



ELSEVIER

Contents lists available at ScienceDirect

Linear Algebra and its Applications

www.elsevier.com/locate/laa



Two finite-time convergent Zhang neural network models for time-varying complex matrix Drazin inverse



Sanzheng Qiao^{a,1}, Xue-Zhong Wang^{b,2}, Yimin Wei^{c,*}

^a Department of Computing and Software, McMaster University, Hamilton, ON L8S 4K1, Canada

^b School of Mathematical Sciences, Fudan University, Shanghai, 200433, PR China

^c School of Mathematical Sciences and Shanghai Key Laboratory of Contemporary Applied Mathematics, Fudan University, Shanghai, 200433, PR China

ARTICLE INFO

Article history:

Received 21 December 2016

Accepted 15 March 2017

Available online 21 March 2017

Submitted by M. Van Barel

MSC:

15A09

65F20

Keywords:

Complex neural network

Drazin inverse

Convergence

Time-varying complex matrix

Li function

ABSTRACT

This paper concerns the computation of the Drazin inverse of a complex time-varying matrix. Based on two Zhang functions constructed from two limit representations of the Drazin inverse, we present two complex Zhang neural network (ZNN) models with the Li activation function for computing the Drazin inverse of a complex time-varying square matrix. We prove that our ZNN models globally converge in finite time. In addition, upper bounds of the convergence time are derived analytically via the Lyapunov theory. Our simulation results verify the theoretical analysis and demonstrate the superiority of our ZNN models over the gradient-based GNN models.

© 2017 Elsevier Inc. All rights reserved.

* Corresponding author.

E-mail addresses: qiao@cas.mcmaster.ca (S. Qiao), 14130180001@fudan.edu.cn (X.-Z. Wang), yimwei@fudan.edu.cn, yimin.wei@gmail.com (Y. Wei).

¹ The author is partially supported by the Shanghai Key Laboratory of Contemporary Applied Mathematics of Fudan University and International Cooperation Project of Shanghai Municipal Science and Technology Commission under grant 16510711200.

² This author is supported by International Cooperation Project of Shanghai Municipal Science and Technology Commission under grant 16510711200.

1. Introduction

The Drazin inverse has many applications, such as singular differential or difference equations and Markov chains. Iterative methods for computing the matrix Drazin inverse can be found in [5,6,11,35–37].

The neural network approach to parallel computing and signal processing has been successfully developed through a variety of neurodynamic models with learning capabilities [14,22]. Various types of neural networks have been introduced to solve systems of linear algebraic equations.

A number of nonlinear and linear Hopfield-type recurrent neural network models were recently developed for computing the regular inverse of a nonsingular matrix and the generalized inverses of a full-rank rectangular matrix, see [17,24,29,30]. A method with high convergence rate for finding approximate inverses of nonsingular matrices was suggested and established analytically in [25]. In particular, a feedforward neural network architecture for computing the Drazin inverse A^D of a square matrix A was proposed in [8]. Various recurrent neural networks (RNN) for computing the generalized inverses of rank-deficient matrices are presented in [31,34]. In [28], an RNN with linear activation functions for computing the Drazin inverse of a square matrix was proposed by Stanimirović, Zivković and Wei. The dynamic equation and its corresponding artificial RNN for computing the Drazin inverse of a square real matrix, with no restrictions on its eigenvalues, were presented in [27]. Four gradient-based recurrent neural networks (RNNs) for computing the Drazin inverse of a square real matrix were developed in [32].

A new type of complex-valued Zhang neural network (ZNN) was proposed and investigated in [20]. The ZNN models for computing the Moore–Penrose inverse of an online time-varying and full-rank matrix are investigated and analyzed in [42]. The design of the ZNN models is based on an indefinite error-monitoring function, called Zhang function (ZF), which can be real, complex, positive, zero, negative or unbounded. The ZF plays an important role in the development of the ZNN, and largely enriches the theory of the ZNN.

Complex matrices occur in situations where the problem contains online frequency domain identification processes, or the input signals incorporate both the magnitude and phase information [15,26]. Thus, problems in the complex domain have attracted extensive attention, [3,12,13,15,16,26,38]. Furthermore, there are applications involving the computation of the Drazin inverse of a time-varying matrix [1,4,41]. Two types of the ZNN models with various activation functions for computing the online time-varying Drazin inverse were proposed, investigated and analyzed by Wang, Wei and Stanimirović in [33].

In this paper, based on the idea of the ZF, we design two ZNNs for the online solution of the Drazin inverse of a time-varying complex matrix. In addition, inspired by the study of finite-time convergence [2,18,19,21,23,39], we use a novel activation function, called the Li activation function [18,19,21], to accelerate the ZNN to finite-time convergence to the Drazin inverse of a time-varying complex matrix.

This paper is organized as follows. In Section 2, we review some preliminary results. Our two complex ZNN models with Li activation functions for online solution of the Drazin inverse of a time-varying complex matrix are presented in Section 3. Convergence properties of our complex ZNN models are discussed in Section 4. Illustrative numerical examples are presented in Section 5.

The main contributions of this paper are:

- (1) This paper focuses on solving time-varying linear matrix equations in the complex domain, rather than static or time-varying linear matrix equations in the real domain.
- (2) Two finite-time convergent neural dynamical models are proposed and investigated for online solution of the Drazin inverse of a time-varying complex matrix.
- (3) The results in this paper are a generalization of those in [32] in two directions, one is real domain investigated in [32] is extended into the complex domain, and the other is Li activation is used and a finite time convergence achieved.
- (4) The paper performs a theoretical analysis of our proposed ZNN models and shows that our models converge to the Drazin inverse of a time-varying complex matrix in finite time. In addition, upper bounds of the convergence time are derived via the Lyapunov theory.

2. Preliminaries

For any matrix $A \in \mathbb{C}^{n \times n}$, we denote its column space and null space by $\mathbf{R}(A)$ and $\mathbf{N}(A)$ respectively. The index of A , denoted by $\text{Ind}(A)$, is the smallest nonnegative integer k for which $\mathbf{N}(A^k) = \mathbf{N}(A^{k+1})$. We write $\|\cdot\|$ for the spectral norm. Given a matrix $A \in \mathbb{C}^{n \times n}$, A^\top , A^H , A_{Re} and A_{Im} are the transpose, the complex conjugated transpose, real part and imaginary part of A , respectively. For given matrices $A, B \in \mathbb{C}^{n \times n}$, $A \circ B$ denotes the Hadamard product of A and B , i.e., $A \circ B = [a_{ij}b_{ij}]$. We use $|E| = (|E_{ij}|)$ and $\Theta(E) = (\Theta(E_{ij}))$ to denote the element-wise modulus and the element-wise argument of the matrix $E \in \mathbb{C}^{n \times n}$, respectively. Thus, for a given $E \in \mathbb{C}^{n \times n}$, $E = |E| \circ \exp(\iota\Theta(E))$, where ι is the imaginary unit $\sqrt{-1}$.

