A multi-observer approach for the state estimation of nonlinear systems

Romain Postoyan, Mohammed Hacene Adnane Hamid, Jamal Daafouz

A multi-observer approach for the state estimation of nonlinear systems

Romain Postoyan, Mohammed H.A. Hamid, and Jamal Daafouz

***Abstract*— We present a prescriptive approach for the state estimation of nonlinear systems. We first assume that we know a local observer, i.e. a dynamical system whose state converges to the plant state when it is initialized nearby the plant initial condition. We sample the set where the plant initial condition is assumed to lie with a finite number of points; noting that this set can be arbitrarily large. A local observer is initialized at each of these sampled points to form a bank of observers called multi-observer. A supervisor is then constructed to select one of these observers at any time instant. The selected state estimate is guaranteed to converge to the state of the plant, provided the number of samples is sufficiently large and a detectability property holds, which is expressed in terms of Lyapunov-based conditions. An explicit lower bound on the required number of observers is given when an estimate of the basin of convergence of the local observer is available. We explain how to apply the approach to nonlinear systems with globally Lipschitz and differentiable nonlinearities. Simulations results are presented for a Wilson-Cowan oscillator.**

I. Introduction

The design of global observers for nonlinear systems is a difficult task in general and available techniques apply to specific classes of systems. Local observers, on the other hand, are usually easier to construct. In this case, we are only interested in the behaviour of the estimation error around the origin. We can thus use first order approximations to analyze it, when the involved vector fields are sufficiently smooth. This leads to a time-varying dynamical system which is linear in the estimation error and hence easier to analyze in general. Various methods have been proposed in the literature to construct local observers, see e.g., [1], [3], [4], [12], [16], [19].

The main drawback of local observers is that they must be initialized nearby the initial condition of the system to work efficiently. However, we do not know the initial condition of the plant, it may therefore be difficult to guarantee this requirement in practice. The aim of this paper is to present a method to overcome this limitation. We consider general nonlinear systems, for which it is assumed that we know a local observer. Any of the techniques cited above can be applied at this step to synthesize the local observer. We assume that the initial condition of the plant lies in a known compact set $\mathcal{H}$. There is no restriction of the 'size' of this set, it can therefore be taken arbitrarily large if needed. We then sample the compact set with N points, and we initialize N versions of the same local observer at each of these points. The sampling of $\mathcal{H}$ has to be such that at least one observer has its initial condition nearby the initial condition of the plant by taking N sufficiently large. We then design a suitable criterion to select a single observer at any time instant, and we define the state estimate of the overall scheme to be the estimate provided by this observer. The convergence of the selected estimate is guaranteed, provided N is sufficiently large and a detectability condition for the state estimation error system holds, which is related to the concept of output-to-state stability [15].

The proposed technique is inspired by the supervisory control framework, see e.g., [6], [8], [9], [10]. Indeed, like in works on supervisory control, the scheme consists of a multi-estimator and a supervisor, which selects one of the observers at each time instant. However, none of the results on supervisory control applies to our problem as, on the one hand, their purpose is to stabilize the origin of the system, and, on the other hand, they typically assume the knowledge of a global estimator that converges to the trajectory of the system when there is no parametric uncertainty, which is very different from the scope of this paper. The supervisory approach has recently been extended for the estimation of the states and the parameters of nonlinear systems in [2]. It is assumed in [2] that a global state estimator can be constructed when the system parameters are known. A method is then proposed to also estimate the parameters on-line. We therefore address a different objective compared to [2], namely the state estimation of nonlinear systems, and we resort to different assumptions, techniques and analysis.

The presented approach is prescriptive. We explain how to apply it to a class of globally Lipschitz nonlinear systems as an example. We provide conditions using matrix inequalities, which guarantee the satisfaction of the assumptions needed for the estimation scheme to work. It has to be noted that the idea to use multiple versions of the same observer initialized at different values reminds of nonlinear filtering techniques, such as particle filters ([13]) or unscented Kalman filters ([7]) for instance, which are often applied in applications, but for which there is no convergence guarantee in general.

The paper is organised as follows. The estimation scheme is presented in Section II and the convergence guarantees are stated in Section III. We explain how to apply the approach to a class of nonlinear systems in Section IV, and we present simulations results for a Wilson-Cowan model [18] in Section V. Section VI concludes the paper. The proofs are omitted for space reasons.

Notation. Let $\mathbb{R} := (-\infty, \infty)$, $\mathbb{R}_{\geq 0} := [0, \infty)$, $\mathbb{R}_{>0} := (0, \infty)$, $\mathbb{Z} := \{\ldots, -2, -1, 0, 1, 2, \ldots\}$, and $\mathbb{Z}_{>0} :=$

The authors are with the Université de Lorraine, CRAN, UMR 7039 and the CNRS, CRAN, UMR 7039, France {mohammed.hamid,romain.postoyan,jamal.daafouz}@univ-lorraine.fr. J. Daafouz is also with the Institut Universitaire de France (IUF). This work was partially supported by the ANR under the grant SEPICOT (ANR 12 JS03 004 01).

