



Design of A Nonlinearly Activated Gradient-Based Neural Network and Its Application to Matrix Inversion

Yongsheng Zhang^a, Lin Xiao^{b,*}, Lei Ding^{a,*}, Zhiguo Tan^c, Ke Chen^c, Yumin Yin^a

^aCollege of Information Science and Engineering, Jishou University, Jishou 416000, China

^bHunan Provincial Key Laboratory of Intelligent Computing and Language Information Processing, Hunan Normal University, Changsha 410081, China

^cSchool of Automation Science and Engineering, South China University of Technology, Guangzhou 510640, China

Abstract. Different from the traditional linearly activated gradient-based neural network model (GNN model), two nonlinear activation functions are presented and investigated to construct two nonlinear gradient-based neural network models (NGNN-1 model and NGNN-2 model) for matrix inversion in this paper. For comparative and illustrative purposes, the traditional GNN model is also used to solve matrix inversion problems under the same circumstance. In addition, the simulation results of the computer finally confirm the validity and superiority of the two nonlinear gradient-based neural network models specially activated by two nonlinear activation functions for matrix inversion, as compared with the traditional GNN model.

1. Introduction

The matrix inversion plays a very important role in the scientific research and practical engineering applications [1–3]; e.g., in MIMO system [1, 2], signal-processing [3], and robotics [4–8]. In the past, some numerical methods have been presented and investigated for matrix inversion [9–13]. In [9], a fast parallel algorithm is proposed for matrix inversion, and the complexity of this algorithm is analyzed. In [10], two new sorting methods are used for matrix inversion. In addition, these two new sorting methods can greatly reduce the amount of calculation as compared with other traditional sorting methods. In [11], an iterative algorithm that can solve the problem within finite iteration steps is proposed for solving the extended Sylvester-conjugate matrix equation.

Different from numerical methods, neural networks have been studied in depth and used to find the inverse of matrices in recent years [14–26]. In [22], a gradient-based neural network is proposed for matrix inversion, and its global exponential convergence performance and stability are analyzed. In [24], the software MATLAB is used to simulate the process of online matrix inversion by gradient-based neural networks, and the efficiency of online matrix inversion is verified by simulation results. In [25],

2010 Mathematics Subject Classification. Primary 92B20

Keywords. Nonlinear activation function, gradient-based neural network, matrix inversion, simulation verification

Received: 20 October 2018; Accepted: 16 January 2019

Communicated by Bolin Liao

This work is supported by the National Natural Science Foundation of China (grant nos. 61866013 and 61503152), the Natural Science Foundation of Hunan Province, China (grant nos. 2019JJ50478, 18A289, JGY201926, 2016JJ2101, 2018TP1018, and 2018RS3065), and the National Natural Science Foundation of China (grant nos. 61563017, 61561022, 61363073 and 61363033).

* Corresponding author

Email addresses: xiaolin860728@163.com (Lin Xiao), yy1xdinglei@126.com (Lei Ding)

Wang proposed a recurrent neural network for matrix inversion in real-time, and the asymptotically stable performance is proven.

Inspired by [22, 23, 27–36], we can find that some specially constructed nonlinear activation functions can improve the convergence performance of neural networks. In this paper, we focus on the study of using different nonlinear activation functions to construct nonlinear gradient-based neural networks for matrix inversion.

2. Model Formulation

In mathematics, the formula of matrix inverse $A^{-1} \in R^{n \times n}$ can be given as :

$$AX(t) = I \in R^{n \times n} \text{ or } X(t)A = I \in R^{n \times n}, \tag{1}$$

where $A \in R^{n \times n}$ denotes a known reversible matrix, $I \in R^{n \times n}$ denotes the identity matrix and $X(t) \in R^{n \times n}$ denotes the unknown matrix to be inverted in this paper. Without loss of generality, let $A^{-1} \in R^{n \times n}$ denotes the corresponding inverse of $A \in R^{n \times n}$. In the following subsections, two different nonlinear gradient-based neural networks that can find the corresponding inverse of $A \in R^{n \times n}$ in a very short period of time are constructed.

