Variable Neural Adaptive Robust Observer for Uncertain Systems

Jianming Lian, Jianghai Hu and Stanislaw H. Żak

***Abstract*— The design of variable neural adaptive robust observer is proposed for the state estimation of a class of uncertain systems. The proposed observer incorporates a variable-structure radial basis function (RBF) network to approximate unknown system dynamics. The RBF network can determine its structure on-line dynamically by adding or removing RBFs. The observer gain matrix is obtained by solving an optimization problem subject to linear matrix inequalities. The structure variation of the RBF network is taken into account in the stability analysis through the use of the piecewise quadratic Lyapunov function. The effectiveness of the proposed observer is illustrated with a simulation example.**

I. INTRODUCTION

State observer can be viewed as a dynamic system that estimates the states of a plant using its input-output measurements. The first observer was proposed by Luenberger [1]. Luenberger-like state observers for a class of nonlinear systems have been presented in [2], where the nonlinearities are assumed to be known and Lipschitz continuous. The asymptotic stability of the observation error of the proposed state observers is guaranteed if the observer gain satisfies a sufficient condition. In [3], new results related to the observer in [2] were presented, where the distance to uncontrollable (unobservable) pair of matrices and the matrix condition number of the eigenvector matrix are considered.

For systems with parametric uncertainties, adaptive observers have been developed for both state estimation and parameter identification. In [4], an adaptive observer was developed for a class of single-input single-output (SISO) linear time-invariant systems, where all the system parameters are unknown. In [5], three adaptive observers with exponential convergence were proposed for the same class of systems as in [4]. More results related to the design of adaptive observers for linear time-invariant systems can be found in [6]. The design of adaptive observers for nonlinear systems with linear parametrization are presented in [7]–[9]. In [7], an adaptive observer was proposed for a class of Lipschitz nonlinear systems. In [8], a new approach to the design of adaptive observers for multi-input multi-output (MIMO) linear time-varying systems has been proposed. In [9], robust adaptive observers were presented for nonlinear systems in the presence of bounded disturbances to prevent parameter drift. The adaptive observers for nonlinear systems with nonlinear parametrization can be found in [10], [11].

J. Lian is with the Center for Advanced Power Systems, Florida State University, Tallahassee, FL 32310, USA. lian@caps.fsu.edu.

J. Hu and S. H. Żak are with the School of Electrical and Computer Engineering, Purdue University, West Lafayette, IN 47907, USA. {jianghai,zak}@purdue.edu.

Neural network based adaptive observers have been proposed in [12]–[16] for systems with dynamic uncertainties, where fixed-structure neural networks are used to approximate unknown system dynamics. In [12], generalized dynamic recurrent neural networks were employed in the adaptive observer design for a class of SISO nonlinear systems. To use the available measurements as feedback signals, a strictly positive real (SPR) filter was applied to the output error equation. In [13], two linearly parameterized neural networks were used in the design of a robust adaptive observer for a general class of systems without the SPR-like condition. In [14], two nonlinearly parameterized neural networks were utilized, and the feedback signals in the adaptation law was generated by a simple linear filter. In [15], a stable nonlinearly parameterized neural network based adaptive observer was proposed and applied to the state estimation of flexible-joint manipulators. In [16], a robust adaptive observer with radial basis function (RBF) networks was proposed for a class of MIMO uncertain systems with both parametric and dynamic uncertainties, where the stability is guaranteed in the case when the linear part of the plant behaves like an SPR filter.

In general, it is impossible to know how the network parameters of fixed-structure neural networks should be selected. They are usually determined off-line by trial and error or using presumptive training data that are actually not available. To overcome this problem, we propose a novel variable neural adaptive robust observer for the same class of uncertain systems as considered in [16] but without parametric uncertainties. The presented adaptive robust observer adopts the variable-structure RBF network proposed by us in [17] for the self-organizing approximation of unknown system dynamics. This variable-structure RBF network avoids selecting basis functions off-line by determining the network structure on-line dynamically. It adds RBFs to improve the approximation accuracy and removes RBFs to prevent the network redundancy. The structure variation of the RBF network is considered in the stability analysis employing a piecewise quadratic Lyapunov function approach. We also discuss several methods for the calculation of the observer gain matrix.

II. SYSTEM DESCRIPTION AND PROBLEM STATEMENT

We consider a class of uncertain systems modeled by

$$\begin{cases} \dot{\boldsymbol{x}} = \boldsymbol{A}\boldsymbol{x} + \boldsymbol{B}\left(\boldsymbol{f}(\boldsymbol{x},\boldsymbol{u}) + \boldsymbol{d}\right) \\ \boldsymbol{y} = \boldsymbol{C}\boldsymbol{x}, \end{cases} \tag{1}$$

where $\boldsymbol{x} \in \mathbb{R}^n$ is the state vector, $\boldsymbol{u} \in \mathbb{R}^m$ is the input vector, $\boldsymbol{y} \in \mathbb{R}^p$ is the output vector, $\boldsymbol{d} \in \mathbb{R}^p$ is the vector of disturbances, $\boldsymbol{A} \in \mathbb{R}^{n \times n}$, $\boldsymbol{B} \in \mathbb{R}^{n \times p}$ and $\boldsymbol{C} \in \mathbb{R}^{p \times n}$ are known constant matrices, and $\boldsymbol{f}(\boldsymbol{x}, \boldsymbol{u}) \in \mathbb{R}^p$ is an unknown vector-valued function. For this system, we make the following three assumptions.

Assumption 1: The system state and input vectors take values in compact sets for all $t \geq t_0$, that is, $(\boldsymbol{x}(t),\ \boldsymbol{u}(t)) \in \boldsymbol{\Omega}_x \times \boldsymbol{\Omega}_u$ for all $t \geq t_0$, where $\boldsymbol{\Omega}_x \subset \mathbb{R}^n$ and $\boldsymbol{\Omega}_u \subset \mathbb{R}^m$ are compact sets.

Assumption 2: The vector-valued function $\boldsymbol{f}(\boldsymbol{x}, \boldsymbol{u})$ is Lipschitz continuous with respect to $\boldsymbol{x}$ and $\boldsymbol{u}$.

Assumption 3: The disturbance $\boldsymbol{d}$ is Lipschitz continuous in $\boldsymbol{x}$ and piecewise continuous in t. In addition, it is globally bounded, that is, $\|\boldsymbol{d}\| \leq d_o$ for some known $d_o > 0$.

