Letter

Recurrent Neural Network Inspired Finite-Time Control Design

Jianan Liu, Shihua Li, *Fellow, IEEE*, and Rongjie Liu, *Senior Member, IEEE*

Dear Editor,

This letter is concerned with the role of recurrent neural networks (RNNs) on the controller design for a class of nonlinear systems. Inspired by the architectures of RNNs, the system states are stacked according to the dynamic along with time while the controller is represented as the neural network output. To build the bridge between RNNs and finite-time controller, a novel activation function is imposed on RNNs to drive the convergence of states at finite-time and propel the overall control process smoother. Rigorous stability proof is briefly provided for the convergence of the proposed finite-time controller. At last, a numerical simulation example is presented to illustrate the efficiency of the proposed strategy.

Neural networks can be classified as static (feedforward) and dynamic (recurrent) nets [1]. The former nets do not perform well in dealing with training data and using any information of the local data structure [2]. In contrast to the feedforward neural networks, RNNs are constituted by high dimensional hidden states with dynamics. More specifically, the hidden state serves as the memory of the neural network that are related to the state of previous instant and the input of current instant [3]. RNNs demonstrate a workable behaviour for sequences modelling due to their self feedback architectures [4]. The specific structures of RNNs enable them to store, acquire memory and process complicated information for long time sequences [5]. Therefore, RNNs are used extensively for tasks like natural language processing, language modeling, speech recognition, and machine translation [6]. In terms of RNNs architectures and their application development, several new advances about RNNs have been summarized and research challenges have been presented [5]. The letter [7] investigated the single-layer RNN with delay between input and output, which can mimic bidirectional networks and solve some acausal tasks. A stacked recurrent neural network (SRNN) involving the dynamic states was proposed, which can be used in short-term wind power forecasting [8].

For the control problem in nonlinear systems, the controller design plays an essential role in the dynamic stability. Further, the finite-time stability of system has a wide range of practical applications in aerospace, military manufacturing and part fabrication due to its good robustness, adaptability and anti interference [9]. However, it is well known that the finite-time controller design is an intricate problem. The two primary obstacles were revealed in [10]. One is constructing a desired Lyapunov function to satisfy finite-time stabilization performance in complicated environment. The other one is its slower convergence than the exponential one when initial state is far away from the origin, which directly leads to a poor robustness.

Fortunately, an elegant framework was provided by neural networks for system identification and control design [11]. Motivated by the above finite-time control problems, it calls a design thinking and practical way for the advisable and flexible controller design. Inspired by specific architectures of RNNs, this letter aims to develop a novel idea of controller design with network structure compared to the traditional control. The main contributions of this letter can be highlighted from three aspects: 1) A new activation function with the nonlinear saturation term is proposed to guarantee the finite-time convergence and propel the overall control process smoother. 2) Both the states of closed-loop system and the training recurrent neural networks converge in finite-time, and the RNNs can achieve convergence with a finite number of iterations on the transverse time sequences determined by the system convergence time. 3) This letter provides a new thinking to the controller design from the structure of neural network. Therefore, the controller constructions of more dynamic systems can be motivated by a variety of neural network structures, which will help us to design controllers with better performance.

Corresponding author: Rongjie Liu.

Citation: J. Liu, S. Li, and R. Liu, "Recurrent neural network inspired finite-time control design," *IEEE/CAA J. Autom. Sinica*, vol. 11, no. 6, pp. 1527–1529, Jun. 2024.

J. Liu and S. Li are with the School of Automation, Southeast University, Nanjing 210096, China (e-mail: ljn@seu.edu.cn; lsh@seu.edu.cn).

R. Liu is with the Department of Statistics, Florida State University, FL 32304 USA (e-mail: rliu3@fsu.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/JAS.2023.123297

Problem formulation: Considering the global finite-time stabilization problem, we focus on a class of nonlinear systems in the form of upper-triangular structure. Here, we adopt the following order example for better illustration:

$$\begin{cases}\dot{x}_1 = x_2 + f_1(x_2, x_3)\\ \dot{x}_2 = x_3 + f_2(x_3)\\ \dot{x}_3 = u\end{cases} \tag{1}$$

where $x = [x_1, x_2, x_3]^T \in \mathbb{R}^3$ is the system state and $u \in \mathbb{R}$ is the control input, respectively. Initial condition is $x(0) = x_0$. The nonlinear functions $f_1(\cdot)$, $f_2(\cdot)$ are continuous and unknown. The objective is to design a controller $u(t)$ such that the state of closed-loop system globally converges to the origin in finite-time for any initial condition $x(0) \in \mathbb{R}^3$.