2.1. Gradient-based Neural Network (GNN)

In this subsection, the traditional GNN model is first exploited for online solution of the above-presented matrix inversion. Inspired by the traditional GNN model’s design methods [22, 24, 36–40], by defining the following norm-based energy function $\mathcal{E}(t) = \|AX - I\|_F^2/2$, and exploiting the negative gradient descent formula $-\frac{\partial \mathcal{E}(t)}{\partial X}$, we can easily obtain the traditional linearly activated gradient-based neural network as:

$$\dot{X} = -\lambda \frac{\partial \mathcal{E}(t)}{\partial X} = -\lambda A^T (AX - I), \tag{2}$$

where X denotes the state matrix corresponding to the theoretical inversion A^{-1} , λ denotes a adjustable positive design parameter, and A^T denotes the transpose of matrix A .

2.2. Nonlinear Gradient-Based Neural Network-1 Model

In the previous subsection, the traditional gradient-based neural network model was presented for matrix inversion. It is worth pointing out that the convergence performance of this traditional GNN model (2) can still be improved by adding a specially constructed nonlinear activation function [41–43]. In this subsection, an improved nonlinear gradient-based neural network model (i.e., the NGNN-1 model) is proposed for matrix inversion by using a nonlinear function (i.e., the power-sigmoid function):

$$\dot{X} = -\lambda A^T \Psi (AX - I), \tag{3}$$

where $\lambda > 0$ denote an adjustable design parameter used to adjust the convergence rate of the NGNN-1 model and the $\Psi(\cdot)$ denotes a specially constructed nonlinear activation function (i.e., the power-sigmoid function) to improve the convergence performance of the NGNN-1 model for matrix inversion. And this nonlinear activation function is defined as follows:

$$\Psi(v) = \begin{cases} v^p, & \text{if } |v| \geq 1; \\ \frac{1+\exp(-\eta)}{1-\exp(-\eta)} \cdot \frac{1-\exp(-\eta v)}{1+\exp(-\eta v)}, & \text{otherwise;} \end{cases}$$

where $v \in R$ and $|v|$ denotes the absolute value of v , and the design parameters (η, p) have to meet this conditions: $\eta \geq 1, p \geq 3$.

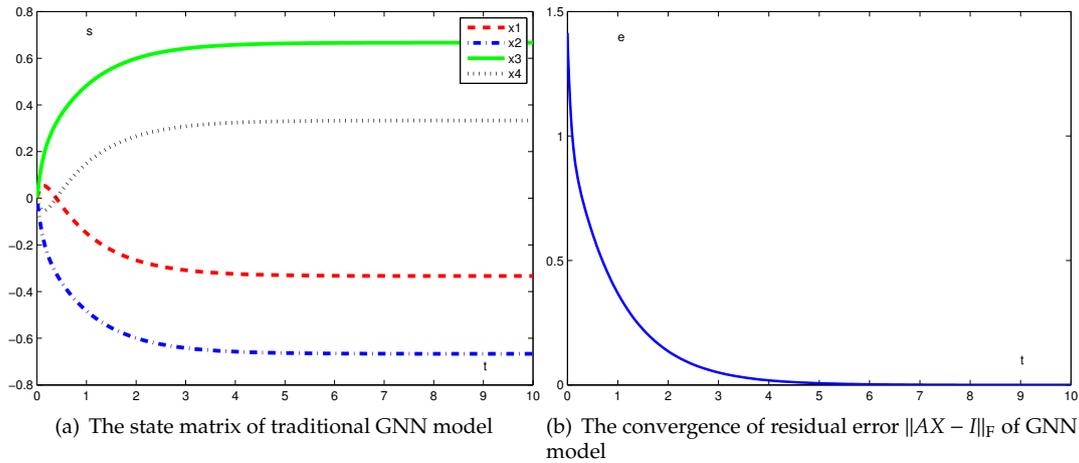


Figure 1: Simulative results synthesized by the traditional GNN model (2) with design parameter $\lambda = 1$ starting from zero initial state.