Assumption 4: The pair $(\boldsymbol{A}, \boldsymbol{C})$ is detectable, and there exist a symmetric positive definite matrix $\boldsymbol{P} \in \mathbb{R}^{n \times n}$ and gain matrices $\boldsymbol{L} \in \mathbb{R}^{n \times p}$ and $\boldsymbol{T} \in \mathbb{R}^{p \times p}$ such that

$$(\boldsymbol{A} - \boldsymbol{L}\boldsymbol{C})^\top \boldsymbol{P} + \boldsymbol{P}(\boldsymbol{A} - \boldsymbol{L}\boldsymbol{C}) = -\boldsymbol{Q} \tag{2}$$

and

$$\boldsymbol{B}^\top \boldsymbol{P} = \boldsymbol{T}\boldsymbol{C} \tag{3}$$

for some symmetric positive definite $\boldsymbol{Q} \in \mathbb{R}^{n \times n}$.

As pointed out in [16], Assumption 4 actually implies that the linear part of the system (1) behaves like an SPR filter. We discuss later methods of solving (2) and (3) for the matrices $\boldsymbol{P}$, $\boldsymbol{L}$ and $\boldsymbol{T}$ for the given triple $(\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C})$.

Our goal is to design an adaptive robust observer for the system (1) in the presence of unknown system dynamics and disturbances. To proceed, let $\hat{\boldsymbol{x}}$ be the estimate of the state vector $\boldsymbol{x}$, and let $\hat{\boldsymbol{y}} = \boldsymbol{C}\hat{\boldsymbol{x}}$ be the estimate of the output vector $\boldsymbol{y}$. Then we define the state estimation error as $\boldsymbol{e} = \boldsymbol{x} - \hat{\boldsymbol{x}}$, and the output estimation error as $\boldsymbol{e}_y = \boldsymbol{y} - \hat{\boldsymbol{y}}$. Let $\boldsymbol{\Omega}_{e_0}$ be a compact set including all possible initial state estimation errors and let

$$c_{e_0} = \max_{\boldsymbol{e} \in \boldsymbol{\Omega}_{e_0}} \frac{1}{2} \boldsymbol{e}^\top \boldsymbol{P} \boldsymbol{e},$$

where $\boldsymbol{P}$ is defined in (2) and (3). Choose $c_e > c_{e_0}$ and let $\boldsymbol{\Omega}_e = \left\{\boldsymbol{e} : \frac{1}{2}\boldsymbol{e}^\top \boldsymbol{P}\boldsymbol{e} \leq c_e\right\}$. Then we define the compact set $\boldsymbol{\Omega}_{\hat{x}} = \{\hat{\boldsymbol{x}} : \hat{\boldsymbol{x}} = \boldsymbol{x} - \boldsymbol{e}, \boldsymbol{x} \in \boldsymbol{\Omega}_x, \boldsymbol{e} \in \boldsymbol{\Omega}_e\}$. We adopt the variable-structure RBF network proposed in [17] to approximate the unknown function $\boldsymbol{f}(\boldsymbol{x}, \boldsymbol{u})$ over the compact set $\boldsymbol{\Omega}_\chi = \boldsymbol{\Omega}_{\hat{x}} \times \boldsymbol{\Omega}_u$.

III. Variable-Structure RBF Network

The variable-structure RBF network we use here has N different admissible structures, where N is a design parameter. For each admissible structure, the RBF network consists of n input neurons, M_v hidden neurons, where $v \in \{1, \ldots, N\}$, and p output neurons. The k-th output of the RBF network with the v-th admissible structure can be represented as

$$\hat{f}_{k,v}(\hat{\boldsymbol{x}}, \boldsymbol{u}) = \sum_{j=1}^{M_v} \omega_{kj,v} \xi_{j,v}(\hat{\boldsymbol{x}}, \boldsymbol{u}), \tag{4}$$

where $\omega_{kj,v}$ is the weight from the j-th hidden neuron to the k-th output neuron and $\xi_{j,v}(\hat{\boldsymbol{x}}, \boldsymbol{u})$ is the radial basis function for the j-th hidden neuron. Let $\boldsymbol{W}_v = [\boldsymbol{\omega}_{1,v} \ \cdots \ \boldsymbol{\omega}_{p,v}]$ with $\boldsymbol{\omega}_{i,v} = [\omega_{i1,v} \ \cdots \ \omega_{iM_v,v}]^\top$ and $\xi_{j,v}(\hat{\boldsymbol{x}}, \boldsymbol{u}) = [\xi_{1,v}(\hat{\boldsymbol{x}}, \boldsymbol{u}) \ \cdots \ \xi_{M_v,v}(\hat{\boldsymbol{x}}, \boldsymbol{u})]^\top$. We have

$$\hat{\boldsymbol{f}}_v(\hat{\boldsymbol{x}}, \boldsymbol{u}) = \boldsymbol{W}_v^\top \boldsymbol{\xi}_v(\hat{\boldsymbol{x}}, \boldsymbol{u}),$$

where $\hat{\boldsymbol{f}}_v(\hat{\boldsymbol{x}}, \boldsymbol{u}) = [\hat{f}_{1,v}(\hat{\boldsymbol{x}}, \boldsymbol{u}) \ \cdots \ \hat{f}_{p,v}(\hat{\boldsymbol{x}}, \boldsymbol{u})]^\top$. We employ the raised-cosine RBF (RCRBF) proposed in [18] instead of the commonly used Gaussian RBF. The one-dimensional RCRBF is defined as

$$\xi(x) = \begin{cases} \frac{1}{2}\left(1 + \cos\left(\frac{\pi(x-c)}{\delta}\right)\right) & \text{if } |x - c| \leq \delta \\ 0 & \text{if } |x - c| > \delta, \end{cases}$$

where c is the center and δ is the radius. The n-dimensional RCRBF can be represented as the product of n one-dimensional RCRBFs. Although the Gaussian RBF is commonly used for the construction of the radial basis function network, we prefer the raised-cosine RBF to the Gaussian RBF because of the compact support of the raised-cosine RBF. As discussed in [18], the compact support of the RCRBF enables fast and efficient training and output evaluation of the RBF network.

In the following, we provide a description of the variable-structure RBF network adopted from [17].

A. Center Grid

Recall that the unknown function $\boldsymbol{f}(\boldsymbol{x}, \boldsymbol{u})$ is approximated over the compact set $\boldsymbol{\Omega}_\chi \subset \mathbb{R}^n \times \mathbb{R}^m$. To locate the centers of RBFs inside the approximation region $\boldsymbol{\Omega}_\chi$, we utilize an n-dimensional center grid with layer hierarchy, where each grid node corresponds to the center of one RBF. The grid nodes of the center grid are located at $S_1 \times \cdots \times S_n$, where S_i is the set of locations of the grid nodes in the i-th coordinate and $\times$ denotes the Cartesian product. The center grid is initialized inside the approximation region $\boldsymbol{\Omega}_\chi$ with $S_i = \{\chi_{li}, \chi_{ui}\}$, $i = 1, \ldots, n+m$, where χ_{li} and χ_{ui} denote the lower and upper bounds in the i-th coordinate. The 2^{n+m} grid nodes of the initial grid are referred to as the boundary grid nodes, and they are non-removable.