A square matrix $X \in \mathbb{C}^{n \times n}$ is called the Drazin inverse of a matrix $A \in \mathbb{C}^{n \times n}$, if X satisfies

$$A^{k+1}X = A^k, \quad XAX = X, \quad AX = XA, \quad \text{for some } k > 0.$$

We write $X = A^D$ for the Drazin inverse of A [9]. Furthermore, if the index $k = 1$, the Drazin inverse is reduced to the group inverse and denoted by A^\sharp .

In the present paper, we solve the time varying Drazin inverse problem by using two limit representations of the Drazin inverse, stated in the following two lemmas. The first limit representation of A^D can be derived from the results in [7].

Lemma 2.1. ([7]) *Let $A \in \mathbb{C}^{n \times n}$, then a closed-form solution of A^D is given by*

$$A^D = \lim_{\lambda \rightarrow 0} (A^l(A^{2l+1})^H A^{l+1} + \lambda I)^{-1} A^l(A^{2l+1})^H A^l, \quad \text{for } l \geq \text{Ind}(A). \quad (2.1)$$

Campbell and Meyer [5] gave another limit representation of the Drazin inverse of a complex matrix.

Lemma 2.2. ([5]) *Let $A \in \mathbb{C}^{n \times n}$, then its Drazin inverse has the limit representation:*

$$A^D = \lim_{\lambda \rightarrow 0} (A^{l+1} + \lambda I)^{-1} A^l, \quad \text{for } l \geq \text{Ind}(A). \tag{2.2}$$

The above two limit presentations are the bases for our two ZNN models proposed in the next section.

3. Neural network models based on the ZNN

Now we are ready for the design of our ZNN models for computing the Drazin inverse of a complex time-varying matrix $A(t)$, for which we make the following assumption.

Assumption 3.1. The time-varying matrix $A(t) \in \mathbb{C}^{n \times n}$ in our discussion has the properties: $\text{Ind}(A(t)) = k$, for $t \in [0, +\infty)$; $A(t)^m$, $m > 0$, are continuously differentiable with respect to the time t ; $A(t)^m$ and their time derivatives are uniformly bounded.

Following the three steps outlined in [20,21,42], we present two ZNN models, named ZNN-I and ZNN-II, based on the two limit representations (2.1) and (2.2).

3.1. Neural network model ZNN-I

Let $\dot{E}(t)$ denote the time derivative of the complex function $E(t)$, our design procedure consists of three steps.

Step 1 (Construct the Zhang Function). In this first step, the ZF, a matrix-valued error-monitoring function, is constructed. In this model we use the limit representation (2.1), which suggests the use of the matrix $G(t) = A(t)^l (A(t)^{2l+1})^H A(t)^l$. Since $A(t)^{l+1} A(t)^D = A(t)^l$ holds for any $l \geq k$, we have the following complex matrix identity:

$$G(t)A(t)A(t)^D = G(t),$$

where $G(t), A(t), A(t)^D \in \mathbb{C}^{n \times n}$.

To convert the above complex matrix identity into a real one, we first rewrite it as

$$((G(t)A(t))_{\text{Re}} + \iota(G(t)A(t))_{\text{Im}}) (A(t)^D_{\text{Re}} + \iota A(t)^D_{\text{Im}}) = (G(t)_{\text{Re}} + \iota G(t)_{\text{Im}}),$$

where $(G(t)A(t))_{\text{Re}}, (G(t)A(t))_{\text{Im}}, A(t)^D_{\text{Re}}, A(t)^D_{\text{Im}}, G(t)_{\text{Re}}, G(t)_{\text{Im}} \in \mathbb{R}^{n \times n}$. Then, equating the real part and the imaginary part and writing it in matrix form, we obtain

$$\begin{bmatrix} (G(t)A(t))_{\text{Re}} & -(G(t)A(t))_{\text{Im}} \\ (G(t)A(t))_{\text{Im}} & (G(t)A(t))_{\text{Re}} \end{bmatrix} \begin{bmatrix} A(t)^D_{\text{Re}} \\ A(t)^D_{\text{Im}} \end{bmatrix} = \begin{bmatrix} G(t)_{\text{Re}} \\ G(t)_{\text{Im}} \end{bmatrix}.$$

In order to ensure the existence of the unique solution at any time instant $t \in [0, +\infty)$, we introduce the bias matrix

$$\begin{bmatrix} \lambda I & 0 \\ 0 & \lambda I \end{bmatrix},$$

where $\lambda \in \mathbb{R}$, $\lambda > 0$ and $I \in \mathbb{R}^{n \times n}$ is the identity matrix, to the above equation and obtain the equation

$$\begin{bmatrix} (G(t)A(t))_{\text{Re}} + \lambda I & -(G(t)A(t))_{\text{Im}} \\ (G(t)A(t))_{\text{Im}} & (G(t)A(t))_{\text{Re}} + \lambda I \end{bmatrix} \begin{bmatrix} V(t)_{\text{Re}} \\ V(t)_{\text{Im}} \end{bmatrix} = \begin{bmatrix} G(t)_{\text{Re}} \\ G(t)_{\text{Im}} \end{bmatrix}. \tag{3.1}$$

Thus, from (2.1), $V(t)_{\text{Re}}$ and $V(t)_{\text{Im}}$ approach to $A(t)_{\text{Re}}^D$ and $A(t)_{\text{Im}}^D$ respectively as $\lambda \rightarrow 0$. This process of replacing a possible singular matrix with a well-conditioned matrix is known as the Tikhonov regularization method [10]. Thus we define the following complex function as the fundamental error-monitoring function, called ZF1:

$$E(t) = \begin{bmatrix} (G(t)A(t))_{\text{Re}} + \lambda I & -(G(t)A(t))_{\text{Im}} \\ (G(t)A(t))_{\text{Im}} & (G(t)A(t))_{\text{Re}} + \lambda I \end{bmatrix} \begin{bmatrix} V(t)_{\text{Re}} \\ V(t)_{\text{Im}} \end{bmatrix} - \begin{bmatrix} G(t)_{\text{Re}} \\ G(t)_{\text{Im}} \end{bmatrix}. \tag{3.2}$$

For simplicity, we express the above error-monitoring ZF as

$$E(t) = U(t)V(t) - D(t), \tag{3.3}$$

where $U(t), V(t), D(t) \in \mathbb{R}^{2n \times n}$ are respectively the first, second, and third matrices in (3.2).

