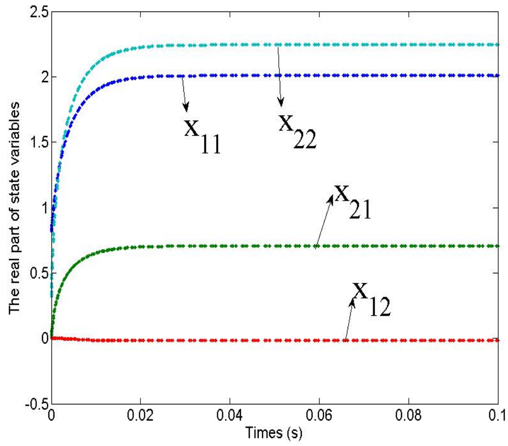
$$\begin{aligned}\frac{\mathrm{d}L(t)}{\mathrm{d}t} &= -\gamma\mathrm{Tr}\left\{\Re\left[\exp\left(-\iota\Theta\left(E(t)^H\right)\circ\left|E(t)^H\right|\right)\right.\right.\\ &\quad\left.\left.\times\ \left(\mathcal{F}\left(|E(t)|\right)\circ\exp\left(\iota\Theta\left(E(t)\right)\right)\right)\right]\right\}.\end{aligned}$$

Since $\mathcal{F}(\cdot)$ is monotonically increasing, we conclude $\mathcal{F}(|E(t)|) \geq 0$ for $E(t) \neq 0$. As a result, $L(t)$ is negative definite. According to the Lyapunov stability theory, the matrix $E(t) = (A(t) + \lambda I)\, X(t) - B(t)$ globally converges to the zero matrix from arbitrary initial value. Similarly as in the proof of Theorem IV.1, we conclude that the state matrix $X(t)$ globally converges to the time-varying theoretical solution of (I.1) starting from arbitrary initial state $X(0)$.

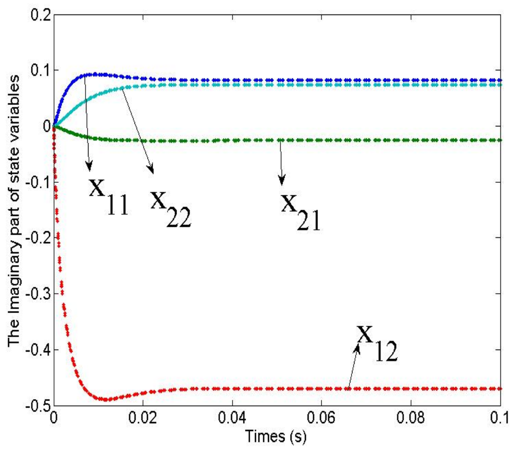
Because ZNN-I model (III.4), for $k = 2$, is derived by using the intrinsically nonlinear method of ZNN design similar to ZNN-I model, we also have

$$\dot{E}_X(t) = \dot{E}(t) = -\gamma\Psi_2(E(t)).$$

Therefore, the proof of finite time convergence can be generalized from the proof of Theorem IV.1 and is thus omitted. □



(a) The real part.



(b) The imaginary part.

Fig. 1. Trajectories of the state variables of the model ZNN-I in Example 1.

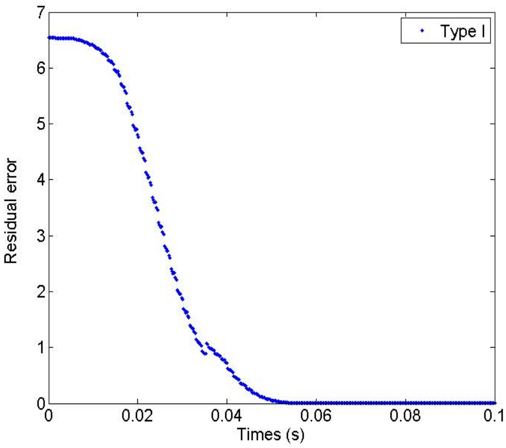
V. Numerical examples

In this section, we show that we can use the neural network (III.4) with different activation functions to solve the matrix principal square root via several examples. The computations are implemented in Matlab Version 2014a on a laptop with Intel Core i5-4200M CPU (2.50GHz) and 7.89GB RAM.