B. Adding RBFs

If the time elapsed since the last operation of adding or removing is greater than the dwell time T_d, and $\|\boldsymbol{e}_y\| > e_{\max}$, where $e_{\max}$ is a prespecified design parameter, for a period of time greater than T_d, then the network attempts to add new RBFs represented by some new grid nodes. First, the nearest neighboring grid node in the center grid, denoted $\boldsymbol{c}_{(\text{nearest})}$, to the current input $\boldsymbol{\chi} = [\hat{\boldsymbol{x}}^\top \boldsymbol{u}^\top]^\top$ is located among existing grid nodes. Then the "nearer" neighboring grid node in the center grid denoted $\boldsymbol{c}_{(\text{nearer})}$ is located, where $c_{i(\text{nearer})}$ is determined so that χ_i is between $c_{i(\text{nearest})}$ and $c_{i(\text{nearer})}$. The adding operation is performed for each coordinate independently. If the following conditions, in the i-th coordinate, are satisfied:

(1) $\left|\chi_i - c_{i(\text{nearest})}\right| > \frac{1}{4}\left|c_{i(\text{nearest})} - c_{i(\text{nearer})}\right|$,
(2) $\left|\chi_i - c_{i(\text{nearest})}\right| > d_{i(\text{threshold})}$,

where $d_{i(\text{threshold})}$ is a design parameter that specifies the minimum grid distance in the i-th coordinate and thus determines the number of admissible structures denoted by N, then a new location at exactly the middle of $c_{i(\text{nearest})}$ and $c_{i(\text{nearer})}$ is added into S_i. Otherwise, no new location is added to S_i. The layer of the newly added location is one level higher than the highest layer of the two adjacent existing locations in the same coordinate.

C. Removing RBFs

If the elapsed time since the last operation of adding or removing is greater than the dwell time T_d, and $\|\boldsymbol{e}_y\| \leq \rho e_{\max}$, $\rho \in (0,\ 1]$, for a period of time greater than T_d, then the network attempts to remove some of the existing RBFs, that is, some of the existing grid nodes, to prevent network redundancy. The RBF removing operation is also implemented for each coordinate independently. If the following conditions, in the i-th coordinate are satisfied:

(1) $c_{i(\text{nearest})} \notin \{\chi_{li}, \chi_{ui}\}$,
(2) the location $c_{i(\text{nearest})}$ is in the higher or the same layer as the highest layer of the two neighboring locations in the same coordinate,
(3) $\left|\chi_i - c_{i(\text{nearest})}\right| < \varrho \left|c_{i(\text{nearest})} - c_{i(\text{nearer})}\right|$, $\varrho \in (0,\ 0.5)$,

then the location $c_{i(\text{nearest})}$ is removed from S_i. Otherwise, no location is removed from S_i.

D. Uniform Grid Transformation

The radius of the RBF is determined using the uniform grid transformation, whose details can be found in [19].

IV. Adaptive Robust Observer Architecture

The proposed adaptive robust observer has the form

$$\begin{cases} \dot{\hat{\boldsymbol{x}}} = \boldsymbol{A}\hat{\boldsymbol{x}} + \boldsymbol{L}\left(\boldsymbol{y} - \boldsymbol{C}\hat{\boldsymbol{x}}\right) + \boldsymbol{B}\left(\hat{\boldsymbol{f}}(\hat{\boldsymbol{x}}, \boldsymbol{u}) + \boldsymbol{u}_s\right) \\ \hat{\boldsymbol{y}} = \boldsymbol{C}\hat{\boldsymbol{x}}, \end{cases} \tag{5}$$

where $\hat{\boldsymbol{f}}_v(\hat{\boldsymbol{x}}, \boldsymbol{u}) = \boldsymbol{W}_v^\top \boldsymbol{\xi}_v(\hat{\boldsymbol{x}}, \boldsymbol{u})$, and $\boldsymbol{u}_s$ is a robustifying component described later. Note that the architecture of the proposed adaptive robust observer (5) varies as the structure of the RBF network changes. A particular observer architecture is referred to as a mode. Because there are N potential network structures, the proposed observer has N different modes.

To proceed, recall that $\boldsymbol{W}_v = [\boldsymbol{\omega}_{1,v} \ \cdots \ \boldsymbol{\omega}_{p,v}]$. For the practical implementation, we constrain the weight vectors $\boldsymbol{\omega}_{i,v}$ to reside in the compact sets

$$\boldsymbol{\Omega}_{i,v} = \left\{\boldsymbol{\omega}_{i,v} : \underline{\omega}_i \leq \omega_{ij,v} \leq \overline{\omega}_i, 1 \leq j \leq M_v\right\},$$

where $\underline{\omega}_i$ and $\overline{\omega}_i$, $i = 1, 2, \ldots, p$, are design parameters. Let $\boldsymbol{W}_v^* = [\boldsymbol{\omega}_{1,v}^* \ \cdots \ \boldsymbol{\omega}_{p,v}^*]$ denote the "optimal" constant weight matrix corresponding to the v-th network structure, that is,

$$\boldsymbol{W}_v^* = \underset{\boldsymbol{\omega}_{i,v} \in \boldsymbol{\Omega}_{i,v}}{\operatorname{argmin}} \ \max_{\boldsymbol{x} \in \boldsymbol{\Omega}_{\hat{x}}, \boldsymbol{u} \in \boldsymbol{\Omega}_u} \left\| \boldsymbol{f}(\boldsymbol{x}, \boldsymbol{u}) - \boldsymbol{W}_v^\top \boldsymbol{\xi}_v(\boldsymbol{x}, \boldsymbol{u}) \right\|.$$

Let $\boldsymbol{\Phi}_v = \boldsymbol{W}_v - \boldsymbol{W}_v^* = [\boldsymbol{\phi}_{1,v} \ \cdots \ \boldsymbol{\phi}_{p,v}]$, where $\boldsymbol{\phi}_{i,v} = \boldsymbol{\omega}_{i,v} - \boldsymbol{\omega}_{i,v}^*$, and let $c = \sum_{i=1}^p c_i$, where

$$c_i = \max_v \left(\max_{\boldsymbol{\omega}_{i,v}, \boldsymbol{\omega}_{i,v}^* \in \boldsymbol{\Omega}_{i,v}} \frac{1}{2\kappa} \boldsymbol{\phi}_{i,v}^\top \boldsymbol{\phi}_{i,v} \right), \tag{6}$$