Neural network representation: The first thing is to merge dynamics into the neural network. The system considered in this letter is continuous while the RNN itself is a discrete system. When the sampling time is small enough, the discrete system can be approximated as a continuous system. An example with multiple hidden layers SRNN diagram is illustrated in Fig. 1. $l = 1, 2$ is the layer number in SRNN except the output layer, h_l^t is the hidden state of the l-th layer at time step t. For the better understanding, the corresponding relationships between system dynamics and notions in Fig. 1 are given as follows. $[x_1^t, x_2^t, x_3^t]^T$ corresponding to x^t in Fig. 1 represents the input at time step t. u^t corresponding to y^t in Fig. 1 denotes the output at time step t.

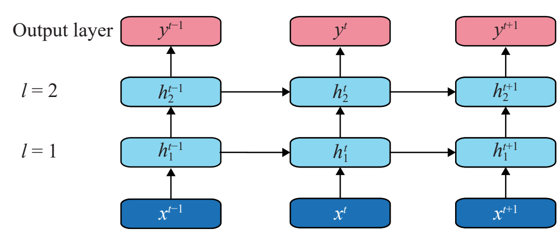

Fig. 1. The architecture of the stacked recurrent neural network.

Activation function design: Up to now, system states and control input have been embedded into the RNNs. The control input u plays a current time output role in RNNs and transfers to the next moment as input. Therefore, it is of great importance to design the activation function that produces the controller u. Next, we propose the following activation function σ:

$$\sigma(x) = \begin{cases}\epsilon \operatorname{sgn}(x), & |x| > \epsilon\\ \dfrac{\operatorname{sgn}(x)|x|^{\delta+1}}{\epsilon^{\delta}}, & |x| \le \epsilon\end{cases} \tag{2}$$

where $\epsilon > 0$ is a small constant to be determined later and $\delta = \frac{\delta_1}{\delta_2} > -1$ with an even integer δ_1 and an odd integer δ_2.

Remark 1: The novel activation function σ will be reduced to the conventional saturation function when $\delta = 0$, that is

$$\bar{\sigma}(x)=\begin{cases}\epsilon\,\text{sgn}(x), & |x|>\epsilon\\ x, & |x|\le\epsilon.\end{cases}$$

What is noteworthy is that the controller with activation function σ is more robust near the equilibrium point as well as stabilizes the system in finite-time.

Finite-time control design: Based on the proposed activation function σ, next we will design the controller. To simplify notations, we define $X_i=(x_1,\dots,x_i)$, $i=1,2,3$. The nonlinear nested saturation controller is given as follows:

$$u=u_3(X_3(t)) \tag{3}$$

$$u_i(X_i(t))=-\beta_i\sigma^{\frac{r_{i+1}}{a}}\left(x_i^{\frac{a}{r_i}}-u_{i-1}^{\frac{a}{r_i}}(X_{i-1})\right) \tag{4}$$

where $u_0=0$, $a=\max_{1\le i\le 3}\{r_i,r_{i+1}\}$, $r_1=1$, $r_2=r_1+\tau$, $r_3=r_2+\tau$, $\tau>-\frac{1}{3}$, $\tau=\frac{\tau_1}{\tau_2}$ with an even integer τ_1 and an odd integer τ_2. The constants r_2, r_3 are ratio of two odd numbers. The gains β_i are select as $\beta_1^*=3$, $\beta_i^*=c_i+\gamma_i+4-i$, $\beta_1>\max\{\beta_1^*,3\cdot 2^{\frac{r_2}{a}}\}$, $\beta_i>\max\{\beta_i^*, 2^{\frac{r_{i+1}}{a}}[4(1+\beta_{i-1})\alpha_{i-1}(\beta_1,\dots,\beta_{i-1})+1]\}$, with $\alpha_1(\beta_1)=a\beta_1^{\frac{a}{r_2}}(2+\beta_1)$, $\alpha_2(\beta_1,\beta_2)=\frac{a\beta_2^{\frac{a}{r_3}}}{r_2}(1+\beta_1)^{\frac{a}{r_2}-1}(2+\beta_2)+\beta_2^{\frac{a}{r_3}}\alpha_1(\beta_1)$. Here, c_i,γ_i are constants will be shown in the proof later.

