Author/s:
Cuevas, L;Nesic, D;Manzie, C;Postoyan, R

Title:
A Multi-observer Approach for Parameter and State Estimation of Nonlinear Systems with Slowly Varying Parameters

Date:
2020-01-01

Citation:
Cuevas, L., Nesic, D., Manzie, C. & Postoyan, R. (2020). A Multi-observer Approach for Parameter and State Estimation of Nonlinear Systems with Slowly Varying Parameters. IFAC PAPERSONLINE, 53, (2), pp.4208-4213. ELSEVIER. https://doi.org/10.1016/j.ifacol.2020.12.2465.

Persistent Link:
https://hdl.handle.net/11343/281464

Available online at www.sciencedirect.com

ScienceDirect

IFAC PapersOnLine 53-2 (2020) 4208–4213

A Multi-observer Approach for Parameter and State Estimation of Nonlinear Systems with Slowly Varying Parameters ⋆

Luis Cuevas * **Dragan Nešić** * **Chris Manzie** *
Romain Postoyan **

* *Department of Electrical and Electronic Engineering, The University of Melbourne, Australia (e-mail: lcuevas@student.unimelb.edu.au, dnesic@unimelb.edu.au, manziec@unimelb.edu.au).*
** *Université de Lorraine, CNRS, CRAN, F-54000 Nancy, France (e-mail: romain.postoyan@univ-lorraine.fr)*

Abstract: This manuscript addresses the parameter and state estimation problem for continuous time nonlinear systems with unknown slowly time-varying parameters, which are assumed to belong to a known compact set. The problem is tackled by using the multi-observer approach under the supervisory framework, which generates parameter and state estimates by using a finite number of sample points of the parameter set, a bank of observers, a set of monitoring signals and a selection criterion. This note proposes a novel dynamic sampling policy for the multi-observer technique and studies its convergence properties. We prove that the parameter and state estimation errors are ultimately bounded where the ultimate bounds can be made arbitrarily small if the parameter varies sufficiently slowly, and the number of samples is sufficiently large.

Keywords: Uncertain nonlinear systems, Multi-observer, Supervisory Framework.

1. INTRODUCTION

The parameter and state estimation problem has been of central importance in control theory. There is a number of approaches and methodologies addressing state estimation e.g., Nijmeijer and Fossen (1999); Besançon (2007); Khalil (2017), and parameter estimation or identification e.g., Ioannou (1996); Ljung (1999); Adetola and Guay (2008), to cite a few. The simultaneous estimation of both parameter and state is commonly tackled by augmenting the state vector with the parameter vector so that the original problem is transformed into a state estimation task. Nevertheless, augmenting the state may lead to a model with nonlinearities, which are difficult to handle, even in the linear case. Adaptive observers for linear and nonlinear systems allow to overcome this issue e.g., Zhang (2002); Tyukin et al. (2007); Farza et al. (2009). However, these results only apply to specific classes of systems, and the construction of an adaptive observer is a challenging task in general.

The multi-observer approach under the supervisory framework is another technique for the joint parameter and state estimation of nonlinear systems with unknown constant parameters, see Chong et al. (2015). This estimation technique was motivated by works on supervisory control for uncertain systems e.g., Morse (1996); Hespanha et al. (2003); Vu and Liberzon (2011). The multi-observer approach relies on the assumption that the unknown parameter belongs to a known compact set and that we know how to design a global state observer when the parameter is known. The technique consists of a hybrid scheme with two main parts, 1) a bank of state observers synthesized for a finite set of nominal parameter samples, and 2) a selection criterion that uses a set of monitoring signals to choose the best parameter and state estimates.

Chong et al. (2015) introduced a static and a dynamic sampling policies to create the set of nominal samples. The static policy samples the parameter set at the initial time ($t = 0$) so that it creates a large number of samples to ensure accurate estimates (accuracy is understood as the estimation errors being sufficiently small). On the other hand, the dynamic sampling policy periodically updates the samples to generate estimates with a comparable accuracy as the static case by using a smaller number of observers. This policy uses a zoom-in procedure inspired by the techniques presented in Liberzon and Nešić (2007). The zoom-in procedure is used to reduce the size of the sampled set and obtain a denser set, which may lose the parameter after a sufficiently large time. However, if the multi-observer is appropriately tuned, the error will be sufficiently small when the sampled set loses the parameter.

Although nonlinear systems with constant unknown parameters are common, there is a number of nonlinear dynamical systems with slowly varying unknown parameters; for instance, the Duffing system and the van der Pol system, see Thompson and Stewart (1986); Rajagopalan et al. (2008). Results in Chong et al. (2015) do not a priori apply to the case when the parameters vary slowly. The static sampling policy proposed by Chong et al. (2015) may exhibit implementation issues when choosing the estimates

⋆ This work was supported by the Melbourne International Research Scholarship scheme of The University of Melbourne, and by the Australian Research Council Discovery Project (DP170104102).

10.1016/j.ifacol.2020.12.2465

based on the most recent data as they assume the unknown parameters are constant. On the other hand, their dynamic sampling policy is unable to provide arbitrarily small estimation errors when the real changing parameters move away from the zoomed-in sampled set.

Here, we address the parameter and state estimation problem for nonlinear systems with slowly time-varying parameters that evolve in a known compact set. We revisit the multi-observer approach and propose a novel dynamic sampling policy to overcome the inability of existing results to work under our setting. The new dynamic policy incorporates zoom-out and zoom-in procedures inspired by quantized control techniques from Liberzon and Nešić (2007). These procedures allow to the sampled set to follow the varying parameter. The zoom-in procedure moves the centre of the "box" where the samples are taken from to the last parameter estimate and uses a zoom-in factor to reduce the size of the sampled set. On the other hand, the zoom-out procedure keeps the centre of the "box" at the same position of the last zoom-in and increases the size of the sampled set when the parameter has potentially left it. In this manuscript, we characterise the convergence properties of the parameter and state estimation errors when using our novel dynamic sampling policy. Moreover, we provide appropriate tuning algorithms that guarantee such convergence properties and an illustrative example.

