HAL
open science

Local adaptive observer for linear time-varying systems with parameter-dependent state matrices

Romain Postoyan, Qinghua Zhang

To cite this version:

Romain Postoyan, Qinghua Zhang. Local adaptive observer for linear time-varying systems with parameter-dependent state matrices. 2018. hal-01869339

HAL Id: hal-01869339

https://hal.science/hal-01869339v1

Preprint submitted on 6 Sep 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Local adaptive observer for linear time-varying systems with parameter-dependent state matrices

Romain Postoyan and Qinghua Zhang

***Abstract*— We present an adaptive observer for linear time-varying systems whose state matrix depends on unknown parameters. We first assume that the state matrix is affine in these parameters. In this case, the proposed observer generates state and parameter estimates, which exponentially converge to the plant state and the true parameter, respectively, provided a persistence of excitation condition holds and the unknown parameters lie in a neighborhood of some known nominal values. Hence, some prior knowledge on the unknown parameters is required, but not on the state. We then modify the adaptive observer and its convergence analysis to systems whose state matrix is smooth, instead of being affine, in the unknown parameters. The convergence is approximate, and no longer exponential, in this case. An example is provided to illustrate the results, for which the required distance between the unknown parameters and their nominal values is investigated numerically.**

I. Introduction

The objective of this work is to estimate both the state and the parameters of the system

$$\begin{array}{rcl} \dot{x}(t) & = & A(t,\theta)x(t) + B(t)u(t) \\ y(t) & = & C(t)x(t), \end{array} \tag{1}$$

where[1] $x(t) \in \mathbb{R}^{n_x}$ is the state, $u(t) \in \mathbb{R}^{n_u}$ is the input, $y(t) \in \mathbb{R}^{n_y}$ is the output, $\theta \in \mathbb{R}^{n_\theta}$ is the vector of unknown constant parameters, $t \geq 0$ is the time, and $n_x, n_u, n_y, n_\theta \in \mathbb{Z}_{>0}$. In other words, we aim at designing an adaptive observer for system (1). This problem arises in many practical applications, including electrochemical batteries [13] or mechanical systems [9] to cite a couple of examples, as soon as we need to estimate on-line both the unmeasured state and unknown parameters of a plant for monitoring purposes for instance.

In the majority of works on adaptive observers, the state matrix is assumed to be fully known contrary to (1), and the unknown parameters enter in the system through an additional term of the form $\Phi(t, u(t), y(t))\theta$ where Φ is known. Solutions for this case can be found in e.g., [3], [4], [20], [28]–[30]. The fact that A depends on θ in (1) is a major difference with these references, which makes the problem difficult because of the cross-terms involving θ and $x(t)$. There are nevertheless results in the literature, which propose adaptive observers for systems with partially unknown state matrix A. In [17], several designs are given for single-input single-output linear time-invariant (LTI) systems. In [18], it is explained how to optimally fit a discrete-time LTI model to any given output signals. In [21], a switched approach is proposed for LTI systems based on linear matrix inequalities (LMI) conditions. More recently, a generic approach for the state and parameter estimation of nonlinear systems was presented in [8] using a multi-observer set-up. On the other hand, adaptive observers for systems with known state matrices with an additive term of the form $\Phi(x)\theta$ (see e.g., [5], [7], [11], [12]) are relevant in this context when A is affine in θ, as a simple manipulation brings system (1) to the right form.

In this work, we propose an approach, which relies on a simple idea. We essentially assume that we know a state-observer for system (1) for some known nominal parameter value θ_0. Because $\theta \neq \theta_0$, the idea is to modify the original state-observer and design an adaptation law to estimate θ by relying on similar techniques as in [30]. Indeed, by adding and subtracting $A(t,\theta_0)x(t)$ in the first equation in (1), we obtain $\dot{x}(t) = A(t,\theta_0)x(t) + B(t)u(t) + [A(t,\theta) - A(t,\theta_0)]\,x(t)$. Assuming $A(t,\theta)$ is affine in θ we can write $A(t,\theta)x(t) - A(t,\theta_0)x(t) = \Lambda(t,x(t))(\theta - \theta_0)$, where $\Lambda(t,x(t))$ is a time-varying matrix linear in its second argument; this is explained in more detailed in the sequel. Hence, $\dot{x}(t) = A(t,\theta_0)x(t) + B(t)u(t) + \Lambda(t,x(t))(\theta - \theta_0)$. Denoting by $\hat{x}$ the state estimate of the adaptive observer to be designed and $\tilde{x} = x - \hat{x}$ the associated state estimation error, we derive $\dot{x}(t) = A(t,\theta_0)x(t) + B(t)u(t) + \Lambda(t,\hat{x}(t))(\theta - \theta_0) + \Lambda(t,\tilde{x}(t))(\theta - \theta_0)$, since Λ is linear in its second argument. Now $\Lambda(t,\hat{x}(t))$ is known since $\hat{x}$ is the state estimate we generate. The term $\Lambda(t,\hat{x}(t))(\theta - \theta_0)$ is therefore of the form $\Phi(t)(\theta - \theta_0)$ where $\Phi(t)$ is known. We are almost back to the same class of systems as in [30] where the unknown parameter is $\theta - \theta_0$ instead of θ, but this is fine since we know θ_0. The difference with [30] is that we have to deal with the extra term $\Lambda(t,\tilde{x}(t))(\theta - \theta_0)$. It appears that when $||\theta - \theta_0||$ is small, so is $\Lambda(t,\tilde{x}(t))(\theta - \theta_0)\tilde{x}(t)$ so that stability of the estimation error may be ensured. The point of these developments is to show that after some simple manipulations, we can write the problem in a similar, though different form, as in [30]. Still, the design of the adaptive observer is not a straightforward application of [30] and requires appropriate modifications as well as specific assumptions. The exponential convergence of the proposed adaptive observer is guaranteed provided a persistence of excitation condition holds and $||\theta - \theta_0||$ is small enough. Such an excitation condition is needed in almost all works

R. Postoyan is with the Université de Lorraine, CNRS, CRAN, F-54000 Nancy, France romain.postoyan@univ-lorraine.fr.

Q. Zhang is with Inria-IFSTTAR, Campus de Beaulieu, 35042 Rennes Cedex, France qinghua.zhang@inria.fr.

This report corresponds to [23] with all the proofs as well as further details.

[1]The notation and the terminology are defined at the end of the introduction.

on adaptive observers. The fact that $||\theta-\theta_0||$ has to be small means that the design is local in the parameter estimate, but we are free to initialize the state estimate as we wish. It can be noted that many estimation or identification schemes are also local, see [18] for instance and all the works on extended Kalman filters, see e.g., [6], [16], [24]. Afterwards, the results are extended to the case where A is only smooth in θ, and not necessarily affine. The adaptive observer and its analysis are modified accordingly and an approximate convergence property is ensured in this case.

Compared to [17], our results apply to multi-input multi-output linear *time-varying* systems, as opposed to single-input single-output LTI systems, and compared to [21], we address LTV systems and our approach relies on different type of conditions (no LMI). The closest work in spirit, although it is written in discrete-time for LTI models, is the one in [18]. A key difference with [18] is that we do not need to use projection techniques to maintain the parameter estimate in some given set, and the estimates we generate cannot be trapped in the border of the given set due to projections, but always converge to the true parameter provided, again, that $||\theta-\theta_0||$ is small enough and a persistence of excitation holds. On the other hand, the adaptive observer is less computationally demanding than the supervisory observer presented in [8], which requires a large number of observers to work, and we ensure the asymptotic convergence of the estimation errors as opposed to a practical property in [8]. Finally, compared to [5], [7], [11], [12]: (i) we do not require any particular triangular or block-diagonal structure of the system ([5], [11], [12]); (ii) we provide results for the case where the state matrix is not affine in θ in which case the results in [5], [7], [11] do not apply; (iii) we do not rely on high-gain techniques ([11], [12]), which are not known to be very sensitive to noise measurements; (iv) the observer is of smaller dimension, which is relevant for computational reasons ([5], [11], [12]). Also, we believe that the way we approach the problem has its own interest, which can be relevant for further extensions.

A similar problem is addressed in [9]. First, exponential convergence is ensured in our case when A is affine in θ as opposed to an ultimate boundedness property in [9]. Second, the proposed solution is computationally lighter, since [9] requires the use of both a state-observer for the nominal parameter values and of an adaptive observer, while we only need the latter.

The rest of the paper is organized as follows. The assumptions we make on system (1) are presented in Section II. The main result is then given in Section III and an illustrative example is provided in Section IV. Section V concludes the paper. A technical lemma is provided in the appendix.

