Adaptive sliding-mode observer for second order discrete-time MIMO nonlinear systems based on recurrent neural-networks

Salgado, I., Ahmed, H., Camacho, O. & Chairez, I.

The final publication is available at Springer via http://dx.doi.org/10.1007/s13042-018-00908-z

Adaptive sliding-mode observer for second order discrete-time MIMO nonlinear systems based on recurrent neural-networks

Iván Salgado · Hafiz Ahmed · Oscar
Camacho · Isaac Chairez

Abstract This manuscript introduces a novel methodology to solve the state estimation of discrete-time multi-input multi-output (MIMO) nonlinear systems with uncertain dynamics. The mathematical model of the nonlinear systems considered in this paper satisfies the usual Lagrangian structure that characterizes many mechanical, electrical or electromechanical systems. A recurrent neural network (RNN) estimates the uncertain dynamics of the MIMO system with an updating law based on a particular variant of the discrete-time version of the super-twisting algorithm (DSTA). A Lyapunov stability analysis defines the convergence zone for the state estimation error throughout the solution of a matrix inequality. The convergence zone for the estimation is smaller when the DSTA and the RNN work together in an observer. Numerical examples demonstrate how the adaptive observer reduces the zone of convergence and the oscillations in the steady state compared with a discrete version of the STA with additional linear correcting terms. An experimental implementation shows how the observer estimates the unknown states of a Van Der Pol Oscillator. A comparison against some variations of the DSTA justifies the advantages of the mixed DSTA-RNN observer.

Isaac Chairez
UPIBI-Instituto Politécnico Nacional
Mexico City, Mexico E-mail: isaac_chairez@hotmail.com

Iván Salgado and Oscar Camacho
CIDETEC-Instituto Politécnico Nacional
Mexico City, Mexico E-mail: isalgador@ipn.mx

Hafiz Ahmed
Coventry University
London, United Kingdom E-mail: ac7126@coventry.ac.uk

1 Introduction

1.1 Preliminaries and motivation

The estimation of non-measurable states for nonlinear systems with uncertain dynamics is a relevant problem in control theory [1,2]. State observers are the main tool aimed to solve the estimation of the non-measurable states. The most common designs introduce a linear correction term depending on the output error (usually known as the Luenberger form). These observers require the mathematical description of the system under study [3,4]. Nevertheless, these models may be inaccurate to represent the whole dynamics of the system. Therefore, the state observer may not produce a reliable estimation of the real non-measurable states [5,6].

Sliding mode observers (SMO) offer a different alternative to solve the problem of state estimation for uncertain systems. SMO may endorse finite-time convergence for the estimation error and robustness against parametric uncertainties as well as bounded external perturbations [7,8]. Nowadays, there is a bunch of design methods for SMO based on the first and higher (equal or greater than two) order sliding mode theories. The main disadvantage of the so-called first order sliding mode (FOSM) solutions is the high frequency oscillations in the estimation error trajectory (known as *chattering*) [9]. This undesired characteristic is partially alleviated by increasing the order of the controller. Despite this operational disadvantage, SMO exhibit better estimation performance than high gain observers when the mathematical description of the system is not available [10,11]. In particular, the super-twisting algorithm (STA) can estimate the uncertain dynamics of the second order system if the upper-bound of the function associated with the time derivative of the second state is known [12]. The success of STA came from its numerous applications in state estimation, feedback control and parameter identification problems. In particular, the STA was successfully applied to solve the state observer design for single-input single-output (SISO) or multi-input multi-output (MIMO) systems [13,14] as well as to design output feedback controllers [15, 16]. Just recently, some results [17–19] have introduced the discrete-time analysis of the STA working as a nonlinear observer, but requiring a suitable mathematical description of the system that should be estimated. Moreover, the application of SMO enforces the presence of quasi-periodic oscillations on the states affected by the controller or by the observer. This movement is known as quasi-sliding mode and its amplitude is proportional to the gains fixed for the sliding mode algorithm [20–23].

Notice that, all the STA applications require a discrete-time implementation in embedded systems (using complex methods of integration). In consequence, there is an evident necessity of developing a discrete-time analysis of stability for the STA. The existing results offered analytic solutions to adjust the gains of the DSTA. Nevertheless, the possibility of selecting the gains value to reduce the size of region that characterize the quasi-sliding mode regimen has not been deeply studied. Adaptive techniques are an alternative to

reduce the zone of convergence as well as new designs that exploit the benefits of a sliding mode observer and the characteristics of adaptive techniques where the mathematical model is inaccurate or even unavailable [24]. When the mathematical description of the system (for continuous and discrete plants) is unavailable, adaptive modeling techniques can reduce the steady state oscillations of controlled systems or estimated states in the case of designing observers [25–28].

Artificial Neural Networks (ANN) may approximate nonlinear functions affected by the presence of bounded perturbations, model uncertainties and signal noises [29,30]. The adaptive modeling by DNN has been developed for both, continuous and discrete-time systems. For discrete-time systems, recurrent neural networks (RNN) [31] are updated by learning laws represented by equations in differences. The work developed in [32] addresses the identification, state estimation and tracking problems for DNN with continuous dynamics (DiffNN) by means of Lyapunov control functions. On the other hand, RNN have employed the backpropagation through-time algorithm to update their free parameters [33]. In the context of discrete-time Lyapunov theory, the work in [34] updates the weights of a RNN by means of an extended Kalman Filter like learning. Besides, the work presented in [35] exploits the concept of Lyapunov controlled functions to develop new learning laws for RNN.

Several results [36–40] explore solutions based on sliding modes (SM) and DNN to overcome the disadvantages of each individual technique. The work developed in [36] solves the adaptive control problem based on the DNN with a first order SM learning law. For second-order systems, the work in [26] suggests a two stages observer: in the first stage, a DNN estimates the dynamics of the mechanical system under analysis, and in the second stage, the STA working as an observer rejects some kinds of bounded perturbations. All these collective solutions based on both DNN and SM work mainly in continuous-time systems. The work in [18] introduces a discrete-time version of the STA (DSTA) with the stability analysis in terms of the discrete-time second Lyapunov method. However, this algorithm requires the complete knowledge of the functions that describe the system dynamics or at least an upper-bound for them.

1.2 Contribution

There are several designs gathering the advantages of DiffNN and SM solutions in continuous time. The lack of results in discrete-time framework that justifies the real-time implementation of such continuous observers constitutes a strong motivation to develop a discrete-time observers for second order nonlinear systems based on an integrating method using RNN and the DSTA. Therefore, the main contributions of this study are

1. The design of a discrete-time adaptive observer for MIMO systems with a adaptive second order sliding mode structure.

2. A novel learning law obtained from a discrete time controlled Lyapunov analysis with two different correction terms: a classical linear term and a nonlinear one based on the DSTA.
3. The comparative analysis of the region of convergence for the estimation error and how this proposal reduces it when the RNN is introduced in the observer design.
4. The method issued to calculate the gains needed in both the RNN and the DSTA.

1.3 Structure of the paper

Section 2 describes the problem formulation, the class of nonlinear systems and the characteristics of the RNN structure used in this study to approximate the uncertain system. Section 3 presents the observer structure implementing the RNN and the DSTA. The main result is summarized in the Theorem included in this section. Section 4 contains numerical results related to the state estimation of a simple pendulum and a two-link robot manipulator. The results and advantages provided by the RNN with DSTA learning law are compared with an observer based on a generalization of the STA working as a feasible realization of a differentiator in discrete-time [41]. Section 5 provides a real experimental setup where the states of a Van Der Pol Oscillator are estimated by the RNN-DSTA observer. Some discussion about the results obtained in this study are formulated in Section 6. Finally, Section 7 presents the conclusions of this work.

2 Problem formulation

The problem solved in this study was to design an adaptive observer based on RNN and DSTA intended to estimate the states of an uncertain and perturbed MIMO system formed by a set of second order systems. The observer design included the adaptive laws that adjusted the weights of the RNN as well as the gains of the DSTA. Consider the class of uncertain discrete-time nonlinear MIMO systems governed by a set of $2n$ equations in differences and an algebraic linear state output mapping:

$$
\begin{aligned}
x_i\left(k+1\right) &= x_i\left(k\right) + \tau x_{n+i}\left(k\right) \\
x_{n+i}\left(k+1\right) &= x_{n+i}\left(k\right) + \tau f_i\left(x\left(k\right), u\left(k\right)\right) + \tau \xi_i\left(k\right) \\
y_i\left(k\right) &= x_i\left(k\right)
\end{aligned}
\tag{1}
$$

with $i = 1, \ldots, n$. $x := [x_1, \ldots, x_{2n}]^\top$, $x \in \mathbb{R}^{2n}$ is the state of the MIMO second-order nonlinear uncertain system. The output of the system is $y \in \mathbb{R}^n$ formed as $y := [y_1, \ldots, y_n]^\top$. The positive scalar τ defines the sampling time for the discrete system. The signal $u \in \mathbb{R}^n$, $u := [u_1, \ldots, u_n]^\top$ is the control action belonging to the admissible control set $U^{adm} := \left\{u(k) : \|u(k)\| \le u_0^+ < \infty\right\}$.

The nonlinear function $f_i : \mathbb{R}^{2n} \times \mathbb{R}^n \to \mathbb{R}^n$ is uncertain but there exists (by assumption) a known constant $f^+ \in \mathbb{R}^+$ such that $\|f_i(\cdot, \cdot)\| \leq f_i^+$. The vector $\xi := [\xi_1, \dots, \xi_n]^\top$ represents the non measurable and bounded system perturbations. The set U^{adm} may contain a large class of stabilizing control designs. Therefore, one may assume that the system trajectories are bounded as follows:

$$\|x(k)\| \leq \varepsilon, \quad 0 < \varepsilon < +\infty, \ \forall k \geq 0 \tag{2}$$

Notice that the assumption given for the function f_i, supports the application of diverse state estimators based on sliding modes [18]. However, the upper bound f_i^+ may overestimate the value of f_i throughout the time which in the case of discrete SMO may lead to larger oscillations in the transient period of the estimation of x_{n+i}. Indeed, this undesirable behavior can affect the design of a possible output feedback controller. For example, the domain where the control action is valid may reach big values that could be unattainable by real actuators.