Notation: Let $\mathbb{R} := (-\infty, \infty)$, $\mathbb{R}_{\geq 0} := [0, \infty)$, $\mathbb{N} := \{0, 1, \ldots\}$, $\mathbb{N}_{\geq 1} := \{1, \ldots\}$. The ∞-norm of $x \in \mathbb{R}^n$ is denoted as $|x|_\infty = \max_i |x_i|$, where $x = (x_1, \ldots, x_n)^\top$ and $|x_i|$ denotes the absolute value of x_i. Let $s : \mathbb{R}_{\geq 0} \to \mathbb{R}$, we say that $s \in \mathcal{L}_\infty$ if $||s||_\infty \leq \infty$, where $||s||_\infty := \text{ess sup}_{t\geq 0}|s(t)|$. Let $\Theta(\theta_c, \Delta) := \{\theta \in \mathbb{R}^m \mid |\theta - \theta_c|_\infty \leq \Delta\}$ be the hypercube centred at $\theta_c \in \mathbb{R}^m$ with distance to the edge $\Delta > 0$. For a vector $\theta \in \mathbb{R}^n$ and a non-empty and closed set $\boldsymbol{\Theta} \subset \mathbb{R}^m$, the distance from θ to $\boldsymbol{\Theta}$ is denoted by $d(\theta, \boldsymbol{\Theta}) := \min_{\tilde{\theta} \in \boldsymbol{\Theta}} |\theta - \tilde{\theta}|_\infty$. A continuous function $\alpha : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ is a function of class-$\mathcal{K}$, if it is strictly increasing and $\alpha(0) = 0$; additionally, if $\alpha(r) \to \infty$ as $r \to \infty$, then α is a function of class-$\mathcal{K}_\infty$. A continuous function $\beta : \mathbb{R}_{\geq 0} \times \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ is a function of class-$\mathcal{KL}$, if: (i) $\beta(\cdot, s)$ is a function of class-$\mathcal{K}$ for each $s \geq 0$; (ii) $\beta(r, \cdot)$ is non-increasing and (iii) $\beta(r, s) \to 0$ as $s \to 0$ for each $r \geq 0$. The left-limit operator is denoted by $(\cdot)^-$.

2. SYSTEM MODEL

Consider the following class of nonlinear systems

$$\dot{x}(t) = f(x(t), \theta(t), u(t)), \tag{1a}$$

$$y(t) = h(x(t), \theta(t), u(t)), \tag{1b}$$

where $x \in \mathbb{R}^n$ is the state of the system, $y \in \mathbb{R}^p$ is the measured output, $u \in \mathbb{R}^r$ is a known input, and $\theta \in \boldsymbol{\Theta}$ is an unknown time-varying parameter vector where $\boldsymbol{\Theta} \subset \mathbb{R}^m$ is assumed to be a known compact set. We make the next assumptions on system (1).

Assumption 1. θ is continuously differentiable and $|\dot{\theta}|_\infty$ is sufficiently small, i.e. there exists $\varepsilon^* \ll 1$ such that, for $\varepsilon \in (0, \varepsilon^*)$, $|\dot{\theta}(t)|_\infty \leq \varepsilon$ for all $t \geq 0$. □

Assumption 2. The maps f and h in (1) are continuously differentiable. □

Assumption 3. Consider the nonlinear system (1) with $\theta(t)$ satisfying Assumption 1. For any $\Delta_x > 0$ and $\Delta_u > 0$, there exists $k_x > 0$ such that for all $|x(0)| \leq \Delta_x$, $||u||_\infty \leq \Delta_u$, the following holds

$$|x(t)|_\infty \leq k_x \ \forall\ t \geq 0. \tag{2}$$

□

Assumption 3 implies that the solutions to (1) are bounded for any bounded initial conditions, bounded inputs and any $\theta(t)$ satisfying Assumption 1. Constant $k_x > 0$ in (2) does not need to be known in order to implement the estimation algorithm presented in the sequel.

3. A NOVEL DYNAMIC SAMPLING POLICY

Let us introduce the hypercube $\overline{\boldsymbol{\Theta}}$ which satisfies that $\boldsymbol{\Theta} \subset \overline{\boldsymbol{\Theta}}$, and discretise it with $N \in \mathbb{N}$ sample points where N is generated by Theorem 10. The sampled set is denoted as

$$\widehat{\boldsymbol{\Theta}} := \{\hat{\theta}_1, \ldots, \hat{\theta}_N\} \mid \hat{\theta}_i \in \overline{\boldsymbol{\Theta}} \text{ for } i \in \{1, \ldots, N\}. \tag{3}$$

The sample points are generated such that the following property is verified

$$\max_{\theta \in \overline{\boldsymbol{\Theta}}} d\left(\theta, \widehat{\boldsymbol{\Theta}}\right) \leq \pi(\Delta, N), \tag{4}$$

where $\pi(\cdot, \cdot) \in \mathcal{KL}$ satisfies that $\pi(\Delta, N) \leq \Delta$ where $\Delta > 0$ is the distance from the center of $\overline{\boldsymbol{\Theta}}$ to its edge. Observe that, for a uniform sampling, $\pi(s, r) = \min\left\{s, \frac{s}{r^{\frac{1}{p}}}\right\}$ ensures (4), see Chong et al. (2015). When implementing a dynamic sampling policy, we work with time-varying sets $\overline{\boldsymbol{\Theta}}(t_k)$ and $\widehat{\boldsymbol{\Theta}}(t_k) = \{\hat{\theta}_1(t_k), \ldots, \hat{\theta}_N(t_k)\}$, where t_k, for $k \in \mathbb{N}$, is the updating time satisfying

$$t_{k+1} - t_k = T_d, \tag{5}$$

with $T_d > 0$ being a design parameter. Hence, $t_k = kT_d$, for $k \in \mathbb{N}$. The full algorithm for the updating of $\overline{\boldsymbol{\Theta}}(t_k)$ is explained in Section 3.3.

3.1 Multi-observer

We work under the assumption that, if the parameter value is *known* and is *constant* for all time, we know how to design a state observer for system (1), as formalized in the following. A state observer for the system (1) is designed for each $\hat{\theta}_i(t_k) \in \widehat{\boldsymbol{\Theta}}(t_k)$, for $i \in \{1, \cdots, N\}$ and $k \in \mathbb{N}$. Hence, consider the following multi-observer

$$\dot{\hat{x}}_i(t) = f_o(\hat{x}_i(t), \hat{\theta}_i(t_k), u(t), y(t)),\ \forall\ t \in [t_k, t_{k+1}), \tag{6a}$$

$$\hat{y}_i(t) = h(\hat{x}_i(t), \hat{\theta}_i(t_k), u(t)), \tag{6b}$$

$$\hat{x}_i(t_k) = \hat{x}_i(t_k^-), \tag{6c}$$

where $\hat{x}_i \in \mathbb{R}^n$, for $i \in \{1, \ldots, N\}$, are the potential estimates of $x \in \mathbb{R}^n$, and $\hat{y}_i \in \mathbb{R}^p$ is the output estimate for each $\hat{\theta}_i(t_k) \in \widehat{\boldsymbol{\Theta}}(t_k)$, for $i \in \{1, \cdots, N\}$ and $k \in \mathbb{N}$.