$\{1, 2, \ldots\}$. For $(x, y) \in \mathbb{R}^{n+m}$, (x, y) stands for $[x^{\mathrm{T}}, y^{\mathrm{T}}]^{\mathrm{T}}$. The notation $\mathcal{L}_{\infty}^{\mathrm{pc}}$ denotes the set of piecewise continuous functions $f : \mathbb{R}_{\geq 0} \to \mathbb{R}^n$, $n \in \mathbb{Z}_{>0}$, such that $\|f\|_{\infty} := \sup_{\tau \geq 0} |f(\tau)| < r$, with $r \in \mathbb{R}_{\geq 0}$. Let $x \in \mathbb{R}^{n_x}$, $|x|$ and $|x|_{\infty}$ respectively stand for the Euclidean norm and the infinity norm of x. We use $\lceil x \rceil$ to denote $\min\{n \in \mathbb{Z} : x \leq n\}$ for $x \in \mathbb{R}$. Let $P \in \mathbb{R}^{n \times n}$ be a real, symmetric matrix, we respectively denote its maximum and minimum eigenvalues by $\lambda_{\max}(P)$ and $\lambda_{\min}(P)$. The notation $\mathbb{I}$ denotes the identity matrix, whose dimensions depend on the context. Let A be a real matrix, $|A|$ is the spectral norm of A, i.e. $|A| := \sqrt{\lambda_{\max}(A^{\mathrm{T}}A)}$. A function $\gamma : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ is of class $\mathcal{K}$ if it is continuous, zero at zero and strictly increasing, and it is of class $\mathcal{K}_{\infty}$, if, in addition, it is unbounded. A continuous function $\gamma : \mathbb{R}^2_{\geq 0} \longrightarrow \mathbb{R}_{\geq 0}$ is of class $\mathcal{KL}$ if for each $t \in \mathbb{R}_{\geq 0}$, $\gamma(\cdot, t)$ is of class $\mathcal{K}$, and, for each $s \in \mathbb{R}_{\geq 0}$, $\gamma(s, \cdot)$ is decreasing to zero.

II. Estimation scheme

The estimation scheme is inspired by works on supervisory control (e.g., [6], [8], [9], [10]). It consists of a bank of local observers, called *multi-observer*, and a *supervisor*, as depicted in Figure 1. We present each of these components in this section.

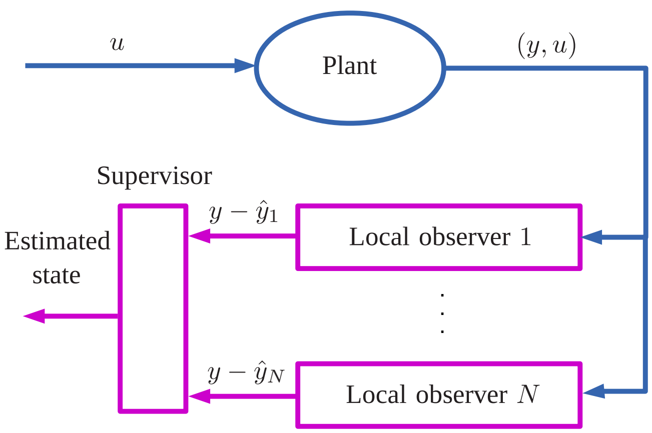

Fig. 1. Estimation scheme.

A. Local observer

Consider the system

$$\begin{array}{rcl} \dot{x} & = & f(x, u) \\ y & = & h(x), \end{array} \tag{1}$$

where $x \in \mathbb{R}^{n_x}$ is the state, $u \in \mathbb{R}^{n_u}$ is a vector of known inputs, $y \in \mathbb{R}^{n_y}$ is the measured output, and $n_x, n_u, n_y \in \mathbb{Z}_{>0}$. The objective is to estimate the state x. We require for this purpose that system (1) verifies the conditions below.

Assumption 1: The following holds.

(i) The initial condition to (1) lies in a known compact set $\mathcal{H}$.

(ii) For any initial condition $x_0 \in \mathcal{H}$, and any input $u \in \mathcal{L}_{\infty}^{\mathrm{pc}}$, system (1) generates a unique solution, which is defined for all positive times.

(iii) The mapping h is continuous. □

Item (i) of Assumption 1 requires some prior knowledge of the initial condition of system (1), which is the case for all stochastic Kalman filtering techniques for instance. It has to be emphasized that there is no restriction on the 'size' of the set $\mathcal{H}$, as it can be taken arbitrarily large. That implies that, when the solutions to (1) are known to take value in a compact operating region, we can define $\mathcal{H}$ as being this region for example. Items (ii)-(iii) of Assumption 1 are reasonable as these hold for many physical and biological systems.

We assume that we know a local observer for system (1) of the form

$$\begin{array}{rcl} \dot{\hat{x}} & = & \hat{f}(\hat{x}, u, y) \\ \hat{y} & = & h(\hat{x}), \end{array} \tag{2}$$

where $\hat{x} \in \mathbb{R}^{n_x}$ is the estimated state, $\hat{y} \in \mathbb{R}^{n_y}$ is the estimated output. In particular, we make the following assumption.

Assumption 2: The following holds.

(i) There exists a neighborhood $\mathcal{B}$ of the origin of $\mathbb{R}^{n_x}$ such that for any initial conditions $x_0 \in \mathcal{H}$ and $\hat{x}_0 \in \mathbb{R}^{n_x}$ of, respectively, systems (1) and (2), for any $u \in \mathcal{L}_{\infty}^{\mathrm{pc}}$, if $x_0 - \hat{x}_0 \in \mathcal{B}$, then the corresponding solutions to (1) and (2) are such that $|x(t) - \hat{x}(t)| \to 0$ as $t \to \infty$.