One of the main objectives of this work is to relax the assumption of the class of systems that can be estimated when the function f_i is unavailable. The assumption presented in (2) justifies the locally nonlinear approximation of system (1) by a RNN. Consider that the states of this RNN as $\hat{x} \in \mathbb{R}^{2n}$ and their dynamics depend on two sets of parameters W_1 and W_2 that must be adjusted in such a way that the estimation error $\Delta = \hat{x} - x$ has its origin as a practical stable equilibrium point with a convergence region characterized by a positive scalar β defined as

$$\beta := \overline{\lim_{k \to \infty}} \|\Delta(k)\| \tag{3}$$

In equation (3), $\overline{\lim_{k \to \infty}} (\cdot) = \limsup_{k \to \infty} (\cdot)$. The learning laws depend only on the available output y. Then, the RNN used to approximate the discrete nonlinear system operates as an state estimator for (1). The key tool to solve the design of the adaptive laws for the weights in the RNN was the Lyapunov stability theory for discrete-time systems. The RNN observer did not consider the application of classical Luenberger correction terms [42,3]. Instead, the RNN observer implements nonlinear gains based on the DSTA structure in both the observer structure and the corresponding learning laws derived from the Lyapunov analysis.

3 Design of the RNN state observer

3.1 System description and main assumptions

Notice that the system in (1) can be rewritten in the following form

$$\begin{bmatrix} x_\alpha(k+1) \\ x_\beta(k+1) \end{bmatrix} = \begin{bmatrix} x_\alpha(k) + \tau x_\beta(k) \\ x_\beta(k) + \tau f(x, u) + \tau \xi(k) \end{bmatrix}$$

$$y(k) = Cx(k) = x_\alpha(k) \tag{4}$$

In equation (4), the following definitions are used, $x := \begin{bmatrix} x_\alpha^\top & x_\beta^\top \end{bmatrix}^\top$, $x_\alpha :=$ $[x_1, \ldots, x_n]^\top$, $x_\beta := [x_{n+1}, \ldots, x_{2n}]^\top$ and $C = \begin{bmatrix} I_{n \times n}, & 0_{n \times n} \end{bmatrix}$. The matrix $I_{n \times n}$ represents the identity matrix of dimension $n \times n$, $0_{n \times n}$ represents a matrix with zero entries of dimension $n \times n$ and $f(x, u) := \begin{bmatrix} f_1(x, u), \cdots, f_n(x, u) \end{bmatrix}^\top$. In this study, the following assumptions are supposed to be fulfilled.

Assumption 1. The nonlinear functions $f_i(\cdot, \cdot)$ satisfy the Lipschitz condition with respect to their first argument, that is:

$$
\begin{aligned}
|f_i(x, u) - f_i(z, u)| \le L_i \|x - z\| \\
x, z \in \mathbb{R}^{2n}; \quad u \in U^{adm} \subset \mathbb{R}^n
\end{aligned}
\tag{5}
$$

Remark 1 Assumption 1 justifies the existence and the uniqueness for the solution of (1). This assumption is needed because RNN can approximate unknown/uncertain nonlinear functions with bounded trajectories and bounded variations. However, when the nonlinear functions in (1) have fast oscillations, the RNN cannot track its states.

Assumption 2. The term ξ has a deterministic nature and it is uniformly bounded as

$$
\begin{aligned}
\|\xi(k)\|_{\Lambda_\Upsilon}^2 \le \Upsilon, \quad \forall k \ge 0, \\
\Lambda_\Upsilon^\top = \Lambda_\Upsilon \in \mathbb{R}^{n \times n}, \quad \Lambda_\Upsilon > 0,
\end{aligned}
\tag{6}
$$

In the last equation, $\|\cdot\|_\Lambda$ represents the weighted norm, that is, for a vector $T \in \mathbb{R}^n$, its weighted norm is given by $\|T\|_\Lambda = T^\top \Lambda T$, where $0 < \Lambda = \Lambda^\top \in \mathbb{R}^{n \times n}$. The adaptive observer proposed in this paper obeys a RNN structure that approximates the nonlinear function $f(\cdot, \cdot)$, that is, there exists a nominal part $f_0(x(k), u(k) \mid \Omega)$ and a modeling error $\tilde{f}(x(k), u(k), \Omega)$ (based on the Stone-Weisstrass theorem [43]) where Ω is the set of parameters used to adjust the nominal part. This approximation model is given by

$$
f(x(k), u(k)) := f_0(x(k), u(k)|\Omega) + \tilde{f}(x(k), u(k), \Omega)
$$

The specific selection of Ω implies the validity of the following assumption:

Assumption 3. The error modeling $\tilde{f}$ satisfies:

$$
\left\| \tilde{f}(x, u, \Omega) \right\|_{\Lambda_{\tilde{f}}}^2 \le n_1, \; n_1 \in \mathbb{R}^+ \; \forall k \ge 0
\tag{7}
$$

with $\Lambda_{\tilde{f}} \in \mathbb{R}^{n \times n}$, $0 < \Lambda_{\tilde{f}} = \Lambda_{\tilde{f}}^\top$.

The nominal part $f_0 := f_0(x(k), u(k)|\Omega)$ satisfies a particular RNN structure

$$
f_0 := Ax(k) + W_1^* \sigma(x(k)) + W_2^* \varphi(x(k)) u(k)
\tag{8}
$$

where $A \in \mathbb{R}^{n \times n}$ is a Hurwitz matrix in the discrete-time sense and $W_1^* \in \mathbb{R}^{n \times 2n}$ and $W_2^* \in \mathbb{R}^{n \times 2n}$ are constants matrices. These parameters can be presented as $\Omega \in \mathbb{R}^{n \times 4n}$, which are defined as $\Omega = [W_1^*, W_2^*]$. The set of parameters satisfies, by assumption, the following condition $\Omega = argmin_{\Omega^0} \|f(x, u) - f_0(x, u, \Omega^0)\|$ for a given pair x and u.

Assumption 4. The weights W_1^* and W_2^* are assumed to be unknown but bounded by known positive constant matrices $\breve{W}_1$ and $\breve{W}_2$, that is,

$$
\begin{aligned}
(W_1^*)^\top M^\top \tilde{\Lambda}_{W_1^*} M W_1^* \leq \breve{W}_1 \\
(W_2^*)^\top M^\top \tilde{\Lambda}_{W_2^*} M W_2^* \leq \breve{W}_2
\end{aligned}
\tag{9}
$$

With $M = \begin{bmatrix} 0_{n\times n} & I_{n\times n} \end{bmatrix}^\top$, $\breve{W}_1, \breve{W}_2 > 0$ and $\tilde{\Lambda}_{W_1}, \tilde{\Lambda}_{W_2}$ are positive definite and symmetric matrices of appropriate dimensions.

In (8), $\sigma(\cdot) : \mathbb{R}^{2n} \to \mathbb{R}^n$ and $\varphi(\cdot) : \mathbb{R}^{2n} \to \mathbb{R}^{n\times m}$ are the so-called activation functions used in the design of RNN. In this study, sigmoid functions defined the set of activation functions. However, they may also be selected as simple polynomials of fixed order, Chebyshev polynomials or even Wavelets [44]. The class of sigmoid functions used in this study is:

$$
\sigma_i(x) = \frac{a_{\sigma_i}}{1 + b_{\sigma_i} e^{-c_{\sigma_i}^\top x}}, \qquad \varphi_{ij}(x) = \frac{a_{\varphi_{ij}}}{1 + b_{\varphi_{ij}} e^{-c_{\varphi_{ij}}^\top x}}
\tag{10}
$$

where $a_{\sigma_i}, a_{\varphi_{ij}}, b_{\sigma_i}, b_{\varphi_{ij}} \in \mathbb{R}$, $c_{\sigma_i}, c_{\varphi_{ij}} \in \mathbb{R}^{2n}$, $i = 1 : n, j = 1 : m$.

Assumption 5. The activation functions must satisfy the conditions of continuity and boundedness, that is,

$$
\begin{aligned}
\|\sigma(x) - \sigma(z)\|^2 \leq L_\sigma \|x - z\|^2 & \qquad \|\sigma(x)\|^2 \leq L_\sigma^+ \\
\|\varphi(x) - \varphi(z)\|^2 \leq L_\varphi \|x - z\|^2 & \qquad \|\varphi(x)\|^2 \leq L_\varphi^+
\end{aligned}
\tag{11}
$$

$\forall x, z \in \mathbb{R}^{2n}$ where L_σ, L_φ, L_σ^+, L_φ^+ are bounded positive scalars.

Clearly, by the characteristics of matrices W_i^* and the boundedness of the activation functions, the following inequalities are valid:

$$
\begin{aligned}
\|M W_i^* \Psi_i(k)\|_{\Lambda_{W_i}}^2 \leq \epsilon_i & \qquad \|W_i^* \Psi_i(k)\|_{\Lambda_{\bar{W}_i}}^2 \leq \bar{\epsilon}_i \\
\epsilon_i, \bar{\epsilon}_i \in \mathbb{R}^+ & \qquad i = 1, 2
\end{aligned}
\tag{12}
$$

3.2 Discrete-Time RNN Observer

The substitution of the RNN representation (8) in equation (4) yields:

$$
\begin{aligned}
x(k+1) &= \begin{bmatrix} x_\alpha(k) + \tau x_\beta(k) \\ x_\beta(k) + \tau\left(f_0 + \tilde{f}(x(k), u(k), \Omega) + \xi(k)\right) \end{bmatrix} \\
y(k) &= x_\alpha(k)
\end{aligned}
\tag{13}
$$

The adaptive observer proposed in this study introduces the output correction terms of the DSTA structure in the RNN design. The results in [45] provide a discrete-time characterization of the convergence zone for the DSTA.