Assumption 4. The map f_o in (6) is continuous and continuously differentiable in its first argument. Furthermore, the solutions to (6) are unique and defined for all positive times, for all initial conditions, any input $u \in \mathbb{R}^r$, any system output $y \in \mathbb{R}^p$ and any sampled parameter point $\hat{\theta}_i(t_k) \in \widehat{\boldsymbol{\Theta}}(t_k)$, for $i \in \{1, \cdots, N\}$ and $k \in \mathbb{N}$. □

Denoting the state estimation error as $e_{x_i} := \hat{x}_i - x$, the output error as $e_{y_i} := \hat{y}_i - y$, and the parameter error as

$e_{\theta_i} := \hat{\theta}_i - \theta$, we obtain the following state estimation error systems for system (1) and observer (6),

$$\dot{e}_{x_i}(t) = f_{e_i}(x(t), \theta(t), e_{x_i}(t), e_{\theta_i}(t), u(t)), \quad \forall\, t \in [t_k, t_{k+1}), \tag{7a}$$

$$e_{y_i}(t) = h_e(x(t), \theta(t), e_{x_i}(t), e_{\theta_i}(t), u(t)), \tag{7b}$$

for $i \in \{1, \cdots, N\}$ and $k \in \mathbb{N}$, where $f_{e_i} = f_o(e_{x_i}(t) + x(t), e_{\theta_i}(t) + \theta(t), u(t), y(t)) - f(x(t), \theta(t), u(t))$ and $h_e = h(e_{x_i}(t) + x(t), e_{\theta_i}(t) + \theta(t), u(t)) - h(x(t), \theta(t), u(t))$. Note that $e_{\theta_i}(t)$ are piecewise continuous functions with jumps (discontinuities) at each $t = t_k$, $k \in \mathbb{N}$. The assumption we make on observer (6) is stated next.

Assumption 5. Consider the state estimation error system (7), for $i \in \{1, \cdots, N\}$ and $k \in \mathbb{N}$. There exists $a_i > 0$, for $i \in \{1, \ldots, 4\}$, $\lambda_0 > 0$, a continuous non-negative function $\tilde{\gamma} : \mathbb{R}^m \times \mathbb{R}^n \times \mathbb{R}^r \to \mathbb{R}_{\geq 0}$ with $\tilde{\gamma}(0, x, u) = 0$ for all $x \in \mathbb{R}^n$, $u \in \mathbb{R}^r$ such that for any $\hat{\theta}_i(t_k) \in \widehat{\boldsymbol{\Theta}}(t_k)$, for $i \in \{1, \cdots, N\}$ and $k \in \mathbb{N}$, there exists a continuously differentiable function $V_i : \boldsymbol{\Theta} \times \mathbb{R}^n \to \mathbb{R}_{\geq 0}$, which satisfies the following for all $e_{x_i} \in \mathbb{R}^n$, $x \in \mathbb{R}^n$, $\theta \in \boldsymbol{\Theta}$, $u \in \mathbb{R}^r$

$$a_1 |e_{x_i}|_\infty^2 \leq V_i(\theta, e_{x_i}) \leq a_2 |e_{x_i}|_\infty^2, \tag{8}$$

$$\frac{\partial V_i(x, e_{x_i})}{\partial e_{x_i}} f_{e_i}(x, \theta, e_{x_i}, e_{\theta_i}, u) \leq -\lambda_0 V_i(\theta, e_{x_i}) + \tilde{\gamma}(e_{\theta_i}, x, u), \tag{9}$$

$$\left| \frac{\partial V_i(\theta, e_{x_i})}{\partial e_{x_i}} \right|_\infty \leq a_3 |e_{x_i}|_\infty, \tag{10}$$

$$\left| \frac{\partial V_i(\theta, e_{x_i})}{\partial \theta} \right|_\infty \leq a_4 |e_{x_i}|_\infty^2. \tag{11}$$

□

Assumption 5 implies that each observer is robust with respect to small parameter errors on compact sets. When $e_{\theta_i} = 0$, (8) and (9) imply that the origin of the state estimation error system (7a) is globally exponentially stable. Observe that the negativity of the derivative would depend on how large is the magnitude of e_{x_i} with respect to e_{θ_i} so that for large e_{θ_i} the error dynamics (7a) may become unstable. Inequalities in (10) and (11) are needed to handle perturbations when the parameter is slowly time-varying, see Section 9.6 in Khalil (2001). For instance, observers based on the circle criterion, e.g. Arcak and Kokotović (2001) and Chong et al. (2012), satisfy Assumption 5.

3.2 Supervisor: Monitoring signals and selection criterion

We next define monitoring signals $\mu_i(\cdot, \cdot)$, for $i \in \{1, \ldots, N\}$, which are used to select the "best" estimate from the potential estimates $\hat{x}_i$, $\hat{\theta}_i$, for $i \in \{1, \ldots, N\}$, produced by the multi-observer (6). The signal associated with each observer is the exponentially weighted $\mathcal{L}_2$-norm of the output error defined as, for any $0 \leq t_1 \leq t_2 < \infty$,

$$\mu_i(t_1, t_2) = \int_{t_1}^{t_2} \exp(-\lambda(t_2 - s)) |e_{y_i}(s)|_\infty^2 ds, \tag{12}$$

for $i \in \{1, \ldots, N\}$, where $\lambda > 0$ is a design parameter, see Vu and Liberzon (2011). We next assume that the output error of each of the observers satisfies the property below.