(ii) For any initial condition $\hat{x}_0 \in \mathbb{R}^{n_x}$, for any $u, y \in \mathcal{L}_{\infty}^{\mathrm{pc}}$, the corresponding solution to (2) is unique and is defined for all positive times. □

Item (i) of Assumption 2 means that the state of a solution to system (2) converges to the state of the solution to system (1) provided their initial conditions are sufficiently close to each other. There exists several methods in the literature to design local observers, i.e. to ensure Assumption 2, see e.g., [1], [3], [4], [12], [16], [19]. We can use any of these methods to construct system (2) at this step. Item (ii) of Assumption 2 is a global property of existence and uniqueness of solutions for system (2).

B. Multi-observer

System (2) is guaranteed to provide 'good' estimates of the state x only when the initial state estimation error is sufficiently small. The problem is that we do not know the initial condition of system (1), which we denote x_0. However, we know that $x_0 \in \mathcal{H}$ according to item (i) of Assumption 1. The idea is to sample the set $\mathcal{H}$ with N points and to initialize N local observers (2) at each of these points. We thus obtain a bank of N observers, which forms the multi-observer. We denote the state of the i^{th} observer as $\hat{x}^i$ and its output as $\hat{y}^i$, $i \in \{1, \ldots, N\}$.

We suppose that the sampling of the set $\mathcal{H}$ is such that the distance between x_0 and the closest sampled point tends to zero as N tends to infinity. In other words, denoting $\hat{x}_0^i$ the initial condition of the i^{th} observer, for $i \in \{1, \ldots, N\}$,

$$\min_{i \in \{1, \ldots, N\}} |x_0 - \hat{x}_0^i| \to 0 \text{ as } N \to \infty. \tag{3}$$

Property (3) can always be ensured by uniformly sampling $\mathcal{H}$. In that way, by selecting N sufficiently large, we are sure that, at least, one observer will be initialized nearby x_0 and that its state will converge to the state of the solution to (1), according to Assumption 2.

It is possible to provide an estimate of the required number of observers N when we have prior knowledge on the basin of attraction of the origin for the estimation error system, i.e. the set $\mathcal{B}$ in Assumption 2, as stated in the lemma below.

Lemma 1: Suppose the following holds.

(i) Item (i) of Assumption 1 is verified with $\mathcal{H}$ an hypercube of edge length $\Omega > 0$.
(ii) There exists $\widehat{\mathcal{B}} := \{\tilde{x} \in \mathbb{R}^{n_x} : |\tilde{x}|_\infty \leq \ell\}$, with known $\ell > 0$, such that $\widehat{\mathcal{B}} \subseteq \mathcal{B}$.

Consider a uniform sampling of $\mathcal{H}$ with step $\delta = \Omega \lceil \frac{\Omega}{2\ell} \rceil^{-1}$. There exists $i \in \{1, \ldots, N\}$ such that $x_0 - \hat{x}_0^i \in \mathcal{B}$ and $N = (\frac{\Omega}{\delta} + 1)^{n_x}$. □

The compact set $\mathcal{H}$ in item (i) of Assumption 1 can always be taken as an hypercube (since we assume that x_0 lies in a known *compact* set, see item (i) of Assumption 1), so that item (i) of Lemma 1 holds. Item (ii) of Lemma 1 requires the knowledge of an estimate of the basin of attraction of the origin for the estimation error system, namely $\widehat{\mathcal{B}}$. The latter is typically obtained from the analysis used to prove that system (2) is local observer for system (1) (see for instance Proposition 3 in [1]). Once the constants Ω and ℓ are derived, we deduce the number of local observers N by using the formula in Lemma 1.

C. *Supervisor*

Now that we have designed the multi-observer, we need a criterion to select one of the local observers at any time instant. To this purpose, we introduce the *monitoring signals* $\mu_i \in \mathbb{R}$, $i \in \{1, \ldots, N\}$, whose dynamics are given by (like in e.g., [2], [17])

$$\dot{\mu}_i = -\lambda \mu_i + \gamma(y, \hat{y}^i) \qquad \mu_i(0) = 0, \tag{4}$$

where $\lambda > 0$ is a design parameter, and $\gamma : \mathbb{R}^{2n_y} \to \mathbb{R}_{\geq 0}$ is a continuous function such that $\gamma(y, y) = 0$ for any $y \in \mathbb{R}^{n_y}$, which we design in the sequel. The idea is to filter $\gamma(y, \hat{y}^i)$ and to use the obtained signal to select one of the local observers. The selection variable $\sigma \in \{1, \ldots, N\}$ is defined as

$$\sigma(t) := \text{argmin } \mu_i(t) \qquad \forall t \geq 0. \tag{5}$$

At any time $t \geq 0$, the supervisor selects the local observer with the smallest monitoring signal (if several observers provide the same minimal value of μ_i, an arbitrary prefixed order is used). In that way, the state estimate is given by $\hat{x}^\sigma$. The signal σ may experience rapid switches, but these do not affect the dynamics of the local observers. If that is an issue for the implementation, hysteresis based switching ([5]) can be used on top of (5) to avoid it. The results derived hereafter only require minor modifications to apply in this case.