Let us propose a discrete-time observer using the robust characteristics of the DSTA and the adaptive advantages obtained with the RNN approximation:

$$\hat{x}(k+1) = \begin{array}{c} \mu_1\hat{x}_\alpha(k) + \tau\hat{x}_\beta(k) + \tau\bar{B}_1(k) \\ \mu_2\hat{x}_\beta(k) + \tau\left(A\hat{x}_\beta(k) + \bar{W}_1(k)\hat{\sigma} + \bar{W}_2(k)\hat{\varphi}u(k) + \bar{B}_2(k)\right) \end{array} \quad (14)$$

with the following definitions

$$\bar{B}_1(k) := B_1\lambda\left(\Delta_\alpha(k)\right)S\left(\Delta_\alpha(k)\right)$$

$$\bar{B}_2(k) := B_2 S\left(\Delta_\alpha(k)\right)$$

$\bar{B}_1(k)$ and $\bar{B}_2(k)$ are the correction terms based on the DSTA. These terms offer extra robustness against external perturbations and modeling uncertainties [13,37]. The gain matrices B_1, B_2 and the function $\lambda\left(\Delta_\alpha(k)\right)$ are

$$B_1 := diag\left[B_{11}, B_{12}, \ldots, B_{1n}\right]$$
$$B_2 := diag\left[B_{21}, B_{22}, \ldots, B_{2n}\right]$$

$$\lambda\left(\Delta_\alpha(k)\right) := diag\left[\lambda_1\left(\Delta_1(k)\right), \ldots, \lambda_n\left(\Delta_n(k)\right)\right]$$
$$\lambda_i\left(\cdot\right) := \left|\cdot\right|^{1/2}, \quad i = 1:n$$
$$S\left(\Delta_\alpha(k)\right) := \left[sign\left(\Delta_1(k)\right), \ldots, sign\left(\Delta_n(k)\right)\right]^\top$$

The sign function $sign\left(\cdot\right)$ in discrete-time is

$$sign\left(a\right) := \begin{cases} -1 & if\ a < 0 \\ 0 & if\ a = 0 \\ 1 & if\ a > 0 \end{cases}$$

This representation for the sign function in the discrete-time domain avoids the discontinuity problem commonly exhibited in continuous sliding modes algorithms. The time varying parameters $\bar{W}_1$ and $\bar{W}_2$ are the weights of the observer to be updated with an appropriate learning law and they are defined as

$$2\bar{W}_1(k) := W_1(k+1) + W_1(k),$$
$$2\bar{W}_2(k) := W_2(k+1) + W_2(k) \quad (15)$$

$\hat{\sigma} := \sigma(\hat{x})$ and $\hat{\varphi} := \varphi(\hat{x})$ are the activation functions evaluated in the estimated states. Notice that equation (12) can be rewritten as

$$\hat{x}(k+1) = A_x\hat{x}(k) + B(k)S(\Delta_\alpha(k)) + M\bar{W}_1(k)\hat{\sigma} + M\bar{W}_2(k)\hat{\varphi}u(k) \quad (16)$$

where

$$A_x := \begin{array}{cc} \mu_1 I_{n\times n} & \tau I_{n\times n} \\ 0 & \mu_2 I + \tau A \end{array}, \quad B(k) := \begin{array}{c} \tau B_1\lambda\left(\Delta_\alpha(k)\right) \\ \tau B_2 \end{array}$$

The values of μ_1, μ_2 in equation (3.2) ensure the stability of matrix A_x placing their poles inside the unitary circle. The term $\Delta_\alpha := x_\alpha - \hat{x}_\alpha$ corresponds to the output error. The observer uses an on-line training to improve the current representation of (1) depending of the current values of Δ_α. The weight matrices are updated by nonlinear learning laws given by

$$W_j(k+1) = \Phi_j\left(W_j(k), \hat{x}(k), y(k), u(k)\right) \quad j = 1, 2. \quad (17)$$

Here, the weights W_j must be adjusted to reduce the approximation error between the nominal part and the uncertain nonlinear model.

3.3 Discrete-time adaptive learning law

The nonlinear updating laws for the observer in (16) are described by the following equations in differences

$$
\begin{aligned}
W_i(k+1) = \Pi_i^{-1} \big[- \big(\varepsilon_i - g_i^{-1} I_{n \times n} \big) W_i(k) \big] + \\
\tau \Pi_i^{-1} M^\top P \big[A_x N_\delta C^\top \Delta_\alpha(k) \big] \Psi_i^\top - \\
\tau \Pi_i^{-1} M^\top P \big[B(k) S(\Delta_\alpha(k)) \big] \Psi_i^\top
\end{aligned}
\tag{18}
$$

where $i = 1, 2$ and

$$
\begin{aligned}
\Pi_i &:= \big(\varepsilon_i + g_i^{-1} I_{n \times n} \big), \quad g_i > 0 \\
\varepsilon_i &:= \tau^2 \big(\Psi_i^\top(k) \Psi_i(k)/4 \big) M^\top \Lambda_{W_i} M \\
\Psi_1(k) &:= \sigma(\hat{x}(k)), \quad \Psi_2(k) := \varphi(\hat{x}(k)) u(k) \\
\Lambda_{W_i} &= \Lambda_{W_i}^\top > 0
\end{aligned}
$$

The matrix $N_\delta \in \mathbb{R}^{n \times n}$ is defined by $N_\delta := \big(CC^\top + \delta I_{n \times n} \big)^{-1}$ with δ being a small positive scalar. The parameters g_i are the learning coefficients for the RNN. The matrix P is the positive definite solution (if there exists) $0 < P = P^\top$, $P \in \mathbb{R}^{2n \times 2n}$ to the next matrix inequality

$$
A_x^\top P A_x + A_x^\top P R_a P A_x + R_b - (1-\alpha) P \leq Q_0
\tag{19}
$$

where

$$
R_a := \sum_{i=1}^{5} \Lambda_i
$$

$$
R_b = \delta N_{A_x} + C N_{A_x} C^\top + \tau^2 u_0^+ L_\varphi \breve{W}_2 + \tau^2 L_\sigma \breve{W}_1
$$

$$
N_{A_x} = N_\delta^\top A_x^\top A_x N_\delta
$$

The matrices $\Lambda_j = \Lambda_j^\top > 0$, $j = 1 : 5$ are positive definite matrices. The main result of the this study is stated in the following theorem.

3.4 Main Result

Theorem 1 *Assume that the assumptions 1-5 are valid for the class of systems represented by (1). Consider the observer (16), supplied by the learning laws (18), which was aimed to estimate the nonlinear trajectories of x. If there exists a positive definite matrix $Q_0 = Q_0^\top$, $Q_0 > 0$, such that the algebraic Riccati-like matrix inequality (19) is feasible for a positive definite and symmetric matrix P, and if the observer gains are selected such as all the components $B_{1,j}$ are positive and $\|B_2\| \geq \eta_1$, then, the state estimation error Δ is ultimately bounded with a bound given by*

$$
\beta = \frac{\rho}{1 - \alpha}
\tag{20}
$$

where $\rho := \frac{1}{2}\lambda_{\max}(\Lambda_B) \left\| Q_0^{-1/2} \right\|^2 + \Gamma,\ 0 < \Lambda_B = \Lambda_B^\top,\ \Lambda_B \in \mathbb{R}^{n \times n},\ \Gamma :=$ $\sum_{i=1}^2 (\epsilon_i + \bar{\epsilon}_i + \vartheta_i) + n_1 + \Upsilon.$ *The constants* n_1 *and* Υ *defined in (6) and (7).* $\epsilon_i, \bar{\epsilon}_i$ *are given in (12),* $\vartheta_1 = \lambda_{max}(Z_{22})\|B_2\|^2,\ \vartheta_2 = \lambda_{max}(\Lambda_{31})\|B_1^\top Z_{12}^\top B_2\|^2$ *with* Z_{22} *and* Λ_{31} *being positive definite matrices of appropriate dimensions.*

Proof The recurrent dynamics of the estimation error Δ satisfies the following form

$$\Delta(k+1) = A_x \Delta(k) - B(k) S(\Delta_\alpha(k)) + \tau M \xi(k) +$$
$$\tau M \sum_{i=1}^2 \left\{ W_i^* \tilde{\Psi}_i + \left[\tilde{W}_i(k+1) + \tilde{W}_i(k) \right] \Psi_i/2 \right\} + \tau M \tilde{f}(x(k), u(k)) \qquad (21)$$

where $\tilde{\Psi}_1(k) := \sigma(x(k)) - \sigma(\hat{x}(k))$ and $\tilde{\Psi}_2(k) := (\varphi(x(k)) - \varphi(\hat{x}(k)))u(k)$. Let define the following Lyapunov-like (energetic) function as

$$V(\Delta, W_1, W_2) = \Delta^\top P \Delta + g_1^{-1}\mathrm{tr}\left\{ W_1^\top W_1 \right\} + g_2^{-1}\mathrm{tr}\left\{ W_2^\top W_2 \right\} \qquad (22)$$

where $\mathrm{tr}(H)$ defines the trace operator of matrix H. Following the second stability method of Lyapunov for discrete systems, one must estimate the first difference of the Lyapunov function, that is,

$$V(k+1) - V(k) = \Delta^\top(k+1)P\Delta(k+1) - \Delta^\top(k)P\Delta(k) +$$
$$g_1^{-1}\mathrm{tr}\left\{ (W_1(k+1) - W_1(k))^\top (W_1(k+1) + W_1(k)) \right\} + \qquad (23)$$
$$g_2^{-1}\mathrm{tr}\left\{ (W_2(k+1) - W_2(k))^\top (W_2(k+1) + W_2(k)) \right\}$$

Substituting the expression (21) in the first term obtained in (23) and applying a number of times the Young's inequality $Y^\top X + X^\top Y \leq X^\top \Lambda X + Y^\top \Lambda^{-1} Y$ [46] valid for any $X, Y \in \mathbb{R}^{r \times s},\ 0 < \Lambda = \Lambda^\top \in \mathbb{R}^{r \times r}$, one may estimate the upper bound of $\Delta^\top(k+1)P\Delta(k+1)$ as

$$\Delta^\top(k+1)P\Delta(k+1) \leq \Delta^\top(k) \left(A_x P A_x + A_x^\top P R_\alpha P A_x \right) \Delta(k) +$$
$$\|B(k) S(\Delta_\alpha(k))\|_Z^2 + \sum_{i=1}^2 \tau^2 \left\| W_i^* \tilde{\Psi}_i(k) \right\|_{\Lambda_{W_i^*}}^2 +$$
$$\tau^2 \left\| \tilde{f}(x(k), u(k)) \right\|_{\Lambda_{\tilde{f}}} + \tau^2 \|\xi(k)\|_{\Lambda_{\tilde{\xi}}} +$$
$$2\tau \sum_{i=1}^2 S(\Delta_\alpha(k))^\top B(k)^\top P M^\top \left[\tilde{W}_i(k+1) - \tilde{W}_i(k) \right] \Psi(k)/2 - \qquad (24)$$
$$2 \sum_{i=1}^2 \Delta^\top(k) A_x^\top P M^\top \left[\tilde{W}_i(k+1) - \tilde{W}_i(k) \right] \Psi/2 +$$
$$\tau^2 \sum_{i=1}^2 \left\| \left[\tilde{W}_i(k+1) - \tilde{W}_i(k) \right] \Psi(k)/2 \right\|_{\Lambda_{\tilde{W}_1}}^2$$