Assumption 6. Let $\dot{\theta} = 0$. Then, for any $\Delta_x > 0$, $\Delta_{e_x} > 0$, and $\Delta_u > 0$, there exist a class-$\mathcal{K}_\infty$ function $\alpha_{e_y}(\cdot)$ and a constant $T_{e_y} = T_{e_y}(\Delta_x, \Delta_{e_x}, \Delta_u) > 0$ such that for all $\hat{\theta}_i(t_k) \in \boldsymbol{\Theta}$, $i \in \{1, \ldots N\}$, $|x(0)| \leq \Delta_x$, $|e_{x_i}(0)| \leq \Delta_{e_x}$, and $||u||_\infty \leq \Delta_u$, the corresponding solution to systems (1) and (7) satisfies

$$\int_{t - T_{e_y}}^{t} |e_{y_i}(\tau)|_\infty^2 d\tau \geq \alpha_{e_y}(|e_{\theta_i}(t_k)|_\infty), \tag{13}$$

for all $t \geq t_k + T_{e_y}$ for any $t_k \geq 0$. □

The inequality (13) is known as a persistence of excitation (PE) condition that appears in identification and adaptive literature. Note that the excitation level grows as the norm of the parameter error increases. Then, the left-hand side of (13) gives quantitative information regarding the parameter estimation error. Assumption 6 holds when the output errors $e_{y_i} \in \mathbb{R}^p$, for $i \in \{1, \ldots, N\}$, satisfy the classical PE condition, see Chong et al. (2015).

We now define the signal $\sigma : \mathbb{R}_{\geq 0} \to \{1, \ldots, N\}$ which is used to choose a parameter estimate and an observer from (6) at every time instant. It is defined as

$$\sigma(t_{k+1}) \in \arg \min_{i \in \{1, \ldots, N\}} \mu_i(t_k, t_{k+1}), \tag{14}$$

for $k \in \mathbb{N}$. Based on the signal (14), the estimated parameter and the estimated state are given by $\hat{\theta}(t) := \hat{\theta}_{\sigma(t_k^-)}(t)$, and $\hat{x}(t) := \hat{x}_{\sigma(t_k^-)}(t)$, for all $t \in [t_k, t_{k+1})$. Hence, $\hat{\theta}(t)$ and $\hat{x}(t)$ are discontinuous in general as they switch among a finite family of continuous trajectories that are in general different at the switching instants.

3.3 Algorithm: Dynamic sampling policy

Let $\Delta_0 \geq 0$ and $\theta_c \in \mathbb{R}^n$ be given such that $\boldsymbol{\Theta} \subset \Theta(\theta_c, \Delta_0)$. Let $a \in (0, 1)$, $b > 1$, $c > 0$, $\delta_0 > 0$, and $\delta_1 \in (0, \delta_0)$, and let $N \in \mathbb{N}$ and $T_d > 0$, where a and b are the zoom-in and zoom-out factors, respectively, c is a threshold for the zoom-out procedure, δ_0 and δ_1 are thresholds for the monitoring signals, N is the number of samples and T_d is the sampling time in (5). We explain how to tune these parameters in the following. In view of (5), let $t_k := kT_d$, for $k \in \mathbb{N}$. Moreover, let $\hat{\theta}_0$ be the initial condition for the parameter estimate and let $m(t_k)$ be a discrete variable which will take values in the set {'zoom-in', 'zoom-out'} with initial value $m(t_0) =$ 'zoom-in'.

i. Set $k = 0$. Let $\theta_c(t_0) = \theta_c$ and $\Delta(t_0) = \Delta_0$ such that $\boldsymbol{\Theta} \subset \Theta(\theta_c(t_0), \Delta(t_0))$ and define $\overline{\boldsymbol{\Theta}}(t_0) = \Theta(\theta_c(t_0), \Delta(t_0)) \cap \boldsymbol{\Theta}$.

ii. Generate the set $\widehat{\boldsymbol{\Theta}}(t_k)$ by using (3) and (4).

iii. Design the multi-observer (6) for the system (1). Then, the monitoring signals (12) are implemented as follows, for $i \in \{1, \ldots, N\}$ and $k \in \mathbb{N}$

$$\dot{\mu}_i(t_k, t) = -\lambda \mu_i(t_k, t) + |e_{y_i}(t)|^2, \ \forall\, t \in [t_k, t_{k+1}),$$
$$\mu_i(t_k, t_k) = 0. \tag{15}$$

The selection criterion signal is as in (14), for $k \in \mathbb{N}$.

iv. Let $\mu_{\sigma(t_{k+1})} = \min_{i \in \{1, \ldots, N\}} \mu_i(t_k, t_{k+1})$, and

$$m(t_{k+1}) = \begin{cases} \text{'zoom-in'} & \text{if } \mu_{\sigma(t_{k+1})} < \delta_1, \\ \text{'zoom-out'} & \text{if } \mu_{\sigma(t_{k+1})} > \delta_0, \\ m(t_k) & \text{if } \mu_{\sigma(t_{k+1})} \in [\delta_1, \delta_0]. \end{cases} \tag{16}$$

v. Implement the following,

- **Zoom-in**: If $m(t_{k+1})$ = 'zoom-in', let

$$\theta_c(t_{k+1}) = \hat{\theta}_{\sigma(t_{k+1}^-)}(t_{k+1}^-), \quad (17)$$
$$\Delta(t_{k+1}) = a\Delta(t_k), \quad (18)$$
$$\overline{\mathbf{\Theta}}(t_{k+1}) = \Theta(\theta_c(t_{k+1}), \Delta(t_{k+1})) \cap \overline{\mathbf{\Theta}}(t_k). \quad (19)$$

- **Zoom-out**: If $m(t_{k+1})$ = 'zoom-out', let

$$\theta_c(t_{k+1}) = \theta_c(t_k), \quad (20)$$
$$\Delta(t_{k+1}) = b\max\{\Delta(t_k), c\}, \quad (21)$$
$$\overline{\mathbf{\Theta}}(t_{k+1}) = \Theta(\theta_c(t_{k+1}), \Delta(t_{k+1})) \cap \overline{\mathbf{\Theta}}(t_0). \quad (22)$$

vi. Let $k = k + 1$. Then, go to step ii.

Step **i** is an initialization step used to generate the first set where the samples will be taken from at t_0. Step **ii** generates the sampled set at the k-th iteration. Step **iii** corresponds to the implementation of the multi-observer technique by itself. Step **iv** uses the information from the monitoring signals to indirectly evaluate if the changing parameter is within the sampled set or not by using the hysteresis switching law in (16). Step **v** implements a zoom-in or a zoom-out procedure. First, a hypercube is defined by specifying its centre $\theta_c(t_{k+1})$ and half of its edge's length $\Delta(t_{k+1})$. Then, the hypercube $\Theta(\theta_c(t_{k+1}), \Delta(t_{k+1}))$ is intersected with the previous set $\overline{\mathbf{\Theta}}(t_k)$ for the zoom-in case, and with the initial set $\overline{\mathbf{\Theta}}(t_0)$ for the zoom-out. The final step closes the loop so that the new set generated in Step **v** is used to create the new set of samples for the next T_d time interval. Note that our proposed algorithm is such that $\hat{\theta}_i(t_k) \in \mathbf{\Theta}$, for $i \in \{1, \ldots, N\}$ and $k \in \mathbb{N}$.