2.3. Nonlinear Gradient-Based Neural Network-2 Model

In this subsection, for the purpose of further improving the convergence performance of NGNN-1 model (3), the other improved nonlinear gradient-based neural network model (i.e., the NGNN-2 model) is proposed for matrix inversion by using another nonlinear function (i.e., the sign-bi-power function). And this sign-bi-power function $\Phi(\cdot)$ is defined as follows:

$$\Phi(v) = \text{sgn}^q(v) + \text{sgn}^{1/q}(v), \tag{4}$$

where the parameter $r \in (0, 1)$ and $\text{sgn}^q(\cdot)$ represents the following formula:

$$\text{sgn}^q(v) = \begin{cases} |v|^q, & \text{if } v > 0; \\ 0, & \text{if } v = 0; \\ -|v|^q, & \text{if } v < 0; \end{cases}$$

where $v \in R$ and $|v|$ means the same as $|v|$ in the above subsection.

Therefore, by adding such a sign-bi-power function, the second nonlinear gradient-based neural network model (i.e., the NGNN-2 model) mentioned above can be obtained as:

$$\dot{X} = -\lambda A^T \Phi(AX - I). \tag{5}$$

Note that the sign-bi-power function $\Phi(\cdot)$ can accelerate NGNN-2 model (5) to finite-time convergence, while the power-sigmoid function $\Psi(\cdot)$ only speed up the infinite convergence rate of NGNN-1 model (3). This is the first time to apply this function to activate traditional GNN model, and accelerate the traditional GNN model from infinite-time convergence to finite-time convergence.

3. Illustrative Example

In the above section, the traditional GNN model (2), NGNN-1 model (3) and NGNN-2 model (5) have been presented and investigated for matrix inversion. In this section, an illustrative example is used to verify the efficacy of such three models [GNN model (2), NGNN-1 model (3) and NGNN-2 model (5)]. Without loss of generality, we set this invertible matrix A as follows:

$$A = \begin{bmatrix} 1 & -2 \\ 2 & -1 \end{bmatrix}.$$

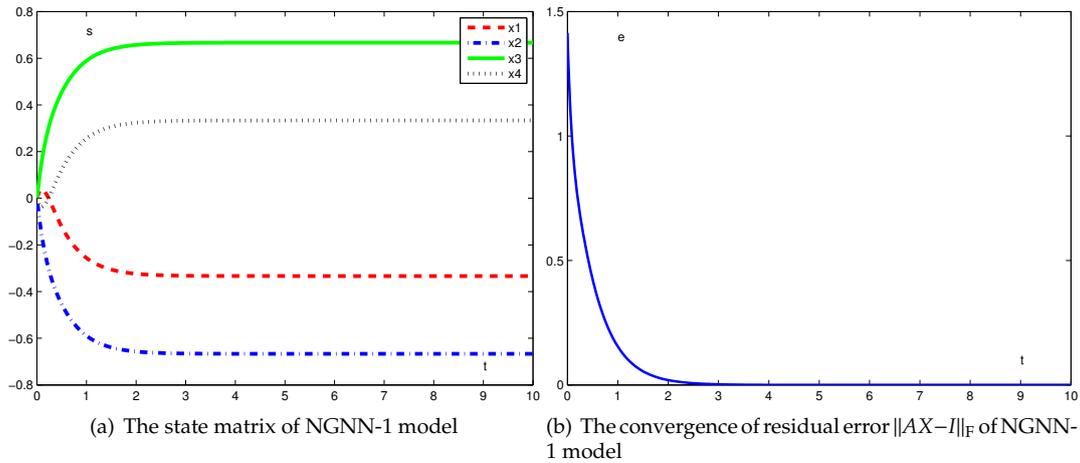


Figure 2: Simulative results synthesized by the NGNN-1 model (3) with design parameter $\lambda = 1$ starting from zero initial state.