In equation (24), $\tilde{W}_i := W_i - W_i^*$ and the matrix R_a corresponds to the definition introduced in (19). Besides, the following identities are valid

$$Z := \Lambda_1^{-1} + P + P\left(\Lambda_{6-9}\right)P$$

$$\Lambda_{W_1^*} := M^\top\left(\Lambda_2^{-1} + \Lambda_6^{-1} + \Lambda_{10-12} + P + P\left(\Lambda_{13-14}\right)P\right)M$$

$$\Lambda_{W_2^*} = M^\top\left(\Lambda_3^{-1} + \Lambda_7^{-1} + \Lambda_{15-16} + P + P\left(\Lambda_{10} + \Lambda_{17-18}\right)P\right)M$$

$$\Lambda_{\tilde{W}_1} := M^\top\left(P\left(\Lambda_{11}^{-1} + \Lambda_{15}^{-1} + \Lambda_{19-21}\right)P + P\right)M$$

$$\Lambda_{\tilde{W}_2} = P + P\left(\Lambda_{12}^{-1} + \Lambda_{16}^{-1} + \Lambda_{19}^{-1} + \Lambda_{22-23}\right)P$$

$$\Lambda_{\tilde{f}} := M^\top\left(\Lambda_4^{-1} + \Lambda_8^{-1} + \Lambda_{13}^{-1} + \Lambda_{17}^{-1} + \Lambda_{20}^{-1} + \Lambda_{22}^{-1} + P + P\Lambda_{24}P\right)M$$

$$\Lambda_{\tilde{\xi}} := M^\top P M + M^\top\left(\Lambda_5^{-1} + \Lambda_9^{-1} + \Lambda_{14}^{-1} + \Lambda_{18}^{-1} + \Lambda_{21-24}^{-1}\right)M$$

The matrices Λ_j are positive definite matrices ($j = \overline{1:24}$). By definition of $\tilde{\sigma}(x,\hat{x})$ and $\tilde{\varphi}(x,\hat{x})$ and the sector conditions requested to construct the activation functions described in (11), the following upper bounds are valid for $\tau^2 \left\|W_i^*\tilde{\Psi}(k)\right\|_{\Lambda_{W_i^*}}^2$

$$\tau^2\left\|W_1^*\tilde{\sigma}\left(x\left(k\right),\hat{x}\left(k\right)\right)\right\|_{\Lambda_{W_1^*}}^2 \leq \tau^2 L_\sigma\left\|W_1^*\Delta\left(k\right)\right\|_{\Lambda_{W_1^*}}^2$$

If we considered the following equations $\Lambda_{13} + \Lambda_{14} = P^{-2}$ and $\Lambda_4^{-1} + \Lambda_8^{-1} + \Lambda_{13}^{-1} + \Lambda_{17}^{-1} + \Lambda_{20}^{-1} + \Lambda_{21} + P = \tilde{\Lambda}_{W_1^*}$ with $\tilde{\Lambda}_{W_1^*} = \tilde{\Lambda}_{W_1^*}^\top > 0$ one gets

$$\tau^2 L_\sigma\left\|W_1^*\Delta\left(k\right)\right\|_{\Lambda_{W_1^*}}^2 \leq \tau^2 L_\sigma\Delta^\top\left(k\right)\left(W_1^*\right)^\top M^\top\tilde{\Lambda}_{W_1^*}MW_1^*\Delta\left(k\right)$$

Considering a similar analysis and defining $\Lambda_{10} + \Lambda_{17} + \Lambda_{18} = P^{-2}$ and $\Lambda_3^{-1} + \Lambda_7^{-1} + \Lambda_{15} + \Lambda_{16} + P = \tilde{\Lambda}_{W_2^*} = \tilde{\Lambda}_{W_2^*}^\top > 0$ the next inequality is obtained

$$\tau^2\left\|W_2^*\tilde{\varphi}\left(x\left(k\right),\hat{x}\left(k\right)\right)u(k)\right\|_{\Lambda_{W_2^*}} \leq \tau^2 L_\varphi u^+\Delta^\top\left(k\right)\left(W_2^*\right)^\top M^\top\tilde{\Lambda}_{W_2^*}MW_2^*\Delta\left(k\right)$$

Assumption 2 implies

$$\begin{aligned}
\tau^2 L_\sigma\left\|W_1^*\Delta\left(k\right)\right\|_{M^\top\tilde{\Lambda}_{W_1^*}M}^2 &\leq \tau^2 L_\sigma\Delta^\top\left(k\right)\breve{W}_1\Delta\left(k\right) \\
\tau^2 L_\varphi u^+\left\|W_2^*\Delta\left(k\right)\right\|_{M^\top\tilde{\Lambda}_{W_2^*}M}^2 &\leq \tau^2 L_\varphi u^+\Delta^\top\left(k\right)\breve{W}_2\Delta\left(k\right)
\end{aligned} \tag{25}$$

Let consider the term $S\left(\Delta_\alpha\left(k\right)\right)^\top B^\top\left(k\right)ZB\left(k\right)S\left(\Delta_\alpha\left(k\right)\right)$, which can be rewritten by a block decomposition for the matrix $Z := \begin{matrix} Z_{11} & Z_{12} \\ Z_{12} & Z_{22} \end{matrix}$ with $Z_{11} \in \mathbb{R}^{n\times n}$, $Z_{12} \in \mathbb{R}^{n\times n}$ and $Z_{22} \in \mathbb{R}^{n\times n}$. Based on this block decomposition, the following expression is valid

$$\begin{aligned}
S^\top\left(\Delta_\alpha(k)\right)B^\top\left(k\right)ZB\left(k\right)S\left(\Delta_\alpha(k)\right) &= \tau^2\left\|B_1\lambda\left(\Delta_a\right)S\left(\Delta_\alpha(k)\right)\right\|_{Z_{11}}^2 + \\
2\tau^2 S^\top\left(\Delta_\alpha\left(k\right)\right)B_2^\top Z_{12}B_1\lambda\left(\Delta_a\right)S\left(\Delta_\alpha(k)\right) &+ \tau^2\left\|B_2 S\left(\Delta_\alpha(k)\right)\right\|_{Z_{22}}^2
\end{aligned}$$

The Cauchy-Schwartz inequality implies that $\|B_1\lambda\left(\Delta_a\right)S\left(\Delta_\alpha(k)\right)\|_{Z_{11}}^2 \leq \|B_1\lambda\left(\Delta_a\right)\|_{Z_{11}}^2$. By the Young's inequality, one gets

$$
\begin{aligned}
S^\top\left(\Delta_\alpha\right)B_2^\top Z_{12}B_1\lambda\left(\Delta_a\right)S\left(\Delta_\alpha\right) \leq \\
\|Z_{12}B_1\lambda\left(\Delta_a\right)S\left(\Delta_\alpha\right)\|_{\Lambda_{31}^{-1}}^2 + \|B_2S\left(\Delta_\alpha\right)\|_{\Lambda_{31}}^2
\end{aligned}
\tag{26}
$$

The last terms of equation (26) are bounded because B_1 and B_2 are finite positive numbers to ensure the convergence of the RNN observer. To introduce the output error in the stability analysis, the following equation is applied

$$
\left(\delta I_{n\times n} + C^\top C\right)\Delta(k) = C^\top\Delta_\alpha\left(k\right) + \delta\Delta(k)
$$

Therefore, the next identity holds

$$
\Delta = N_\delta\left(C^\top\Delta_\alpha + \delta\Delta\right)
\tag{27}
$$

Applying again, the Young's inequality after the substitution of (27) and the definition of $\tilde{W}\left(k\right)$ in equation (24), with the results given in (25) and, if $\Lambda_{25} + \Lambda_{26}$ are selected as P^{-2}, the next equation is valid

$$
\begin{aligned}
\Delta^\top\left(k+1\right)P\Delta\left(k+1\right) \leq \Delta^\top\left(k\right)\left(A_xPA_x + A_x^\top PR_aPA_x + R_b\right)\Delta\left(k\right) - \\
2\sum_{i=1}^{2}\tau\Delta_\alpha^\top\left(k\right)CN_\delta^\top A_x^\top PM\left[W_i\left(k+1\right) + W_i\left(k\right)\right]\Psi_i(k)/2 + \\
2\tau\sum_{i=1}^{2}S\left(\Delta_\alpha\left(k\right)\right)^\top B\left(k\right)^\top PM\left[\tilde{W}_i\left(k+1\right) + \tilde{W}_i\left(k\right)\right]\Psi_i(k)/2 + \\
\tau^2\sum_{i=1}^{2}\|M\left[W_i\left(k+1\right) + W_i\left(k\right)\right]\Psi_i(k)/2\|_{\Lambda_{W_1}} + \\
\lambda_{\max}\left(\Lambda_B\right)\|\Delta_\alpha\left(k\right)\| + \left\|\tilde{f}\left(x(k),u(k)\right)\right\|_{\Lambda_{\tilde{f}}}^2 + \|\xi\left(k\right)\|_{\Lambda_{\tilde{\xi}}}^2 + \\
\tau^2\sum_{i=1}^{2}\|MW_i^*\Psi_i(k)\|_{\Lambda_{W_i}}^2 + \tau^2\sum_{i=1}^{2}\|W_i^*\Psi_i(k)\|_{\Lambda_{\tilde{W}_i}}^2 + \sum_{i=1}^{2}\vartheta_i
\end{aligned}
$$

Here $\Lambda_{\tilde{W}_1} := \Lambda_{27}^{-1} + \Lambda_{28}^{-1}$, $\Lambda_{\tilde{W}_2} := \Lambda_{29}^{-1} + \Lambda_{30}^{-1}$ and $i = 1, 2$. By the assumptions 3 and 4, the upper-bound of $\Delta^\top\left(k+1\right)P\Delta\left(k+1\right)$ in (23), adding and subtracting the term $\alpha V(k)$ and if there exists a solution for the discrete-time Riccati like equation given in (19), the next inequality is valid for $\tilde{V}(k)$ in

$$
\begin{aligned}
\tilde{V}\left(k\right) \leq -\alpha V(k) - \|\Delta_\alpha\left(k\right)\|_{Q_0}^2 + \lambda_{\max}\left(\Lambda_B\right)\|\Delta_\alpha\left(k\right)\| + \Gamma + \\
g_i^{-1}tr\left\{\left[W_i\left(k+1\right) + W_i\left(k\right)\right]^\top\left(\Theta_i\left(\Delta_\alpha\left(k\right), W_i\right)\right)\right\}
\end{aligned}
$$