Step 2 (Choose an activation-function). To ensure the convergence of the error-monitoring function $E(t)$ to zero, we assume its derivative

$$\dot{E}(t) := \frac{dE(t)}{dt} = -\gamma \Psi(E(t)), \tag{3.4}$$

where the design parameter $\gamma \in \mathbb{R}$, $\gamma > 0$, corresponds to the inductance parameter or the reciprocal of the capacitance parameter, and $\Psi(\cdot)$ is a specially constructed activation-function, a matrix mapping of the neural network. The design parameter γ should be as large as the hardware permits [20]. The following four real-valued functions are widely used as the activation-function. For more details see, for example [40].

The linear function:

$$f(x) = x.$$

The Power-sigmoid function:

$$f(x) = \begin{cases} x^p, & \text{if } |x| \geq 1 \\ \frac{1+\exp(-q)}{1-\exp(-q)} \cdot \frac{1-\exp(-qx)}{1+\exp(-qx)}, & \text{otherwise} \end{cases}, \quad q > 2, p \geq 3.$$

The Bipolar-sigmoid function:

$$f(x) = \frac{1 + \exp(-q)}{1 - \exp(-q)} \cdot \frac{1 - \exp(-qx)}{1 + \exp(-qx)}, \quad q > 2.$$

The Smooth power-sigmoid function:

$$f(x) = \frac{1}{2}x^p + \frac{1 + \exp(-q)}{1 - \exp(-q)} \cdot \frac{1 - \exp(-qx)}{1 + \exp(-qx)}, \quad p \geq 3, q > 2.$$

To accelerate our ZNN to finite-time convergence, we use the Li activation-function matrix $\Psi(A) = [\psi(a_{ij})]$ adopted from [21], whose (i, j) -entry is given by

$$\psi(e(t)_{ij}) = \text{Lip}^\sigma + \text{Lip}^{1/\sigma}, \tag{3.5}$$

where $e(t)_{ij}$ denotes the (i, j) -entry of $E(t)$ and the function Lip^σ with the parameter σ is defined by

$$\text{Lip}^\sigma = \begin{cases} |e(t)_{ij}|^\sigma, & e(t)_{ij} > 0, \\ 0, & e(t)_{ij} = 0, \\ -|e(t)_{ij}|^\sigma, & e(t)_{ij} < 0, \end{cases}$$

where $0 < \sigma < 1$.

Step 3 (Derive the dynamic equation of the ZNN model). Combining the following time derivative of (3.3):

$$\dot{E}(t) = \dot{U}(t)V(t) + U(t)\dot{V}(t) - \dot{D}(t) \tag{3.6}$$

and (3.4), we obtain the implicit dynamic equation of our ZNN-I model:

$$U(t)\dot{V}(t) = \dot{D}(t) - \dot{U}(t)V(t) - \gamma\Psi(U(t)V(t) - D(t)). \tag{3.7}$$

3.2. Neural network model II

Analogous to the design of the ZNN-I model, from the limit representation (2.2), we propose the complex matrix equation

$$(A(t)^{l+1} + \lambda I)A(t)^D = A(t)^l, \quad \text{for } l \geq k = \text{Ind}(A(t)),$$

where $A(t), A(t)^D \in \mathbb{C}^{n \times n}$, $\lambda \in \mathbb{R}$, $\lambda > 0$, and I is the identity matrix of order n . Equating the real part and the imaginary part of the both sides of the above equation, we get its corresponding real matrix equations in the matrix form

$$\begin{bmatrix} A(t)_{\text{Re}}^{l+1} + \lambda I & -A(t)_{\text{Im}}^{l+1} \\ A(t)_{\text{Im}}^{l+1} & A(t)_{\text{Re}}^{l+1} + \lambda I \end{bmatrix} \begin{bmatrix} V(t)_{\text{Re}} \\ V(t)_{\text{Im}} \end{bmatrix} = \begin{bmatrix} A(t)_{\text{Re}}^l \\ A(t)_{\text{Im}}^l \end{bmatrix}. \tag{3.8}$$

Since the nonzero eigenvalues of $A(t)^{l+1}$, $l \geq k$, lie in the open right-half plane, we define

$$E_1(t) = H(t)V(t) - Q(t), \tag{3.9}$$

where

$$H(t) = \begin{bmatrix} A(t)_{\text{Re}}^{l+1} + \lambda I & -A(t)_{\text{Im}}^{l+1} \\ A(t)_{\text{Im}}^{l+1} & A(t)_{\text{Re}}^{l+1} + \lambda I \end{bmatrix}, \quad V(t) = \begin{bmatrix} V(t)_{\text{Re}} \\ V(t)_{\text{Im}} \end{bmatrix}, \quad Q(t) = \begin{bmatrix} A(t)_{\text{Re}}^l \\ A(t)_{\text{Im}}^l \end{bmatrix},$$

as the fundamental error-monitoring function. Its derivative with respect to the time t , then equals to

$$\dot{E}_1(t) = H(t)\dot{V}(t) + \dot{H}(t)V(t) - \dot{Q}(t).$$

Finally, applying the above derivative to the design pattern (3.4), we obtain the implicit dynamic equation

$$H(t)\dot{V}(t) = \dot{Q}(t) - \dot{H}(t)V(t) - \gamma\Psi(H(t)V(t) - Q(t)) \tag{3.10}$$

for our second ZNN model, named ZNN-II. Thus, from (2.2), the state matrix $V(t)$ approaches the Drazin inverse as λ in $H(t)$ tends to zero.

4. Convergence analyses

In this section, we show that the both neural network models ZNN-I and ZNN-II globally converge to the theoretical time-varying solutions for (3.1) and (3.8) respectively in finite time. Thus, they lead to the Drazin inverse as $\lambda \rightarrow 0$.

4.1. Convergence of the model ZNN-I

We first state the following lemma, which is useful for the convergence analysis of the model ZNN-I.

Lemma 4.1. [33] *Let $A(t) \in \mathbb{C}^{n \times n}$ be of index k , then the eigenvalues λ of the matrix $G(t)A(t) = A(t)^l (A(t)^{2l+1})^H A(t)^{l+1}$, $l \geq k$ satisfy $\text{Re}(\lambda) \geq 0$.*

We then have the following result of the convergence of the ZNN-I model.