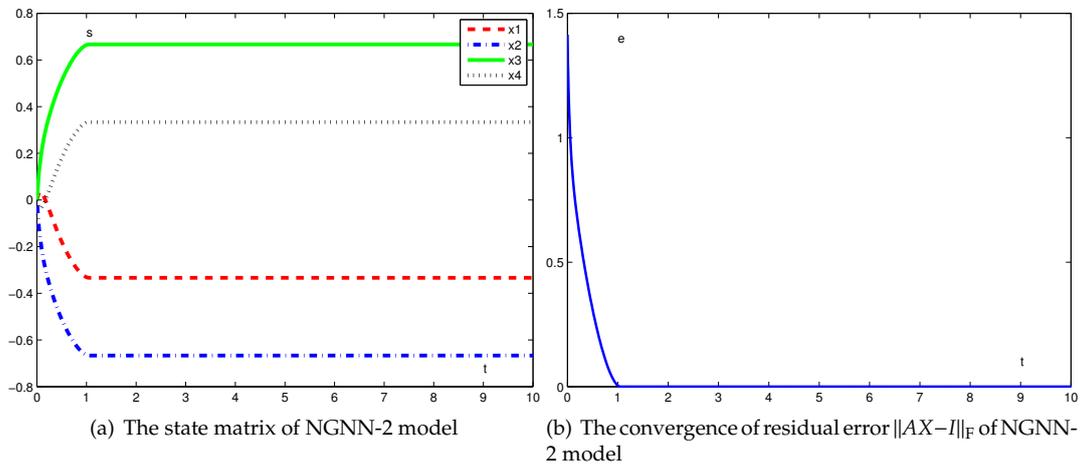


Figure 3: Simulative results synthesized by NGNN-2 model (5) with design parameter $\lambda = 1$ starting from zero initial state.

through mathematical methods, its corresponding inversion A^{-1} can be obtained as follows:

$$A^{-1} = \begin{bmatrix} -0.3333 & 0.6667 \\ -0.6667 & 0.3333 \end{bmatrix}.$$

First of all, let us set the design parameter $\lambda = 1$. Starting from zero initial state $X(0) \in R^{2 \times 2}$, the traditional GNN model is first presented to find the inversion of the matrix A , and the corresponding computer simulative results are shown in Fig. 1, which show the transient behavior of state matrix and the transient behavior of $\|AX - I\|_F$ synthesized by GNN model for matrix A inversion. As seen from Fig. 1(a), state matrix $X(t) \in R^{2 \times 2}$ of the traditional GNN model can converge to the theoretical solution tardily. Beside, from Fig. 1(b), we can see the transient behavior of $\|AX - I\|_F$ of the traditional GNN model can converge to zero about 6.7 seconds.

It is worth noting that although the traditional GNN model can converge to zero for matrix inversion, it takes longer time (about 6.7 seconds). For the purpose of improving the convergence rate of the traditional GNN model, a nonlinear activation function (power-sigmoid function) is added to the traditional GNN model, so a nonlinear gradient-based neural network model (NGNN-1 model) can be obtained for matrix inverse. The corresponding computer simulative results are shown in Fig. 2. Fig. 2(a) displays the transient

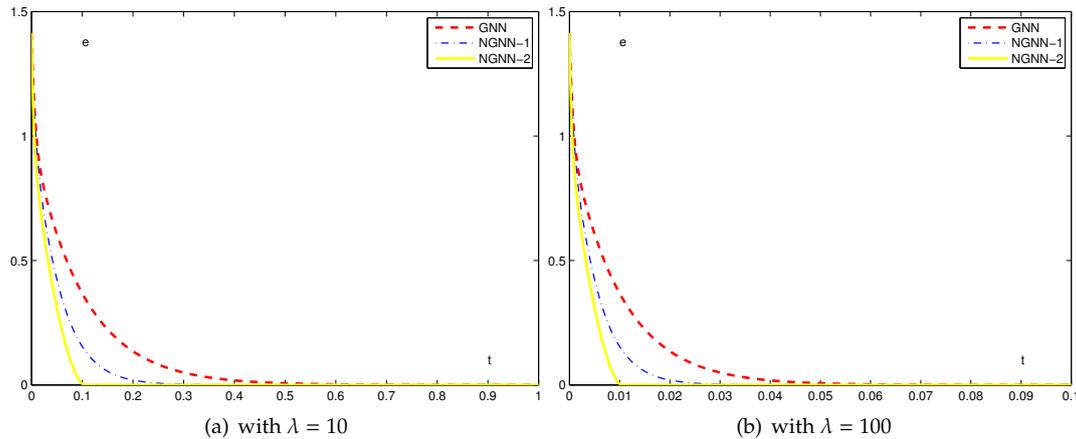


Figure 4: The convergence of the residual error $\|AX - I\|_F$ synthesized by GNN model (2), NGNN-1 model (3) and NGNN-2 model (5) with different λ .

behavior of state matrix $X(t) \in R^{2 \times 2}$ of the NGNN-1 model for matrix A inversion. Besides, from Fig. 2(b), we can see that the residual error $\|AX - I\|_F$ of the NGNN-1 model can converge to zero in only 3.7 seconds.