Notation and terminology. Let $\mathbb{R} := (-\infty,\infty)$, $\mathbb{R}_{\geq 0} := [0,\infty)$, $\mathbb{R}_{>0} := (0,\infty)$, $\mathbb{Z}_{\geq 0} := \{0,1,2,\ldots\}$, and $\mathbb{Z}_{>0} := \{1,2,\ldots\}$. The notation (x,y) stands for $[x^\top,\, y^\top]^\top$, where $x\in\mathbb{R}^n$, $y\in\mathbb{R}^m$, and $n,m\in\mathbb{Z}_{>0}$. The identity matrix of size $n\in\mathbb{Z}_{>0}$ is denoted $\mathbb{I}_n$ or simply $\mathbb{I}$ when the dimension is clear from the context. Given two matrices $Q,R\in\mathbb{R}^{n\times n}$, we write $Q\leq R$ if $R-Q$ is positive semi-definite, $n\in\mathbb{Z}_{>0}$. The symbol $\star$ stands for symmetric blocks in matrices. A time-varying matrix $M(t)\in\mathbb{R}^{n\times m}$ is said to be bounded if there exists $c\geq 0$ such that $||M(t)||\leq c$ for any $t\geq 0$, where $||M(t)|| := \sup_{x\in\mathbb{R}^m\setminus\{0\}}\frac{||M(t)x||}{||x||}$, $||\cdot||$ standing for the Euclidean norm when the argument is a vector. We write $N(x)=\mathcal{O}(||x||^2)$ for $N:\mathbb{R}^n\to\mathbb{R}^m$ with $n,m\in\mathbb{Z}_{>0}$, when there exists $c\geq 0$ such that $||N(x)||\leq c||x||^2$ for any $x\in\mathbb{R}^n$. The set of piecewise continuous functions from $\mathbb{R}_{\geq 0}$ to $\mathbb{R}^n$ is denoted $\mathcal{PC}(\mathbb{R}^n)$, $n\in\mathbb{Z}_{>0}$. Let $p\in\mathbb{R}^n$ and $\delta\in\mathbb{R}_{>0}\cup\{\infty\}$, the closed ball of radius δ centered at p is denoted $\mathcal{B}(p,\delta)$, with $B(p,\delta)=\mathbb{R}^n$ when $\delta=\infty$.

II. Assumptions

The matrices $A(t,\theta)$, $B(t)$ and $C(t)$ in (1) are assumed to be continuous and bounded with respect to the time t. We make the next assumption on how $A(t,\theta)$ depends on θ.

Assumption 1: The matrix $A(t,\theta)$ is affine in θ, in the sense that there exist matrices $A_0(t),\ldots,A_{n_\theta}(t)\in\mathbb{R}^{n_x\times n_x}$ such that, for any $t\geq 0$, $A(t,\theta)=A_0(t)+\sum_{i=1}^{n_\theta}A_i(t)\theta_i$, where $\theta=(\theta_1,\ldots,\theta_{n_\theta})$. □

When Assumption 1 is not satisfied, we may redefine the vector of unknown parameters to enforce it, by eventually over-parameterizing the original system. This assumption will be relaxed in Section III-C. It is important to note that the matrices $A_0(t),\ldots,A_{n_\theta}(t)$ in Assumption 1 are known, as these can be directly derived from the expression of $A(t,\theta)$. These matrices are continuous and bounded with respect to the time since so is $A(t,\theta)$.

Our design implicitly relies on the knowledge of a state observer for system (1) if θ would be known. More precisely, we make the following assumption.

Assumption 2: There exist $\delta_1\in\mathbb{R}_{>0}\cup\{\infty\}$ and $K:\mathbb{R}_{\geq 0}\times\mathcal{B}(\theta,\delta_1)\to\mathbb{R}^{n_x\times n_y}$, which is continuous and bounded, such that for any $\theta_0\in\mathcal{B}(\theta,\delta_1)$ the origin of the system $\dot{\tilde{x}}(t)=[A(t,\theta_0)-K(t,\theta_0)C(t)]\,\tilde{x}(t)$ is uniformly globally exponentially stable. □

Assumption 2 essentially means that we know a set where parameter θ lies and, if we knew θ, we would be able to design a state-observer for the corresponding system. More precisely, Assumption 2 states that there exists a neighborhood of parameter θ, corresponding to $\mathcal{B}(\theta,\delta_1)$, such that we can design a standard (Luenberger) state observer with gain $K(t,\theta_0)$ for any θ_0 in this set, whose solutions uniformly, globally, and exponentially converge to solutions to (1) when $\theta=\theta_0$. To see it, let $\dot{\hat{x}}(t)=A(t,\theta_0)x(t)+B(t)u(t)+K(t,\theta_0)(y(t)-C(t)\hat{x}(t))$. If $\theta=\theta_0$, then the estimation error system with variable $\tilde{x}=x-\hat{x}$ is $\dot{\tilde{x}}(t)=[A(t,\theta_0)-K(t,\theta_0)C(t)]\,\tilde{x}(t)$ as in Assumption 2.

A necessary and sufficient condition for Assumption 2 to hold is provided next.

Lemma 1: Under Assumption 1, Assumption 2 holds if and only if there exists a continuous and bounded mapping $K^*:\mathbb{R}_{\geq 0}\to\mathbb{R}^{n_x\times n_y}$ such that the origin of $\dot{\tilde{x}}(t)=[A(t,\theta)-K^*(t)C(t)]\,\tilde{x}(t)$ is uniformly globally exponentially stable for the (unknown) true parameter value θ. □

Proof. The necessity part immediately follows from Assumption 2. For the sufficiency, we consider the func-

tion $W(P^*,\tilde{x}) := \tilde{x}^\top P^*\tilde{x}$ where P^* is the solution to $-\dot{P}^*(t) = P^*(t)\big[A(t,\theta) - K^*(t)C(t)\big] + \big[A(t,\theta) - K^*(t)C(t)\big]^\top P^*(t) + \mathbb{I}$. According to Theorem 4.12 in [15], such a matrix $P(\cdot)$ exists on $\mathbb{R}_{\geq 0}$, is continuously differentiable, bounded, symmetric, positive definite and there exist $\underline{a}_{P^*}, \overline{a}_{P^*} > 0$ such that $\underline{a}_{P^*}\mathbb{I} \leq P^*(t) \leq \overline{a}_{P^*}\mathbb{I}$ for any $t \geq 0$. Let $\theta_0 \in \mathbb{R}^{n_\theta}$ and $t \geq 0$. Along the solutions to $\dot{\tilde{x}}(t) = [A(t,\theta_0) - K^*(t)C(t)]\,\tilde{x}(t)$, $\dot{W}(P^*(t),\tilde{x}(t)) = -||\tilde{x}(t)||^2 + 2\tilde{x}(t)^\top P^*(t)\,[A(t,\theta_0) - A(t,\theta)]\,\tilde{x}(t)$. Since $A(t,\theta)$ is bounded in t and affine in θ according to Assumption 1, there exists $\delta_1 > 0$ such that for any θ_0 such that $||\theta_0 - \theta|| < \delta_1$, $\dot{W}(P^*(t),\tilde{x}(t)) \leq -\frac{1}{2}||\tilde{x}(t)||^2$. Consequently, the origin of $\dot{\tilde{x}}(t) = [A(t,\theta_0) - K^*(t)C(t)]\,\tilde{x}(t)$ is uniformly globally exponentially stable according to Theorem 4.10 in [15]. Hence, Assumption 2 holds with $K(t,\theta_0) = K^*(t)$. ■

Lemma 1 involves θ, which we do not know. The following lemma provides exploitable conditions for the satisfaction of Assumption 2.

Lemma 2: Suppose the following holds.

(i) Assumption 1 holds.

(ii) There exists $\delta_1 \in \mathbb{R}_{\geq 0} \cup \{\infty\}$ such that the pair $(A(t,\theta_0), C(t,\theta_0))$ is uniformly completely observable[2] for any $\theta_0 \in \mathcal{B}(\theta,\delta_1)$.