Theorem 4.1. *Given a time-varying complex matrix $A(t) \in \mathbb{C}^{n \times n}$. Under Assumption 3.1, if the Li activation function (3.5) is used, then the state matrix $V(t) \in \mathbb{R}^{2n \times n}$ of the dynamic equation (3.7) of the ZNN model ZNN-I, starting from an arbitrary initial state $V(0)$, converges to the theoretical solution of (3.1) in finite time*

$$t_f < \max \left\{ \frac{|e^-(0)|^{1-\sigma}}{\gamma(1-\sigma)}, \frac{|e^+(0)|^{1-\sigma}}{\gamma(1-\sigma)} \right\}, \tag{4.1}$$

where $e^+(0)$ and $e^-(0)$ are the largest and the smallest elements in the matrix $E(0) = U(0)V(0) - D(0)$, respectively.

Proof. Let the real matrix $V(t)$ be the state matrix generated by the dynamic equation (3.7) and $V_*(t)$ the solution of the matrix equation (3.1). We consider the difference $E_*(t) = V(t) - V_*(t)$. The time derivative of

$$U(t)V_*(t) - D(t) = 0 \tag{4.2}$$

is

$$\dot{U}(t)V_*(t) + U(t)\dot{V}_*(t) - \dot{D}(t) = 0. \tag{4.3}$$

Since $V(t) = V_*(t) + E_*(t)$, (4.2) implies

$$E(t) = U(t)V(t) - D(t) = U(t)E_*(t).$$

Applying (4.3) and $V(t) = V_*(t) + E_*(t)$ to the model (3.7), we obtain

$$\dot{U}(t)E_*(t) + U(t)\dot{E}_*(t) = -\gamma \Psi(U(t)E_*(t)),$$

that is,

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

whose elementwise expression is

$$\dot{e}_{ij}(t) = -\gamma \psi(e_{ij}(t)), \quad i, j = 1, 2, \dots, n.$$

In order to show the global convergence of $e_{ij}(t)$, we consider the Lyapunov function $L(t) = [l_{ij}(t)]$, where

$$l_{ij}(t) = \frac{e_{ij}^2(t)}{2}.$$

It then follows that

$$\dot{l}_{ij}(t) = e_{ij}(t)\dot{e}_{ij}(t) = -\gamma e_{ij}(t)\psi(e_{ij}(t)).$$

Since

$$\psi(e_{ij}(t)) \begin{cases} > 0, & e_{ij} > 0, \\ = 0, & e_{ij} = 0, \\ < 0, & e_{ij} < 0, \end{cases}$$

and $\gamma > 0$, we get

$$\dot{l}_{ij}(t) \begin{cases} < 0, & e_{ij}(t) \neq 0, \\ = 0, & e_{ij}(t) = 0. \end{cases}$$

Following the argument in the proof of Theorem 1 in [21], by the Lyapunov stability theory, $e_{ij}(t)$, $1 \leq i, j \leq n$, globally converges to zero. Thus, from the nonsingularity of $U(t)$ in $E(t) = U(t)V(t) - D(t)$, the state matrix $V(t)$ globally converges to the solution of (3.1), as the time t increases, with any initial state $V(0)$, and approaches the Drazin inverse, as $\lambda \rightarrow 0$.

Next, we prove that the ZNN-I model converges in finite time.

Let $e^+(t)$ be the entry in $E(t)$ such that $e^+(0) = \max(E(0))$ and let $e^-(t)$ be the entry in $E(t)$ such that $e^-(0) = \min(E(0))$. Following the proof of Theorem 1 in [21], we have

$$e^-(t) \leq e_{ij}(t) \leq e^+(t), \quad \text{for } 1 \leq i, j \leq n.$$

Thus it remains to show that both $e^-(t)$ and $e^+(t)$ converge to zero in finite time. We first investigate the convergence time of $e^+(t)$, which satisfies

$$\dot{e}^+ = -\gamma\psi(e^+(t)) \quad \text{with } e^+(0) = \max(E(0)),$$

where $\psi(\cdot)$ is the Li activation function (3.5). Considering an alternative Lyapunov function $l_+(t) = |e^+(t)|^2 \geq 0$, we have

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

Solving the differential inequality

$$\dot{l}_+(t) \leq -2\gamma l_+^{(\sigma+1)/2}(t)$$

with the initial condition $l_+(0) = |e^+(0)|^2$, we obtain

$$l_+^{(1-\sigma)/2} \begin{cases} \leq |e^+(0)|^{1-\sigma} - \gamma(1-\sigma)t, & \text{when } t \leq \frac{|e^+(0)|^{1-\sigma}}{\gamma(1-\sigma)}, \\ 0, & \text{otherwise.} \end{cases}$$

That is, $l_+(t) = |e^+(t)|^2$ converges to zero before the time $t = |e^+(0)|^{1-\sigma}/(\gamma(1-\sigma))$. Similarly, $|e^-(t)|^2$ converges to zero before the time $t = |e^-(0)|^{1-\sigma}/(\gamma(1-\sigma))$. Putting things together, we have the convergence time (4.1). \square

4.2. Convergence of the model ZNN-II

Analogous to Theorem 4.1, we can prove the following theorem of the finite-time convergence of the model ZNN-II.

Theorem 4.2. *Let $A(t) \in \mathbb{C}^{n \times n}$ be a time-varying complex matrix. Under the Assumption 3.1. If the Li activation function $\Psi(\cdot)$ is used, then the state matrix $V(t) \in \mathbb{R}^{2n \times n}$ of the dynamic equation (3.10) of the neural network model ZNN-II, starting from any initial state $V(0)$, globally converges to the theoretical solution of (3.8) in finite time*

$$t_f < \max \left\{ \frac{|e_1^-(0)|^{1-\sigma}}{\gamma(1-\sigma)}, \frac{|e_1^+(0)|^{1-\sigma}}{\gamma(1-\sigma)} \right\},$$

where $e_1^+(0)$ and $e_1^-(0)$ are the largest and the smallest elements in the matrix $E_1(0) = H(0)V(0) - Q(0)$, respectively.