Besides, the sgn-bi-power function is used to construct the second nonlinear gradient-based neural network for matrix inverse. In addition, the transient behavior of state matrix $X(t) \in R^{2 \times 2}$ and the residual error $\|AX - I\|_F$ of the NGNN-2 model are show in Fig. 3, from which we can see the state matrix $X(t) \in R^{2 \times 2}$ can converge to the theoretical inverse A^{-1} for a short time and the residual error $\|AX - I\|_F$ can converge to zero only within finite time 1.03 seconds.

In addition, the numerical value of the tunable design parameter will also have an impact on the convergence rate of the presented three gradient-based neural network models [GNN model (2), NGNN-1 model (3) and NGNN-2 model (5)] for matrix inversion. The transient behavior of the residual error $\|AX - I\|_F$ synthesized by GNN model (2), NGNN-1 model (3) and NGNN-2 model (3) with design parameter $\lambda = 10$ for online matrix inversion is shown in figure Fig. 4(a), from which, we can see that the convergence time of GNN model (2), NGNN-1 model (3) and NGNN-2 model (5) can be shortened from 6.7 seconds to 0.67 seconds, from 3.7 seconds to 0.37 seconds, and from 1.03 seconds to 0.103 seconds respectively as compared with these in the case of the parameter $\lambda = 1$. Fig. 4(b) displays the transient behavior of the residual error $\|AX - I\|_F$ synthesized by GNN model (2), NGNN-1 model (3) and NGNN-2 model (5) with design parameter $\lambda = 100$ for online matrix inversion. From such a subfigure, the convergence time of GNN model (2), NGNN-1 model (3) and NGNN-2 model (5) can be further shortened from 0.67 seconds to 0.067 seconds, from 0.37 seconds to 0.037 seconds, and from 0.1 seconds to 0.01 seconds respectively as compared with these in the case of the parameter $\lambda = 10$. Obviously, with the increase of the adjustable design parameters λ , the convergence rates of these three models [GNN model (2), NGNN-1 model (3) and NGNN-2 model (5)] are correspondingly accelerated.

In summary, from the above computer simulation results, we can see that the traditional GNN model (2), the NGNN-1 model (3) and the NGNN-2 model (5) are valid for matrix inversion. Besides, the two nonlinear gradient-based neural network models [NGNN-1 model (3) and NGNN-2 model (5)] require shorter time as compared with the traditional GNN model for matrix inversion.

4. Conclusions

In this paper, two nonlinear activation functions are exploited to construct two different nonlinear gradient-based neural network models [NGNN-1 model (3) and NGNN-2 model (5)] for matrix inversion. In addition, the traditional gradient-based neural network (GNN) model has also been presented for matrix inversion under the same conditions. Computer simulation results have verified the effectiveness of the above three models for matrix inversion, and also show the superiority of the two nonlinear gradient-based

neural network models [NGNN-1 model (3) and NGNN-2 model (3)] for matrix inversion as compared to the GNN model (2).