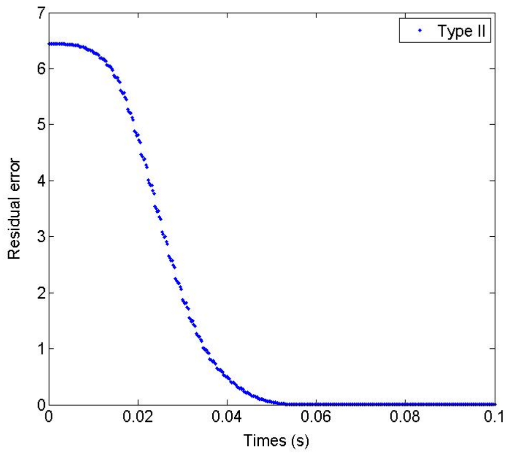
A. Numerical tests based on ZNN

Example 1. Consider the following matrix A with

$$A = \begin{pmatrix} 4 & -2\iota \\ 3 & 5 \end{pmatrix}.$$



(a) ZNN-I model.



(b) ZNN-II model.

Fig. 2. Trajectories of the residual errors of the model ZNN in Example 1.

It is easy to check that matrix $\frac{A+A^{\mathrm{H}}}{2}$ is an Hermite positive definite matrix. The theoretical principal square root X^* of the matrix A is presented as below for comparative purposes:

$$X^* = \begin{pmatrix} 2.0078 + 0.0824i & -0.0173 - 0.4699i \\ 0.7049 - 0.0259i & 2.2427 + 0.0737i \end{pmatrix}.$$

Starting from a randomly generated vector $u = rand(2,1)$, take $\gamma = 100$ and initial state $X(0) =$diag(u). If we choose the sign-bi-power activation function as $f(\cdot)$ with $\sigma = 0.2$, then state variables trajectories of real part generated by ZNN-I and imaginary part corresponding the model ZNN-II are shown in Figures 1 (a) and (b), respectively.

Trajectories of residual errors $\|X^2(t) - A\|_F$, generated by using the complex models ZNN-I and ZNN-II with $\gamma = 100$ are shown in Figure 2 (a) and (b), respectively.

Example 2. Consider the Toeplitz matrix A and has the following form:

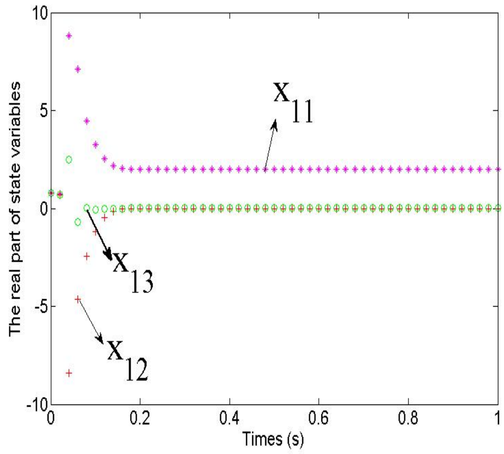
$$A = \begin{pmatrix} a_1 & a_2 & a_3 & \cdots & a_n \\ a_2 & a_1 & a_2 & \cdots & a_{n-1} \\ a_3 & a_2 & a_1 & \cdots & a_{n-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_n & a_{n-1} & a_{n-2} & \cdots & a_1 \end{pmatrix},$$

where $a_1 = n + 1$ and a_j $(j = 2, 3, \ldots, n)$ with the form

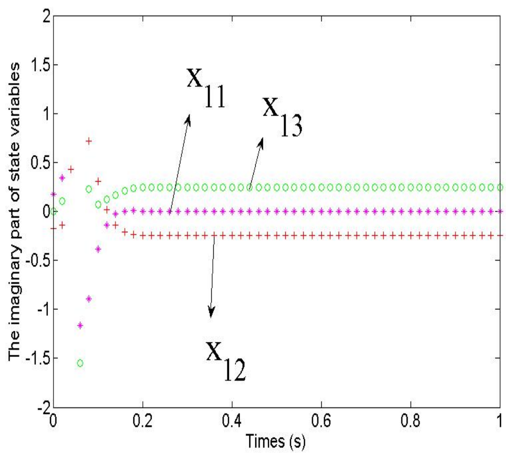
$$a_j = \begin{cases} 1, & \text{if } \mathrm{mod}(j,3) = 1, \\ -\iota, & \text{if } \mathrm{mod}(j,3) = 2, \\ \iota, & \text{if } \mathrm{mod}(j,3) = 0. \end{cases}$$