5. Numerical examples

For the purpose of comparison, we state the following two gradient-based neural networks (GNN) [27,28]

$$\dot{V}(t) = -\gamma U(t)^\top \Psi(U(t)V(t) - D(t)), \quad \text{with } V(0) = 0, \tag{5.1}$$

and

$$\dot{V}(t) = -\gamma H(t)^\top \Psi(H(t)V(t) - Q(t)), \quad \text{with } V(0) = 0, \tag{5.2}$$

where

$$U(t) = \begin{bmatrix} (G(t)A(t))_{\text{Re}} + \lambda I & -(G(t)A(t))_{\text{Im}} \\ (G(t)A(t))_{\text{Im}} & (G(t)A(t))_{\text{Re}} + \lambda I \end{bmatrix},$$

$$V(t) = \begin{bmatrix} V(t)_{\text{Re}} \\ V(t)_{\text{Im}} \end{bmatrix}, \quad D(t) = \begin{bmatrix} G(t)_{\text{Re}} \\ G(t)_{\text{Im}} \end{bmatrix},$$

and

$$H(t) = \begin{bmatrix} A(t)_{\text{Re}}^{l+1} + \lambda I & -A(t)_{\text{Im}}^{l+1} \\ A(t)_{\text{Im}}^{l+1} & A(t)_{\text{Re}}^{l+1} + \lambda I \end{bmatrix}, \quad Q(t) = \begin{bmatrix} A(t)_{\text{Re}}^l \\ A(t)_{\text{Im}}^l \end{bmatrix},$$

and $\Psi(\cdot)$ is the Li activation function defined in (3.5).

We first present our test results on the ZNNs, then we present our test results on the GNNs for comparison.

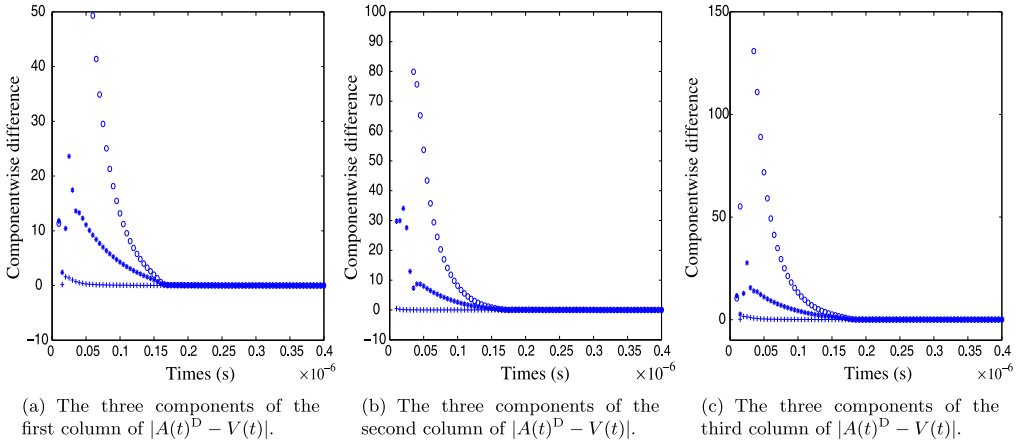


Fig. 1. Componentwise convergence of the model ZNN-I in Example 1.

5.1. Numerical tests based on ZNN

Example 1. Consider the complex time-varying matrix

$$A(t) = \begin{bmatrix} \iota \sin(2t) & \iota \cos(2t) & -\iota \sin(2t) \\ -\iota \cos(2t) & \iota \sin(2t) & \iota \cos(2t) \\ \iota \sin(2t) & \iota \cos(2t) & 0 \end{bmatrix}.$$

Numerical values of $A(t)^D$ at time t are computed numerically via the formula

$$A(t)^D = A(t)^k (A(t)^{2k+1})^\dagger A(t)^k, \tag{5.3}$$

where $k = \text{Ind}(A(t)) = 1$ and A^\dagger denotes the Moore–Penrose inverse of A . The exact Drazin inverse of $A(t)$ is

$$A(t)^D = \begin{bmatrix} \iota \cos(2t) \cot(2t) & \iota \cos(2t) & -\iota \csc(2t) \\ -\iota \cos(2t) & -2\iota \cos(t) \sin(t) & 0 \\ \iota \csc(2t) & 0 & -\iota \csc(2t) \end{bmatrix}.$$

We set $l = 2$, the initial state $V(0)$ as the matrix of appropriate dimensions whose entries are ones and chose the Li activation function as $\Psi(\cdot)$ defined in (3.5) with $\sigma = 0.8$. In our ZNN-I model, we set $\gamma = 3 \times 10^5$ and $\lambda = 10^{-6}$.

To show the convergence, Fig. 1 depicts the componentwise absolute difference $|A(t)^D - V(t)|$ between the Drazin inverse $A(t)^D$ computed using (5.3) and the state matrix $V(t)$ produced by our ZNN-I model. In particular, Fig. 1 (a) shows the three components of the first column of $|A(t)^D - V(t)|$, where the first component is represented by ‘*’, the second component by ‘o’, and the third by ‘+’. Fig. 1 (b) and (c)

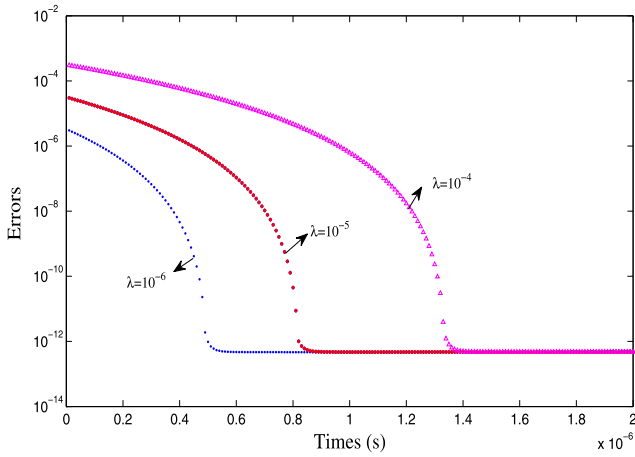


Fig. 2. Trajectory of Frobenius norm of the residual (5.4) in Example 1.

depict the second and third columns respectively. The figures show that our ZNN-I model converges to the Drazin inverse $A(t)^D$.