4. MAIN RESULT

We first present three useful lemmas which are the key ingredients to state our main result. In Lemma 7, we state an input-to-state stability property with respect to the parameter error e_{θ_i}, for $i \in \{1, \ldots, N\}$, for each of the error systems (7a). Then, we show in Lemma 8 that Assumption 6 leads to a weaker persistence of excitation condition when $\dot{\theta}(t) \neq 0$ and $|\theta(t)|_\infty \leq \varepsilon$. Furthermore, we state in Lemma 9 that the monitoring signals $\mu_i(t_k, t_{k+1})$, for $i \in \{1, \ldots, N\}$, are lower and upper bounded by class-$\mathcal{K}_\infty$ functions of the parameter estimation error. In Theorem 10, we characterise the convergence properties of the multi-observer (6) and our dynamic sampling policy. We do not present any proofs due to the page limitation.

Lemma 7. Consider the system (1), and the estimation error systems (7a). Let Assumptions 1 - 5 hold. Then, there exist $\kappa > 0$, $\lambda_{L1} > 0$ and $\tilde{\varepsilon}^* > 0$ such that for any $\overline{\Delta}_x > 0$, $\overline{\Delta}_{e_x} > 0$ and $\overline{\Delta}_u > 0$, there exists $\gamma_L(\cdot) \in \mathcal{K}_\infty$ such that the corresponding solutions to (7a) satisfy, for $i \in \{1, \ldots, N\}$,

$$|e_{x_i}(t)|_\infty \leq \kappa \exp(-\lambda_{L1} t)\, |e_{x_i}(0)|_\infty + \gamma_L(||e_{\theta_i}||_\infty), \quad (23)$$

for all $\varepsilon \in (0, \tilde{\varepsilon}^*)$, $\theta, \hat{\theta}_i \in \mathbf{\Theta}$, $|x(0)|_\infty \leq \overline{\Delta}_x$, $|e_{x_i}(0)|_\infty \leq \overline{\Delta}_{e_x}$, $||u||_\infty \leq \overline{\Delta}_u$, and $t \geq 0$. □

Lemma 8. Consider the error systems (7a) and let Assumptions 1 - 6 hold. Then, for any $\underline{\Delta}_x > 0$, $\underline{\Delta}_{e_x} > 0$ and $\underline{\Delta}_u > 0$, there exist a class-$\mathcal{K}_\infty$ function $\alpha_L(\cdot)$, a constant $T_f = T_f(\underline{\Delta}_x, \underline{\Delta}_{e_x}, \underline{\Delta}_u) > 0$, $\hat{\kappa} > 0$ and $\tilde{\varepsilon}^* > 0$, such that, for $i \in \{1, \ldots N\}$ and for $k \in \mathbb{N}$,

$$\int_{t-T_f}^{t} |e_{y_i}(\tau)|_\infty^2 d\tau \geq \max\left\{\alpha_L(|e_{\theta_i}(t_k)|_\infty) - \varepsilon^2\hat{\kappa}, 0\right\}, \quad (24)$$

for all $t \geq t_k + T_f$, for any $t_k \geq 0$, $\varepsilon \in (0, \tilde{\varepsilon}^*)$, $|x(0)|_\infty \leq \underline{\Delta}_x$, $|e_{x_i}(0)|_\infty \leq \underline{\Delta}_{e_x}$, and $||u||_\infty \leq \underline{\Delta}_u$. □

Lemma 9. Consider the system (1), the error system (7a), and the monitoring signals (12). Let Assumptions 1 - 6 hold. For any $\tilde{\Delta}_x > 0$, $\tilde{\Delta}_{e_x} > 0$, $\tilde{\Delta}_u > 0$ and $\nu > 0$, there exist class-$\mathcal{K}_\infty$ functions $\underline{\chi}(\cdot)$ and $\overline{\chi}(\cdot)$ independent of ν, $k_{LM} > 0$, a constant $T = T(\tilde{\Delta}_x, \tilde{\Delta}_{e_x}, \tilde{\Delta}_u, \nu) > 0$, $T_d \geq T$ and $\overline{\varepsilon}^* > 0$ such that the monitoring signals $\mu_i(t_k, t)$ satisfy, for $i \in \{1, \ldots, N\}$ and for $k \in \mathbb{N}$,

$$\max\left\{\underline{\chi}(|e_{\theta_i}(t_k)|_\infty) - \varepsilon^2 k_{LM}, 0\right\} \leq \mu_i(t_k, t) \leq \overline{\chi}(|e_{\theta_i}(t_k)|_\infty) + \nu, \quad (25)$$

for all $t \in [t_k + T, t_{k+1})$, and for all $\varepsilon \in (0, \overline{\varepsilon}^*)$, $\theta, \hat{\theta}_i \in \mathbf{\Theta}$, $|x(0)|_\infty \leq \tilde{\Delta}_x$, $|e_{x_i}(0)|_\infty \leq \tilde{\Delta}_{e_x}$, $||u||_\infty \leq \tilde{\Delta}_u$. □

We are now ready to present the main result of this note.

Theorem 10. Consider the nonlinear system (1), the dynamic sampling policy, and the error systems (7). Let Assumptions 1 - 6 hold. Then, for any given $\hat{\Delta}_x > 0$, $\hat{\Delta}_{e_x} > 0$, $\hat{\Delta}_u > 0$, $\hat{\nu}_{e_\theta} > 0$, $\hat{\nu}_{e_x} > 0$, zooming factors $a \in (0, 1)$ and $b > 1$, a constant $c \in (0, \min\left\{\hat{\nu}_{e_\theta}, \gamma_L^{-1}(\hat{\nu}_{e_x})\right\}/2b\sqrt{m})$, where $\gamma_L \in \mathcal{K}_\infty$ comes from Lemma 7 and $m > 0$ is the dimension of $\mathbf{\Theta}$, $\delta_0 \in \left(0, \underline{\chi}\left((1-\vartheta)c\right)\right)$, for $\vartheta \in (0, 1)$, and $\delta_1 \in (0, \delta_0)$, there exists $\widehat{K}_{e_\theta} > 0$, $\widehat{K}_{e_x} > 0$, sufficiently large $T^* > 0$ and $N^* \in \mathbb{N}$ such that for any $T_d \geq T^*$ and $N \geq N^*$ there exists $\varepsilon^* > 0$, constructed according to Algorithm 1 below, such that the following holds