It is easy to check that matrix $\frac{A+A^{\mathrm{H}}}{2}$ is a complex symmetric positive definite matrix. We test the performance of models ZNN-I and ZNN-II for solving above matrix principal square root problem by choosing the sign-bi-power activation function as $f(\cdot)$ with $\sigma = 0.2$ and the initial state $X(0) = U + \iota V$ for randomly generated diagonal matrices U and V, respectively.

Firstly, we set $n = 3$ and adopt ZNN-II model to solve such a matrix square root problem under the conditions of $\gamma = 10$. The state variables trajectories of real part and imaginary part of the model ZNN-II are shown in Figures 3 (a) and (b), respectively.



(a) The first line.

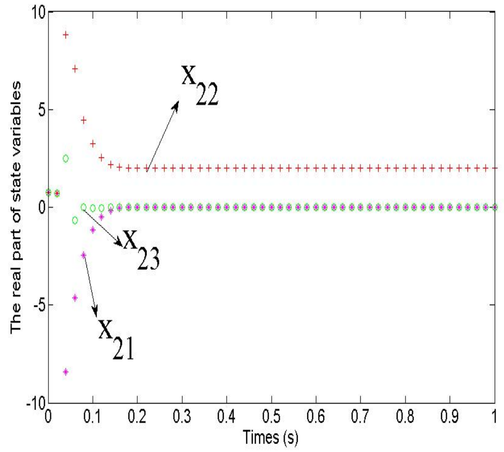


(b) The first line.

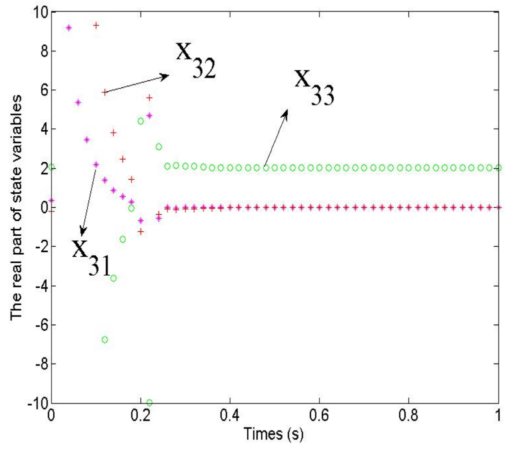
Secondly, we take $n = 11$, the trajectories of residual errors $\|X^2(t) - A\|_F$, generated by using the complex model ZNN-II with $\gamma = 10$, is shown in Figure 4 (a).

Finally, we take $n = 20$, the trajectories of residual errors $\|X^2(t) - A\|_F$, generated by using the complex model ZNN-I with $\gamma = 100$, is shown in Figure 4 (b).

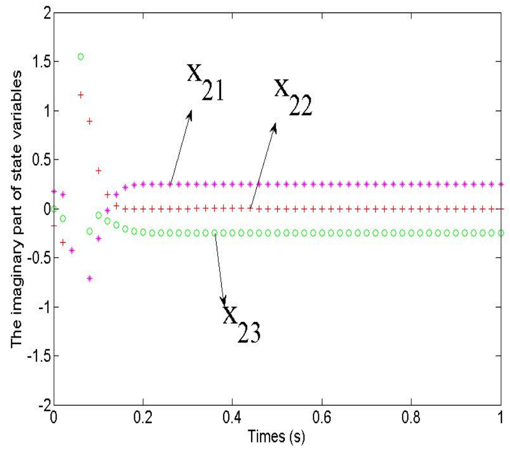
It is seen from Figures 1 and 3 that state variables $X(t)$ of ZNN converges directly and accurately to the principal square root after a very short finite time. In addition,