III. Convergence analysis

A. *A detectability assumption*

By taking N sufficiently large, we know from (3) that at least one local observer, whose index is denoted by $i^\star \in \{1, \ldots, N\}$, will be such that $x_0 - \hat{x}_0^{i^\star} \in \mathcal{B}$. Consequently, the state estimation error $x - \hat{x}^{i^\star}$ will converge to zero (according to Assumption 2), and so will $y - \hat{y}^{i^\star}$ by continuity of h (see item (iii) of Assumption 1). The latter implies that $\mu_{i^\star}$ converges to zero, in view of (4). We would therefore expect that σ selects observer $i^\star$ after some time. The issue is that other observers may generate output estimation errors converging to zero (and so, monitoring signals converging to zero), while the state estimation errors do not, see the example below. We may then not be able to distinguish these observers from the 'right' observers.

Example 1: Consider the system $\dot{x}_1 = x_2$, $\dot{x}_2 = -x_2 - 2(x_1^3 - x_1)$ and $y = x_2$. This system has three equilibrium points: $(0, 0)$, which is unstable, and $(-1, 0)$ and $(1, 0)$, which are both locally exponentially stable. We denote the basin of attraction of the latter by $\mathcal{A}_{(-1,0)}$ and $\mathcal{A}_{(1,0)}$, respectively. Assume that the initial condition x_0 lies in a compact set $\mathcal{H} \subset \mathcal{A}_{(1,0)}$. Consequently, Assumption 1 holds. A copy of the system is a local observer in this case, and Assumption 2 is verified with $\mathcal{B} = \{\tilde{x} \in \mathbb{R}^2 : |\tilde{x}| \leq \varepsilon\}$ for $\varepsilon > 0$ sufficiently small such that $x_0 \in \mathcal{H}$ and $x_0 - \hat{x}_0 \in \mathcal{B}$ implies $\hat{x}_0 \in \mathcal{A}_{(1,0)}$. Take the initial condition of the local observer $\hat{x}_0$ in $\mathcal{A}_{(-1,0)}$. The output estimation error $\tilde{y} = x_2 - \hat{x}_2$ converges to zero as so do x_2 and $\hat{x}_2$ because $x_0 \in \mathcal{A}_{(1,0)}$ and $\hat{x}_0 \in \mathcal{A}_{(-1,0)}$, respectively. However, the state estimation error $\tilde{x} := x - \hat{x}$ does not converge to zero because x_1 and $\hat{x}_1$ respectively converge to 1 and -1. ■

We need an additional condition to guarantee that, if $y - \hat{y}^i$ converges to zero, then so does $x - \hat{x}^i$, for $i \in \{1, \ldots, N\}$. This is the purpose of the assumption below.

Assumption 3: There exist a continuously differentiable function $V : \mathbb{R}^{2n_x} \to \mathbb{R}_{\geq 0}$, $\underline{\alpha}_V, \overline{\alpha}_V \in \mathcal{K}_\infty$, $\theta > 0$, a continuous function $\gamma : \mathbb{R}^{2n_y} \to \mathbb{R}_{\geq 0}$ with $\gamma(y, y) = 0$ for any $y \in \mathbb{R}^{n_y}$, such that for any $x, \hat{x} \in \mathbb{R}^{n_x}$

$$\underline{\alpha}_V(|x - \hat{x}|) \leq V(x, \hat{x}) \leq \overline{\alpha}_V(|x - \hat{x}|), \tag{6}$$

and for any $x, \hat{x} \in \mathbb{R}^{n_x}$ and $u \in \mathbb{R}^{n_u}$

$$\begin{aligned} &\Big\langle \nabla V(x, \hat{x}), (f(x, u), \hat{f}(\hat{x}, u, h(x))) \Big\rangle \\ &\qquad\qquad \leq -\theta V(x, \hat{x}) + \gamma(h(x), h(\hat{x})). \end{aligned} \tag{7}$$

□

Assumption 3 is a *global* detectability property of the estimation error system (and not of system (1)) with respect to the output $y - \hat{y}$. It covers the situation where the estimation error system is output-to-state stable ([15]) as a special case for which $\gamma(h(x), h(\hat{x})) = \chi(|h(x) - h(\hat{x})|)$ with $\chi \in \mathcal{K}_\infty$. It is the function γ in (7) that is implemented in (4) to generate the monitoring signals. In Section IV, we explain how to verify Assumption 3 for a class of nonlinear systems. The example below shows that the system in Example 1 satisfies Assumption 3 when it is equipped with a different output map.