Here Γ is defined in the main theorem and Θ_i as

$$\Theta_i \left(\Delta_\alpha \left(k \right), W_i \right) = -\tau M^\top P A_x N_\delta C^\top \Delta_\alpha \left(k \right) \Psi_i^\top \left(k \right) +$$

$$\tau M^\top P B \left(k \right) S \left(\Delta_\alpha \left(k \right) \right) \Psi_i^\top \left(k \right) + g_i^{-1} \left[W_i \left(k + 1 \right) - W_i \left(k \right) \right]$$

$$+ \tau^2 M^\top \Lambda_{W_i} M \left[W_i \left(k + 1 \right) + W_i \left(k \right) \right] \left(\Psi_i^\top \left(k \right) \Psi_i \left(k \right) / 4 \right)$$

With the application of the Choleskii decomposition [46] and the learning law defined in (17) the next inequality is obtained

$$\Delta V \left(k \right) \leq -\alpha V(k) - \left(Q_{\Delta_a} \right)^T \left(Q_{\Delta_a} \right) + \frac{1}{4} \lambda_{\max}^2 \left(\Lambda_B \right) \left\| Q_0^{-1/2} \right\|^2 + \Gamma$$

Where $Q_{\Delta_a} := \left\| Q_0^{1/2} \Delta_\alpha \left(k \right) - \frac{1}{2} \lambda_{\max} \left(\Lambda_B \right) Q_0^{-1/2} \right\|$. Finally, with the definition of ρ in Theorem 1,

$$V(k + 1) \leq \left(1 - \alpha \right) V(k) + \rho$$

The recursion process on this inequality leads to

$$V(k + 1) \leq \left(1 - \alpha \right)^k V(0) + \sum_{i=1}^{k} \left(1 - \alpha \right)^{i-1} \rho$$

If the limit when k goes to infinity is considered, one has

$$\overline{\lim_{k \to \infty}} V(k) \leq \frac{\rho}{1 - \alpha} \tag{28}$$

This last inequality concludes the proof.

Remark 2 The solution of the matrix inequality (MI) described in equation (19) seems to be a restrictive condition. However, this MI can be transformed into two Linear MI's (LMI). The MI in (19) is rewritten as

$$A_x^\top P A_x + A_x^\top P R_a P A_x + R_b - \left(1 - \alpha \right) P \leq Q_0 \tag{29}$$

With R_a defined in (19). If the next inequality is fulfilled,

$$P + P R_a P \leq G \tag{30}$$

which is equivalent (by Shur complement [46]) to

$$\begin{matrix} G - P & P \\ P & R_a^{-1} \end{matrix} \geq 0 \tag{31}$$

The MI in (29) can be rewritten as

$$A_x^\top G A_x - (1 - \alpha) P + R_b \leq Q_0 \tag{32}$$

Then, the solution of (19) is relaxed to the solution of LMI's (31) and (32).

Remark 3 The regular DSTA applied as robust discrete approximation of derivative function only requires that $\|f(\cdot,\cdot,\cdot))\|^2 \leq f^+$. In consequence, the gains of the STA can be adjusted using only this value. The actual value of the chattering, obtained when the DSTA is realized in practice, depends on this gain. The application of DNN introduces a formal method to reduce the DSTA values and in consequence, the chattering values are also smaller.

Let consider a RNN structure with a given number of components in the weights W_1 and W_2 such that:

$$\|\tilde{f}(x,u,\Omega)\| + \|f_o(x,u|\Omega)\| \leq \|F(x,u,\xi)\| \tag{33}$$

Notice that this inequality also includes the modeling strategy based on RNN, that is $\|F(x,u,\xi)\| \geq \|\tilde{f}(x,u,\Omega) + f_o(x,u|\Omega)\|$. Assume that we selected the number of components in the weights W_1 and W_2 in such a way that

$$\|f_o(x,u|\Omega)\|^2 \leq \|Ax(k)\|^2 + \|W_1^* \sigma(x(k))\|^2 + \|W_1^* \varphi(x(k))u(k)\|^2 \leq \eta_2$$

where $\eta_2 := \lambda_{\min}\{A\}\varepsilon^+ + (W_1)^+ \sigma^+ + (W_2)^+ \varphi^+ u_0^+$. The Stone-Weierstrass theorem [43] justifies the value of η_2 which satisfies $\eta_2 \leq f^+$,

$$\|\tilde{f}(x,u,\Omega)\| + \eta_2 \leq \|F(x,u,\xi)\| \leq f^+ \tag{34}$$

The upper bound f^+ is given in (20). We can conclude that

$$\|\tilde{f}(x,u,\Omega)\| \leq f^+ - \eta_2 \tag{35}$$

.

Therefore according to Theorem 1, $\|B_1\| \geq 0$ and $\|B_2\| \geq (f^+ - \eta_2)$ when the RNN is included in the model. From classical results applying the DSTA as a robust exact differentiator [13],[45], the gains needed to ensure the convergence of the DSTA are smaller than the case when the RNN is not in the modeling process. The results in Theorem 1 of this manuscript and work reported in [45] show that the boundary layer also depends on the gain selection. As a consequence the boundary layer where the estimation error converge with the scheme DSTA-RNN is smaller than the use of the DSTA working as a RED.

4 Numerical Results

4.1 Simple pendulum

As an illustration of the results presented in this paper, an Euler (explicit) discretization of the generalized STA (GSTA) presented in [41] and a DSTA reinforced with the RNN scheme are compared. Let consider a simple pendulum described by the following equations in differences

$$
\begin{aligned}
x_1(k+1) &= x_1(k) + \tau x_2(k) \\
x_2(k+1) &= x_2(k) + \tau \frac{1}{J} u_k - \tau \frac{mgl}{2J}\sin(x_1(k)) - \tau \frac{V_s}{J} x_2(k) + \tau \psi(k) \\
y(k) &= x_1(k)
\end{aligned}
\tag{36}
$$

where $x_1 = \theta$ is the angle of oscillation, x_2 is the angular velocity, m is the pendulum mass, g is the gravitational force, l is the pendulum length, $J = ml^2$ is the arm of inertia, V_s is the pendulum's viscous friction coefficient. For the simulation the bounded perturbation is expressed as $\psi_1(k) = 0.5\sin(2k) + 0.5\cos(5k)$. Here, the upper-bound for the control signal is $u^+ = 1$. For simulation, the initial conditions were $x_{1,0} = -1$ and $x_{2,0} = 3$ for the model and $\hat{x}_{1,0} = 0$, and $\hat{x}_{2,0} = 0$ for the observer. The following numeric values were applied to simulate the pendulum model: $m = 1.1kg$, $l = 1m$, $g = 9.81\left(\frac{m}{s^2}\right)$ and $V_s = 0.18\frac{kg \cdot m}{s^2}$. The input applied into the system was $u = \sin(2k)\cos(5k)$. The observer parameters were chosen as $B_1 = 5$ and $B_2 = 2$ and $A = -0.012$ The activation functions were

$$a_\sigma = a_\phi = \begin{bmatrix} 0.25 \\ 0.25 \\ 0.25 \end{bmatrix}, \qquad b_\sigma = b_\phi = \begin{bmatrix} 0.05 \\ 0.05 \\ 0.05 \end{bmatrix}, \qquad c_\sigma = c_\phi = \begin{matrix} 0.05 \\ 0.05 \end{matrix}$$

Selecting $\mu_1 = \mu_2 = 0.95$ the value of $\delta = 0.001$, $R_b = \begin{matrix} 2.0005 & 2.0005 \\ 2.0005 & 2.0005 \end{matrix}$ and $Q_0 = 10 * I_{2 \times 2}$ the positive definite solutions P and G in Remark 1 were

$$P = \begin{matrix} 1.3613 & 2.15 \times 10^{-5} \\ 2.15 \times 10^{-5} & 1.3613 \end{matrix} \qquad G = \begin{matrix} 7.825 & 0.011 \\ 0.011 & 7.826 \end{matrix}$$

With these values the boundary layer defined in equation (20) is defined as

$$\beta = \frac{10}{1 - 0.019} = 10.1937$$

The discrete structure for the GSTA working as a robust exact differentiator is

$$\hat{x}_1(k+1) = \hat{x}_1(k) + \tau\hat{x}_2(k) - \tau B1|\Delta_1(k)|^{\frac{1}{2}}sign(\Delta_1(k)) - \tilde{B}_1\Delta_1(k)$$
$$\hat{x}_2(k+1) = \hat{x}_2(k) - \tau B2 sign(\Delta_1(k)) + \tilde{B}_2\Delta_1(k)$$

Notice that the GSTA has a linear gain that improves the performance of the estimation process when the estimation error has a big amplitude, while the sign function is more effective when the estimation error has a small amplitude. The gains of the GSTA were chosen similar to the gains used in the DSTA-RNN. The linear gains were $\tilde{B}_1 = 2$ and $\tilde{B}_2 = 1.5$. Figure 1 shows the results of the state estimation process. The first graphic represents the estimation of the measurable state, i.e., the position of the pendulum. Both observers (GSTA and the DSTA-RNN) reproduced the nonlinear trajectories of the system. The DSTA-RNN presented a delay before reproducing the position of the pendulum, which is a direct consequence of the learning process of the RNN structure. The second graph of this Figure shows a better performance in the state estimation process when the RNN is introduced in the observer structure. The GSTA observer did not converge to the real states with small gains. This problem was corrected by the adaptive contribution of the RNN. Notice that the DSTA-RNN observer was designed assuming that the function $f(\cdot, \cdot)$

was unknown. The system given in (36) was used only as a data generator to train the RNN section. If the gains corresponding to the GSTA observer are increased to fulfill the condition $B_2 > f^+$, the GSTA can reproduce the nonlinear dynamics of the simple pendulum. In this example, the condition was not accomplished only to show how the RNN structure improved the performance of the DSTA with smaller gains B_1 and B_2. The increment of B_2 in order to fulfill the condition $B_2 > f^+$ in the GSTA implied a considerable increment of the chattering amplitude in the estimation process. With the gains chosen as 10 and 7, and $\tilde{B}_1 = 2$ and $\tilde{B}_2 = 1.5$ the GSTA observer presented and adequate performance. The performance of DSTA-RNN and the GSTA observers were compared throughout the euclidean norm of the estimation error. Figure 2 presents this comparison. The DSTA-RNN had a
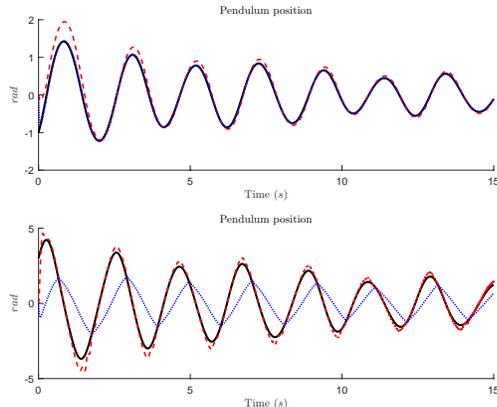
Fig. 1 Simple pendulum state estimation process a) Position state estimation process b) Velocity state estimation process

faster convergence and its steady state response remains below the euclidean norm of the estimation error obtained with the GSTA. The obtained value of β coincided with the simulation results.