Assumption 2 is verified with $K(t,\theta_0) = P(t,\theta_0)C(t)^\top R^{-1}(t)$ where $P(t,\theta_0)$ is the solution to $\dot{P}(t,\theta_0) = P(t,\theta_0)A(t,\theta_0)^\top + A(t,\theta_0)P(t,\theta_0) - P(t,\theta_0)C(t)^\top R^{-1}(t)C(t)P(t,\theta_0) + Q(t)$ with $P(0) = P_0$, where P_0 is any symmetric, positive definite matrix, and $R(t), Q(t)$ are any symmetric, positive definite, continuous and bounded matrices. □

Proof. Let $\theta_0 \in \mathcal{B}(\theta,\delta_1)$. Since $R(t)$ is positive definite and the pair $(A(t,\theta_0), C(t,\theta_0))$ is uniformly completely observable, the pair $(A(t,\theta_0), C(t,\theta_0)R^{\frac{1}{2}}(t))$ is also uniformly completely observable. Moreover, the matrix $W(t_0,t)$ in [1] is nonsingular at t_0 since P_0 is symmetric and positive definite, it is therefore nonsingular at some $t_1 > t_0$ by continuity. Consequently, all the conditions of Theorem 3.1 in [1] are verified, which ensures the uniform global exponential stability of the origin of $\dot{\tilde{x}}(t) = [A(t,\theta) - K^*(t)C(t)]\,\tilde{x}(t)$. Since $A(t,\theta_0)$ is continuous in θ_0 in view of Assumption 1, so is $P(t,\theta_0)$ according to Corollary 6 in Chapter 2 in [2]. Therefore, $K(t,\theta_0)$ is continuous in θ_0 for $\theta_0 \in \mathcal{B}(\theta,\delta_1)$. Boundedness of $K(t,\theta_0)$ follows from the boundedness of $P(t,\theta_0)$, which is ensured in [1]. We have proved that Assumption 2 is satisfied. ■

When the matrices A and C in (1) are time-invariant, it suffices to have the pair $(A(\theta_0), C(\theta_0))$ observable for any θ_0 with $||\theta - \theta_0|| \leq \delta_1$ for some $\delta_1 \in \mathbb{R}_{>0} \cup \{\infty\}$ to ensure Assumption 2. Indeed, in this case, Bass-Gura formula[3] leads to a gain $K(\theta_0)$, which is continuous in θ_0. The same applies when $L(\theta_0)$ is selected to minimize a linear quadratic cost, i.e. when $L(\theta_0) = P(\theta_0)C^\top R^{-1}$ and $P(\theta_0)$ is the solution to the Riccati equation $P(\theta_0)A(\theta_0)^\top + A(\theta_0)P(\theta_0) - P(\theta_0)C^\top R^{-1}CP(\theta_0) + Q$. It suffices to select Q and R symmetric and positive definite to ensure the continuity of $P(\theta_0)$ in θ_0 according to Proposition 1 in [10], which in turn implies the continuity of $K(\theta_0)$ in θ_0.

We finally make the next boundedness assumption on system (1).

Assumption 3: There exists a hyper-rectangle $\mathcal{X} = \{x = (x_1,\ldots,x_{n_x}) \in \mathbb{R}^{n_x} : x_i \in [\underline{c}_i, \overline{c}_i]\}$ where $\underline{c}_i \leq \overline{c}_i \in \mathbb{R}$, $i \in \{1,\ldots,n_x\}$, a non-empty set of inputs $\mathcal{M}_u \subset \mathcal{PC}(\mathbb{R}^{n_u})$, and a non-empty set of initial conditions $\mathcal{S}_\mathcal{X} \subset \mathbb{R}^{n_x}$, such that any solution x to system (1) initialized in $\mathcal{S}_\mathcal{X}$ with input $u \in \mathcal{M}_u$ satisfies $x(t) \in \mathcal{X}$ for all $t \geq 0$. □

Assumption 3 means that the solutions to (1) lie in a known compact set, which can always be embedded in a closed hyper-rectangle, whenever its input lies in set $\mathcal{M}_u$. Note that there is no restriction of the "size" of $\mathcal{X}$.

III. Main result

A. Design and analysis

Like in [25], we introduce the element-wise saturation function $\sigma_\mathcal{X} : \mathbb{R}^{n_x} \to \mathbb{R}^{n_x}$ defined as $\sigma_\mathcal{X}(x) := (\sigma_{1,\mathcal{X}}(x_1),\ldots,\sigma_{n_x,\mathcal{X}}(x_{n_x}))$ with $\sigma_{i,\mathcal{X}}(x_i) = x_i$ when $x_i \in [\underline{c}_i, \overline{c}_i]$, $\sigma_{i,\mathcal{X}}(x_i) = \underline{c}_i$ when $x_i \leq \underline{c}_i$ and $\sigma_{i,\mathcal{X}}(x_i) = \overline{c}_i$ when $x_i \geq \overline{c}_i$, for any $x_i \in \mathbb{R}$, where $\mathcal{X}, \underline{c}_i, \overline{c}_i$ come from Assumption 3, $i \in \{1,\ldots,n_x\}$. Based on Assumptions 1-2 and inspired by [30], we propose the adaptive observer

$$\begin{aligned}
\dot{\hat{x}}(t) &= A(t,\theta_0)\hat{x}(t) + B(t)u(t) + \Lambda(t,\sigma_\mathcal{X}(\hat{x}(t)))\bar{\theta}(t) \\
&\quad + \big[K(t,\theta_0) + \gamma\Upsilon(t)\Upsilon^\top(t)C^\top(t)\Sigma(t)\big] \\
&\quad \times (y(t) - C(t)\hat{x}(t)) \\
\dot{\bar{\theta}}(t) &= \gamma\Upsilon^\top(t)C^\top(t)\Sigma(t)(y(t) - C(t)\hat{x}(t)) \\
\dot{\Upsilon}(t) &= [A(t,\theta_0) - K(t,\theta_0)C(t)]\,\Upsilon(t) + \Lambda(t,\sigma_\mathcal{X}(\hat{x}(t))) \\
\hat{\theta}(t) &= \bar{\theta}(t) + \theta_0,
\end{aligned} \tag{2}$$

where $\bar{\theta}(t) \in \mathbb{R}^{n_\theta}$ is the estimate of $\theta - \theta_0$, $\hat{\theta}(t) \in \mathbb{R}^{n_\theta}$ is therefore the estimate of θ, $\Upsilon(t) \in \mathbb{R}^{n_x \times n_\theta}$, $\Lambda(t,z) := [A_1(t)z \;\ldots\; A_{n_\theta}(t)z] \in \mathbb{R}^{n_x \times n_\theta}$ for $z \in \mathbb{R}^{n_x}$ with $A_1(t),\ldots,A_{n_\theta}(t)$ coming from Assumption 1, $\Sigma(t) : \mathbb{R} \to \mathbb{R}^{n_y \times n_y}$ is any symmetric, positive definite, bounded, continuous matrix, $\gamma \in \mathbb{R}_{>0}$, and $\mathcal{X} \subset \mathbb{R}^n$ comes from Assumption 3. We have two degree of freedom when designing (2): $\Sigma(t)$ and γ. The former has to be designed to verify a persistence of excitation condition stated next. In practice, we typically take it diagonal. Parameter γ tunes the convergence speed of $\bar{\theta}$, the bigger γ, the faster the convergence, but the more sensitive to noise.

Remark 1: The use of the saturation function $\sigma_\mathcal{X}$ in (2) is essential to guarantee the boundedness of Υ in the proof of the next theorem, which is needed to establish the desired convergence property. This is the main difference with [30] in terms of design. Other differences reside in the assumptions we rely on. □

The next theorem ensures that the state and the parameter estimates provided by (2) exponentially converges to x and θ, provided a persistence of condition holds and $||\theta - \theta_0||$ is sufficiently small. Its proof is provided in Section III-B.

[2] See [26] for a definition of uniform complete observability as well as conditions to ensure it.

[3] See Chapter 4 in [14] for instance.

Theorem 1: Consider systems (1) and (2) and suppose the following holds.

(i) Assumptions 1-3 hold.