Example 2: Consider the same system as in Example 1 but with[1] $y = x_1$ (and not $y = x_2$). The local observer (2) is still given by a copy of the plant's dynamics. Let $V(x, \hat{x}) = \frac{1}{2}(x_1 - \hat{x}_1)^2 + \frac{1}{2}(x_2 - \hat{x}_2)^2$ for $x_1, x_2, \hat{x}_1, \hat{x}_2 \in \mathbb{R}$, hence (6) holds with $\underline{\alpha}_V(s) = \overline{\alpha}_V(s) = \frac{1}{2}s^2$ for $s \in \mathbb{R}_{\geq 0}$. For any $x_1, x_2, \hat{x}_1, \hat{x}_2 \in \mathbb{R}$,

$$\langle \nabla V(x, \hat{x}), F(x, u, \hat{x}) \rangle = -\tilde{x}_2^2 + 3\tilde{x}_1\tilde{x}_2 - 2\tilde{x}_2(x_1^3 - \hat{x}_1^3), \tag{8}$$

where $F(x, u, \hat{x}) = (x_2, -x_2 - 2(x_1^3 - x_1), \hat{x}_1, -\hat{x}_2 - 2(\hat{x}_1^3 - \hat{x}_1))$, $\tilde{x}_1 = x_1 - \hat{x}_1$, and $\tilde{x}_2 = x_2 - \hat{x}_2$. Using[2] $3\tilde{x}_1\tilde{x}_2 \leq 9\tilde{x}_1^2 + \frac{1}{4}\tilde{x}_2^2$ and $2\tilde{x}_2(x_1^3 - \hat{x}_1^3) \leq \frac{1}{2}\tilde{x}_2^2 + 2(x_1^3 - \hat{x}_1^3)^2$, we derive

$$\begin{aligned} \langle \nabla V(x, \hat{x}), F(x, u, \hat{x}) \rangle &\leq -\tfrac{1}{4}\tilde{x}_2^2 + 9\tilde{x}_1^2 + 2(x_1^3 - \hat{x}_1^3)^2 \\ &= -\tfrac{1}{4}\tilde{x}_2^2 + 9\tilde{x}_1^2 + 2(x_1^3 - \hat{x}_1^3)^2 \\ &\quad -\tfrac{1}{4}\tilde{x}_1^2 + \tfrac{1}{4}\tilde{x}_1^2. \end{aligned} \tag{9}$$

Consequently, (7) is verified with $\theta = \frac{1}{2}$ and $\gamma(x_1, \hat{x}_1) = 2(x_1^3 - \hat{x}_1^3)^2 + (9 + \frac{1}{4})(x_1 - \hat{x}_1)^2$, which verifies the required properties. ■

B. Main result

We are ready to state the main result.

Theorem 1: Consider system (1), the estimation scheme described in Section II, and suppose Assumptions 1-3 hold. There exists $N^\star \in \mathbb{Z}_{>0}$ such that for any $N \geq N^\star$, $\lambda \leq \theta$ where θ comes from Assumption 3, $x_0 \in \mathcal{H}$, $u \in \mathcal{L}_\infty^{\text{pc}}$, $|x(t) - \hat{x}^{\sigma(t)}(t)| \to 0$ as $t \to \infty$. □

Theorem 1 ensures that $x - \hat{x}^\sigma$ converges to zero when the number of local observers N is sufficiently large. We recall that Lemma 1 provides an estimate of $N^\star$, when an estimate of the basin of attraction $\mathcal{B}$ in Assumption 2 is available. The influence of λ on the convergence of $|x - \hat{x}^\sigma|$ is studied in simulations in Section V.

C. Discussions

The proposed estimation scheme may be computationally demanding when $N^\star$ is large in Theorem 1. When we have prior quantitative knowledge on the convergence of $y - \hat{y}^i$ for the local observers which are initialized nearby x_0, we can use this information to reduce the number of local observers in the multi-observer, and thus reduce the computational cost. To be more precise, suppose that there exists a known function $\beta \in \mathcal{KL}$ such that for any local observer with $x_0 - \hat{x}_0^i \in \mathcal{B}$, and $u \in \mathcal{L}_\infty^{\text{pc}}$, the corresponding solutions to (1) and (2) verify

$$|x(t) - \hat{x}^i(t)| \leq \beta(|x_0 - \hat{x}_0^i|, t) \qquad \forall t \geq 0. \tag{10}$$

When $y = Cx$, with C is real matrix, this implies $|y(t) - \hat{y}^i(t)| \leq |C|\beta(|x_0 - \hat{x}_0^i|, t)$ for any $t \geq 0$. Suppose that we know an estimate of the basin of attraction $\mathcal{B}$ in Assumption 2 of the form $\widehat{\mathcal{B}}' = \{\tilde{x} \in \mathbb{R}^{n_x} : |\tilde{x}| \leq \ell'\}$ with $\ell' > 0$. By taking $N^\star$ sufficiently large, there exists $i^\star \in \{1, \ldots, N\}$, with any $N \geq N^\star$, such that $x_0 - \hat{x}_0^{i^\star} \in \widehat{\mathcal{B}}'$ in view of (3). Hence, we only have to run the local observers $i \in \{1, \ldots, N\}$, which verify

$$|y(t) - \hat{y}^i(t)| \leq |C|\beta(\ell', t) \qquad \forall t \geq 0, \tag{11}$$

and not the others. Condition (11) can be evaluated on-line as it only depends on known quantities.

IV. Application

We study the case where system (1) has the form below

$$\begin{aligned} \dot{x} &= Ax + \phi(x, u) \\ y &= Cx, \end{aligned} \tag{12}$$

where A and C are real matrices, and ϕ is globally Lipschitz and differentiable on $\mathbb{R}^{n_x + n_u}$. The local observer (2) is taken as

$$\dot{\hat{x}} = A\hat{x} + \phi(\hat{x}, u) + K(\hat{x}, u)C\tilde{x}, \tag{13}$$

where $K(\hat{x}, u)$ is a correction matrix, and $\tilde{x} = x - \hat{x}$. We assume that $K(\hat{x}, u)$ is designed such that the condition below holds.