$$|e_{\theta\sigma(t)}(t)|_\infty \leq \widehat{K}_{e_\theta}, \quad (26)$$
$$|e_{x\sigma(t)}(t)|_\infty \leq \widehat{K}_{e_x}, \quad (27)$$
$$\limsup_{t\to\infty} |e_{\theta\sigma(t)}(t)|_\infty \leq \hat{\nu}_{e_\theta}, \quad (28)$$
$$\limsup_{t\to\infty} |e_{x\sigma(t)}(t)|_\infty \leq \hat{\nu}_{e_x}, \quad (29)$$

for all $\varepsilon \in (0, \varepsilon^*)$, $|x(0)|_\infty \leq \hat{\Delta}_x$, $|e_{x_i}(0)|_\infty \leq \hat{\Delta}_{e_x}$ for $i \in \{1, \ldots, N\}$, $||u||_\infty \leq \hat{\Delta}_u$, and $t \geq 0$. □

Theorem 10 states that the parameter and state estimation errors are bounded for all time. Moreover, this result ensures that the estimated parameter $\hat{\theta}(t)$ and state $\hat{x}(t)$, are respectively guaranteed to converge to their true values within some selected margins $\hat{\nu}_{e_\theta}$ and $\hat{\nu}_{e_x}$. Furthermore, it provides information on how to select and determine the parameters of the algorithm in Section 3.3. We next provide the algorithm for the construction of $T^* > 0$, $N^* \in \mathbb{N}$, $N \geq N^*$, $T_d > 0$, and $\varepsilon^* > 0$.

Algorithm 1. Let $\hat{\Delta}_x > 0$, $\hat{\Delta}_{e_x} > 0$, $\hat{\Delta}_u > 0$, $\hat{\nu}_{e_\theta} > 0$, $\hat{\nu}_{e_x} > 0$, $a \in (0, 1)$, $b > 1$, $c \in (0, \min\left\{\hat{\nu}_{e_\theta}, \gamma_L^{-1}(\hat{\nu}_{e_x})\right\}/2b\sqrt{m})$, $\delta_0 \in \left(0, \underline{\chi}\left((1-\vartheta)c\right)\right)$, for $\vartheta \in (0, 1)$, and $\delta_1 \in (0, \delta_0)$ be given. Consider $\gamma_L(\cdot) \in \mathcal{K}_\infty$ and $\tilde{\varepsilon}^* > 0$ generated by Lemma 7, $\underline{\chi}(\cdot), \overline{\chi}(\cdot) \in \mathcal{K}_\infty$, $k_{LM} > 0$ and $\overline{\varepsilon}^* > 0$ generated by Lemma 9 and $\pi(\cdot, \cdot) \in \mathcal{KL}$ satisfying (4). Then, we have the following.

(1) Define $\hat{\eta} > 0$ as $\hat{\eta} := \min\left\{\hat{\nu}_{e_\theta}, \gamma_L^{-1}(\hat{\nu}_{e_x})\right\}$.

(2) Select $\Delta_0 > 0$ as stated in Section 3.3, $\Delta_2 \in (0, \Delta_0)$ and $\Delta_1 \in (\Delta_2, \Delta_0)$ such that $\Delta_2 = c$ and $\Delta_1 = 2bc\sqrt{m}$.

(3) Select $T^* > 0$, $N^* \in \mathbb{N}$ sufficiently large and $T_d \geq T^*$ such that Lemma 9 holds with $\nu > 0$ sufficiently small such that

$$\underline{\chi}^{-1}(\overline{\chi}(\pi(s, N^*)) + 2\nu) \leq as, \quad (30)$$

for all $s \in [\Delta_2, \Delta_0]$.

(4) Select $N > 0$ sufficiently large such that $N \geq N^*$,

$$\overline{\chi}(\pi(\Delta_0, N)) + \nu \leq \delta_0, \quad (31)$$

and

$$\overline{\chi}(\pi(\Delta_1, N)) + \nu < \delta_1. \quad (32)$$

(5) Define $\varepsilon^* > 0$ as follows

$$\varepsilon^* := \min\{\varepsilon_1^*, \varepsilon_2^*, \varepsilon_3^*, \varepsilon_4^*, \varepsilon_5^*, \varepsilon_6^*\}. \quad (33)$$

where $\varepsilon_1^* = \tilde{\varepsilon}^*$, $\varepsilon_2^* = \overline{\varepsilon}^*$, $\varepsilon_3^* = \frac{\pi(\Delta_2, N)}{T_d}$, $\varepsilon_4^* = \sqrt{\frac{\nu}{k_{LM}}}$, $\varepsilon_5^* = \sqrt{\frac{\underline{\chi}((1-\vartheta)\Delta_2) - \delta_0}{k_{LM}}}$, $\varepsilon_6^* = \frac{(b-1)\Delta_2}{2T_d}$.

Remark 11. It is always possible to ensure (30) since $\underline{\chi}(\cdot), \overline{\chi}(\cdot) \in \mathcal{K}_\infty$ and $\pi(\cdot,\cdot) \in \mathcal{KL}$. Furthermore, the choice of δ_0 and Δ_2 ensures a positive numerator in the argument of the square root of ε_5^*. Observe that we can always choose $\delta_0 \in (0, \underline{\chi}((1-\vartheta)c))$, for $\vartheta \in (0,1)$, $\delta_1 \in (0, \delta_0)$ and $N \geq \mathbb{N}^*$ such that (31) and (32) hold due to the properties of the functions $\overline{\chi}(\cdot)$ and $\pi(\cdot,\cdot)$ and because $\Delta_0 > \Delta_1$.