References

- [1] J. A. Zhang, X. Huang, and H. Suzuki, Gaussian Approximation Based Interpolation for Channel Matrix Inversion in MIMO-OFDM Systems, *IEEE Transactions on Wireless Communications*, 12, (2013) 1407–1417.
- [2] J. Chen, K. Li, et al., A parallel random forest algorithm for big data in spark cloud computing environment, *IEEE Trans. Parall. Distrib. Syst.*, 28(4) (2017) 919–933.
- [3] R. J. Steriti and M. A. Fiddy, Regularized image reconstruction using SVD and a neural network method for matrix inversion, *IEEE Trans. Signal Proc.*, 41, (1993) 3074–3077.
- [4] L. Xiao, Z. Zhang, et al., Design, verification and robotic application of a novel recurrent neural network for computing dynamic Sylvester equation, *Neural Netw.* 105 (2018) 185–196.
- [5] D. Guo, K. Li, L. Yan, Z. Nie, F. Jin, The application of Li-function activated RNN to acceleration-level robots kinematic control via timevarying matrix inversion, in : *Proceedings of the Chinese Control and Decision Conference.* (2016) 3455–3460.
- [6] L. Xiao, B. Liao, S. Li, et al., Design and analysis of FTZNN applied to real-time solution of nonstationary Lyapunov equation and tracking control of wheeled mobile manipulator, *IEEE Trans. Industr. Inform.* 14(1) (2018) 98–105.
- [7] L. Xiao, S. Li, et al., Zeroing neural dynamics for control design: comprehensive analysis on stability, robustness, and convergence speed, *IEEE Trans. Ind. Inform.*, to be published, 2018.
- [8] L. Xiao, S. Li, et al., Co-design of finite-time convergence and noise suppression: a unified neural model for time varying linear equations with robotic applications, *IEEE Trans. Syst., Man, Cybern., Syst.*, to be published, 2018.
- [9] L. Csanky, Fast parallel matrix inversion algorithms, *Foundations of Computer Science*, 1975. Symposium on IEEE, (1975) 11–12.
- [10] L. Xiao, A finite-time convergent Zhang neural network and its application to real-time matrix square root finding, *Neural Comput. Appl.*, to be published, 2018.
- [11] A. G. Wu, L. L. Lv, and M. Z. Hou, Finite iterative algorithms for extended Sylvester-conjugate matrix equations, *Mathematical and Computer Modelling*, 54, (2011) 2363–2384.
- [12] Y. Zhang, W. Ma, and B. Cai, From Zhang Neural Networks to Newton Iteration for Matrix Inversion, *IEEE Trans. Circuits Syst. I*, 56, (2009) 1405–1415.
- [13] P. S. Stanimirović, M. D. Petković, Gradient neural dynamics for solving matrix equations and their applications, *Neurocomputing*, 306, (2018) 200–212.
- [14] L. Xiao, Z. Zhang, and S. Li, Solving Time-varying system of nonlinear equations by finite-time recurrent neural networks with application to motion tracking of robot manipulators, *IEEE Trans. Syst., Man, Cybern. Syst.*, to be published, 2018.
- [15] B. Qiu, and Y. Zhang, Two New Discrete-Time Neurodynamic Algorithms Applied to Online Future Matrix Inversion With Nonsingular or Sometimes-Singular Coefficient, *IEEE Trans. Cybern.*, (2018) 1–14.
- [16] B. Qiu, Y. Zhang, and Z. Yhang, New Discrete-Time ZNN Models for Least-Squares Solution of Dynamic Linear Equation System With Time-Varying Rank-Deficient Coefficient, *IEEE Trans. Neural Netw. Learn. Syst.*, 29, (2018) 5767–5776.
- [17] D. Chen, S. Li, and Q. Wu, Rejecting Chaotic Disturbances Using a Super-Exponential-Zeroing Neurodynamic Approach for Synchronization of Chaotic Sensor Systems, *Sensors*, 19, (2019).
- [18] D. Chen, Y. Zhang, and S. Li, Tracking control of robot manipulators with unknown models: A Jacobian-matrix-adaption method, *IEEE Trans. Ind. Informat.*, 14(7), (2018) 3044–3053.
- [19] L. Jin, Y. Zhang, and S. Li, Integration-enhanced Zhang neural network for real-time-varying matrix inversion in the presence of various kinds of noises, *IEEE Trans. Neural Netw. Learn. Syst.*, 27(12), (2016) 2615–2627.
- [20] L. Xiao, B. Liao, S. Li, and K. Chen, Nonlinear recurrent neural networks for finite-time solution of general time-varying linear matrix equations, *Neural Netw.*, 98, (2018) 102–113.
- [21] Y. Zhang, and 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.
- [22] Y. Zhang, Y. Shi, and K. Chen, Global exponential convergence and stability of gradient-based neural network for online matrix inversion, *Appl. Math. Comput.*, 215, (2009) 1301–1306.
- [23] S. Li, F. Qin, A dynamic neural network approach for solving nonlinear inequalities defined on a graph and its application to distributed, routing-free, range-free localization of WSNs, *Neurocomputing*, 117, (2013) 72–80.
- [24] Y. Zhang, K. Zhang, W. Ma, MATLAB simulation of gradient-based neural network for online matrix inversion, in: *Proceedings of Advanced Intelligent Computing Theories and Applications.* (2007) 98–109.
- [25] J. Wang, A recurrent neural network for real-time matrix inversion, *Appl. Math. Comput.*, 55(1), (1993) 89–10.
- [26] L. Xiao, A finite-time convergent neural dynamics for online solution of time-varying linear complex matrix equation, *Neurocomputing*, 167, (2015) 254–259.
- [27] S. Li, Y. Li, and W. Zheng, A class of finite-time dual neural networks for solving quadratic programming problems and its k-winners-take-all application, *Neural Netw.*, 39, (2013) 27–39.
- [28] L. Jin, S. Li, H.M. La, et al, Manipulability optimization of redundant manipulators using dynamic neural networks, *IEEE Trans. Ind. Electron.*, 64, (2017) 4710–4720.
- [29] L. Xiao, Y. Zhang, Solving time-varying inverse kinematics problem of wheeled mobile manipulators using Zhang neural network with exponential convergence, *Nonlinear Dyn.*, 76(2) (2014) 1543–1559.
- [30] 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.