In order to describe characteristics of neural network structure more intuitively, we introduce the transformation $\varrho_i=-\beta_i\sigma^{\frac{r_{i+1}}{a}}$, and for $i=1,2,3$, the functions are defined as $z_i=\hbar(x_i)=x_i^{\frac{a}{r_i}}$, $v_{i-1}=\rho(u_{i-1})=u_{i-1}^{\frac{a}{r_i}}$. Then, the controller can be re-expressed as below:

$$u_i=\varrho_i(z_i-v_{i-1})=\varrho_i(\hbar(x_i)-\rho(u_{i-1})),\quad i=1,2,3. \tag{5}$$

The Fig. 2 illustrates an example of the neural network structure after embedding dynamic system at two time points. $[x_1^t,x_2^t,x_3^t]^T$, u_3^t represent the input vector and output at time step t, respectively. And u_3^t will transfer to the next time as the initial control input. u_l^t, $l=1,2$ are hidden states of the layer l at time step t. ϱ_1, ϱ_2, ϱ_3 are activation functions that can be well-designed.

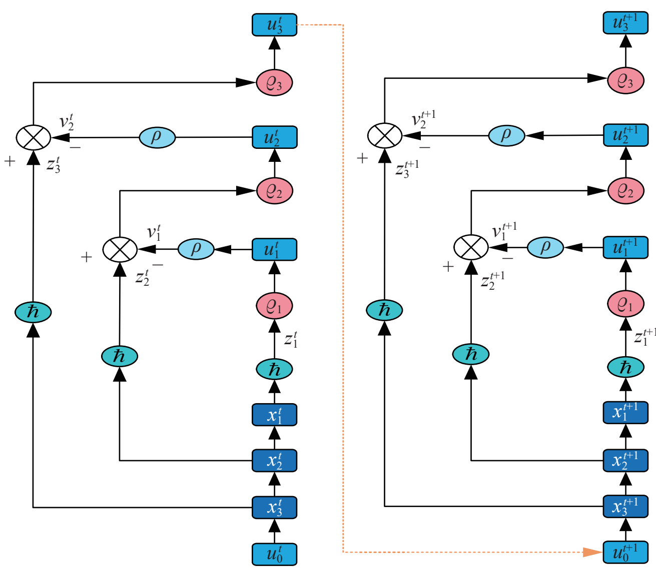

Fig. 2. Recurrent neural network architecture about proposed finite-time controller with new activation function.

Stability analysis: At first, we introduce the following assumption and important lemma used in control design and theoretical analysis. After that, the main result in Theorem 1 can be formally stated.

Assumption 1: In a neighborhood of the origin, there holds $|f_i(\cdot)|\le b(|x_{i+1}|^{q_{i,i+1}}+\cdots+|x_3|^{q_{i,3}})$ for positive constants b and $q_{i,j}$ satisfying $q_{i,j}>(r_{i+1})/r_j>0$, $i=1,2$, $j=2,3$.

Lemma 1 [12]: For system (1), under control law (3), there exist a constant $0<\epsilon_1<1$ and functions $\alpha_1(\beta_1)$, $\alpha_2(\beta_1,\beta_2)$ such that $|f_i(x_{i+1},\dots,x_3)|\le\epsilon^{\frac{r_{i+1}}{a}}$, $|u_i^{\frac{a}{r_{i+1}}}(X_i(\bar{t}))-u_i^{\frac{a}{r_{i+1}}}(X_i(t))|\le\alpha_i(\beta_1,\dots,\beta_i)\epsilon^{\frac{a+\tau}{a}}(\bar{t}-t)$, provided that $|x_j|\le\epsilon^{\frac{r_j}{a}}(1+\beta_{j-1})$, $j=2,3$, $i=1,2,3$, for any $0<\epsilon\le\epsilon_1$, $-1<\delta\le 0$ and $\forall\bar{t}\ge t$.

Theorem 1: Under Assumption 1, for $-\frac{1}{3}<\tau<0$ and $-1<\delta\le 0$, there exists a constant $\epsilon\in(0,\epsilon_1]$ such that the control law (3) will globally stabilize the upper-triangular system (1).

Proof: The complete proof is divided into three steps.