(ii) Let $\delta_2 \in \mathbb{R}_{>0} \cup \{\infty\}$, $\hat{x}_0 \in \mathbb{R}^{n_x}$, $\bar{\theta}_0 \in \mathbb{R}^{n_\theta}$, $\Upsilon_0 \in \mathbb{R}^{n_x \times n_\theta}$, there exist $\underline{\alpha}, T > 0$, which depend on $\hat{x}_0, \bar{\theta}_0, \Upsilon_0$, such that, for any $x(0) \in \mathcal{S}_\mathcal{X}$ and input $u \in \mathcal{M}_u$, any θ_0 such that $||\theta - \theta_0|| \leq \delta_2$, the solution $\hat{x}, \bar{\theta}, \Upsilon$ to system (2) initialized at $(\hat{x}_0, \bar{\theta}_0, \Upsilon_0)$ with $||\theta_0 - \theta|| \leq \delta_2$, verifies for any $t \geq 0$,

$$\underline{\alpha}\mathbb{I} \leq \int_t^{t+T} \Upsilon^\top(\tau) C^\top(\tau) \Sigma(\tau) C(\tau) \Upsilon(\tau) d\tau. \quad (3)$$

There exist $\delta \in (0, \min\{\delta_1, \delta_2\})$ and $c_1, c_2 > 0$, which depend on $\mathcal{X}, \mathcal{M}_u, \hat{x}_0, \bar{\theta}_0, \Upsilon_0$, such that, if $||\theta - \theta_0|| \leq \delta$, the solution $(\hat{x}, \bar{\theta}, \Upsilon)$ to (2) initialized at $(\hat{x}_0, \bar{\theta}_0, \Upsilon_0)$ and any solution to system (1) with $x(0) \in \mathcal{S}_\mathcal{X}$ and input $u \in \mathcal{M}_u$, verify $||(x(t) - \hat{x}(t), \hat{\theta}(t) - \theta)|| \leq c_1 e^{-c_2 t} ||(x(0) - \hat{x}_0, \hat{\theta}(0) - \theta)||$ for any $t \geq 0$. □

Item (ii) of Theorem 1 is a persistence of excitation condition on the Υ-system, like in [30]. It actually is a *uniform* persistence of excitation property, see Definition 2 in [19], as that the constants $\underline{\alpha}$ and T are independent of θ_0, which is essential. Indeed, if the bound δ_2, and thus δ, on the norm of $\theta - \theta_0$ would depend on θ_0, there will be no guarantee that we can always select θ_0 sufficiently close to θ such that $||\theta - \theta_0|| \leq \delta$. That is also the reason why we assume the existence of the gain $K(t, \theta_0)$ for any $\theta_0 \in \mathcal{B}(\theta, \delta_1)$ in Assumption 2, and not for a given θ_0, otherwise δ_1 in Assumption 2 may depend on θ_0 and the same issue would arise. On the other hand, δ depends on the initial conditions of observer (2). It is possible to ensure that δ is uniform over these initial conditions over given sets, when (3) holds for any initial conditions in these sets.

Property (3) is difficult to verify as it involves any θ_0 in a set, and any solution initialized in $\mathcal{S}_\mathcal{X}$ with input $u \in \mathcal{M}_u$. Even if we could verify it, that would not mean that the estimates converge to the true values, as we also need for that $||\theta - \theta_0||$ to be sufficiently small, which is something we cannot check in practice. It is important to note that this difficulty is generic to *any* local estimation schemes, including works on extended Kalman filters where the initial conditions x_0 and $\hat{x}_0$ have to be close to each other (among other conditions), see e.g., [6], [16], [24]. The only way to avoid it is when we have the property that the observation error $y(t) - C(t)\hat{x}(t)$ converging to zero implies that the estimation errors $x(t) - \hat{x}(t)$ and $\theta - \hat{\theta}(t)$ also converging to zero, see Section III.A in [22]. In practice, we can check numerically on-line whether the inequality in (3) holds for a given θ_0 and a given solution x to (1), in which case the non-satisfaction of (3) indicates that the obtained estimate are not valid.

The fact that we need to know a set where the unknown parameters lie is a valid assumption in various applications, like electrochemical batteries for which the parameters are accurately known when the battery is produced, but slowly vary with time. Then, the interest of the adaptive observer is to track these slow variations for monitoring purposes. If we only have some vague knowledge about θ, we may overcome this issue by employing a multi-observer architecture as explained in [22].

Remark 2: Observer (2) can easily be shown to be robust in the following sense, under the conditions of Theorem 1. Consider system (1) perturbed as $\dot{x}(t) = A(t, \theta)x(t) + B(t)u(t) + w_x(t)$, $y(t) = C(t)x(t) + w_z(t)$, where w_x and w_z are disturbances and noise signals respectively. The Lyapunov analysis in Section III-B almost immediately allows to conclude that the estimation error system satisfies an input-to-state stability property [27] in this case. □ provided the conditions of Theorem 1 hold.

B. Proof of Theorem 1

We introduce the estimation errors $\tilde{x} := x - \hat{x}$ and $\tilde{\theta} := \theta - \theta_0 - \bar{\theta}$. Let $x(0) \in \mathbb{R}^{n_x}$, $t \geq 0$ for which the solution to (1), (2) is defined, and $\theta_0 \in \mathcal{B}(\theta, \min\{\delta_1, \delta_2\})$ where δ_1, δ_2 come from Assumptions 2 and item (ii) of Theorem 1, respectively. In view of (1) and (2),

$$\begin{aligned} \dot{\hat{x}}(t) = & A(t,\theta)x(t) + B(t)u(t) - A(t,\theta_0)\hat{x}(t) - B(t)u(t) \\ & -\Lambda(t, \sigma_\mathcal{X}(\hat{x}(t)))\bar{\theta}(t) - [K(t,\theta_0) \\ & +\gamma\Upsilon(t)\Upsilon^\top(t)C^\top(t)\Sigma(t)]\,(y(t) - C(t)\hat{x}(t)), \end{aligned} \quad (4)$$

we add and subtract $A(t, \theta_0)x(t)$ and we obtain

$$\begin{aligned} \dot{\tilde{x}}(t) = & [A(t,\theta_0) - K(t,\theta_0)C(t)]\,\tilde{x}(t) \\ & + [A(t,\theta) - A(t,\theta_0)]\,x(t) - \Lambda(t, \sigma_\mathcal{X}(\hat{x}(t)))\bar{\theta}(t) \\ & -\gamma\Upsilon(t)\Upsilon^\top(t)C^\top(t)\Sigma(t)(y(t) - C(t)\hat{x}(t)). \end{aligned} \quad (5)$$

In view of Assumption 1 and the definition of $\Lambda(\cdot, \cdot)$ given after (2),

$$[A(t,\theta) - A(t,\theta_0)]\,x(t) = \Lambda(t, x(t))(\theta - \theta_0). \quad (6)$$

We now restrict our attention to the case where $x(0) \in \mathcal{S}_\mathcal{X}$ with $u \in \mathcal{M}_u$. Hence, $x(t) \in \mathcal{S}_\mathcal{X}$ and $x(t) = \sigma_\mathcal{X}(x(t))$ in view of Assumption 3 and the definition of σ, respectively. Thus $[A(t,\theta) - A(t,\theta_0)]\,x(t) = \Lambda(t, \sigma_\mathcal{X}(x(t)))(\theta - \theta_0)$. Consequently, in view of (5),

$$\begin{aligned} \dot{\tilde{x}}(t) = & [A(t,\theta_0) - K(t,\theta_0)C(t)]\,\tilde{x}(t) \\ & +\Lambda(t, \sigma_\mathcal{X}(x(t)))(\theta - \theta_0) - \Lambda(t, \sigma_\mathcal{X}(\hat{x}(t)))\bar{\theta}(t) \\ & -\gamma\Upsilon(t)\Upsilon^\top(t)C^\top(t)\Sigma(t)(y(t) - C(t)\hat{x}(t)). \end{aligned} \quad (7)$$

We add and subtract $\Lambda(t, \sigma_\mathcal{X}(\hat{x}(t)))(\theta - \theta_0)$ and we obtain, by exploiting the linearity of $\Lambda(\cdot, \cdot)$ in its second argument and using $\tilde{\theta}(t) = \theta - \theta_0 - \bar{\theta}(t)$,

$$\begin{aligned} \dot{\tilde{x}}(t) = & [A(t,\theta_0) - K(t,\theta_0)C(t)]\,\tilde{x}(t) \\ & +\Lambda(t, \sigma_\mathcal{X}(x(t)) - \sigma_\mathcal{X}(\hat{x}(t)))(\theta - \theta_0) \\ & +\Lambda(t, \sigma_\mathcal{X}(\hat{x}(t)))\tilde{\theta}(t) \\ & -\gamma\Upsilon(t)\Upsilon^\top(t)C^\top(t)\Sigma(t)(y(t) - C(t)\hat{x}(t)). \end{aligned} \quad (8)$$

Like in [30], we introduce $\eta := \tilde{x} - \Upsilon\tilde{\theta}$. We deduce from

the last equation and (2),

$$\begin{aligned}\dot{\eta}(t)=&[A(t,\theta_0)-K(t,\theta_0)C(t)]\,\tilde{x}(t)\\&+\Lambda(t,\sigma_{\mathcal{X}}(x(t))-\sigma_{\mathcal{X}}(\hat{x}(t)))(\theta-\theta_0)\\&+\Lambda(t,\sigma_{\mathcal{X}}(\hat{x}(t)))\tilde{\theta}(t)\\&-\gamma\Upsilon(t)\Upsilon^\top(t)C^\top(t)\Sigma(t)(y(t)-C(t)\hat{x}(t))\\&-\left[(A(t,\theta_0)-K(t,\theta_0)C(t))\,\Upsilon(t)+\Lambda(t,\sigma_{\mathcal{X}}(\hat{x}(t)))\right]\tilde{\theta}(t)\\&+\Upsilon(t)\left[\gamma\Upsilon^\top(t)C^\top(t)\Sigma(t)(y(t)-C(t)\hat{x}(t))\right]\\=&[A(t,\theta_0)-K(t,\theta_0)C(t)]\,\eta(t)\\&+\Lambda(t,\sigma_{\mathcal{X}}(x(t))-\sigma_{\mathcal{X}}(\hat{x}(t)))(\theta-\theta_0).\end{aligned}\tag{9}$$