Our ZNN-I model produces an approximation of the solution $V(t)$ for (3.1). Thus, to verify the upper bound for the convergence time given by Theorem 4.1, in Fig. 2, for given different λ , we plot the trajectory of the Frobenius norm for the residual error

$$\left\| \begin{bmatrix} (G(t)A(t))_{\text{Re}} + \lambda I & -(G(t)A(t))_{\text{Im}} \\ (G(t)A(t))_{\text{Im}} & (G(t)A(t))_{\text{Re}} + \lambda I \end{bmatrix} \begin{bmatrix} V(t)_{\text{Re}} \\ V(t)_{\text{Im}} \end{bmatrix} - \begin{bmatrix} G(t)_{\text{Re}} \\ G(t)_{\text{Im}} \end{bmatrix} \right\|_{\text{F}}.$$

Note that the theoretical upper bound for the convergence time of this example, where $\gamma = 3 \times 10^5$, $\lambda = 10^{-6}$, and $\sigma = 0.8$, can be obtained by

$$\max \left\{ \frac{|e^-(0)|^{1-\sigma}}{\gamma(1-\sigma)}, \frac{|e^+(0)|^{1-\sigma}}{\gamma(1-\sigma)} \right\} = \frac{\lambda^{0.2}}{0.2\gamma} \approx 1.0 \times 10^{-6}. \tag{5.4}$$

As expected, in this example, our model converges before the above theoretical upper bound. It is worth pointing out that, as shown in Fig. 2, the convergence time of ZNN-I model can be shortened from 1.4×10^{-6} s to 0.9×10^{-6} s and even to 0.6×10^{-6} s when the value of parameter λ decreases from 10^{-4} to 10^{-5} and to 10^{-6} . These simulative observations have shown different values of λ can influence the speed of convergence for ZNN-I model.

Example 2. Consider the time-varying matrix

$$A(t) = \begin{bmatrix} 2 \sin(t) + \iota \cos(t) & \iota \sin(t) & 1 & \sin(t) + \iota \cos(t) \\ \iota \sin(t) & \cos(t)(\iota + 1) & 2 \sin(t) + 2\iota \cos(t) & \iota \sin(t) + 1 \\ \cos(t) + \iota \sin(t) & \iota t + 1 & 0 & 2 \cos(t) + 2\iota \sin(t) \\ \iota + 1 & 2 \cos(t) + 2\iota \sin(t) & \sin(t) + \iota \cos(t) & 0 \end{bmatrix}.$$

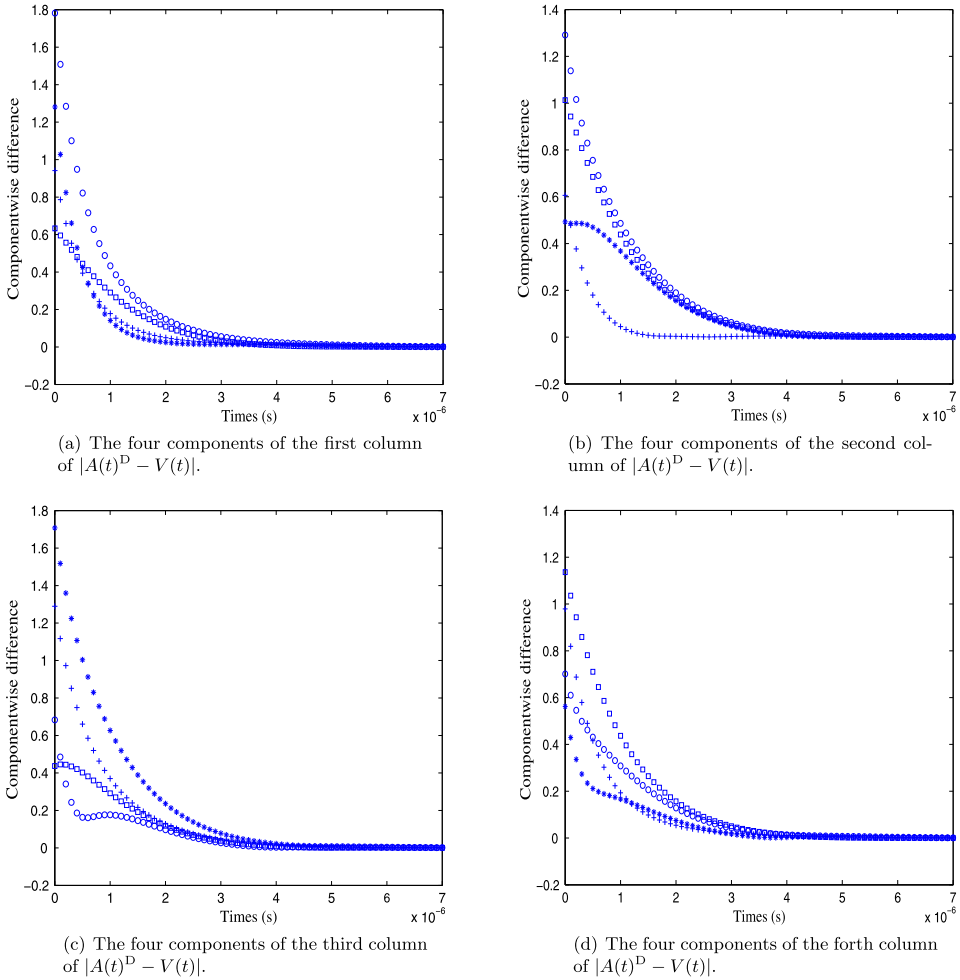


Fig. 3. Componentwise convergence of the model ZNN-II in Example 2.

For any $t \geq 0$, $A(t)$ is an invertible matrix. Thus $A(t)^D = A(t)^{-1}$. Numerical values of $A(t)^{-1}$ at time t were computed by calling the Matlab function `inv(A(t))`. We set $l = 2$, the initial state $V(0)$ as the matrix of appropriate dimensions whose entries are ones and chose the Li activation function as $\Psi(\cdot)$ defined in (3.5) with $\sigma = 0.8$. In our ZNN-II model, we set $\gamma = 10^6$ and $\lambda = 10^{-6}$.

Fig. 3 shows the convergence by depicting the componentwise absolute difference $|A(t)^{-1} - V(t)|$ between $A(t)^{-1}$ computed by Matlab and the state matrix $V(t)$ produced by our ZNN-II model. In particular, Fig. 3 (a) shows the four components of the first column of $|A(t)^{-1} - V(t)|$, where the first component is represented by ‘*’, the second component by ‘◊’, the third by ‘+’, and the fourth by ‘◻’. Fig. 3 (b), (c) and (d) depict the second, third and fourth columns respectively. The figures show that our ZNN-II model converges to the inverse $A(t)^{-1}$.

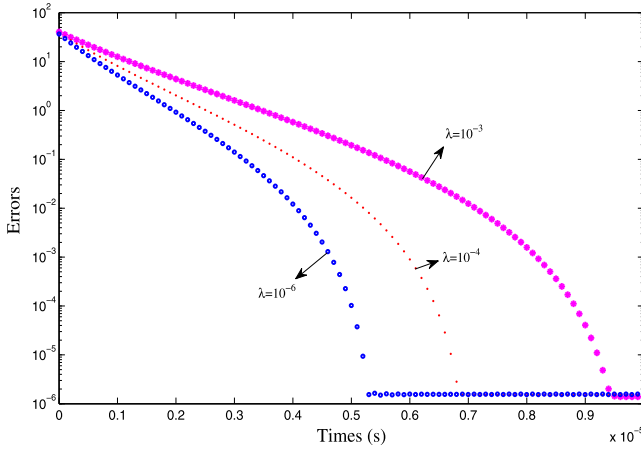


Fig. 4. Trajectory of the Frobenius norm of the residual $\|H(t)V(t) - Q(t)\|_F$ in Example 2.