Assumption 4: There exist real, symmetric, positive definite matrices P and Q such that, for any $\hat{x} \in \mathbb{R}^{n_x}$ and $u \in \mathbb{R}^{n_u}$,

$$\begin{aligned} &(\mathcal{A}(\hat{x}, u) - K(\hat{x}, u)C)^{\mathrm{T}} P + P(\mathcal{A}(\hat{x}, u) - K(\hat{x}, u)C) \\ &\qquad\qquad \leq -Q, \end{aligned} \tag{14}$$

where $\mathcal{A}(\hat{x}, u) := A + \left.\frac{\mathrm{d}\phi(x,u)}{\mathrm{d}x}\right|_{\hat{x}}$. In addition, item (ii) of Assumption 2 is verified. □

The inequality (14) is a priori difficult to verify as it needs to hold for any $\hat{x} \in \mathbb{R}^{n_x}$ and $u \in \mathbb{R}^{n_u}$. To overcome this issue, linear matrix inequalities have been proposed in[3] [3], which may allow designing the nonlinear gain $K(\hat{x}, u)$ so that (14) holds. We note that the gain $K(\hat{x}, u)$ constructed in [3] is such that item (ii) of Assumption 2 is verified. We will see below that the assumptions made so far ensure the satisfaction of Assumptions 1-2 (only item (i) of Assumption 1 will need to be explicitly assumed).

We need to introduce some notations before stating the main result of this section. Let $x, \hat{x} \in \mathbb{R}^{n_x}$ and $u \in \mathbb{R}^{n_u}$, we first write $\left(\left.\frac{\mathrm{d}\phi(x,u)}{\mathrm{d}x}\right|_x - \left.\frac{\mathrm{d}\phi(x,u)}{\mathrm{d}x}\right|_{\hat{x}}\right)\tilde{x}$ as

$$\begin{aligned} \left(\left.\frac{\mathrm{d}\phi(x,u)}{\mathrm{d}x}\right|_x - \left.\frac{\mathrm{d}\phi(x,u)}{\mathrm{d}x}\right|_{\hat{x}}\right)\tilde{x} &= \phi_1(x, \hat{x}, u)\tilde{x} \\ &\quad + \phi_2(x, \hat{x}, u)C\tilde{x}, \end{aligned} \tag{15}$$

where ϕ_1 and ϕ_2 are such that $|\phi_i(x, \hat{x}, u)| \leq m_i$ with $m_i \geq 0$ for any $x, \hat{x}, u$ and $i \in \{1, 2\}$. This is always possible as we can take $\phi_1(x, \hat{x}, u) = \left.\frac{\mathrm{d}\phi(x,u)}{\mathrm{d}x}\right|_x - \left.\frac{\mathrm{d}\phi(x,u)}{\mathrm{d}x}\right|_{\hat{x}}$ and $\phi_2(x, \hat{x}, u) = 0$, noting that ϕ_1 is bounded since ϕ is globally Lipschitz (and differentiable). We further write $\phi_1(x, \hat{x}, u)$ as

$$\phi_1(x, \hat{x}, u) = M\Delta(x, \hat{x}, u)R, \tag{16}$$

[1] In this case, it is easy to construct a global observer since the system is linear up to an output injection term. Nevertheless, the objective of this example is only to demonstrate that Assumption 3 holds.

[2] These inequalities are obtained by applying the formula $2ab \leq \epsilon a^2 + \frac{1}{\epsilon}b^2$ for any $a, b \in \mathbb{R}$ and $\epsilon \in \mathbb{R}_{>0}$.

[3] The results in [3] are written for systems of the form of (12) with ϕ which only depends on x. They also apply when ϕ depends on both x and u.

with $\Delta(x,\hat{x},u)$ diagonal and $|\Delta(x,\hat{x},u)| \leq 1$, which is always possible since ϕ_1 is bounded.

The proposition below provides conditions tailored to systems (12) and (13) under which the results of Sections II-III can be applied.

Proposition 1: Consider systems (12) and (13) and suppose the following holds.

(i) Item (i) of Assumption 1 is verified.

(ii) Assumption 4 is verified.

(iii) There exist $\eta > 0$ and $\nu \geq 0$ such that

$$\begin{bmatrix} -Q + \eta R^{\mathrm{T}} R - \nu C^{\mathrm{T}} C & PM \\ M^{\mathrm{T}} P & -\eta \mathbb{I} \end{bmatrix} < 0, \quad (17)$$

where P and Q come from Assumption 4, and M and R are defined in (16).