On the other hand, in view of (2) and since $\tilde{x}(t)=\eta(t)+\Upsilon(t)\tilde{\theta}(t)$,

$$\begin{aligned}\dot{\tilde{\theta}}(t)&=-\gamma\Upsilon^\top(t)C^\top(t)\Sigma(t)C(t)\tilde{x}(t)\\&=-\gamma\Upsilon^\top(t)C^\top(t)\Sigma(t)C(t)\left(\Upsilon(t)\tilde{\theta}(t)+\eta(t)\right).\end{aligned}\tag{10}$$

The dynamics of the $(\eta,\tilde{\theta})$-system in (9), (10) is the same as in (14)-(15) in [30] up to the perturbative term $\Lambda(t,\sigma_{\mathcal{X}}(x(t))-\sigma_{\mathcal{X}}(\hat{x}(t)))(\theta-\theta_0)$ in (9). Because of the latter, we can no longer invoke cascade arguments to conclude about the stability of the $(\eta,\tilde{\theta})$-system as in [30]. We will use small-gain arguments instead, and the small-gain condition will be enforced by imposing $\theta-\theta_0$ to be small, which explains why the convergence property we will guarantee is local and not global (in general).

Let $(\hat{x},\bar{\theta},\Upsilon)$ be the solution to system (2) initialized at $(\hat{x}_0,\bar{\theta}_0,\Upsilon_0)$. We define $P:\mathbb{R}_{\geq 0}\times\mathbb{R}^{n_\theta}\to\mathbb{R}^{n_x\times n_x}$ as the solution to $-\dot{P}(t,\theta_0)=P(t,\theta_0)\,[A(t,\theta_0)-K(t,\theta_0)C(t)]+[A(t,\theta_0)-K(t,\theta_0)C(t)]^\top P(t,\theta_0)+\mathbb{I}$, with $P(0,\theta_0)=P_0$ and P_0 is symmetric and positive definite. Since θ_0 belongs to the closed ball $\mathcal{B}(\theta,\delta_1)$, Assumption 2 holds and $A(t,\theta_0)-K(t,\theta_0)C(t)$ is continuous and bounded, Lemma 1 in [19] ensures that $P(t,\theta_0)$ exists for all $t\geq 0$, is continuously differentiable, bounded, symmetric, positive definite and there exist $\underline{a}_P,\overline{a}_P>0$, which are independent of θ_0, such that $\underline{a}_P\mathbb{I}\leq P(t,\theta_0)\leq\overline{a}_P\mathbb{I}$ for any $t\geq 0$. We define $V:(P,\eta)\mapsto\eta^\top P\eta$. Let[4] $t\geq 0$, in view of (9) and the definition of $P(t,\theta_0)$ (we omit the time and θ_0 dependencies of the solutions), $\dot{V}(P,\eta)=-||\eta||^2+2\eta^\top P\Lambda(t,\sigma_{\mathcal{X}}(x)-\sigma_{\mathcal{X}}(\hat{x}))(\theta-\theta_0)$, where $\hat{x}=x-\eta-\Upsilon\tilde{\theta}$. The definition of $\Lambda(\cdot,\cdot)$ and the fact that the matrices $A_1(t),\ldots,A_{n_\theta}(t)$ are bounded according to Assumption 1 imply that there exists a constant $a_\Lambda>0$ such that $||\Lambda(t,z)||\leq a_\Lambda||z||$ for any $z\in\mathbb{R}^{n_x}$. Consequently, since $P\leq\overline{a}_P\mathbb{I}$,$\dot{V}(P,\eta)\leq-||\eta||^2+2\overline{a}_Pa_\Lambda||\eta||\cdot||\sigma_{\mathcal{X}}(x)-\sigma_{\mathcal{X}}(\hat{x})||\times||\theta-\theta_0||$. From the definition of σ, $||\sigma_{\mathcal{X}}(x)-\sigma_{\mathcal{X}}(\hat{x})||\leq n_x||x-\hat{x}||=n_x||\tilde{x}||\leq n_x||\eta||+n_x||\Upsilon||\cdot||\tilde{\theta}||$. Thus,

$$\begin{aligned}\dot{V}(P,\eta)\leq&-||\eta||^2+2n_x\overline{a}_Pa_\Lambda||\eta||^2||\theta-\theta_0||\\&+2n_x\overline{a}_Pa_\Lambda||\eta||\cdot||\Upsilon||\cdot||\tilde{\theta}||\cdot||\theta-\theta_0||\\=&-\left[1-2n_x\overline{a}_Pa_\Lambda||\theta-\theta_0||\right]||\eta||^2\\&+2n_x\overline{a}_Pa_\Lambda||\eta||\cdot||\Upsilon||\cdot||\tilde{\theta}||\\&\times||\theta-\theta_0||.\end{aligned}\tag{11}$$

[4]We can take any $t\geq 0$ here as the solutions we consider are defined for all positive times since $x(t)\in\mathcal{X}$ and in view of (2).

We now investigate the $\tilde{\theta}$-system in (10). We need the next claim for that purpose.

Claim 1: There exists a continuous function $\nu:\mathbb{R}^{n_x\times n_\theta}\to[1,\infty)$, which is independent of θ_0, such that for any continuous function $\hat{x}:\mathbb{R}_{\geq 0}\to\mathbb{R}^{n_x}$, the corresponding solution Υ to the Υ-system in (2) initialized at Υ_0 is such that $||\Upsilon(t)||\leq\nu(\Upsilon_0)$ for any $t\geq 0$. □

Proof of Claim 1. The result follows from Assumption 2, Lemma 1 in [19] and the fact that $\Lambda(t,\sigma_{\mathcal{X}}(\hat{x}))$ is bounded uniformly with respect to $\theta_0\in\mathcal{B}(\theta,\min\{\delta_1,\delta_2\})$. The fact that the image of ν is in $[1,\infty)$ can always be ensured by adding 1 to any bound on $||\Upsilon(t)||$. ■

The matrix $\gamma\Upsilon^\top(t)C^\top(t)\Sigma(t)C(t)\Upsilon(t)$ is continuous and bounded since so are $C(t)$ by assumption, $\Sigma(t)$ by design, and $\Upsilon(t)$ according to Claim 1 for the boundedness. Therefore, since $\theta\in\mathcal{B}(\theta,\delta_2)$, item (ii) of Theorem 1 is assumed to hold, $\Sigma(t)$ and $C(t)$ are bounded and Claim 1 applies, according to Lemmas 1 and 5 in [19], the solution $\widetilde{P}(t)$ to $-\dot{\widetilde{P}}(t)=-\gamma\widetilde{P}(t)\Upsilon^\top(t)C^\top(t)\Sigma(t)C(t)\Upsilon(t)-\left[\gamma\Upsilon^\top(t)C^\top(t)\Sigma(t)C(t)\Upsilon(t)\right]^\top\widetilde{P}(t)+\mathbb{I}$, with $\widetilde{P}(0)=\widetilde{P}_0$ and $\widetilde{P}_0$ is symmetric and positive definite, exists for all $t\geq 0$, is continuously differentiable, bounded, symmetric, positive definite and there exist $\underline{a}_{\widetilde{P}},\overline{a}_{\widetilde{P}}>0$, which are independent of θ_0, such that $\underline{a}_{\widetilde{P}}\mathbb{I}\leq\widetilde{P}(t)\leq\overline{a}_{\widetilde{P}}\mathbb{I}$ for any $t\geq 0$. Let $\widetilde{V}:(\widetilde{P},\tilde{\theta})\mapsto\tilde{\theta}^\top\widetilde{P}\tilde{\theta}$. Let $t\geq 0$, in view of (10) and the definition of $\widetilde{P}(t)$, $\dot{\widetilde{V}}(\widetilde{P},\tilde{\theta})=-||\tilde{\theta}||^2-2\gamma\tilde{\theta}^\top\widetilde{P}\Upsilon^\top(t)C^\top(t)\Sigma(t)C(t)\eta$. By following similar steps as above, we derive

$$\dot{\widetilde{V}}(\widetilde{P},\tilde{\theta})\leq-||\tilde{\theta}||^2+2\gamma\overline{a}_{\widetilde{P}}a_C||\tilde{\theta}||\cdot||\Upsilon||\cdot||\eta||,\tag{12}$$

where $a_C>0$ is a constant such that $||C^\top(t)\Sigma(t)C(t)||\leq a_C$ for all $t\geq 0$, which exists since $C(t)$ and $\Sigma(t)$ are bounded.