$\kappa > 0$ is a design parameter, and $\max_v(\cdot)$ denotes the maximization taken over all the potential structures of the RBF network. It is obvious that c_i decreases as κ increases. Let $\boldsymbol{\sigma} = \boldsymbol{B}^\top \boldsymbol{P} \boldsymbol{e}$. It follows from (3) that $\boldsymbol{\sigma} = \boldsymbol{T}\boldsymbol{C}\boldsymbol{e} = \boldsymbol{T}\boldsymbol{e}_y$. We employ the following projection based weight matrix adaptation law,

$$\dot{\boldsymbol{W}}_v = \operatorname{Proj}\left(\boldsymbol{W}_v, \kappa \boldsymbol{\xi}_v(\hat{\boldsymbol{x}}, \boldsymbol{u}) \boldsymbol{\sigma}^\top\right), \tag{7}$$

where $\operatorname{Proj}(\boldsymbol{W}_v, \boldsymbol{\Theta}_v)$ denotes $\operatorname{Proj}(\omega_{ij,v}, \theta_{ij,v})$ for $i = 1, \ldots, p$ and $j = 1, \ldots, M_v$ and $\operatorname{Proj}(\omega_{ij,v}, \theta_{ij,v})$ is the discontinuous projection operators defined in [17]. It follows from the definition of the projection operator that

$$\frac{1}{\kappa} \operatorname{trace}\left(\boldsymbol{\Phi}_v^\top \left(\dot{\boldsymbol{W}}_v - \kappa \boldsymbol{\xi}_v(\hat{\boldsymbol{x}}, \boldsymbol{u}) \boldsymbol{\sigma}^\top\right)\right) \leq 0. \tag{8}$$

Furthermore, the adaptation law (7) guarantees that $\boldsymbol{\omega}_{i,v}(t) \in \boldsymbol{\Omega}_{i,v}$ for $t \geq t_0$ if $\boldsymbol{\omega}_{i,v}(t_0) \in \boldsymbol{\Omega}_{i,v}$.

It follows from (1) and (5) that the dynamics of the state estimation error are given by

$$\begin{aligned} \dot{\boldsymbol{e}} &= \left(\boldsymbol{A} - \boldsymbol{L}\boldsymbol{C}\right)\boldsymbol{e} + \boldsymbol{B}\left(\boldsymbol{f}(\boldsymbol{x}, \boldsymbol{u}) + \boldsymbol{d} - \hat{\boldsymbol{f}}_v(\hat{\boldsymbol{x}}, \boldsymbol{u}) - \boldsymbol{u}_s\right) \\ &= \left(\boldsymbol{A} - \boldsymbol{L}\boldsymbol{C}\right)\boldsymbol{e} + \boldsymbol{B}\left(\boldsymbol{f}(\boldsymbol{x}, \boldsymbol{u}) - \boldsymbol{W}_v^{*\top} \boldsymbol{\xi}_v(\boldsymbol{x}, \boldsymbol{u}) + \boldsymbol{d}\right) \\ &\quad + \boldsymbol{B}\left(\boldsymbol{W}_v^{*\top} \boldsymbol{\xi}_v(\boldsymbol{x}, \boldsymbol{u}) - \boldsymbol{W}_v^{*\top} \boldsymbol{\xi}_v(\hat{\boldsymbol{x}}, \boldsymbol{u})\right) \\ &\quad + \boldsymbol{B}\left(\boldsymbol{W}_v^{*\top} \boldsymbol{\xi}_v(\hat{\boldsymbol{x}}, \boldsymbol{u}) - \hat{\boldsymbol{f}}_v(\hat{\boldsymbol{x}}, \boldsymbol{u}) - \boldsymbol{u}_s\right) \\ &= (\boldsymbol{A} - \boldsymbol{L}\boldsymbol{C})\boldsymbol{e} - \boldsymbol{B}\boldsymbol{\Phi}_v^\top \boldsymbol{\xi}_v(\hat{\boldsymbol{x}}, \boldsymbol{u}) - \boldsymbol{B}\boldsymbol{u}_s \\ &\quad + \boldsymbol{B}\left(\boldsymbol{f}(\boldsymbol{x}, \boldsymbol{u}) - \boldsymbol{W}_v^{*\top} \boldsymbol{\xi}_v(\boldsymbol{x}, \boldsymbol{u}) + \boldsymbol{d}\right) \\ &\quad + \boldsymbol{B}\left(\boldsymbol{W}_v^{*\top} \boldsymbol{\xi}_v(\boldsymbol{x}, \boldsymbol{u}) - \boldsymbol{W}_v^{*\top} \boldsymbol{\xi}_v(\hat{\boldsymbol{x}}, \boldsymbol{u})\right). \end{aligned} \tag{9}$$

As can be seen from (9), the robustifying component $\boldsymbol{u}_s$ has to counteract the effects of the approximation error as well as the bounded disturbance to force the state estimation error $\boldsymbol{e}$ to converge or, at least, be bounded. We propose a robustifying component $\boldsymbol{u}_s$ of the form

$$\boldsymbol{u}_{s,v} = \begin{cases} k_s \dfrac{\boldsymbol{\sigma}}{\|\boldsymbol{\sigma}\|} & \text{if } \|\boldsymbol{\sigma}\| \geq \nu \\ k_s \dfrac{\boldsymbol{\sigma}}{\nu} & \text{if } \|\boldsymbol{\sigma}\| < \nu, \end{cases} \tag{10}$$

where $\nu > 0$ is a design parameter and k_s is to be specified.

Remark 1: Let the increasing sequence $\{t_i\}_{i=0}^\infty$ be a partition of the interval $[t_0, \infty)$ such that $v = v_i$ over $[t_i, t_{i+1})$. During the i-th time interval $[t_i, t_{i+1})$, the observer given by (5) has a fixed architecture. Thus, as discussed in [20], there exists a unique solution $\hat{\boldsymbol{x}}_{v_i}(t)$ to (5) starting at $\hat{\boldsymbol{x}}_{v_i}(t_i)$ over $[t_i, t_{i+1})$. On the other hand, we have imposed the dwell time requirement on each mode so that $t_{i+1} - t_i \geq T_d$. Therefore, we can piece together the solutions $\hat{\boldsymbol{x}}_{v_i}(t)$ over $[t_i, t_{i+1})$ to establish the existence of a unique solution $\hat{\boldsymbol{x}}(t)$ to (5) starting at $\hat{\boldsymbol{x}}(t_0)$ over $[t_0, \infty)$, where $\hat{\boldsymbol{x}}_{v_0}(t_0) = \hat{\boldsymbol{x}}(t_0)$ and $\hat{\boldsymbol{x}}_{v_{i+1}}(t_{i+1}) = \hat{\boldsymbol{x}}_{v_i}(t_{i+1}^-)$. Here $\hat{\boldsymbol{x}}_{v_i}(t_{i+1}^-) = \lim_{t \to t_{i+1}^-} \hat{\boldsymbol{x}}_{v_i}(t)$.