Step 1: We prove that there exists a time instance t_1 such that

$$X_3(t)\in Q_3=\{X_3:|x_3^{\frac{a}{r_3}}(t)-u_2^{\frac{a}{r_3}}(X_2(t))|<\epsilon\},\quad t\ge t_1. \tag{6}$$

Before demonstrate that, let us use a contradiction argument to prove that there exists a time instance t_1 such that $|x_3^{\frac{a}{r_3}}(t_1)-u_2^{\frac{a}{r_3}}(X_2(t_1))|\le\frac{\epsilon}{2}$. Otherwise, it can be assumed that for all $t\ge 0$ such that $|x_3^{\frac{a}{r_3}}(t)-u_2^{\frac{a}{r_3}}(X_2(t))|>\frac{\epsilon}{2}$. The following case is taken in consideration first:

$$x_3^{\frac{a}{r_3}}(t)-u_2^{\frac{a}{r_3}}(X_2(t))>\frac{\epsilon}{2}. \tag{7}$$

According to (1) and (3), for $t\ge 0$, there holds

$$\dot{x}_3(t)=-\beta_3\sigma^{\frac{r_4}{a}}\left(x_3^{\frac{a}{r_3}}(t)-u_2^{\frac{a}{r_3}}(X_2(t))\right)<-\mu_3\epsilon^{\frac{r_4}{a}}$$

with $\mu_3=\beta_3(\frac{1}{2})^{\frac{r_4}{a}}>0$ determined by the definition of β. By integrating both sides of the above inequality for $t\ge 0$, one can obtains that $x_3(t)<x_3(0)-\mu_3\epsilon^{\frac{r_4}{a}}t$. This, together with the fact that $|u_2^{\frac{a}{r_3}}(X_2(t))|\le\beta_2^{\frac{a}{r_3}}\epsilon$, one can get $x_3^{\frac{a}{r_3}}(t)-u_2^{\frac{a}{r_3}}(X_2(t))\le(x_3(0)-\mu_3\epsilon^{\frac{r_4}{a}}t)^{\frac{a}{r_3}}+\beta_2^{\frac{a}{r_3}}\epsilon$, which leads to a contradiction disavowing (7) when $t\to\infty$. Likewise, note that in the occasion that $x_3^{\frac{a}{r_3}}(t)-u_2^{\frac{a}{r_3}}(X_2(t))<-\frac{\epsilon}{2}$, $t\ge 0$ is also impossible. Therefore, the inequality $|x_3^{\frac{a}{r_3}}(t_1)-u_2^{\frac{a}{r_3}}(X_2(t_1))|\le\frac{\epsilon}{2}$ holds for a time instance t_1. Next, using a contradiction argument again to demonstrate that $|x_3^{\frac{a}{r_3}}(t)-u_2^{\frac{a}{r_3}}(X_2(t))|<\epsilon$, $t\ge t_1$. Hence, for $t\ge t_1$, the proof of (6) is complete.

On the basis of the above discussion, we can further prove that there exists $t\ge t_2\ge t_1\ge 0$ such that $X_2(t)\in Q_2=\left\{X_2:|x_2^{\frac{a}{r_2}}(t)-u_1^{\frac{a}{r_2}}(X_1(t))|<\epsilon\right\}$. Moreover, one has $x_1(t)\in Q_1=\{x_1:|x_1^{\frac{a}{r_1}}(t)|<\epsilon\}$, $t\ge t_3$. $X_3(t)$ will enter and stay in the set

$$\begin{aligned}Q=\{X_3:&|x_1^{\frac{a}{r_1}}(t)|<\epsilon,|x_2^{\frac{a}{r_2}}(t)-u_1^{\frac{a}{r_2}}(X_1(t))|<\epsilon,\\ &|x_3^{\frac{a}{r_3}}(t)-u_2^{\frac{a}{r_3}}(X_2(t))|<\epsilon\},\quad t\ge t_3.\end{aligned} \tag{8}$$