Define $U:(P,\widetilde{P},\eta,\tilde{\theta})\mapsto V(P,\eta)+\mu\widetilde{V}(\widetilde{P},\tilde{\theta})$ with $\mu>0$ a constant, which will be selected below. In view of (11) and (12), for any $t\geq 0$,

$$\begin{aligned}\dot{U}(P,\widetilde{P},\eta,\tilde{\theta})\leq&-\left[1-2n_x\overline{a}_Pa_\Lambda||\theta-\theta_0||\right]||\eta||^2\\&+2n_x\overline{a}_Pa_\Lambda||\Upsilon||\cdot||\tilde{\theta}||\cdot||\eta||\cdot||\theta-\theta_0||\\&-\mu||\tilde{\theta}||^2+2\mu\gamma\overline{a}_{\widetilde{P}}a_C||\tilde{\theta}||\cdot||\Upsilon||\cdot||\eta||.\end{aligned}\tag{13}$$

According to Claim 1, $||\Upsilon||\leq\nu(\Upsilon_0)$, thus

$$\begin{aligned}\dot{U}(P,\widetilde{P},\eta,\tilde{\theta})\leq&-\left[1-2n_x\overline{a}_Pa_\Lambda||\theta-\theta_0||\right]||\eta||^2\\&+2n_x\overline{a}_Pa_\Lambda\nu(\Upsilon_0)||\tilde{\theta}||\cdot||\eta||\cdot||\theta-\theta_0||\\&-\mu||\tilde{\theta}||^2+2\mu\gamma\overline{a}_{\widetilde{P}}a_C\nu(\Upsilon_0)||\tilde{\theta}||\cdot||\eta||\\=&-\left\langle(||\eta||,||\tilde{\theta}||),\mathcal{M}(||\eta||,||\tilde{\theta}||)\right\rangle,\end{aligned}\tag{14}$$

where $\mathcal{M}:=\begin{pmatrix}\mathcal{M}_{11}&\mathcal{M}_{12}\\\star&\mathcal{M}_{22}\end{pmatrix}$ with $\mathcal{M}_{11}:=1-2n_x\overline{a}_Pa_\Lambda||\theta-\theta_0||$, $\mathcal{M}_{12}:=-n_x\overline{a}_Pa_\Lambda\nu(\Upsilon_0)||\theta-\theta_0||-\mu\gamma\overline{a}_{\widetilde{P}}a_C\nu(\Upsilon_0)$ and $\mathcal{M}_{22}:=\mu$. Matrix $\mathcal{M}$ is positive definite if and only if

$$\begin{cases}1-2n_x\overline{a}_Pa_\Lambda||\theta-\theta_0||>0\\\left[1-2n_x\overline{a}_Pa_\Lambda||\theta-\theta_0||\right]\mu\\-\left[n_x\overline{a}_Pa_\Lambda\nu(\Upsilon_0)||\theta-\theta_0||+\mu\gamma\overline{a}_{\widetilde{P}}a_C\nu(\Upsilon_0)\right]^2>0.\end{cases}\tag{15}$$

The first inequality above is ensured as long as $||\theta-\theta_0||$ is sufficiently small, in particular we assume that $||\theta-\theta_0||<$

$\dfrac{1}{4n_x\overline{a}_P a_\Lambda}$ so that $1-2n_x\overline{a}_P a_\Lambda||\theta-\theta_0||>\frac{1}{2}$. The second inequality also holds with $||\theta-\theta_0||$ sufficiently small, by taking $\mu<\dfrac{1}{2\left(\gamma\overline{a}_{\widetilde{P}}a_C\nu(\Upsilon_0)\right)^2}$; note that the denominator is well-defined since $\nu(\Upsilon_0)\geq 1$ in view of Claim 1. As a result, there exists $\delta\in(0,\min\{\delta_1,\delta_2\})$, which is independent of θ_0, such that $||\theta-\theta_0||\leq\delta$ implies the existence of $\varepsilon>0$ such that, for $t\geq 0$,

$$\dot{U}(P,\widetilde{P},\eta,\tilde{\theta}) \quad\leq\quad -\varepsilon||(\eta,\tilde{\theta})||^2. \tag{16}$$

Since $\min\{\underline{a}_P,\mu\underline{a}_{\widetilde{P}}\}||(\eta,\tilde{\theta})||^2\leq U(P,\widetilde{P},\eta,\tilde{\theta})\leq(\overline{a}_P+\mu\overline{a}_{\widetilde{P}})||(\eta,\tilde{\theta})||^2$, we deduce that $(\eta(t),\tilde{\theta}(t))$ exponentially converges to 0 as $t\to\infty$. More precisely, there exist $c_1',c_2'>0$, which depend on $\mathcal{X},\mathcal{M}_u,\hat{x}_0,\bar{\theta}_0,\Upsilon_0$, such that $||(\eta(t),\tilde{\theta}(t))||\leq c_1'e^{-c_2't}||(\eta(0),\tilde{\theta}(0)||$. Moreover, since $||\tilde{x}(t)||\leq||\eta(t)||+||\Upsilon(t)||\cdot||\tilde{\theta}(t)||\leq||\eta(t)||+\nu(\Upsilon_0)||\tilde{\theta}(t)||$ according to the definition of η and Claim 1, we deduce that there exist $c_1,c_2>0$, which depend on $\mathcal{X},\mathcal{M}_u,\hat{x}_0,\bar{\theta}_0,\Upsilon_0$, such that $||(\tilde{x}(t),\tilde{\theta}(t))||\leq c_1e^{-c_2t}||(\tilde{x}(0),\tilde{\theta}(0))||$, which ends the proof.

C. When $A(t,\theta)$ is not affine in θ

We study the case where $A(t,\theta)$ is smooth in θ but not necessarily affine in θ, as required in Assumption 1. Let $t\geq 0$, $\theta\in\mathbb{R}^{n_\theta}$ and denote $A(t,\theta)=[a_{ij}(t,\theta)]_{(i,j)\in\{1,\ldots,n_x\}^2}$. The Taylor expansion of each a_{ij} with respect to θ gives $a_{ij}(t,\theta)=a_{ij}(t,\theta_0)+\dfrac{\partial a_{ij}}{\partial\theta}(t,\theta_0)(\theta-\theta_0)+\rho_{ij}(t,\theta,\theta_0)$ where[5] $\rho_{ij}(t,\theta,\theta_0)=\mathcal{O}(||\theta-\theta_0||^2)$. As a result,

$$A(t,\theta)=A(t,\theta_0)+\sum_{i=1}^{n_\theta}\overline{A}_i(t,\theta_0)(\theta_i-\theta_{0,i})+R(t,\theta,\theta_0), \tag{17}$$

where $\theta=(\theta_1,\ldots,\theta_{n_\theta})$, $\theta_0=(\theta_{0,1},\ldots,\theta_{0,n_\theta})$, $\overline{A}_k(t,\theta_0):=\left[\frac{\partial a_{ij}}{\partial\theta_k}(t,\theta_0)\right]_{(i,j)\in\{1,\ldots,n_x\}^2}$ for $k\in\{1,\ldots,n_\theta\}$, $R(t,\theta,\theta_0):=[\rho_{ij}(t,\theta,\theta_0)]_{(i,j)\in\{1,\ldots,n_x\}^2}$.

In view of (17), we modify the observer in (2) as

$$\begin{aligned}
\dot{\hat{x}}(t) &= A(t,\theta_0)\hat{x}(t)+B(t)u(t)+\overline{\Lambda}(t,\theta_0,\sigma_{\mathcal{X}}(\hat{x}(t)))\bar{\theta}(t)\\
&\quad+\left[K(t,\theta_0)+\gamma\Upsilon(t)\Upsilon^\top(t)C^\top(t)\Sigma(t)\right]\\
&\quad\times(y(t)-C(t)\hat{x}(t))\\
\dot{\bar{\theta}}(t) &= \gamma\Upsilon^\top(t)C^\top(t)\Sigma(t)(y(t)-C(t)\hat{x}(t))\\
\dot{\Upsilon}(t) &= [A(t,\theta_0)-K(t,\theta_0)C(t)]\,\Upsilon(t)+\overline{\Lambda}(t,\theta_0,\sigma_{\mathcal{X}}(\hat{x}(t)))\\
\hat{\theta}(t) &= \bar{\theta}(t)+\theta_0,
\end{aligned} \tag{18}$$

where $\overline{\Lambda}(t,\theta_0,z):=\left[\overline{A}_1(t,\theta_0)z\ \ldots\ \overline{A}_{n_\theta}(t,\theta_0)z\right]$ for $z\in\mathbb{R}^{n_x}$. The other matrices and parameters are selected as in (2).