Then Assumptions 1-3 hold and the conclusion of Theorem 1 applies. □

V. Illustrative example

In this section, we apply the results of Section IV to the following Wilson-Cowan oscillator

$$\begin{array}{rcl} \dot{x}_1 & = & -d_1 x_1 - d_2 x_2 + S(C_1 x_1 + C_2 x_2 + u) \\ \dot{x}_2 & = & -d_3 x_2 - d_4 x_1 \\ \dot{x}_3 & = & -f_0 x_3 + x_1 \\ y & = & -f_0 x_3 + x_1, \end{array} \quad (18)$$

where $x_1, x_2, x_3 \in \mathbb{R}$ is the state, $u \in \mathbb{R}$ is the input, $y \in \mathbb{R}$ is the output, $d_1, d_2, d_3, d_4, f_0 > 0$, $d_4 < 0$ and $C_1, C_2 \in \mathbb{R}$ are parameters. The function S is a sigmoid, which we define as $S(z) = \arctan(z) + \frac{\pi}{2}$ for $z \in \mathbb{R}$. The (x_1, x_2)-system is similar to system (1) in [14] for instance. Wilson-Cowan oscillators are commonly used in computational neurosciences to investigate neurophysiological phenomena. The x_3-system is introduced to generate an output signal y, which is centered at 0. This is justified by the fact that the electrophysiological signals recorded in practice often have a zero offset; this is usually the case with electroencephalography for example (see Section II in [3]).

System (18) can be written as (12) with

$$\begin{array}{rcl} A & = & \begin{bmatrix} -d_1 & -d_2 & 0 \\ -d_4 & -d_3 & 0 \\ 1 & 0 & -f_0 \end{bmatrix} \\ \phi(x,u) & = & [S(C_1 x_1 + C_2 x_2 + u) \quad 0 \quad 0]^{\mathrm{T}} \\ C & = & [1 \quad 0 \quad -f_0]\,. \end{array} \quad (19)$$

We note that ϕ is globally Lipschitz and differentiable as required.

We consider the local observer (13). We design the gain $K(\hat{x},u)$ by applying the technique in [3]. We have taken the following parameters values: $d_1 = d_2 = d_3 = d_4 = 10$, $C_1 = 100$, $C_2 = -100$, $f_0 = 10$. The obtained gain is, for $\hat{x}, u \in \mathbb{R}$,

$$\begin{array}{rcl} K(\hat{x},u) & = & \frac{1}{2}\left(1 - S'(C_1 \hat{x}_1 + C_2 \hat{x}_2 + u)\right) K_1 \\ & & + \frac{1}{2}\left(1 + S'(C_1 \hat{x}_1 + C_2 \hat{x}_2 + u)\right) K_2, \end{array} \quad (20)$$

where $S'(z) = \frac{1}{1+z^2}$ for $z \in \mathbb{R}$, and

$$\begin{array}{rcl} K_1 & := & [433.32 \quad -87.00 \quad -7.06] \\ K_2 & := & [424.33 \quad -89.04 \quad -7.30]. \end{array} \quad (21)$$

Thus, Assumption 4 holds with

$$P = \begin{bmatrix} 0.0184 & -0.0035 & -0.0976 \\ -0.0035 & 1.1741 & -11.7681 \\ -11.7681 & -0.0976 & 122.2508 \end{bmatrix}. \quad (22)$$

To evaluate (17), we need to know R and M. For this purpose, we define the functions ϕ_1 and ϕ_2 in (15) as $\phi_1(x,\hat{x},u) = \begin{pmatrix} 0 & C_2\Delta(x,\hat{x},u) & C_1 f_0 \Delta(x,\hat{x},u) \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$ and $\phi_2(x,\hat{x},u) = \begin{pmatrix} C_1 \Delta(x,\hat{x},u) \\ 0 \\ 0 \end{pmatrix}$ where

$$\begin{array}{rcl} \Delta(x,\hat{x},u) & = & S'(C_1 x_1 + C_2 x_2 + u) \\ & & -S'(C_1 \hat{x}_1 + C_2 \hat{x}_2 + u), \end{array} \quad (23)$$

with $x, \hat{x}, u \in \mathbb{R}$. We derive that $M = [1 \quad 0 \quad 0]^{\mathrm{T}}$ and $R = [0 \quad C_2 \quad C_1 f_0]$ in (16). We have verified that LMI (17) holds and we have obtained $\theta = 0.11$ in (7). We assume that the initial condition x_0 to system (18) lies in $\mathcal{H} = [-10, 10]^3$. Hence, all the conditions of Proposition 1 are satisfied.

We have uniformly sampled the hyperrectangle $\mathcal{H}$ with $N = 27$ points. The input u is given by a continuous Gaussian signal with mean 3 and variance 10, and the initial condition to (18) is $x_0 = (-5, 5, 9)$. The value of λ is set to 0.1. Figure 2 confirms that $\hat{x}^{\sigma}$ converges to the state of system (18). The signal σ is plotted in Figure 3, which shows that the switches only occur during the first instants and a single observer is selected afterwards. Simulations have revealed that no significant difference appears when taking smaller values of λ.

VI. Conclusions

We have presented an estimation scheme for nonlinear systems, which consists of a bank of N identical local observers initialized at different values, and a supervisor which selects one of these observers at any time instant. The convergence of the estimated state to the true state of the plant is guaranteed for N sufficiently large, provided a detectability condition holds. An explicit lower bound on N is provided when we know an estimate of the basin of attraction of the origin for the estimation error system. It is also explained how to reduce the potential computation burden of the scheme.