Step 2: In this step, we provide the proof of $Q\subset\Omega$ (Ω is the abstraction domain). For $i=1,2,3$, let $x_1^*=0$, $x_i^*=-\beta_{i-1}(x_{i-1}^{\frac{a}{r_{i-1}}}-x_{i-1}^{*\frac{a}{r_{i-1}}})^{\frac{r_i(1+\delta)}{a}}\epsilon^{-\frac{r_i\delta}{a}}$, $\xi_i=x_i^{\frac{a}{r_i}}-x_i^{*\frac{a}{r_i}}$, $V_i(X_i)=\sum_{k=1}^{i}\int_{x_k^*}^{x_k}(s^{\frac{a}{r_k}}-x_k^{*\frac{a}{r_k}})^{\frac{2a-r_{k+1}}{a}}ds$, $\Omega=\{X_3|V_3(X_3)\le\lambda_0\}$, $\lambda_0>0$. It can be followed from Step 1 that:

$$u_i=x_{i+1}^*=-\beta_i(x_i^{\frac{a}{r_i}}-x_i^{*\frac{a}{r_i}})^{\frac{r_{i+1}(1+\delta)}{a}}\epsilon^{-\frac{r_{i+1}\delta}{a}},\quad t\ge t_3. \tag{9}$$

So, (8) can be reformulated as $Q=\{X_3:|x_1^{\frac{a}{r_1}}(t)|<\epsilon,|x_2^{\frac{a}{r_2}}(t)-x_2^{*\frac{a}{r_2}}(t)|<\epsilon,|x_3^{\frac{a}{r_3}}(t)-x_3^{*\frac{a}{r_3}}(t)|<\epsilon\}$. By $|x_k-x_k^*|\le 2^{1-\frac{r_k}{a}}\epsilon^{\frac{r_k}{a}}$, one get $V_3(X_3)\le\sum_{k=1}^{3}2^{1-\frac{r_k}{a}}\epsilon^{2-\frac{\tau}{a}}$. Without loss of generality, an appropriate ϵ can be selected to meet the inequality $\sum_{k=1}^{3}2^{1-\frac{r_k}{a}}\epsilon^{2-\frac{\tau}{a}}\le\lambda_0$. With $-\frac{1}{3}<\tau<0$, we can obtain that $\sum_{k=1}^{3}2^{1-\frac{r_k}{a}}\epsilon^{2-\frac{\tau}{a}}\le 6\epsilon^{2-\tau}\le 6\epsilon^2$. Then, we denote by $\epsilon=\min\{\epsilon_1,(\frac{\lambda_0}{6})^{\frac{1}{2}}\}$. In conclusion, we essentially have $Q\subset\Omega$.

Step 3: In this step, we will prove that the system (1) is locally stable in the abstraction domain Ω. Consider the following system:

$$\dot{x}_1=x_2,\ \dot{x}_2=x_3,\ \dot{x}_3=u.$$

By the definition of V_i, one obtains that $\dot{V}_1(x_1)=x_1^{\frac{2a-r_2}{r_1}}x_2^*+x_1^{\frac{2a-r_2}{r_1}}(x_2-x_2^*)$, where $x_2^*=-\beta_1\xi_1^{\frac{r_2(1+\delta)}{a}}\epsilon^{-\frac{r_2\delta}{a}}$. With $\beta_1>\beta_1^*=3$ and $|\xi_1|=|x_1^{\frac{a}{r_1}}|<\epsilon$, we get $\dot{V}_1(x_1)\le-3\xi_1^2+\xi_1^{\frac{2a-r_2}{a}}(x_2-x_2^*)$.

Choose $V_2(X_2)=V_1(x_1)+\int_{x_2^*}^{x_2}(s^{\frac{a}{r_2}}-x_2^{*\frac{a}{r_2}})^{\frac{2a-r_3}{a}}ds$, by addressing indefinite terms and choosing gain $\beta_2\ge\beta_2^*(\beta_1)=[c_2+\gamma_2(\beta_1)+2]$, one obtains $\dot V_2(X_2)\le-2(\xi_1^2+\xi_2^2)+\xi_2^{\frac{2a-r_3}{a}}(x_3-x_3^*)$. Similarly, we can deduce that $\dot V_3(X_3)\le-(\xi_1^2+\xi_2^2+\xi_3^2)$.

Define H_1 and H_2 as $H_1(X_3)=\sum_{i=1}^3\xi_i{}^2$, $H_2(X_3)=\sum_{i=1}^3\frac{\partial W_i}{\partial x_i}\times\sum_{j=i+1}^4|x_j|^{\frac{r_{i+1}}{r_j}}$, respectively. According to homogeneous theory, it is not hard to testify that H_1, H_2 and $V_3^{\frac{2}{2-\tau}}$ are homogeneous of degree 2 with respect to the dilation (r_1,r_2,r_3). And there exist positive constants d_1 and d_2 such that $H_1\ge d_1V_3^{\frac{2}{2-\tau}}$, $H_2\le d_2V_3^{\frac{2}{2-\tau}}$. We can find a region $\Omega=\{X_3|V_3(X_3)\le\lambda_0\}$ with $\lambda_0>0$ such that $\sum_{i=1}^3\sum_{j=i+1}^4|x_j|^{q_{i,j}-\frac{r_{i+1}}{r_j}}\le\frac{d_1}{2d_2b}$, for $\forall X_3\in\Omega$. It is clearly to see that $\dot V_3(X_3)\le-\frac{1}{2}d_1V_3^{\frac{2}{2-\tau}}(X_3)$, which implies that the system (1) is locally finite-time stable. ■