Now we consider the piecewise quadratic Lyapunov function candidate whenever the proposed observer (5) is in the v-th mode,

$$V_v = \frac{1}{2}\boldsymbol{e}^\top \boldsymbol{P}\boldsymbol{e} + \frac{1}{2\kappa} \operatorname{trace}\left(\boldsymbol{\Phi}_v^\top \boldsymbol{\Phi}_v\right). \tag{11}$$

This Lyapunov function candidate has jump discontinuities when the proposed observer switches between different modes. Evaluating the time derivative of V_v along the solutions to (9) and taking into account (8), we obtain

$$\begin{aligned}\dot{V}_v &= \boldsymbol{e}^\top \boldsymbol{P}\dot{\boldsymbol{e}} + \frac{1}{\kappa}\operatorname{trace}\left(\boldsymbol{\Phi}_v^\top \dot{\boldsymbol{\Phi}}_v\right) \\ &= -\boldsymbol{e}^\top \boldsymbol{Q}\boldsymbol{e} - \boldsymbol{\sigma}^\top \boldsymbol{u}_s \\ &\quad + \boldsymbol{\sigma}^\top\left(\boldsymbol{f}(\boldsymbol{x},\boldsymbol{u}) - \boldsymbol{W}_v^{*\top}\boldsymbol{\xi}_v(\boldsymbol{x},\boldsymbol{u}) + \boldsymbol{d}\right) \\ &\quad + \boldsymbol{\sigma}^\top\left(\boldsymbol{W}_v^{*\top}\boldsymbol{\xi}_v(\boldsymbol{x},\boldsymbol{u}) - \boldsymbol{W}_v^{*\top}\boldsymbol{\xi}_v(\hat{\boldsymbol{x}},\boldsymbol{u})\right) \\ &\quad + \frac{1}{\kappa}\operatorname{trace}\left(\boldsymbol{\Phi}_v^\top \dot{\boldsymbol{\Phi}}_v\right) - \boldsymbol{\sigma}^\top \boldsymbol{\Phi}_v^\top \boldsymbol{\xi}_v(\hat{\boldsymbol{x}},\boldsymbol{u}) \\ &\le -\lambda_{\min}(\boldsymbol{Q})\|\boldsymbol{e}\|^2 + \left(d_f^* + d_o\right)\|\boldsymbol{\sigma}\| - \boldsymbol{\sigma}^\top \boldsymbol{u}_s \\ &\quad + \boldsymbol{\sigma}^\top\left(\boldsymbol{W}_v^{*\top}\boldsymbol{\xi}_v(\boldsymbol{x},\boldsymbol{u}) - \boldsymbol{W}_v^{*\top}\boldsymbol{\xi}_v(\hat{\boldsymbol{x}},\boldsymbol{u})\right) \\ &\quad + \frac{1}{\kappa}\operatorname{trace}\left(\boldsymbol{\Phi}_v^\top\left(\dot{\boldsymbol{W}}_v - \kappa\boldsymbol{\xi}_v(\hat{\boldsymbol{x}},\boldsymbol{u})\boldsymbol{\sigma}^\top\right)\right) \end{aligned} \tag{12}$$

where

$$d_f^* = \max_v\left(\max_{\boldsymbol{x}\in\boldsymbol{\Omega}_{\hat{x}},\boldsymbol{u}\in\boldsymbol{\Omega}_u}\left\|\boldsymbol{f}(\boldsymbol{x},\boldsymbol{u}) - \boldsymbol{W}_v^{*\top}\boldsymbol{\xi}_v(\boldsymbol{x},\boldsymbol{u})\right\|\right).$$

It follows from the Lipschitz continuity of the RCRBF that

$$\left\|\boldsymbol{W}_v^{*\top}\boldsymbol{\xi}_v(\boldsymbol{x},\boldsymbol{u}) - \boldsymbol{W}_v^{*\top}\boldsymbol{\xi}_v(\hat{\boldsymbol{x}},\boldsymbol{u})\right\| \le L_v\|\boldsymbol{x}-\hat{\boldsymbol{x}}\| \tag{13}$$

for some $L_v > 0$. Let $L = \max_v L_v$. It follows from (12) and (13) that

$$\begin{aligned}\dot{V}_v &\le -\lambda_{\min}(\boldsymbol{Q})\|\boldsymbol{e}\|^2 + \left(d_f^* + L + d_o\right)\|\boldsymbol{\sigma}\| - \boldsymbol{\sigma}^\top\boldsymbol{u}_s \\ &= -\lambda_{\min}(\boldsymbol{Q})\|\boldsymbol{e}\|^2 + k_s^*\|\boldsymbol{\sigma}\| - \boldsymbol{\sigma}^\top\boldsymbol{u}_s,\end{aligned} \tag{14}$$

where $k_s^* = d_f^* + L + d_o$. Let

$$k_s = d_f + L + d_o \tag{15}$$

with $d_f \ge d_f^*$ so that $k_s^* \le k_s$. If $\|\boldsymbol{\sigma}\| > \nu$, then we have

$$k_s^*\|\boldsymbol{\sigma}\| - \boldsymbol{\sigma}^\top\boldsymbol{u}_s \le k_s^*\|\boldsymbol{\sigma}\| - k_s\|\boldsymbol{\sigma}\| \le 0; \tag{16}$$

if $\|\boldsymbol{\sigma}\| \le \nu$, then

$$k_s^*\|\boldsymbol{\sigma}\| - \boldsymbol{\sigma}^\top\boldsymbol{u}_s \le k_s^*\|\boldsymbol{\sigma}\| - k_s^*\frac{\|\boldsymbol{\sigma}\|^2}{\nu} \le \frac{k_s^*}{4}\nu. \tag{17}$$

Combining (16) and (17), we obtain

$$k_s^*\|\boldsymbol{\sigma}\| - \boldsymbol{\sigma}^\top\boldsymbol{u}_s \le \frac{k_s^*}{4}\nu. \tag{18}$$

It follows from (6) that

$$\frac{1}{2\kappa}\operatorname{trace}\left(\boldsymbol{\Phi}_v^\top\boldsymbol{\Phi}_v\right) = \sum_{i=1}^{p}\frac{1}{2\kappa}\boldsymbol{\phi}_{i,v}^\top\boldsymbol{\phi}_{i,v} \le \sum_{i=1}^{p} c_i = c. \tag{19}$$