5. ILLUSTRATIVE EXAMPLE

Consider the neural mass model of Jansen and Rit (1995) in the following form

$$\dot{x} = A(\theta(t))x + G(\theta(t))\gamma(Hx) + B(\theta(t))\sigma(u, y), \quad (34a)$$
$$y = C(\theta(t))x, \quad (34b)$$

where $\theta \in \mathbf{\Theta} \subset \mathbb{R}^m$, $x \in \mathbb{R}^n$, $y \in \mathbb{R}^p$, $u \in \mathbb{R}^u$, $\gamma : \mathbb{R}^m \to \mathbb{R}^s$ and $\sigma : \mathbb{R}^r \times \mathbb{R}^p \to \mathbb{R}^q$. The system matrices are defined as $A = \mathrm{diag}(A_a, A_a, A_b)$, $G(\theta(t)) = \begin{bmatrix} 0 & 0 & 0 & \theta_1(t)ac_2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \theta_2(t)bc_4 \end{bmatrix}^T$, $H = \begin{bmatrix} c_1 & 0_{1\times5} \\ c_3 & 0_{1\times5} \end{bmatrix}$, $B(\theta(t)) = \begin{bmatrix} 0 & \theta_1(t)a & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \theta_1(t)a & 0 & 0 \end{bmatrix}^T$, $C = [0\ 0\ 1\ 0\ -1\ 0]$, where $A_a = \begin{bmatrix} 0 & 1 \\ -a^2 & -2a \end{bmatrix}$ and $A_b = \begin{bmatrix} 0 & 1 \\ -b^2 & -2b \end{bmatrix}$, where a, b, c_1, c_2, c_3 and c_4 are assumed to be known. The nonlinear terms in (34) are $\gamma = (S, S)$ and $\sigma(u, y) = (S(y), u)$ where the function S denotes the sigmoid function $S(v) := \frac{2e_0}{1+\exp[r(v_0, v)]}$ for $v \in \mathbb{R}$ with known constants e_0, v_0 and $r \in \mathbb{R}_{\geq 0}$. The states x_1, x_3 and x_5 are the membrane potential contributions of the pyramidal neurons, the excitatory and the inhibitory inter-neurons respectively, and x_2, x_4 and x_6 are their respective time derivatives. The unknown parameters θ_1 and θ_2 represent the synaptic gains of excitatory and inhibitory neuronal populations.

We assume the unknown parameter vector is slowly time-varying and that it belongs to $\mathbf{\Theta} := [4, 8] \times [22, 28]$. Importantly, we consider that

$$\theta_1(t) = \begin{cases} 6.5 + 0.01t & \text{if } \theta_1 \in (4, 8), \\ 8 & \text{otherwise}, \end{cases} \quad (35a)$$

$$\theta_2(t) = \begin{cases} 25.5 + 0.015t & \text{if } \theta_2 \in (22, 28), \\ 28 & \text{otherwise}, \end{cases} \quad (35b)$$

so that $|\dot{\theta}(t)| \leq \varepsilon$ where $\varepsilon = 0.015$. Let us consider the state observer introduced in Chong et al. (2012) to construct the following multi-observer

$$\dot{\hat{x}}_i = A(\hat{\theta}_i)\hat{x}_i + G(\hat{\theta}_i)\gamma(H\hat{x}_i + K(\hat{\theta}_i)(C(\hat{\theta}_i)\hat{x}_i - y)) + B(\hat{\theta}_i)\sigma(u, y) + L(\hat{\theta}_i)(C(\hat{\theta}_i)\hat{x}_i - y), \quad (36a)$$
$$\hat{y}_i = C(\hat{\theta}_i)\hat{x}_i, \quad (36b)$$

for $i \in \{1, \ldots, N\}$, where $K(\hat{\theta}_i)$ and $L(\hat{\theta}_i)$ are the observer gain matrices which are computed as described in the following. Suppose there exist real matrices $P_i = P_i^T > 0$, $M_i = \mathrm{diag}(m_{i1}, \ldots, m_{in}) > 0$ and scalars ν_i, μ_i such that the following holds

$$\begin{bmatrix} \mathbf{A}(P_i, L(\hat{\theta}_i), \nu_i) & \mathbf{B}(P_i, M_i, K(\hat{\theta}_i)) & P_i \\ \star & \mathbf{E}(M_i) & 0 \\ \star & \star & -\nu_i\mathbb{I} \end{bmatrix} \leq 0, \quad (37)$$

where the elements are

$$\mathbf{A}(P_i, L(\hat{\theta}_i), \nu_i) = P_i(A(\hat{\theta}_i) + L(\hat{\theta}_i)C(\hat{\theta}_i)) + \nu_i, \mathbb{I} + (A(\hat{\theta}_i) + L(\hat{\theta}_i)C(\hat{\theta}_i))^T P_i$$
$$\mathbf{B}(P_i, M_i, K(\hat{\theta}_i)) = P_i G(\hat{\theta}_i) + (H + K(\hat{\theta}_i)C(\hat{\theta}_i))^T M_i,$$
$$\mathbf{E}(M_i) = -2M_i \mathrm{diag}\left(\frac{1}{b_{\gamma_1}}, \cdots, \frac{1}{b_{\gamma_n}}\right),$$

where $\mathbb{I}$ is the identity matrix and $b_{\gamma_k} \in \mathbb{R}^n \backslash 0$ is such that $\frac{\partial \gamma_k(v_k)}{\partial v_k} \leq b_{\gamma_k} < \infty$ for all $v_k \in \mathbb{R}$ where $\gamma = (\gamma_1, \ldots, \gamma_n)$.

We have checked that Assumptions 1 - 5 hold and assumed that the PE condition stated in Assumption 6 is satisfied. To perform simulations, we consider the following values: $a = 100$, $b = 50$, $c_1 = 135$, $c_2 = 108$, $c_3 = 33.75$, $c_4 = 33.75$, $e_0 = 2.5$, $v_0 = 6$, $r = 0.56$. Table 1 summarises simulation results for our dynamic sampling policy as well as for the dynamic policy from Chong et al. (2015). It validates the good performance of our proposed dynamic sampling policy. Moreover, it demonstrates the inability of the dynamic policy from Chong et al. (2015) to deal with slowly varying parameters as the parameter estimation error cannot be made arbitrarily small. Note that the policy from Chong et al. (2015) eventually loses the parameter even for a large number of observers.

Table 1. Simulation Results ($t_f = 200s$).

	$N = 100$	$N = 256$	$N = 400$
Dynamic Policy: $\lvert e_{\theta_{\sigma(t_f)}}(t_f)\rvert$	0.003	0.001	0.001
Chong et al. (2015) results: $\lvert e_{\theta_{\sigma(t_f)}}(t_f)\rvert$	1.79	1.74	1.45

We provide snapshots of the sampled set for $N = 100$ and $t \in [0, 200)s$ at each updating time t_k, for $k \in \{1, 2, 9, 10, 11, 12, 13, 19\}$, in Fig. 1. For $t \in [80, 90)s$, the system parameter is not longer in the sampled set. This is expected as we are using a zoom-in procedure and the parameter is changing. Although the parameter is no longer in the sampled set at $t = 90s$ (iteration 9), the zoom-out procedure is executed until $t = 100s$ (iteration 10). This is a natural behaviour produced by the hysteresis switching law (16) and the parameter tuning of the algorithm in Section 3.3, which may not immediately identify when the parameter has left the sampled set. The parameter estimation error becomes arbitrarily small after a sufficiently large time.