4.2 Flexible Link Robot Manipulator

Let consider a 4-dimensional nonlinear system that describes the dynamics of a flexible link robot manipulator. The nonlinear dynamic equations for a single-link manipulator with flexible joint and negligible damping are given by

$$I\ddot{q}_1 + MgL\sin(q_1) + \kappa(q_1 - q_2) = 0$$
$$J\ddot{q}_2 - \kappa(q_1 - q_2) = u$$

where q_1 and q_2 are the angular positions, I and J are the moments of inertia, κ is the spring constant, M is the total mass, L is the length of the link in
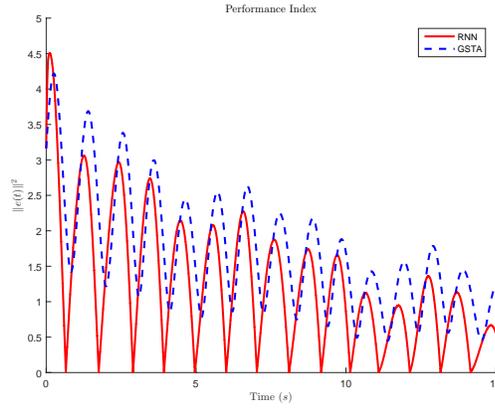
Fig. 2 Performance index to compare the simulation results between the DSTA and the DSTA-RNN observer

the robot manipulator and u is the corresponding input. By the state variable approach, defining $q_1 = x_1$ and $q_2 = x_3$, the last equation can be rewritten as

$$\dot{x} = f(x) + g(x)u$$

where $x = \begin{bmatrix} x_1 & x_2 & x_3 & x_4 \end{bmatrix}^\top$ and

$$f(x) = \begin{bmatrix} x_2 \\ -a\sin(x_1) - b(x_1 - x_3) \\ x_3 \\ -c(x_1 - x_3) \end{bmatrix} \quad g(x) = \begin{bmatrix} 0 \\ 0 \\ 0 \\ d \end{bmatrix} \tag{37}$$

where $a = \frac{MgL}{I}$, $b = \frac{\kappa}{I}$, $c = \frac{\kappa}{J}$ and $d = \frac{1}{J}$. For simulation $a = 2$, $b = 2$, $c = 1$ and $d = 1$. The input signal is $u(k) = -0.1\sin(2\tau k)$. The parameters of the DSTA-RNN used in simulation were

$$B_1 = \begin{matrix} 2 & 0 \\ 0 & 2 \end{matrix}, \qquad B_2 = \begin{matrix} 5 & 0 \\ 0 & 2.1 \end{matrix}, \qquad A = \begin{matrix} -2\tau & -1.1\tau \\ -1.5\tau & -3.1\tau \end{matrix}$$

The activation functions were chosen as

$$a_\sigma = a_\phi = \begin{bmatrix} 0.025 \\ 0.025 \\ 0.025 \end{bmatrix}, \qquad b_\sigma = b_\phi = \begin{bmatrix} 0.05 \\ 0.05 \\ 0.05 \end{bmatrix}, \qquad c_\sigma = c_\phi = \begin{matrix} 0.05 \\ 0.05 \end{matrix},$$

And the linear gains for the GSTA were

$$\tilde{B}_1 = \begin{matrix} 2 & 0 \\ 0 & 1 \end{matrix}, \qquad \tilde{B}_2 = \begin{matrix} 1.5 & 0 \\ 0 & 1 \end{matrix}$$

The sampling period was established as 0.001. The parameters of inequalities in Remark 1 were chosen as $R_a = I_{4 \times 4}$ and R_b as

$$R_b = \begin{bmatrix} 1.0931 & 0 & 1.0460 \times 10^{-4} & 0 \\ 1.045 \times 10^{-7} & 09.094 \times 10^{-6} & 0 & \\ 0 & 1.045 \times 10^{-7} & 0 & 9.0940 \times 10^{-6} \\ 0 & 1.045 \times 10^{-7} & 0 & 9.0940 \times 10^{-6} \end{bmatrix},$$

where $\alpha = 0.006$, $\delta = 0.1$ and $Q = 10 \times I_{4 \times 4}$. The solution of the LMI's yielded to the following results

$$P = \begin{bmatrix} 0.2501 & 0.0478 & 0.0705 & 0.0691 \\ 0.0477 & 0.2501 & 0.0691 & 0.0705 \\ 0.0705 & 0.0691 & 0.2498 & 0.0491 \\ 0.0691 & 0.0705 & 0.0491 & 0.2499 \end{bmatrix},$$

$$G = \begin{bmatrix} 1.0697 & 0.3639 & 0.5177 & 0.5086 \\ 0.3639 & 1.0697 & 0.5086 & 0.5177 \\ 0.5177 & 0.5086 & 1.0685 & 0.3747 \\ 0.5086 & 0.5177 & 0.3747 & 1.0685 \end{bmatrix} \times 10^7$$

With these solutions, the upper bound for the convergence zone was $\beta = 20.1207$. Figure 3 presents the estimation results for the positions and velocities of each link of the flexible robot. Similar to the previous example, a comparison between the DSTA-RNN and the GSTA observer was proposed. In Figure 3, the solid line represents the real dynamics obtained with the nonlinear model (37), the dashed line depicts the trajectories provided by the DSTA-RNN and the dotted line represents the results obtained with the GSTA observer. The DSTA-RNN reached the real trajectories before the GSTA observer. GSTA observer exhibited an overshoot before it reached the real trajectories of the flexible robot manipulator. This disadvantage decreased while using the DSTA-RNN observer. To explore with detail the convergence of both observers, figure 4 presents a closer view of the first two seconds of simulation. Figure 5 illustrates the performance index selected as the euclidean norm of the estimation error. In this figure, the value of the estimation error trajectories remained inside the boundary layer delimited by the value of β. The DSTA-RNN scheme converged faster than the GSTA and with less overshoot.

5 Experimental results

Let us consider the Van Der Pol Oscillator given by the following set of differential equations

$$\dot{x}_1 = x_2,$$
$$\dot{x}_2 = -x_1 + \kappa \left(1 - x_1^2\right) + u, \tag{38}$$
$$y = x_1$$
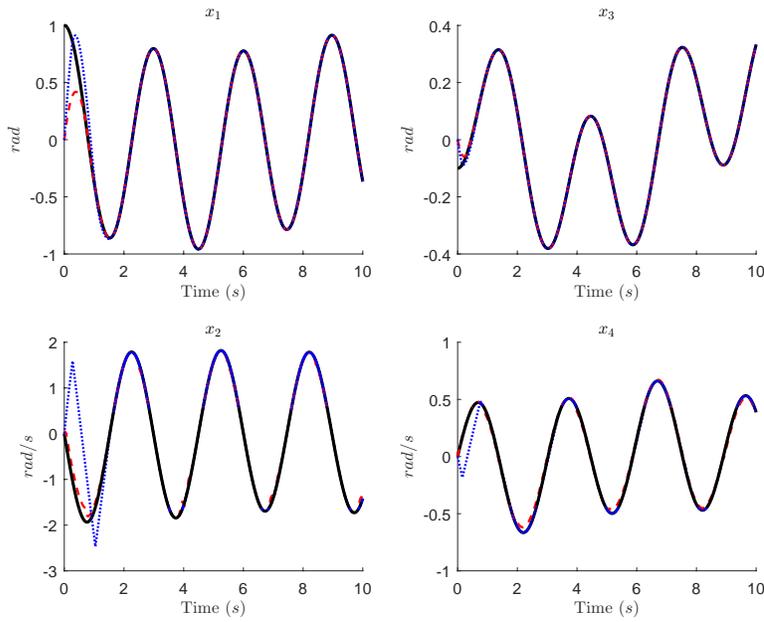
Fig. 3 Comparisson between the DSTA-RNN and the GSTA. The solid lines represent the real dynamics. The dashed lines correspond to the trajectories obtained with the DSTA-RNN observer and the dotte lines are the estimated travectories with the GSTA
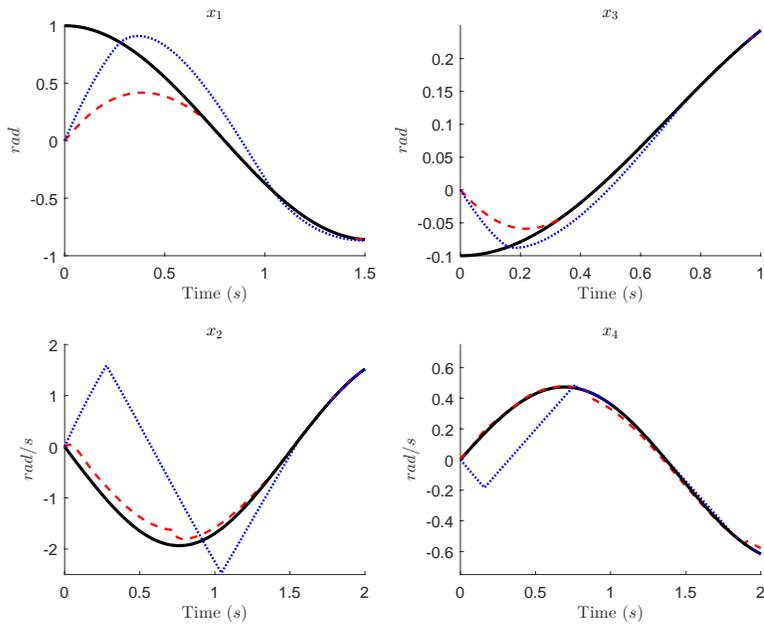
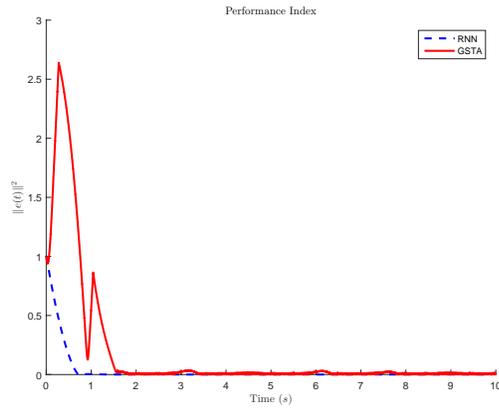**Fig. 4** Closer view of the estimated trayectories obtained by the DSTA-STA and the GSTA