Numerical example: Consider the following nonlinear system:

$$\dot x_1=x_2+x_2^2+x_3^2,\ \dot x_2=x_3,\ \dot x_3=u.$$

For our proposed controller u with proposed activation function σ, select $\tau=-\frac{2}{15}$, so $r_1=1$, $r_2=\frac{13}{15}$, $r_3=\frac{11}{15}$, $r_4=\frac{3}{5}$. According to the Theorem 1, a global finite-time stabilizer can be designed as

$$u=-\beta_3\sigma^{\frac{3}{5}}\left\{x_3^{\frac{15}{11}}-\{-\beta_2\sigma^{\frac{11}{15}}\{x_2^{\frac{15}{13}}-\{-\beta_1\sigma^{\frac{13}{15}}(x_1)\}^{\frac{15}{13}}\}\}^{\frac{15}{11}}\right\} \tag{10}$$

with appropriate gains $\beta_1=1$, $\beta_2=3$, $\beta_3=100$ and $\epsilon=0.5$ that also apply to the other two controllers. In simulation, choose initial values as $x_1(0)=1$, $x_2(0)=2$, $x_3(0)=1$. The alternative nonlinear controller $\bar u$ and linear controller $\tilde u$ with $\bar\sigma$ are given as

$$\bar u=-\beta_3\bar\sigma^{\frac{3}{5}}\left\{x_3^{\frac{15}{11}}-\{-\beta_2\bar\sigma^{\frac{11}{15}}\{x_2^{\frac{15}{13}}-\{-\beta_1\bar\sigma^{\frac{13}{15}}(x_1)\}^{\frac{15}{13}}\}\}^{\frac{15}{11}}\right\} \tag{11}$$

and

$$\tilde u=-\beta_3\bar\sigma\left\{x_3-\{-\beta_2\bar\sigma\{x_2-\{-\beta_1\bar\sigma(x_1)\}\}\}\right\}. \tag{12}$$

It should be noticed that the state convergence speed under the proposed controller (10) is fastest compared to (11) and (12). The convergence time of x are 11.6 s, 15.2 s and 35.1 s for same initial values under three different controllers u, $\bar u$, and $\tilde u$, which can be seen from Figs. 3−5 in details. There are no dramatic changes for control magnitudes between u and $\bar u$, and the former has faster convergent speed, which can be seen from Fig. 6.

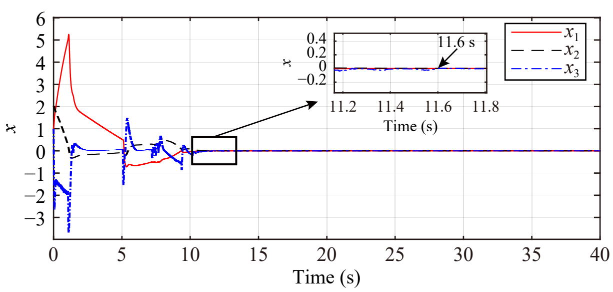

Fig. 3. The trajectories of the state x under proposed controller u.

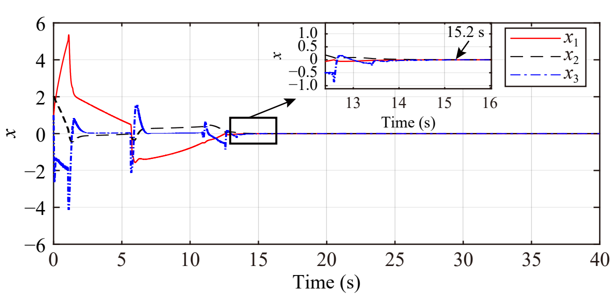

Fig. 4. The trajectories of the state x under nonlinear controller $\bar u$.