Our ZNN-II model produces an approximation of the solution $V(t)$ for (3.8). Thus, to verify the upper bound for the convergence time given by Theorem 4.2, in Fig. 4, we plot the trajectory of the Frobenius norm of the residual error $\|H(t)V(t) - Q(t)\|_F$ in (3.9) by taking different values of λ .

Note that the theoretical upper bound for the convergence time of this example, where $\gamma = 10^6$, $\lambda = 10^{-6}$, and $\sigma = 0.8$, can be obtained by

$$\max \left\{ \frac{|e_1^-(0)|^{1-\sigma}}{\gamma(1-\sigma)}, \frac{|e_1^+(0)|^{1-\sigma}}{\gamma(1-\sigma)} \right\} = \frac{18^{0.2}}{0.2\gamma} \approx 8.9 \times 10^{-6}. \tag{5.5}$$

As expected, in this example, our model converges before the above theoretical upper bound. Similar to the Example 1, we can see different values of λ can influence the speed of convergence for ZNN-II model.

5.2. Numerical tests based on GNN

Now we compare our ZNN-I model with the GNN model (5.2) to show the superiority of the ZNN model over the GNN model.

Example 3. We apply the GNN model (5.2) to Example 1 with the same Li activation function and the same parameters.

Fig. 5 displays the componentwise absolute difference $|A(t)^D - V(t)|$ between $A(t)^D$ computed using (5.3) and the state matrix $V(t)$ produced by the GNN model (5.2). In particular, Fig. 5 (a) shows the three components of the first column of $|A(t)^D - V(t)|$, where the first component is represented by ‘*’, the second component by ‘o’, and the third by ‘+’. Fig. 5 (b) and (c) depict the second and third columns respectively. From

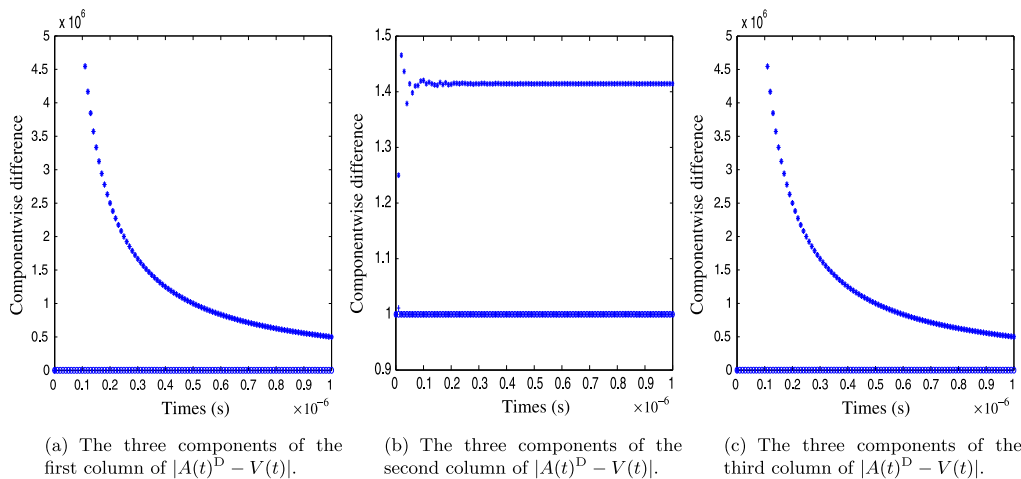


Fig. 5. Componentwise absolute error in the GNN model (5.2) in Example 3.

the figures, we can see that especially the first component of each column of $V(t)$ does not converge to its counterpart in the Drazin inverse $A(t)^D$.

From the Figs. 1, 3 and 5, we can see that the ZNN state matrix $V(t)$ converges to the Drazin inverse, whereas the GNN state matrix $V(t)$ diverges from the Drazin inverse.

The reason for the superiority of the ZNN over the GNN is that the ZNN exploits the time-derivative information during the real-time inverting process, which ensures its global exponential convergence to the solution of the time-varying Drazin inverse problem. In contrast, the GNN, which does not exploit the time-derivative information, is ineffective on solving such time-varying Drazin inverse problem.

6. Conclusion

Two ZNN models (3.7) and (3.10) are developed for computing the Drazin inverse of a complex time-varying matrix. The proposed network models are based on the two limit representations (2.1) and (2.2) of the Drazin inverse. Applying the Lyapunov theory, we prove the finite-time global convergence of our models and drive upper bounds for the convergence time in Theorem 4.1 and Theorem 4.2. Our experiments verify our theoretical analysis and demonstrate the superiority of the ZNN over the gradient-based GNN. Since the ZNN exploits the time-derivative information, it is more effective on the time-varying problems than the GNN.

References

[1] K.E. Avrachenkov, J.B. Lasserre, Analytic perturbation of generalized inverses, *Linear Algebra Appl.* 438 (2013) 1793–1813.
 [2] S.P. Bhat, D.S. Bernstein, Finite-time stability of continuous autonomous systems, *SIAM J. Control Optim.* 38 (2000) 751–766.