Substituting (18) into (14) and taking into account (19) gives

$$\begin{aligned}\dot{V}_v &\le -\lambda_{\min}(\boldsymbol{Q})\|\boldsymbol{e}\|^2 + \frac{k_s^*}{4}\nu \\ &\le -2\mu\frac{1}{2}\boldsymbol{e}^\top\boldsymbol{P}\boldsymbol{e} + \frac{k_s^*}{4}\nu \\ &\le -2\mu V_v + 2\mu c + \frac{k_s^*}{4}\nu \\ &\le -\mu V_v - \mu(V_v - 2\bar{c}),\end{aligned} \tag{20}$$

where

$$\mu = \frac{\lambda_{\min}(\boldsymbol{Q})}{\lambda_{\max}(\boldsymbol{P})} \quad \text{and} \quad \bar{c} = c + \frac{k_s^*\nu}{8\mu_m}.$$

Let $t_{0,v}$ and $t_{f,v}$ denote the initial and the final time, respectively, of the period during with the observer is with the v-th mode. It follows from (20) that if $V_v(t) \ge 2\bar{c}$ for $t \in [t_{0,v}, t_{f,v}]$, then $\dot{V}_v(t) \le -\mu_m V_v(t)$, which implies that

$$V_v(t) \le \exp(-\mu(t - t_{0,v}))V_v(t_{0,v}) \tag{21}$$

for $t \in [t_{0,v}, t_{f,v}]$. Applying the same arguments as in the proof of Theorem 1 and 2 in [17], we can prove the following two theorems.

Theorem 1: Let t_1, t_2 and t_3 be three consecutive switching time instants so that $v = v_1$ for $t \in [t_1, t_2)$ and $v = v_2$ for $t \in [t_2, t_3)$. Suppose that $V_v(t) \ge 2\bar{c}$ for $t \in [t_1, t_3)$. If the dwelling time T_d of the variable-structure RBF network is selected such that

$$T_d \ge \frac{1}{\mu_m}\ln\left(\frac{3}{2}\right), \tag{22}$$

then $V_{v_2}(t_2) < V_{v_1}(t_1)$ and $V_{v_2}(t_3^-) < V_{v_1}(t_2^-)$.

Theorem 2: Consider the system (1) and the variable neural adaptive robust observer (5) with the robustifying component (10) and (15) and the adaptation law (7). Suppose that T_d satisfies (22). If $c_e \ge \max\{c_{e_0} + c, 2\bar{c} + c\}$, then $\boldsymbol{e}(t) \in \boldsymbol{\Omega}_e$ and $\hat{\boldsymbol{x}}(t) \in \boldsymbol{\Omega}_{\hat{x}}$ for $t \ge t_0$. Moreover, there exists a finite time $T_1 \ge t_0$ such that

$$\frac{1}{2}\boldsymbol{e}^\top(t)\boldsymbol{P}\boldsymbol{e}(t) \le 2\bar{c} + c \tag{23}$$

for $t \ge T_1$. If, in addition, there exists a finite time $T_s \ge t_0$ such that $v = v_s$ for $t \ge T_s$, then there exists a finite time $T_2 \ge T_s$ such that

$$\frac{1}{2}\boldsymbol{e}^\top(t)\boldsymbol{P}\boldsymbol{e}(t) \le 2\bar{c} \tag{24}$$

for $t \ge T_2$.

Remark 2: It can be seen from (23) and (24) that the state estimation performance of the proposed variable neural adaptive robust observer is directly proportional to the magnitude of c and ν. Recall that the magnitude of c is inversely proportional to κ. Therefore, the larger κ and the smaller ν, the better state estimation performance.

V. Gain Matrix Computation

The design of the proposed variable neural adaptive robust observer requires the computation of the gain matrices $\boldsymbol{L}$ and $\boldsymbol{T}$. In [21], the problem of computing $\boldsymbol{P}$, $\boldsymbol{L}$ and $\boldsymbol{T}$ for a given triple $(\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C})$ was formulated as a constrained optimization problem of the form,

$$\begin{aligned}&\text{minimize} && \delta \\ &\text{subject to} && \boldsymbol{P} > 0 \\ &&& \boldsymbol{A}^\top\boldsymbol{P} + \boldsymbol{P}\boldsymbol{A} - \boldsymbol{M}\boldsymbol{C} - (\boldsymbol{M}\boldsymbol{C})^\top < 0 \\ &&& \begin{bmatrix} \delta\boldsymbol{I} & \boldsymbol{B}^\top\boldsymbol{P} - \boldsymbol{T}\boldsymbol{C} \\ (\boldsymbol{B}^\top\boldsymbol{P} - \boldsymbol{T}\boldsymbol{C})^\top & \delta\boldsymbol{I}\end{bmatrix} \ge 0,\end{aligned} \tag{25}$$

where $\boldsymbol{M} = \boldsymbol{P}\boldsymbol{L}$. The existence of $\boldsymbol{P}$, $\boldsymbol{L}$ and $\boldsymbol{T}$ is guaranteed if the optimization problem has the minimum of $\delta = 0$.

However, it is often the case that we obtain an observer gain matrix $\boldsymbol{L}$ of very large size by solving the optimization problem (25). To alleviate this, we can apply the method proposed in [22] to limit the size of the observer gain matrix $\boldsymbol{L}$. We can choose two positive constants κ_P and κ_M to constrain $\boldsymbol{P}$ and $\boldsymbol{M}$ so that

$$\boldsymbol{M}^\top \boldsymbol{M} < \kappa_M \boldsymbol{I} \tag{26}$$

and

$$\boldsymbol{P}^{-1} < \kappa_P \boldsymbol{I}. \tag{27}$$

Combining the constraints (26) and (27) and the optimization problem (25), we obtain a modified optimization problem,

$$\begin{array}{ll} \text{minimize} & \delta \\ \text{subject to} & \boldsymbol{P} > 0 \\ & \boldsymbol{A}^\top \boldsymbol{P} + \boldsymbol{P}\boldsymbol{A} - \boldsymbol{M}\boldsymbol{C} - (\boldsymbol{M}\boldsymbol{C})^\top < 0 \\ & \begin{bmatrix} \delta \boldsymbol{I} & \boldsymbol{B}^\top \boldsymbol{P} - \boldsymbol{T}\boldsymbol{C} \\ (\boldsymbol{B}^\top \boldsymbol{P} - \boldsymbol{T}\boldsymbol{C})^\top & \delta \boldsymbol{I} \end{bmatrix} \geq 0, \\ & \begin{bmatrix} -\kappa_M \boldsymbol{I} & \boldsymbol{M}^\top \\ \boldsymbol{M} & -\boldsymbol{I} \end{bmatrix} < 0 \\ & \begin{bmatrix} \boldsymbol{P} & \boldsymbol{I} \\ \boldsymbol{I} & \kappa_p \boldsymbol{I} \end{bmatrix} > 0. \end{array} \tag{28}$$

Therefore, we have $\boldsymbol{L}^\top \boldsymbol{L} < \kappa_M \kappa_P^2 \boldsymbol{I}$, which limits the size of the observer gain matrix $\boldsymbol{L}$.