Conclusion: In this letter, the recurrent neural networks effect in the controller design has been investigated. A neural network structure has been presented for a specific nested saturation controller. We genuinely hope to make it clear to readers that this work is more than a simple finite-time control design. There are lots of open directions following our work, especially how we can get a completely different controller by changing one component of the neural network architecture, for instance, an arrow direction, an input sequence or one neuron. This idea can be used to develop the effective controllers for more generalized control systems.

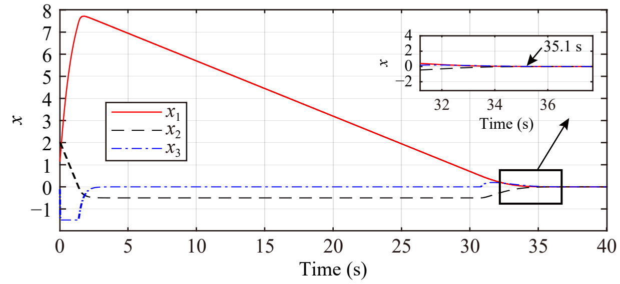

Fig. 5. The trajectories of the state x under linear controller $\tilde u$.

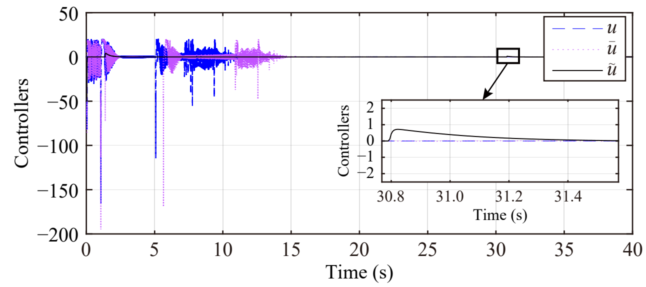

Fig. 6. The comparison of three different controllers.

References

[1] W. Yu, "Nonlinear system identification using discrete-time recurrent neural networks with stable learning algorithms," *Inf. Sci.*, vol. 158, pp. 131–147, 2004.

[2] W. Yu and X. Li, "Some new results on system identification with dynamic neural networks," *IEEE Trans. Neural Netw.*, vol. 12, no. 2, pp. 412–417, 2001.

[3] T. Mikolov, A. Joulin, S. Chopra, M. Mathieu, and M. Ranzato, "Learning longer memory in recurrent neural networks," arXiv preprint arXiv: 1412.7753, 2014.

[4] Y. Bengio, P. Simard, and P. Frasconi, "Learning long-term dependencies with gradient descent is difficult," *IEEE Trans. Neural Netw.*, vol. 5, no. 2, pp. 157–166, 1994.

[5] H. Salehinejad, S. Sankar, J. Barfett, E. Colak, and S. Valaee, "Recent advances in recurrent neural networks," arXiv preprint arXiv: 1801.01078, 2017.

[6] J. Wang, L. Zhang, Q. Guo, and Z. Yi, "Recurrent neural networks with auxiliary memory units," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 29, no. 5, pp. 1652–1661, 2018.

[7] J. Turek, S. Jain, V. Vo, M. Capotă, A. Huth, and T. Willke, "Approximating stacked and bidirectional recurrent architectures with the delayed recurrent neural network," in *Proc. Int. Conf. Mach. Learn.*, PMLR, 2020, pp. 9648−9658.

[8] X. Liu, J. Zhou, and H. Qian, "Short-term wind power forecasting by stacked recurrent neural networks with parametric sine activation function," *Electr. Power Syst. Res.*, vol. 192, p. 107011, 2021.

[9] Y. Liu, H. Li, Z. Zuo, X. Li, and R. Lu, "An overview of finite/fixed-time control and its application in engineering systems," *IEEE/CAA J. Autom. Sinica*, vol. 9, no. 12, pp. 2106–2120, Dec. 2022.

[10] S. P. Bhat and D. S. Bernstein, "Finite-time stability of continuous autonomous systems," *SIAM J. Control Optim.*, vol. 38, pp. 751–766, 2000.

[11] A. Perrusquía and W. Yu, "Identification and optimal control of nonlinear systems using recurrent neuralnetworks and reinforcement learning: An overview," *Neurocomput.*, vol. 438, pp. 145–154, 2021.

[12] R. Liu, S. Li, and S. Ding, "Nested saturation control for overhead crane systems," *Trans. Inst. Meas. Control*, vol. 34, no. 7, pp. 862–875, 2012.