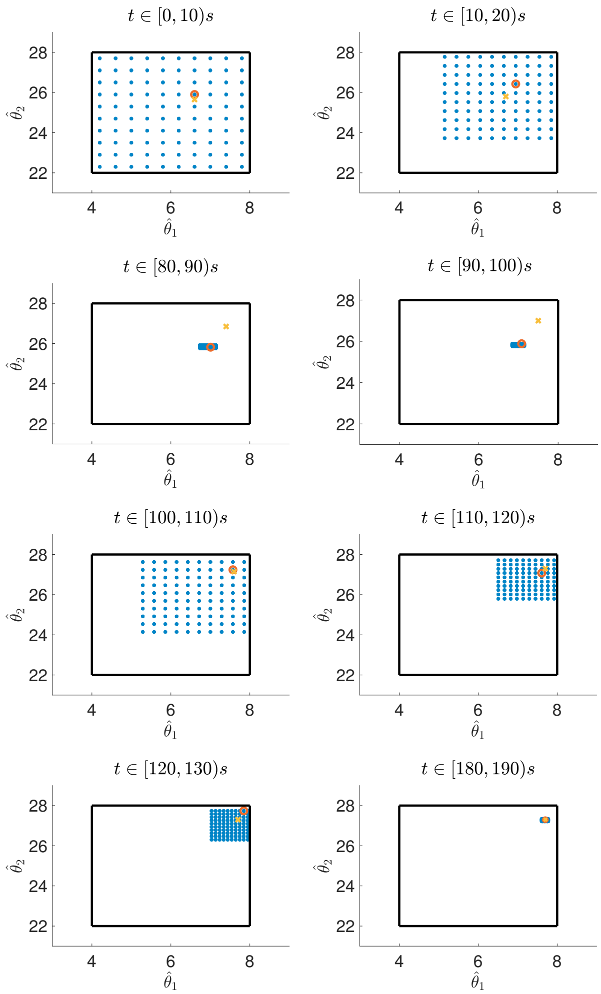

Fig. 1. The dynamic sampling of the parameter set for the updating times t_k, for $k \in \{1, 2, 9, 10, 11, 12, 13, 19\}$, for $N = 100$. Legends: Yellow - real parameter, Blue - Parameter sample points, Red - Parameter estimate, Black - Boundaries of the compact set $\mathbf{\Theta}$.

6. CONCLUSIONS

We generalised the multi-observer approach under the supervisory framework for nonlinear systems to handle slowly varying unknown parameters. We delivered convergence results for the multi-observer approach when using a novel static sampling policy. Furthermore, we proposed a novel dynamic sampling policy that uses zoom-in and zoom-out procedures to guarantee the same accuracy as the static sampling policy with a reduced number of observers. We stated convergence results that guarantee that our proposed technique produces parameter and state estimation errors that can be made as small as desired if the slowly time-varying parameter moves sufficiently slow and if the observer is carefully tuned.

REFERENCES

Adetola, V. and Guay, M. (2008). Finite-time parameter estimation in adaptive control of nonlinear systems. *IEEE Transactions on Automatic Control*, 53(3), 807–811.

Arcak, M. and Kokotović, P. (2001). Observer-based control of systems with slop-restricted nonlinearities. *Automatica*, 37(12), 1923–1930.

Besançon, G. (2007). *Nonlinear Observers and Applications*. Springer-Verlag, Berlin Heidelberg.

Chong, M.S., Nešić, D., Postoyan, R., and Kuhlmann, L. (2015). Parameter and state estimation of nonlinear systems using a multi-observer under the supervisory framework. *IEEE Transactions on Automatic Control*, 60(9), 2336–2349.

Chong, M.S., Postoyan, R., Nešić, D., Kuhlmann, L., and Varsavcky, A. (2012). A robust circle criterion observer with application to neural mass models. *Automatica*, 48(11), 2986–2989.

Farza, M., M'Saad, M., Maatoug, T., and Kamoun, M. (2009). Adaptive observers for nonlinearly parameterized class of nonlinear systems. *Automatica*, 45(10), 2292–2299.

Hespanha, J.P., Liberzon, D., and Morse, A.S. (2003). Hysteresis-based switching algorithms for supervisory control of uncertain systems. *Automatica*, 39, 263–272.

Ioannou, P.A. (1996). *Robust adaptive control*. Prentice Hall.

Jansen, B. and Rit, V.G. (1995). Electroencephalogram and visual evoked potential generation in a mathematical model of coupled cortical columns. *Biological Cybernetics*, 73, 357–366.

Khalil, H. (2001). *Nonlinear Systems*. Prentice Hall, Upper Saddle River, New Jersey, 3rd edition.

Khalil, H.K. (2017). *High-Gain Observers in Nonlinear Feedback Control*. Society for Industrial and Applied Mathematics.

Liberzon, D. and Nešić, D. (2007). Input-to-state stabilization of linear systems with quantized state measurements. *IEEE Transactions on Automatic Control*, 52(5), 767–781.

Ljung, L. (1999). *System Identification*. Wiley Online Library.

Morse, A.S. (1996). Supervisory control of families of linear set-point controllers - part 1: Exact matching. *IEEE Transactions on Automatic Control*, 41(10), 1413–1431.

Nijmeijer, H. and Fossen, T.I. (1999). *New Directions in Nonlinear Observer Design*. Springer-Verlag, London.

Rajagopalan, V., Chakraborty, S., and Ray, A. (2008). Estimation of slowly varying parameters in nonlinear systems via symbolic dynamic filtering. *Signal Processing*, 88, 339–348.

Thompson, J. and Stewart, H. (1986). *Nonlinear Dynamics and Chaos*. Wiley, Chichester, UK.

Tyukin, I.Y., Prokhorov, D.V., and van Leeuwen, C. (2007). Adaptation and parameter estimation in systems with unstable target dynamics and nonlinear parametrization. *IEEE Transactions on Automatic Control*, 52(9), 1543–1559.

Vu, L. and Liberzon, D. (2011). Supervisory control of uncertain linear time-varying systems. *IEEE Transactions on Automatic Control*, 56(1), 27–42.

Zhang, Q. (2002). Adaptive observer for multiple-input-multiple-output (mimo) linear time-varying systems. *IEEE Transactions on Automatic Control*, 47(3), 525–529.