On the other hand, it has been pointed out in [21] that Assumption 4 is equivalent to the following two conditions,

- $\operatorname{rank}(\boldsymbol{C}\boldsymbol{B}) = \operatorname{rank} \boldsymbol{B}$, which is referred to as observer matching condition;
- the system zeros of the system model given by the triple $(\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C})$ are located in the open left-hand complex plane, that is,

  $$\operatorname{rank} \begin{bmatrix} s\boldsymbol{I} - \boldsymbol{A} & \boldsymbol{B} \\ \boldsymbol{C} & \boldsymbol{O} \end{bmatrix} = n + \operatorname{rank} \boldsymbol{B}$$

  for all s such that $\Re(s) \geq 0$.

Therefore, if, in addition, the matrix $\boldsymbol{B}$ is also of full rank, that is, $\operatorname{rank} \boldsymbol{B} = p$, then we can apply the design procedures proposed in [23], [24] to compute $\boldsymbol{P}$, $\boldsymbol{L}$ and $\boldsymbol{T}$.

VI. Example

In this section, we use an example system adapted from [16] to illustrate the performance of the proposed variable neural adaptive robust observer. The example system is modeled by

$$\dot{\boldsymbol{x}} = \boldsymbol{A}\boldsymbol{x} + \boldsymbol{B}\left(f(\boldsymbol{x}) + g(\boldsymbol{x})\theta + \eta\right),$$

where $\boldsymbol{x} = [x_1\ x_2]^\top$, $(\boldsymbol{A}, \boldsymbol{B})$ is the canonical controllable pair, $\boldsymbol{C} = [1\ \ 8]^\top$, $f(\boldsymbol{x}) = -0.5 \tanh(3x_1) + 0.1x_2^2$, $g(\boldsymbol{x}) = 0.6 \sin(-1.8x_1 - 3x_2^3)$, $\theta = 1 - 0.4 \exp(-0.5t)$ and $\eta = \sin(0.5t)$. Let $d = g(\boldsymbol{x})\theta + \eta$. Solving the modified optimization problem (28) with $\kappa_M = \kappa_P = 50$, we obtain $\boldsymbol{L} = [0.1184\ 0.2883]^\top$ and $\boldsymbol{T} = 1$.

We choose the grid boundaries for both x_1 and x_2 to be $[-0.1\ \ 0.1]$. For the observer, we choose $k_s = 20$ and $\nu = 0.01$. The remaining design parameters of the variable-structure RCRBF network are selected as $e_{\max} = 0.0002$, $\rho = 1.0$, $\varrho = 0.3$, $\boldsymbol{d}_{\text{threshold}} = [0.1\ 0.1]$, $\underline{\omega}_1 = \underline{\omega}_2 = -5$, $\overline{\omega}_1 = \overline{\omega}_2 = 5$, $\kappa = 500$. We choose $T_d = 3.0$ so that

$$T_d = 3.0 > \frac{1}{\mu_m} \ln\left(\frac{3}{2}\right) = 1.6572,$$

where $\mu_m = 0.2447$. We can see from Fig. 1 that the estimation errors eventually enter a small neighborhood of the origin.. The variation of the number of hidden neurons is shown in Fig. 2. When the output estimation performance is poor, the network adds more RBFs to improve the approximation accuracy. When the output estimation performance is acceptable, the network removes RBFs to avoid network redundancy. Thus, the proposed variable neural adaptive robust observer performs as predicted by Theorem 2. The variation of the center grid is shown in Fig. 3.

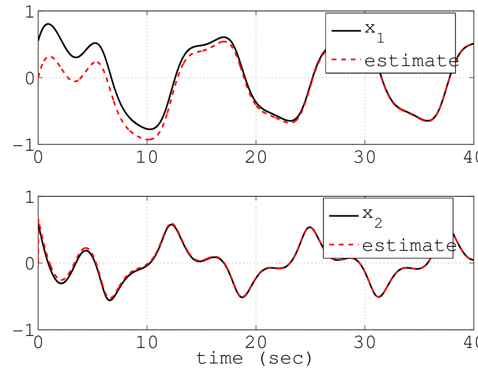


Fig. 1. State estimation performance.

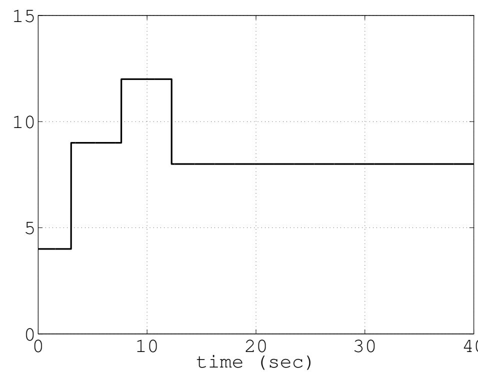
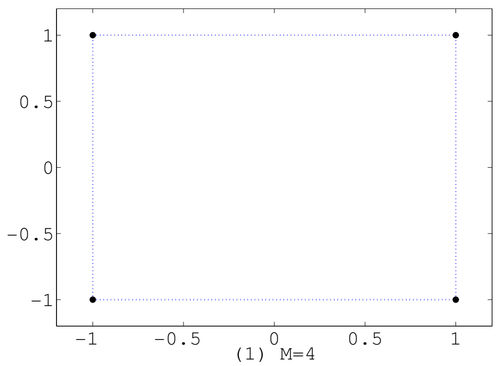


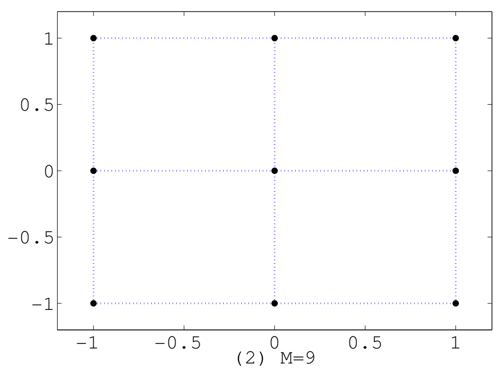
Fig. 2. Variation of the number of grid nodes.