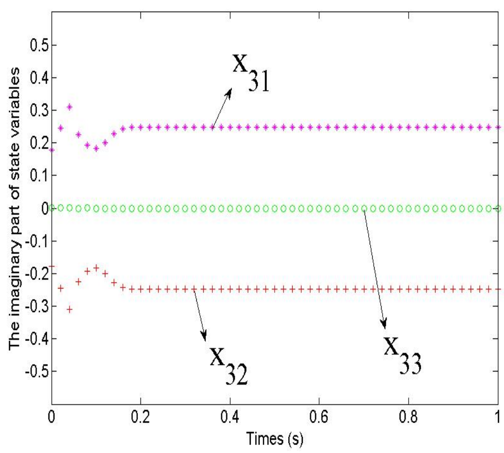
(c) The second line.



(e) The second line.



(d) The second line.



(f) The second line.

Fig. 3. Trajectories of the state variables of the model ZNN-II in Example 2.

Figures 2 and 4 show the transient convergence behavior of $\|X^2(t) - A\|_F$ synthesized by ZNN.

VI. Conclusion

In this paper, we have proposed two complex-valued neural networks for solving the matrix principal square root problem under certain conditions. To achieve this goal, we have designed two new complex-valued activation functions based on the sign-bi-power activation function. We have proved that the state of our neural networks can converge to the matrix principal square root in finite times.

References

[1] E. D. Denman, *The matrix sign function and computations in systems*. Elsevier Science Inc., 1976.

[2] M. A. Hasan, A. A. Hasan, and S. Rahman, "Fixed point iterations for computing square roots and the matrix sign function of complex matrices," in *Decision and Control, 2000. Proceedings of the IEEE Conference on*, 2000, pp. 4253–4258 vol.5.

[3] N. J. Higham, "Bibliography of functions of matrices: Theory and computation siam, 2008," 2008.

[4] B. Iannazzo, "A note on computing the matrix square root," *Calcolo*, vol. 40, no. 4, pp. 273–283, 2003.

[5] C. R. Johnson, K. Okubo, and R. Reams, "Uniqueness of matrix square roots and an application," *Linear Algebra and Its Applications*, vol. 323, no. 1, pp. 51–60, 2001.

[6] D. W. Masser and M. Neumann, "On the square roots of strictly quaslaccretlve complex matrlces," *Linear Algebra and its Applications*, vol. 28, pp. 135–140, 1979.

[7] B. Meini, "The matrix square root from a new functional perspective: Theoretical results and computational issues," *Siam Journal on Matrix Analysis and Applications*, vol. 26, no. 2, pp. 362–376, 2006.

[8] H. Wang and GaoTonghui, "Training algorithm of adaptive neural fuzzy inference system based on improved srukf," *IAENG International Journal of Computer Science*, vol. 44, no. 4, pp. 396–403, 2017.

[9] X. Lin, "Accelerating a recurrent neural network to finite-time convergence using a new design formula and its application to time-varying matrix square root ," *Journal of the Franklin Institute*, 2017.

[10] J. H. Long, X. Y. Hu, and L. Zhang, "Newton's method with exact line search for the square root of a matrix," 2008, p. 012034.

[11] L. Xiao, "A finite-time convergent zhang neural network and its application to real-time matrix square root finding," *Neural Computing and Applications*, no. 10, pp. 1–8, 2017.

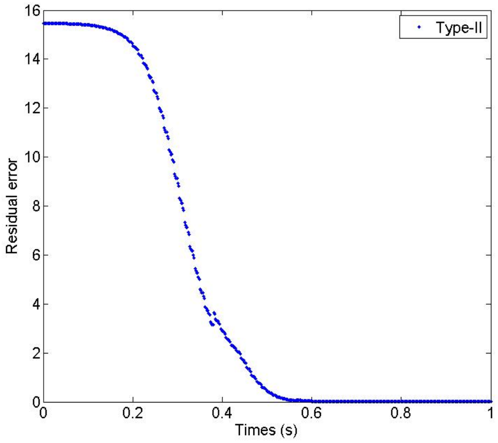
[12] Y. Zhang, D. Chen, D. Guo, B. Liao, and Y. Wang, "On exponential convergence of nonlinear gradient dynamics system with application to square root finding," *Nonlinear Dynamics*, vol. 79, no. 2, pp. 983–1003, 2015.