Proposition 1: Consider systems (1) and (18). Suppose the following holds.

(i) Assumptions 2 and 3 hold.
(ii) Item (ii) of Theorem 1 holds along the solutions to (1) and (18).

[5]We can write that $\rho_{ij}(t,\theta,\theta_0)=\mathcal{O}(||\theta-\theta_0||^2)$ even though ρ_{ij} depends on the time t, since $A(t,\theta)$ is assumed to be bounded with respect to the time, see Section II.

There exist $\delta,c_1,c_2,c_3>0$, which depend on $\mathcal{X},\mathcal{M}_u,\hat{x}_0,\bar{\theta}_0,\Upsilon_0$ such that, if $||\theta-\theta_0||\leq\delta$, the solution $(\hat{x},\bar{\theta},\Upsilon)$ to (18) initialized at $(\hat{x}_0,\bar{\theta}_0,\Upsilon_0)$ and any solution to system (1) with $x(0)\in\mathcal{X}$ and input $u\in\mathcal{M}_u$ are such that $||(x(t)-\hat{x}(t),\hat{\theta}(t)-\theta)||\leq c_1e^{-c_2t}||(\eta(0),\tilde{\theta}(0))||+c_3||\theta-\theta_0||^2$. □

Sketch of proof. The proof follows the same steps as the proof of Theorem 1. Let $t\geq 0$ and $\theta_0\in\mathcal{B}(\theta,\min\{\delta_1,\delta_2\})$ where δ_1,δ_2 come from Assumptions 2 and item (ii) of Theorem 1, respectively. We first notice that instead of (6), we have $[A(t,\theta)-A(t,\theta_0)]\,x(t)=\sum_{i=1}^{n_\theta}\overline{A}_i(t,\theta_0)(\theta_i-\theta_{0,i})x(t)+R(t,\theta,\theta_0)x(t)=\overline{\Lambda}(t,\theta_0,x(t))(\theta-\theta_0)+R(t,\theta,\theta_0)x(t)$ in view of (17) and the definition of $\overline{\Lambda}$. Thus, by following the proof of Theorem 1, we obtain the next equation instead of (9) for $t\geq 0$, $x(t)\in\mathcal{X}$ with $u\in\mathcal{M}_u$,

$$\begin{aligned}
\dot{\eta}(t) &= [A(t,\theta_0)-K(t,\theta_0)C(t)]\,\eta(t)\\
&\quad+\overline{\Lambda}(t,\theta_0,\sigma_{\mathcal{X}}(x(t))-\sigma_{\mathcal{X}}(\hat{x}(t)))(\theta-\theta_0)\\
&\quad+R(t,\theta,\theta_0)x(t),
\end{aligned} \tag{19}$$

and (10) still holds.

Let x belong to $\mathcal{S}_{\mathcal{X}}$ with input $u\in\mathcal{M}_u$, and $(\hat{x},\bar{\theta},\Upsilon)$ be the solution to system (2) initialized at $(\hat{x}_0,\bar{\theta}_0,\Upsilon_0)$. Let $t\geq 0$. The same Lyapunov analysis as in the proof of Theorem 1, leads to the next equation instead of (16) when $||\theta-\theta_0||\leq\delta$ for some $\delta>0$

$$\dot{U}(P,\widetilde{P},\eta,\tilde{\theta}) \quad\leq\quad -\varepsilon||(\eta,\tilde{\theta})||^2+2\eta^TP(t,\theta_0)R(t,\theta,\theta_0)x. \tag{20}$$

Since $P(t,\theta_0)\leq\overline{a}_P\mathbb{I}$, $R(t,\theta,\theta_0)=[\rho_{ij}(t,\theta,\theta_0)]_{(i,j)\in\{1,\ldots,n_x\}^2}$ and $\rho_{ij}(t,\theta,\theta_0)=\mathcal{O}(||\theta-\theta_0||^2)$, and $x(t)\in\mathcal{X}$ according to Assumption 3, there exists $\overline{a}\geq 0$ such that $2||P(t,\theta_0)R(t,\theta,\theta_0)x||\leq\overline{a}||\theta-\theta_0||^2$. Hence,

$$\begin{aligned}
\dot{U}(P,\widetilde{P},\eta,\tilde{\theta}) &\leq -\varepsilon||(\eta,\tilde{\theta})||^2+\overline{a}||\eta||\cdot||\theta-\theta_0||^2\\
&\leq -\varepsilon||(\eta,\tilde{\theta})||^2+\overline{a}||(\eta,\tilde{\theta})||\cdot||\theta-\theta_0||^2
\end{aligned} \tag{21}$$

from which we deduce, using $\overline{a}||(\eta,\tilde{\theta})||\cdot||\theta-\theta_0||^2\leq\frac{\varepsilon}{2}||(\eta,\tilde{\theta})||^2+\frac{2}{\varepsilon}\overline{a}^2||\theta-\theta_0||^4$,

$$\dot{U}(P,\widetilde{P},\eta,\tilde{\theta}) \quad\leq\quad -\tfrac{\varepsilon}{2}||(\eta,\tilde{\theta})||^2+\tfrac{2}{\varepsilon}\overline{a}^2||\theta-\theta_0||^4. \tag{22}$$

Since $\min\{\underline{a}_P,\mu\underline{a}_{\widetilde{P}}\}||(\eta,\tilde{\theta})||^2\leq U(P,\widetilde{P},\eta,\tilde{\theta})\leq(\overline{a}_P+\mu\overline{a}_{\widetilde{P}})||(\eta,\tilde{\theta})||^2$, we deduce from (22) by applying the comparison lemma (see Lemma 3.4 in [15]) that $||(\eta(t),\tilde{\theta}(t))||\leq\overline{c}_1e^{-\overline{c}_2t}||(\eta(0),\tilde{\theta}(0))||+\overline{c}_3||\theta-\theta_0||^2$ for some $\overline{c}_1,\overline{c}_2,\overline{c}_3>0$. The desired result for $\tilde{x}$ then follows by noting $\tilde{x}=\eta+\Upsilon\tilde{\theta}$ and that Υ is bounded according to Claim 1. ■

As already mentioned, Assumption 2 implies that we know a state observer for system (1) when $\theta=\theta_0$. If we would implement this classical (non-adaptive) observer on system (1) with $\theta\neq\theta_0$, the state estimate would converge to x up to an error of the order of $||\theta-\theta_0||$, and, trivially, the parameter estimation error would be $\theta-\theta_0$. Proposition 1 shows that these properties can be improved by employing the adaptive observer (18), which provides estimates with errors of the order of $||\theta-\theta_0||^2$, after a sufficiently long time. We are not

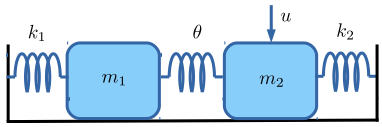

Fig. 1. Coupled mass-spring system.

able to ensure the asymptotic convergence of the estimates of the true values, contrary to Section III-A, because of the perturbative term $R(t, \theta, \theta_0)$ in (17).

IV. Illustrative Example

Consider the mass-spring system with two elements depicted in Figure 1. The system is be modeled by (1) with: $x = (x_1, x_2, x_3, x_4) \in \mathbb{R}^4$, where x_1, x_3 are the displacements of the first and the second mass from their equilibrium and x_2, x_4 are the velocity of the first and the second mass, respectively; u is in the input applied to the second mass; $y = (x_1, x_2) \in \mathbb{R}^2$, which means that only the variables of the first mass are measured. The state matrix is $A = \begin{pmatrix} 0 & 1 & 0 & 0 \\ -\frac{k_1+\theta}{m_1} & -\frac{b_1}{m_1} & \frac{\theta}{m_1} & 0 \\ 0 & 0 & 0 & 1 \\ \frac{\theta}{m_2} & 0 & -\frac{k_2+\theta}{m_2} & -\frac{b_2}{m_2} \end{pmatrix}$, $m_1 = 100$ and $m_2 = 100$ are the masses, $k_1 = 5$ and $k_2 = 10$ are the spring stiffness as shown in Figure 1, $\theta > 0$ is the stiffness of the middle spring, which is assumed to be unknown and $b_1 = 0.1$ and $b_2 = 0.4$ are the damping coefficients. Matrix B is $(0\,0\,0\,\frac{1}{m_2})^\top$ and $C = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}$. Assumption 1 holds with $A_0 = \begin{pmatrix} 0 & 1 & 0 & 0 \\ -\frac{k_1}{m_1} & -\frac{b_1}{m_1} & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -\frac{k_2}{m_2} & -\frac{b_2}{m_2} \end{pmatrix}$ and $A_1 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ -\frac{1}{m_1} & 0 & \frac{1}{m_1} & 0 \\ 0 & 0 & 0 & 0 \\ \frac{1}{m_2} & 0 & -\frac{1}{m_2} & 0 \end{pmatrix}$. Assumption 2 is verified as (A, C) is observable as $\theta \neq 0$. Taking $u(t) = 10\sin(10t)$ for any $t \geq 0$, Assumption 3 is satisfied since the infinity norm of this input is bounded and A is Hurwitz.