VII. Conclusions

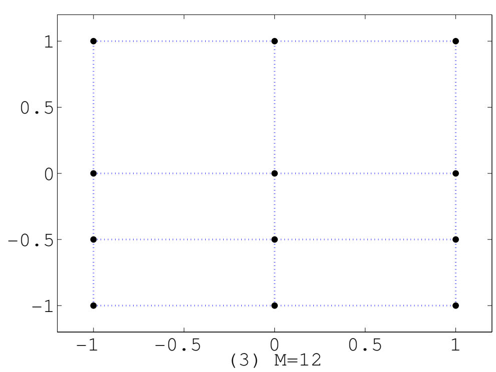
The design of a variable neural adaptive robust observer has been presented for the state estimation of a class of uncertain systems. The variable-structure RBF network employed can grow or shrink on-line dynamically according to the output estimation performance. To account for the effects of the structure variation of the RBF network in the stability analysis of the state estimation error, the piecewise quadratic Lyapunov function approach for switched and hybrid systems was used. Simulation results confirm the effectiveness of the proposed adaptive robust observer.



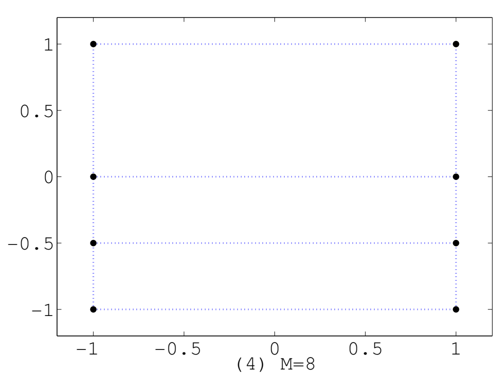
(a)



(b)



(c)



(d)

Fig. 3. Variation of the center grid.

REFERENCES

[1] D. G. Luenberger, "Observers for multivariable systems," *IEEE Trans. Autom. Control*, vol. AC-11, no. 2, pp. 190–197, Apr. 1966.

[2] F. E. Thau, "Observing the state of nonlinear dynamical systems," *Int. J. Contr.*, vol. 17, no. 3, pp. 471–479, Mar. 1973.

[3] R. Rajamani, "Observers for Lipschitz nonlinear systems," *IEEE Trans. Autom. Control*, vol. 43, no. 3, pp. 397–401, Mar. 1998.

[4] G. Lüders and K. S. Narendra, "An adaptive observer and identifier for a linear system," *IEEE Trans. Autom. Control*, vol. AC-18, no. 5, pp. 496–499, Oct. 1973.

[5] G. Kreisselmeier, "Adaptive observers with exponential rate of convergence," *IEEE Trans. Autom. Control*, vol. AC-22, no. 1, pp. 2–8, Feb. 1977.

[6] K. S. Narendra and A. M. Annaswamy, *Stable Adaptive Systems*. Englewood Cliffs, NJ: Prentice Hall, 1989.

[7] Y. M. Cho and R. Rajamani, "A systematic approach to adaptive observer synthesis for nonlinear systems," *IEEE Trans. Autom. Control*, vol. 42, no. 4, pp. 534–537, Apr. 1997.

[8] Q. Zhang and B. Delyon, "A new approach to adaptive observer design for MIMO systems," in *Proc. American Control COnference*, Arlington, VA, Jun. 2001, pp. 1545–1550.

[9] R. Marino, G. L. Santosuosso, and P. Tomei, "Robust adaptive observers for nonlinear systems with bounded disturbances," *IEEE Trans. Autom. Control*, vol. 46, no. 6, pp. 967–972, Jun. 2001.

[10] Y. Liu, "Robust adaptive observer for nonlinear systems with unmodeled dynamics," *Automatica*, vol. 45, no. 8, pp. 1891–1895, Aug. 2009.

[11] M. Farza, M. M'Saad, T. Maatoug, and M. Kamoun, "Adaptive observers for nonlinearly parameterized class of nonlinear systems," *Automatica*, vol. 45, no. 10, pp. 2292–2299, Oct. 2009.

[12] Y. H. Kin, F. L. Lewis, and C. T. Abdallah, "A dynamic recurrent neural-network-based adaptive observer for a class of nonlinear systems," *Automatica*, vol. 33, no. 8, pp. 1539–1543, 1997.

[13] J. A. R. Vargas and E. M. Hemerly, "Neural adaptive observer for general nonlinear systems," in *Proc. American Control Conference*, Chicago, IL, Jun. 2000, pp. 708–712.

[14] N. Hovakimyan, A. J. Calise, and V. K. Madyastha, "An adaptive observer design methodology for bounded nonlinear processes," in *Proc. 41st IEEE Conference on Decision and Control*, Las Vegas, NV, Dec. 2002, pp. 4700–4705.

[15] A. Abdollahi, H. A. Talebi, and R. V. Patel, "A stable neural network-based observer with application to flexible-joint manipulators," *IEEE Trans. Neural Netw.*, vol. 17, no. 1, pp. 118–129, Jan. 2006.

[16] V. Stepanyan and N. Hovakimyan, "Robust adaptive observer design for uncertain systems with bounded disturbances," *IEEE Trans. Neural Netw.*, vol. 18, no. 5, pp. 1392–1403, Sep. 2007.

[17] J. Lian, W. Zhang, J. Hu, and S. H. Żak, "Variable neural adaptive robust output feedback control of uncertain systems," in *Proc. American Control Conference*, Baltimore, MD, Jun. 2010, pp. 99–104.

[18] R. J. Schilling, J. J. Carroll, and A. F. Al-Ajlouni, "Approximation of nonlinear systems with radial basis function neural network," *IEEE Trans. Neural Netw.*, vol. 12, no. 1, pp. 1–15, Jan. 2001.

[19] J. Lian, Y. Lee, S. D. Sudhoff, and S. H. Żak, "Self-organizing radial basis function network for real-time approximation of continuous-time dynamical systems," *IEEE Trans. Neural Netw.*, vol. 19, no. 3, pp. 460–474, Mar. 2008.

[20] M. M. Polycarpou and P. A. Ioannou, "On the existence and uniqueness of solutions in adaptive control systems," *IEEE Trans. Autom. Control*, vol. 38, no. 3, pp. 474–479, Mar. 1993.

[21] M. Corless and J. Tu, "State and input estimation for a class of uncertain systems," *Automatica*, vol. 34, no. 6, pp. 757–764, 1998.

[22] D. D. Šiljak and D. M. Stipanovic, "Robust stabilization of nonlinear systems: The LMI approach," *Mathematical Problems in Engineering*, vol. 6, no. 5, pp. 461–493, 2000.

[23] C. Edwards, S. K. Spurgeon, and R. J. Patton, "Sliding mode observers for fault detection and isolation," *Automatica*, vol. 36, pp. 541–553, 2000.

[24] S. Hui and S. H. Żak, "Observer design for systems with unknown inputs," *Int. J. Appl. Math. Comput. Sci*, vol. 15, no. 4, pp. 431–446, 2005.