- [3] S.M. Booker, A family of optimal excitations for inducing complex dynamics in planar dynamical systems, *Nonlinearity* 13 (2000) 145–163 (19 pp.).
- [4] S.L. Campbell, *Singular Systems of Differential Equations*, Res. Notes Math., vol. 5, Pitman Advanced Publishing Program, Boston, Mass.–London, 1980, pp. 97–107.
- [5] S.L. Campbell, C.D. Meyer Jr., *Generalized Inverses of Linear Transformations*, Dover Publications, Inc., New York, 1991, corrected reprint of the 1979 original.
- [6] N. Castro González, J.J. Koliha, Y. Wei, Perturbation of the Drazin inverse for matrices with equal eigenprojections at zero, *Linear Algebra Appl.* 312 (2000) 181–189.
- [7] Y. Chen, Representation and approximation for the Drazin inverse $a^{(d)}$, *Appl. Math. Comput.* 119 (2001) 147–160.
- [8] A. Cichocki, T. Kaczorek, A. Stajniak, Computation of the Drazin inverse of a singular matrix making use of neural networks, *Bull. Pol. Acad. Sci., Tech. Sci.* 40 (1992) 387–394.
- [9] R.E. Cline, T.N.E. Greville, A Drazin inverse for rectangular matrices, *Linear Algebra Appl.* 29 (1980) 53–62.
- [10] G.H. Golub, P.C. Hansen, D.P. O’Leary, Tikhonov regularization and total least squares, *SIAM J. Matrix Anal. Appl.* 21 (1999) 185–194.
- [11] N.C. González, J. Koliha, Y. Wei, Error bounds for perturbation of the Drazin inverse of closed operators with equal spectral projections, *Appl. Anal.* 81 (2002) 915–928.
- [12] A. Hirose, S. Yoshida, Generalization characteristics of complex-valued feedforward neural networks in relation to signal coherence, *IEEE Trans. Neural Netw. Learn. Syst.* 23 (2012) 541–551.
- [13] X. Hong, S. Chen, Modeling of complex-valued wiener systems using B-spline neural network, *IEEE Trans. Neural Netw.* 22 (2011) 818–825.
- [14] J.J. Hopfield, Neurons with graded response have collective computational properties like those of two-state neurons, *Proc. Natl. Acad. Sci.* 81 (1984) 3088–3092.
- [15] J. Hu, J. Wang, Global stability of complex-valued recurrent neural networks with time-delays, *IEEE Trans. Neural Netw. Learn. Syst.* 23 (2012) 853–865.
- [16] Y. Ilyashenko, Some open problems in real and complex dynamical systems, *Nonlinearity* 21 (2008) T101–T107.
- [17] J.S. Jang, S.Y. Lee, S.Y. Shin, J.S. Jang, S.Y. Shin, An optimization network for matrix inversion, *Neural Inf. Process. Syst.* (1987) 397–401.
- [18] S. Li, S. Chen, B. Liu, Accelerating a recurrent neural network to finite-time convergence for solving time-varying Sylvester equation by using a sign-bi-power activation function, *Neural Process. Lett.* 37 (2013) 189–205.
- [19] S. Li, Y. Li, Z. Wang, A class of finite-time dual neural networks for solving quadratic programming problems and its k-winners-take-all application, *Neural Networks* 39 (2013) 27–39.
- [20] B. Liao, Y. Zhang, Different complex ZFs leading to different complex ZNN models for time-varying complex generalized inverse matrices, *IEEE Trans. Neural Netw. Learn. Syst.* 25 (2014) 1621–1631.
- [21] B. Liao, Y. Zhang, From different ZFs to different ZNN models accelerated via Li activation functions to finite-time convergence for time-varying matrix pseudoinversion, *Neurocomputing* 133 (2014) 512–522.
- [22] R.P. Lippmann, An introduction to computing with neural nets, *IEEE ASSP Mag.* 4 (1987) 4–22.
- [23] Q. Liu, J. Wang, Finite-time convergent recurrent neural network with a hard-limiting activation function for constrained optimization with piecewise-linear objective functions, *IEEE Trans. Neural Netw.* 22 (2011) 601–613.
- [24] F.L. Luo, Z. Bao, Neural network approach to computing matrix inversion, *Appl. Math. Comput.* 47 (1992) 109–120.
- [25] F. Soleymani, P.S. Stanimirović, A higher order iterative method for computing the Drazin inverse, *Sci. World J.* 2013 (2013) 206–232.
- [26] J. Song, Y. Yam, Complex recurrent neural network for computing the inverse and pseudo-inverse of the complex matrix, *Appl. Math. Comput.* 93 (1998) 195–205.
- [27] P.S. Stanimirović, I.S. Živković, Y. Wei, Recurrent neural network approach based on the integral representation of the Drazin inverse, *Neural Comput.* 27 (2015) 2107–2131.
- [28] P.S. Stanimirović, I.S. Živković, Y. Wei, Recurrent neural network for computing the Drazin inverse, *IEEE Trans. Neural Netw. Learn. Syst.* 26 (2015) 2830–2843.
- [29] J. Wang, A recurrent neural network for real-time matrix inversion, *Appl. Math. Comput.* 55 (1993) 89–100.
- [30] J. Wang, Recurrent neural networks for solving linear matrix equations, *Comput. Math. Appl.* 26 (1993) 23–34.

- [31] J. Wang, Recurrent neural networks for computing pseudoinverses of rank-deficient matrices, *SIAM J. Sci. Comput.* 18 (1997) 1479–1493.
- [32] X.Z. Wang, H. Ma, P.S. Stanimirović, Nonlinearly activated recurrent neural network for computing the Drazin inverse, *Neural Process. Lett.* (2017), <http://dx.doi.org/10.1007/s11063-017-9581-y>.
- [33] X.Z. Wang, Y. Wei, P.S. Stanimirović, Complex neural network models for time-varying Drazin inverse, *Neural Comput.* 28 (2016) 2790–2824.
- [34] Y. Wei, Recurrent neural networks for computing weighted Moore–Penrose inverse, *Appl. Math. Comput.* 116 (2000) 279–287.
- [35] Y. Wei, The Drazin inverse of a modified matrix, *Appl. Math. Comput.* 125 (2002) 295–301.
- [36] Y. Wei, G. Wang, The perturbation theory for the Drazin inverse and its applications, *Linear Algebra Appl.* 1 (1996) 118–120.
- [37] Y. Wei, H. Wu, Challenging problems on the perturbation of Drazin inverse, *Ann. Oper. Res.* 103 (2001) 371–378.
- [38] Y. Xia, B. Jelfs, M.M. Van Hulle, J.C. Principe, D.P. Mandic, An augmented echo state network for nonlinear adaptive filtering of complex noncircular signals, *IEEE Trans. Neural Netw.* 22 (2011) 74–83.
- [39] L. Xiao, A finite-time convergent neural dynamics for online solution of time-varying linear complex matrix equation, *Neurocomputing* 167 (2015) 254–259.
- [40] Y. Zhang, S.S. Ge, Design and analysis of a general recurrent neural network model for time-varying matrix inversion, *IEEE Trans. Neural Netw.* 16 (2005) 1477–1490.
- [41] Y. Zhang, B. Qiu, L. Jin, D. Guo, Z. Yang, Infinitely many Zhang functions resulting in various ZNN models for time-varying matrix inversion with link to Drazin inverse, *Inform. Process. Lett.* 115 (2015) 703–706.
- [42] Y. Zhang, Y. Yang, N. Tan, B. Cai, Zhang neural network solving for time-varying full-rank matrix Moore–Penrose inverse, *Computing* 92 (2011) 97–121.