We have designed the adaptive observer as in (2) with $\gamma = 10^3$, $\Sigma(t) = 10^3$ and K such that the eigenvalues of $A - KC$ are $(-1, -1.5, -2, -2.5)$. We have run simulations with $\theta = 15$, and we have considered different values of θ_0, $x(0)$, $\hat{x}(0)$, $\bar{\theta}(0)$ and $\Upsilon(0)$. Figures 2 and 3 show that the state and parameter estimates do track the state and parameter of system (1), respectively, when $\theta_0 = 20$, $x(0) = (1, 0, 2, 0)$, $\hat{x}(0) = (0, 0, 0, 0)$, $\bar{\theta}(0) = 0$ and $\Upsilon_0 = 0$. We have then varied the latter. Simulations suggest that the convergence of the adaptive observer may be independent of the initial conditions of x, $\hat{x}$, $\bar{\theta}$ and Υ: only θ_0 seems to matter. The asymptotic convergence of the estimation errors is seen whenever $\theta_0 \in [\theta - 11, \theta + 75]$. This means that θ_0 does not need to be very close to θ for the adaptive observer to work for this example.

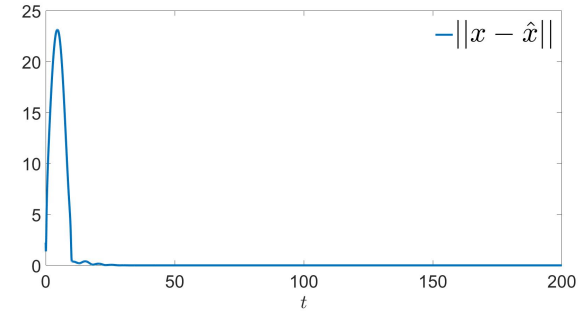

Fig. 2. Norm of the state estimation error.

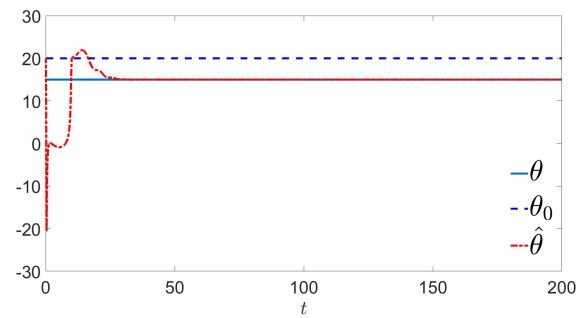

Fig. 3. Parameter θ, θ_0 and the estimate $\hat{\theta}$.

V. Conclusion

We presented an adaptive observer for linear time-varying system whose state matrix $A(t, \theta)$ depends on unknown parameters. When $A(t, \theta)$ is affine in θ, the proposed scheme ensures the exponential convergence of the estimates to the true values provided some initial guess of the unknown parameter is sufficiently closed to the latter and a persistence of excitation holds. When $A(t, \theta)$ is only smooth in θ, a modified version of the observer has been proposed, which ensures the approximate convergence to zero of the estimation errors.

References

[1] B.D.O. Anderson. Stability properties of Kalman-Bucy filters. *Journal of the Franklin Institute*, 291(2):137–144, 1971.

[2] V.I. Arnold. *Ordinary differential equations*. Springer-Verlag, Berlin Heidelberg, Germany, 1992.

[3] G. Besançon. Remarks on nonlinear adaptive observer design. *Systems & Control Letters*, 41(7):271–280, 2000.

[4] G. Besançon, J. De León-Morales, and O. Huerta-Guevara. On adaptive observers for state affine systems. *Int. J. of Control*, 79(6):581–591, 2006.

[5] G. Besançon and A. Ţiclea. On adaptive observers for systems with state and parameter nonlinearities. In *IFAC World Congress, Toulouse, France*, volume 50, pages 15416–15421, 2017.

[6] 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.

[7] Y. M. Cho and R. Rajamani. A systematic approach to adaptive observer synthesis for nonlinear systems. *IEEE Transactions on Automatic Control*, 42:534–537, 1997.

[8] 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.

[9] M. Döhler, Q. Zhang, and L. Mevel. Vibration monitoring by eigenstructure change detection based on perturbation analysis. In *IFAC Symp. on System Identif., Beijing, China*, pages 999–1004, 2015.

[10] L.E. Faibusovich. Algebraic Riccati equation and symplectic algebra. *International Journal of Control*, 43(3):781–792, 1986.

[11] M. Farza, M. M'Saad, T. Maatoug, and M. Kamoun. Adaptive observers for nonlinearly parameterized class of nonlinear systems. *Automatica*, 45:2292–2299, 2009.

[12] M. Farza, M. M'Saad, T. Ménard, A. Ltaief, and T. Maatoug. Adaptive observer design for a class of nonlinear systems. Application to speed sensorless induction motor. *Automatica*, 90:239–247, 2018.

[13] H. He, R. Xiong, X. Zhang, F. Sun, and J. Fan. State-of-charge estimation of the lithium-ion battery using an adaptive extended Kalman filter based on an improved Thevenin model. *IEEE Transactions on Vehicular Technology*, 60(4):1461–1469, 2011.

[14] T. Kailath. *Linear systems*, volume 156. Prentice-Hall Englewood Cliffs, NJ, 1980.

[15] H.K. Khalil. *Nonlinear systems*. Prentice-Hall, Englewood Cliffs, New Jersey, U.S.A., 3rd edition, 2002.

[16] A.J. Krener. The convergence of the extended Kalman filter. In *Lect. Notes in Contr. and Inf. Sciences*, pages 173–182. Springer, 2003.

[17] G. Kresselmeier. Adaptive observers with exponential rate of convergence. *IEEE Transactions on Automatic Control*, AC-22:2–8, 1977.

[18] L. Ljung. Analysis of a general recursive prediction error identification algorithm. *Automatica*, 17(1):89–99, 1981.

[19] A. Loría and E. Panteley. Uniform exponential stability of linear time-varying systems: revisited. *Syst. & Control Lett.*, 47(1):13–24, 2002.

[20] R. Marino and P. Tomei. *Nonlinear control design*. Prentice Hall, New Jersey, U.S.A., 1995.

[21] L. P. Grala Pinto and A. Trofino. State and parameter estimation based on switched observers - an LMI approach. In *American Control Conference, Portland, U.S.A.*, pages 3249–3254, 2014.

[22] R. Postoyan, M.H.A. Hamid, and J. Daafouz. A multi-observer approach for the state estimation of nonlinear systems. In *IEEE Conf. on Dec. and Control, Osaka, Japan*, pages 1793–1798), 2015.

[23] R. Postoyan and Q. Zhang. Local adaptive observer for linear time-varying systems with parameter-dependent state matrices. In *IEEE Conference on Decision and Control (CDC), Miami: U.S.A.*, 2018.

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

[25] H. Shim, Y.I. Son, and J.H. Seo. Semi-global observer for multi-output nonlinear systems. *Systems & Control Letters*, 42(3):233–244, 2001.

[26] L.M. Silverman and B.D.O. Anderson. Controllability, observability and stability of linear systems. *SIAM J. on Contr.*, 6(1):121–130, 1968.

[27] E.D. Sontag and Y. Wang. On characterizations of the input-to-state stability property. *Systems & Control Letters*, 24(5):351–359, 1995.

[28] A. Ţiclea and G. Besançon. Adaptive observer design for discrete time LTV systems. *Int. J. of Control*, 89(12):2385–2395, 2016.

[29] I.Y. Tyukin, E. Steur, H. Nijmeijer, and C. Van Leeuwen. Adaptive observers and parameter estimation for a class of systems nonlinear in the parameters. *Automatica*, 49(8):2409–2423, 2013.

[30] Q. Zhang. Adaptive observer for multi-input-multi-ouput (MIMO) linear time-varying systems. *IEEE Transactions on Automatic Control*, 47(3):525–529, 2002.