Fig. 5 Performance index comparison between the GSTA and the DSTA-RNN observers

where κ is a model parameter equal to 0.1. Figure 6 shows the electronic circuit and the computer interface. The circuit parameters can be found in [47]. The computer interface employed a dSPACE 1104 board. The Van Der Pol oscillator was internally controlled by a Continuous Singular Terminal Sliding-Mode (CSTSM) controller given by

$$
\begin{aligned}
\varphi &= x_2 - \dot{y}_d + k_2 \lceil x_1 - y_d \rfloor^{\frac{2}{3}} \\
u &= \ddot{x}_d + x_1 - \kappa x_2 \left(1 - x_1^2\right) - k_1 \lceil \varphi \rfloor^{\frac{1}{2}} + z \\
\dot{z} &= -k_3 \lceil \varphi \rfloor^0
\end{aligned}
\tag{39}
$$

where y_d is the desired trajectory to follow, z is an extended variable, $k_1 = 4, k_2 = 3$ and $k_3 = 2$ were the design parameters. The dSPACE board ran at $5KH$ and the controller was implemented with an Euler discretization method. Notice that, the discrete measurements of the available output and control input were the ones used to test the DSTA-RNN and the GSTA.
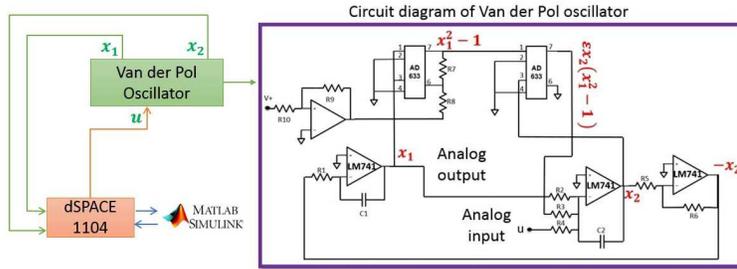

Fig. 6 Schematic overview of the practical implementation and circuit diagram of an autonomous Van Der Pol oscillator.

The values used in the activation functions in the DNN observer were

$$a_{\sigma_1} = 40, \quad b_{\sigma_1} = 1, \quad c_{\sigma_1} = \begin{matrix} 0.5 \\ 0.5 \end{matrix} \ ,$$

$$a_{\sigma_2} = 12, \quad b_{\sigma_2} = 1, \quad c_{\sigma_2} = \begin{matrix} 25 \\ 25 \end{matrix} \ ,$$

$$a_{\phi_1} = 10, \quad b_{\phi_1} = 1, \quad c_{\phi_1} = \begin{matrix} 0.2 \\ 0.7 \end{matrix} \ ,$$

$$a_{\phi_2} = 15, \quad b_{\phi_2} = 1, \quad c_{\phi_2} = \begin{matrix} 12.3 \\ 12.3 \end{matrix}$$

The matrices A and P were selected as

$$A = \begin{matrix} 0 & 1 \\ -0.01 & -1000 \end{matrix} \ , \qquad P = \begin{matrix} 60 & 40 \\ 40 & 106 \end{matrix}$$

The learning coefficients are chosen as $g_1 = g_2 = 0.8$ and $\alpha = 0.01$, $\mu = 0.5$ $\tilde{\Lambda}_{\Psi_1} = Q_{W_i} = I_{2\times2}$ and $B_1 = 5$ and $B_2 = 10$. The linear gains in the GSTA were $\tilde{B}_1 = -0.5$ and $\tilde{B}_2 = -1$. Figure 7 shows the comparison between the states estimated by the DSTA-RNN and the GSTA. The continuous red line is the data measurement of the Van Der Pol Oscillator. The dashed blue lines are the estimation provided by the DSTA-RNN and the dotted black line is the estimation obtained by the GSTA. Figure 7 presented a closer view into the first 0.2 seconds of experimentation. One can appreciate the learning period where the RNN-DSTA had an oscillatory behavior. After the DSTA-RNN reached the steady state, the oscillations disappeared. The main problem with the GSTA is usually the chattering effect in the second state. In the state estimation of x_2, the GSTA presented a bigger overshot than the DSTA-RNN. Moreover, in steady state, the GSTA presented high frequency oscillations. Figure 8 depicts the Euclidean norm of the estimation error. The black line is the result of choosing $B_1 = 0$ and $B_2 = 0$. Therefore, just the linear gains were activated and the GSTA worked just as a classical Luenberger observer. Without the non linear gains B_1 and B_2, the algorithm did not reach the real states. The GSTA presented bigger oscillations and it is represented by a red line. The experimental setup brings a problem when the available measurement had some small but high frequency variations. The GSTA produced more oscillations because of the noisy available output. The DSTA-RNN reduced the oscillations and provided a faster convergence (blue continuous line).

6 Discussion

The simulation results in this study showed some advantages of the DSTA-RNN observer over classical results vailable in literature such as the ones described in [41,45]. This study proposes a discrete-time analysis to guarantee the convergence of the estimation error to a neighborhood around the origin. The main advantages of the observer desing presented in this manuscript are:
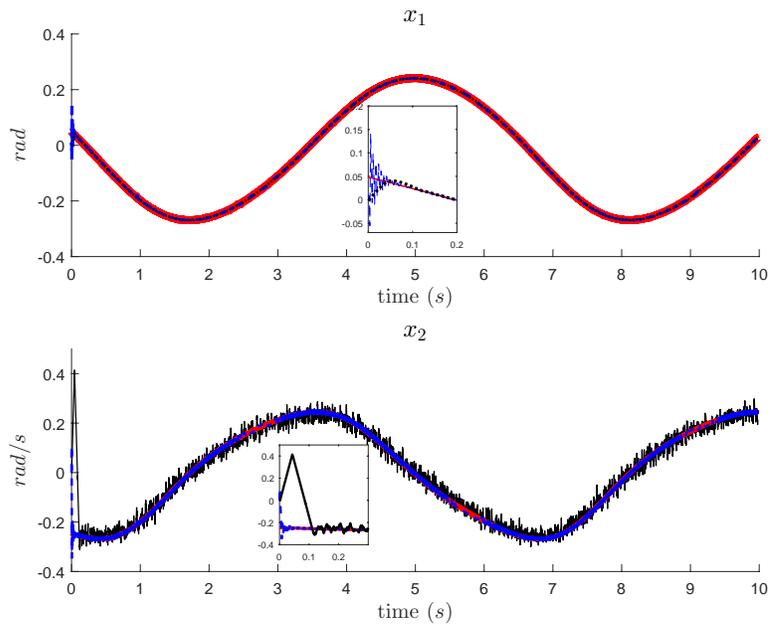
Fig. 7 Estimation procedure obtained with the DSTA-RNN and the GSTA implemented with real data
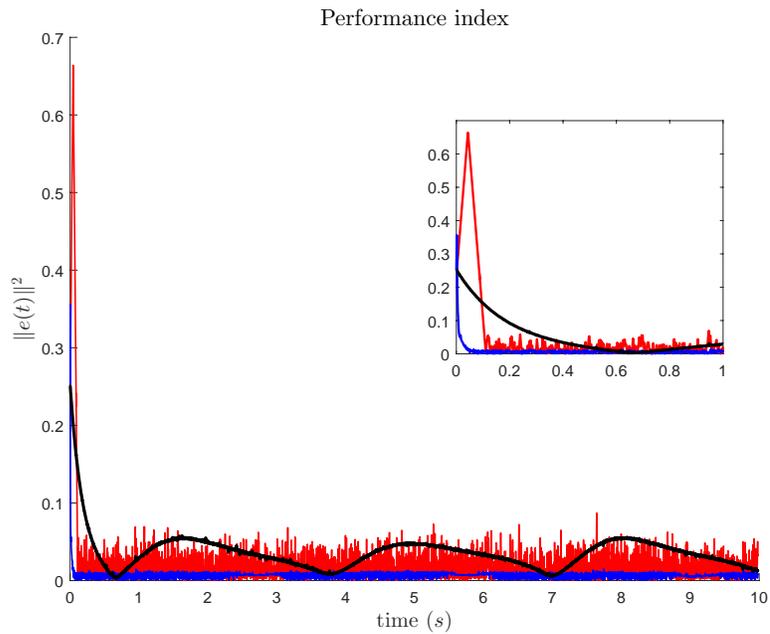
Fig. 8 Euclidean norm of the estimation error.

a) the Lyapunov analysis showed that the DSTA-RNN observer can be applied in real applications because it can be implemented in embeded systems through equations in differences. This is one of the main advantages obtained with a discrete scheme over continuous schemes that require some numerical integration methods, b) without the mathematical description of the nonlinear MIMO system, the DSTA-RNN seemed to be a better option to estimate the unknown dynamics of MIMO systems represented by a second order general form. Indeed, the differentiators in [13, 41] and [45] have higher frequency oscillations in the second state. The results presented in this manuscript reinforce the state estimation process of second-order systems. Moreover, the gains needed to enforce the convergence are smaller than those used in the classical GSTA observer, c) in a possible output feedback control design, the DSTA-RNN produces smaller overshoot.

Some possible disadvantages are a) in terms of complexity, the number of operations needed in the RNN observer is significantly increased. At least, the RNN introduces two new equations in differences for W_1 and W_2. However, the estimation process is better with the adaptive adjustment. It is clear that there exists a compromise between the reduction of conservatism introduced with the kind of systems that the STA can deal with and the increase of the computational complexity introducing an adaptive observer but with the possibility of manage a more general class of MIMO second-order systems, b) the complexity of tuning the gains for each observer is quite different. The GSTA requires to choose only four principal parameters (see the works in [13, 45, 41]), the DSTA-RNN observer introduces an additional MI to obtain the value of matrix P in the learning laws. The complexity of an adaptive scheme is always bigger, but in terms of estimation quality, it presents more advantages, c) the selection of the corresponding values of the activation functions which still remains as an open problem because their estimation is made by try to test process.

7 Conclusions

In this paper, a second-order algorithm based on the DSTA and RNN solved the problem of state estimation of nonlinear MIMO systems. The Lyapunov stability was used as the main tool to derive the learning laws and the stability proofs of the proposed state estimator. With this technique, the dynamics of the DSTA are incorporated into the learning laws of the RNN. This learning procedure provided robustness against perturbations. A possible drawback with the present approach is regarding the computational complexity in the simulation and the implementation. However, in the discrete-time domain, the implementation of the RNN just includes two additional differences equations to upload the weights W_1 and W_2. Another difference with the continuous case is the practical stability achieved in this study against the finite-time stability offered by the GSTA in continuous time. Unfortunately, finite-time convergence requires theoretically *infinite switching frequency*, that cannot be

obtained in real implementations. Numerical results characterized the performance of the observer designed in this manuscript. Further research must include the design of a controller based on the estimated states provided by the DSTA-RNN.