In future work, we will extend the presented results to nonlinear systems affected by exogenous disturbances and measurement noises. The multi-observer approach adopted in this paper might also be useful to improve the performance of existing observers in terms of speed of convergence, in the same spirit as what is done for the control of nonlinear systems with the united controllers technique, see [11] for instance. Contrary to the control problem, we can directly initialize one observer nearby the plant by sufficiently sampling

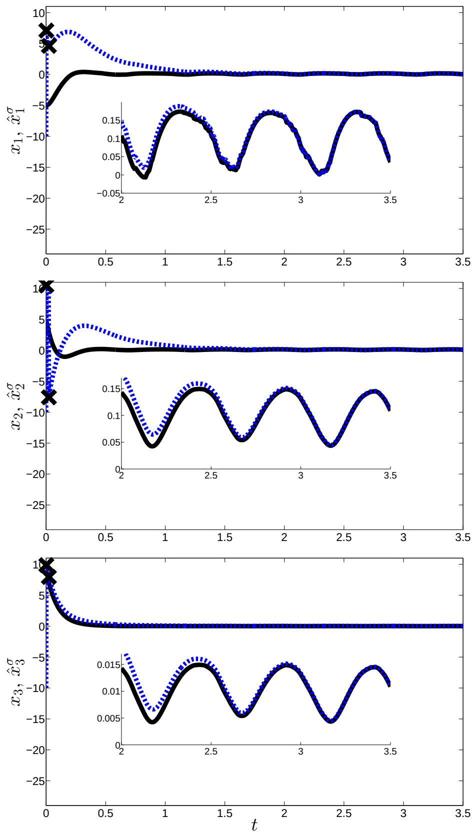

Fig. 2. State x (solid black line) and $\hat{x}^\sigma$ (dashed blue lines).

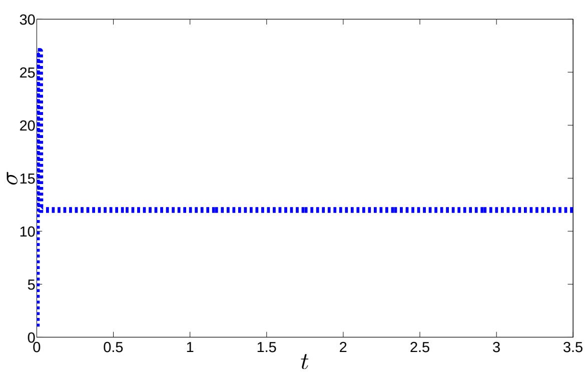

Fig. 3. Variations of σ.

the set where the plant initial condition is. The challenge is then to find an appropriate criterion to select this observer (or one which is initialized nearby), as it is expected to converge faster than those initialized further. Arguments like those invoked in Section III-C should probably be used to solve this problem.

References

[1] S. Bonnabel and J.-J. Slotine. A contraction theory-based analysis of the stability of the deterministic extended Kalman filter. *IEEE Transactions on Automatic Control*, 60(2):565–569, 2015.

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

[3] M.H.A. Hamid, R. Postoyan, and J. Daafouz. Local observers design for a class of neural mass models. In *European Control Conference, Linz, Austria*, 2015 (accepted).

[4] H. Hammouri, H.G. Vu, and H. Yahoui. Local observer for infinitesimally observable nonlinear systems. *International Journal of Control*, 86(4):579–590, 2013.

[5] J. Hespanha, D. Liberzon, A.S. Morse, B.D.O. Anderson, T.S. Brinsmead, and F. De Bruyne. Multiple model adaptive control. Part 2: switching. *International Journal of Robust and Nonlinear Control*, 11(5):479–496, 2001.

[6] J.P. Hespanha, D. Liberzon, and A.S. Morse. Hysteresis-based switching algorithms for supervisory control of uncertain systems. *Automatica*, 39(2):263–272, 2003.

[7] S.J. Julier and J.K. Uhlmann. Unscented filtering and nonlinear estimation. *Proceedings of the IEEE*, 92(3):401–422, 2004.

[8] D. Liberzon. *Switching in systems and control*. Springer, 2003.

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

[10] C. De Persis, R. De Santis, and A.S. Morse. Supervisory control with state-dependent dwell-time logic and constraints. *Automatica*, 40(2):269–275, 2004.

[11] C. Prieur. Uniting local and global controllers with robustness to vanishing noise. *Mathematics of Control, Signals and Systems*, 14(2):143–172, 2001.

[12] K. Reif, F. Sonnemann, and R. Unbehauen. An EKF-based nonlinear observer with a prescribed degree of stability. *Automatica*, 34(9):1119–1123, 1998.

[13] B. Ristic, S. Arulampalam, and N. Gordon. *Beyond the Kalman filter: Particle filters for tracking applications*, volume 685. Artech house Boston, 2004.

[14] T.D. Sauer and S.J. Schiff. Data assimilation for heterogeneous networks: the consensus set. *Physical Review E*, 79(5):051909, 2009.

[15] E.D. Sontag and Y. Wang. Output-to-state stability and detectability of nonlinear systems. *Systems & Control Letters*, 29(5):279–290, 1997.

[16] V. Sundarapandian. Local observer design for nonlinear systems. *Mathematical and computer modelling*, 35(1):25–36, 2002.

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

[18] H.R. Wilson and J.D. Cowan. Excitatory and inhibitory interactions in localized populations of model neurons. *Biophysical journal*, 12(1):1–24, 1972.

[19] X.H. Xia and W.-B. Gao. On exponential observers for nonlinear systems. *Systems & Control Letters*, 11(4):319–325, 1988.