[13] Y. Zhang, Z. Ke, P. Xu, and C. Yi, "Time-varying square roots finding via zhang dynamics versus gradient dynamics and the former's link and new explanation to newton-raphson iteration," *Information Processing Letters*, vol. 110, no. 24, pp. 1103–1109, 2010.

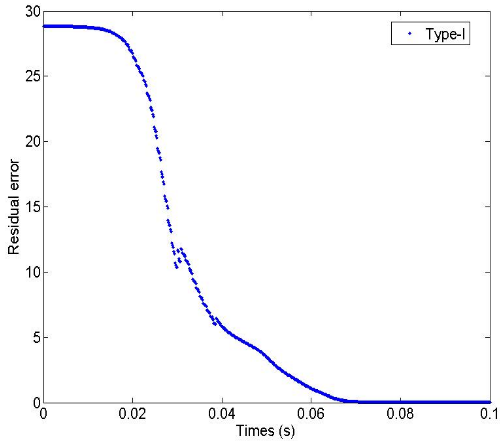
[14] Y. Zhang, W. Li, D. Guo, and Z. Ke, "Different zhang functions leading to different znn models illustrated via time-varying matrix square roots finding," *Expert Systems with Applications An International Journal*, vol. 40, no. 11, pp. 4393–4403, 2013.

[15] Y. Zhang, Y. Yang, B. Cai, and D. Guo, "Zhang neural network and its application to newton iteration for matrix square root estimation," *Neural Computing and Applications*, vol. 21, no. 3, pp. 453–460, 2012.

[16] N. J. Higham, "Stable iterations for the matrix square root," *Numerical Algorithms*, vol. 15, no. 2, pp. 227–242, Sep 1997. [Online]. Available: https://doi.org/10.1023/A:1019150005407



(a) ZNN-II with $\gamma = 10$.



(b) ZNN-I with $\gamma = 100$

Fig. 4. Trajectories of the residual errors of the model ZNN in Example 2.

[17] T. Kato, *Perturbation theory for linear operators.* Springer-Verlag,, 1966.
[18] X. Wang and Q. Zhang, "Complex-valued neural network for hermitian matrices," *Engineering Letters*, vol. 25, no. 3, pp. 312–320, 2017.
[19] F. Ding and T. Chen, "Gradient based iterative algorithms for solving a class of matrix equations," *IEEE Transactions on Automatic Control*, vol. 50, no. 8, pp. 1216–1221, 2005.
[20] F. Ding and H. Zhang, "Gradient-based iterative algorithm for a class of the coupled matrix equations related to control systems," *Iet Control Theory and Applications*, vol. 8, no. 15, pp. 1588–1595, 2014.
[21] D. Guo, C. Yi, and Y. Zhang, "Zhang neural network versus gradient-based neural network for time-varying linear matrix equation solving," *Neurocomputing*, vol. 74, no. 17, pp. 3708–3712, 2011.
[22] B. Zhou and G. Duan, "An explicit solution to the matrix equation AX-XF = BY," *Linear Algebra and its Applications*, vol. 402, pp. 345–366, 2005.
[23] X. Wang, L. Liang, and M. Che, "Finite-time convergent complex-valued neural networks for the time-varying complex linear matrix equations," *Engineering Letters*, vol. 26, no. 4, pp. 432–440, 2018.
[24] A. Hjorungnes and D. Gesbert, "Complex-valued matrix differentiation: Techniques and key results," *IEEE Trans. Signal Process.*, vol. 55, no. 6p1, pp. 2740–2746, 2007.
[25] Y. Shen, P. Miao, Y. Huang, and Y. Shen, "Finite-time stability and its application for solving time-varying sylvester equation by recurrent neural network," *Neural Processing Letters*, vol. 42, no. 3, pp. 763–784, 2015.