- [31] L. Xiao, R. Lu, Finite-time solution to nonlinear equation using recurrent neural dynamics with a specially-constructed activation function, *Neurocomputing*, 151, (2015) 246–251.
- [32] S. Li, Y. Li, Nonlinearly activated neural network for solving time-varying complex Sylvester equation, *IEEE Trans. Cybern.*, 44, (2014) 1397–1407.
- [33] L. Xiao, A nonlinearly-activated neurodynamic model and its finite-time solution to equality-constrained quadratic optimization with nonstationary coefficients, *Appl. Soft Comput.*, 40, (2016) 252–259.
- [34] L. Jin, S. Li, B. Liao, Z. Zhang, Zeroing neural networks: A survey, *Neurocomputing*, 267, (2017) 597–604.
- [35] L. Xiao, A new design formula exploited for accelerating Zhang neural network and its application to time-varying matrix inversion, *Theor. Computer Sci.*, 647, (2016) 50–58.
- [36] L. Xiao, A finite-time recurrent neural network for solving online time-varying Sylvester matrix equation based on a new evolution formula, *Nonlinear Dyn.*, 90(3) (2017) 1581–1591.
- [37] L. Xiao, and Y. Zhang, Zhang neural network versus gradient neural network for solving time-varying linear inequalities, *IEEE Trans. Neural Netw.*, 22, (2011) 1676–1684.
- [38] Y. Zhang, Y. Shi, K. Chen, C. Wang, Global exponential convergence and stability of gradient-based neural network for online matrix inversion, *Appl. Math. Comput.*, 215 (2009) 1301–1306.
- [39] P.S. Stanimirovic, M.D. Petkovic, D. Gerontitis, Gradient neural network with nonlinear activation for computing inner inverses and the Drazin inverse, *Neural Process. Lett.*, 48, (2018) 109–133.
- [40] L. Xiao, and Y. Zhang, Two new types of zhang neural networks solving systems of time-varying nonlinear inequalities, *IEEE Trans. Circuits Syst. I*, 59, (2012) 2363–2373.
- [41] L. Jin, Y. Zhang, S. Li, and Y. Zhang, Noise-tolerant ZNN models for solving time-varying zero-finding problems: A control-theoretic approach, *IEEE Trans. Autom. Contr.*, 62(2), (2017) 992–997.
- [42] L. Xiao, and Y. Zhang, A new performance index for the repetitive motion of mobile manipulators, *IEEE Trans. Cybern.*, 44(2) (2014) 280–292.
- [43] L. Jin, S. Li, L. Xiao, R. Lu, and B. Liao, Cooperative motion generation in a distributed network of redundant robot manipulators with noises, *IEEE Trans. Syst. Man Cybern. Syst.*, 48(10), (2017) 1715–1724.