References

1. A. Oveisi, M. B. Jeronimo, and T. Nestorovic, "Nonlinear observer-based recurrent wavelet neuro-controller in disturbance rejection control of flexible structures," *Engineering Applications of Artificial Intelligence*, vol. 69, pp. 50 – 64, 2018. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0952197617303044
2. X. Liu, C. Yang, Z. Chen, M. Wang, and C.-Y. Su, "Neuro-adaptive observer based control of flexible joint robot," *Neurocomputing*, vol. 275, pp. 73 – 82, 2018. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0925231217308184
3. H. Khalil, *Nonlinear Systems*. Prentice Hall U.S., 2002.
4. R. Ortega, L. Praly, S. Aranovskiy, B. Yi, and W. Zhang, "On dynamic regressor extension and mixing parameter estimators: Two luenberger observers interpretations," *Automatica*, 2018. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0005109818303030
5. C. Faria, "Sensor fusion and rotational motion reconstruction via nonlinear state-observers," *Mechanical Systems and Signal Processing*, vol. 114, pp. 571 – 578, 2019. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0888327018302747
6. S. F. Ardabili, B. Najafi, S. Shamshirband, B. M. Bidgoli, R. C. Deo, and K. wing Chau, "Computational intelligence approach for modeling hydrogen production: a review," *Engineering Applications of Computational Fluid Mechanics*, vol. 12, no. 1, pp. 438–458, 2018.
7. S. Spurgeon, "Sliding mode observers: a survey," *International Journal of Systems Science*, vol. 39, no. 8, pp. 751–764, 2008.
8. G. Palli, S. Strano, and M. Terzo, "Sliding-mode observers for state and disturbance estimation in electro-hydraulic systems," *Control Engineering Practice*, vol. 74, pp. 58 – 70, 2018. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0967066118300236
9. I. Boiko, L. Fridman, A. Pisano, and E. Usai, "Analysis of chattering in systems with second-order sliding modes," *IEEE Transactions on Automatic Control*, vol. 52, no. 11, pp. 2085–2102, 2007.
10. A. Levant, "Sliding order and sliding accuracy in sliding mode control," *International Journal of Control*, vol. 58, no. 6, pp. 1247–1263, 1993.
11. ——, "Principles of 2-sliding mode design," *Automatica*, vol. 43, no. 4, pp. 576–586, 2007.
12. ——, "Robust exact differentiation via sliding mode tecnique," *Automatica*, vol. 34, no. 3, pp. 379–384, 1998.
13. J. Davila, L. Fridman, and A. Levant, "Second-order sliding-mode observer for mechanical systems," *IEEE Transactions on Automatic Control*, vol. 50, no. 11, pp. 1785–1789, 2005.
14. J. Davila, L. Fridman, and A. Poznyak, "Observation and identification of mechanical systems via second-order sliding modes," *International Journal of Control*, vol. 79, no. 10, pp. 1251–1262, 2006.
15. T. Gonzlez, J. Moreno, and L. Fridman, "Variable gain super-twisting sliding mode control," *IEEE Tran. Aut. Cont.*, vol. 57, no. 8, pp. 2100–2105, 2012.

16. C.-S. Tseng and B.-S. Chen, "Fuzzy control for uncertain vehicle active suspension systems via dynamic sliding-mode approach," *IEEE Transactions on Systems, Man and Cybernetics: Systems*, vol. 47, no. 1, pp. 24–32, 2017.

17. A. Levant, "Finite differences in homogeneous discontinuous control," *IEEE Transactions on Automatic Control*, vol. 52, no. 7, pp. 1208–1217, 2007.

18. I. Salgado, S. Kamal, B. Bandyopadhyay, I. Chairez, and L. Fridman, "Control of discrete time systems based on recurrent super-twisting-like algorithm," *ISA Transactions*, vol. 64, pp. 47 – 55, 2016. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0019057816300702

19. Y. Yan, Z. Galias, X. Yu, and C. Sun, "Eulers discretization effect on a twisting algorithm based sliding mode control," *Automatica*, vol. 68, pp. 203 – 208, 2016. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0005109816000522

20. S. Z. Sarpturk, Y. Istefanopulos, and O. Kaynak, "On the stability of discrete-time sliding mode control systems," *IEEE Transactions on Automatic Control*, vol. 32, no. 10, pp. 930–932, 1987.

21. S. Govindaswamy, T. Floquet, and S. Spurgeon, "On output sampling based sliding mode control for discrete time," in *Proceedings of the 47th IEEE Conference on Decision and Control*, 2008, pp. 2190–2195.

22. H. Sira-Ramirez, "Nonlinear discrete variable structure systems in quasi-sliding mode," *International Journal of Control*, vol. 54, no. 5, pp. 1171–1187, 1991.

23. P. L. Priya and B. Bandyopadhyay, "Discrete time sliding mode control for uncertain delta operator systems with infrequent output measurements," *European Journal of Control*, vol. 33, pp. 52 – 59, 2017. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0947358016301017

24. H. Liu, Y. Pan, S. Li, and Y. Chen, "Synchronization for fractional-order neural networks with full/under-actuation using fractional-order sliding mode control," *International Journal of Machine Learning and Cybernetics*, vol. 9, no. 7, pp. 1219–1232, Jul 2018. [Online]. Available: https://doi.org/10.1007/s13042-017-0646-z

25. S. Sastry and M. Bodson, *Adaptive Contol. Stability, Convergence and Robustness*, T. Kailath, Ed. Prentice Hall, 1989.

26. J. Resendiz, W. Yu, and L. Fridman, "Two-stage neural observer for mechanical systems," *IEEE Transactions on Circuits and Systems*, vol. 55, no. 10, pp. 1076–1080, 2008.

27. V. Gholami, K. Chau, F. Fadaee, J. Torkaman, and A. Ghaffari, "Modeling of groundwater level fluctuations using dendrochronology in alluvial aquifers," *Journal of Hydrology*, vol. 529, pp. 1060 – 1069, 2015.

28. S. M. R. Kazemi, B. M. Bidgoli, S. Shamshirband, S. M. Karimi, M. A. Ghorbani, K. wing Chau, and R. K. Pour, "Novel genetic-based negative correlation learning for estimating soil temperature," *Engineering Applications of Computational Fluid Mechanics*, vol. 12, no. 1, pp. 506–516, 2018.

29. L. Chen, M. Liu, X. Huang, S. Fu, and J. Qiu, "Adaptive fuzzy sliding mode control for network-based nonlinear systems with actuator failures," *IEEE Transactions on Fuzzy Systems*, vol. 26, no. 3, pp. 1311–1323, June 2018.

30. R. Taormina, K.-W. Chau, and B. Sivakumar, "Neural network river forecasting through baseflow separation and binary-coded swarm optimization," *Journal of Hydrology*, vol. 529, pp. 1788 – 1797, 2015. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0022169415005673

31. C.-S. Tseng and B.-S. Chen, "Multiobjective pid control design in uncertain robotic systems using neural network elimination scheme," *IEEE Transactions on Systems, Man and Cybernetics: Systems*, vol. 31, no. 6, pp. 632–644, 2001.

32. A. Poznyak, E. Sánchez, and Y. Wen, *Differential Neural Networks for Robust Nonlinear Control (Identification, State Estimation and Trajectory Tracking)*. World Scientific Press, 2001.

33. S. Haykin, *Neural Networks, Comprehensive Foundation*. Prentice Hall U.S., 1999.

34. N. Sanchez, Y. Alanis, and A. Loukianov, *Discrete-Time, High Order Neural Control:Trained with Kalman Filtering*. Springer Verlag, London, 2008.

35. I. Salgado and I. Chairez, "Discrete time recurrent neural network observer," in *2009 International Joint Conference on Neural Networks*, no. 909-916, 2009.

36. I. Chairez, A. Poznyak, and T. Poznyak, "New sliding-mode learning law for dynamic neural network observer," *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 53, no. 12, pp. 1338–1342, 2006.

37. ——, "High order dynamic neuro observer: application for ozone generator," in *International Workshop on Variable Structure Systens (VSS '08)*, 2008.

38. S. Wen, M. Chen, Z. Zeng, T. Huang, and C. Li, "Adaptive neural-fuzzy sliding-mode fault-tolerant control for uncertain nonlinear systems," *IEEE Transactions on Systems, Man and Cybernetics: Systems*, vol. PP, no. 99, pp. 1–11, 2017.

39. Y. Yildiz, A. Sabanovic, and A. Khalid, "Sliding mode neuro-controller for uncertain systems," *IEEE Transactions on Industrial Electronics*, vol. 54, no. 3, pp. 1676–1685, 2007.

40. M. R. Fajani, A. Izadbakhsh, and A. H. Ghazvinipour, "Recurrent neural network based second order sliding mode control of redundant robot manipulators," in *2018 IEEE International Conference on Industrial Technology (ICIT)*, Feb 2018, pp. 298–303.

41. A. Moreno and M. Osorio, "Strict lyapunov funtions for the super-twisting algorithm," *IEEE Tran. Aut. Cont.*, vol. 57, no. 4, pp. 1035–1040, 2012.

42. A. N. Atassi and H. Khalil, "Separation results for the stabilization of nonlinear systems using different high-gain observer design," *Systems and Control Letters*, vol. 39, pp. 183–191, 2000.

43. N. E. Cotter, "The stone-weierstrass theorem and its application to neural networks," *IEEE Transactions on Nueral Networks*, vol. 1, no. 4, pp. 290–295, 1990.

44. I. Chairez, "Wavelet differential neural network observer," *IEEE Transactions on Neural Networks*, vol. 20, no. 9, pp. 1439–1449, 2009.

45. I. Salgado, S. Kamal, I. Chairez, B. Bandyopadhyay, and L. Fridman, "Super-twisting-like algorithm in discrete time nonlinear systems," in *Proceedings of the 2011 International Conference on Advanced Mechatronic Systems*, 2011.

46. A. Poznyak, *Advanced Mathematical Tools for Automatic Control Engineers: Volume 1: Deterministic Systems*.   Elsevier Science, 2008.

47. H. Ahmed, I. Salgado, and H. Ríos, "Robust synchronization of master-slave chaotic systems using approximate model: An experimental study," *ISA transactions*, 2018.