Adaptive Observer for MIMO Linear Time Varying Systems

Qinghua Zhang

INSTITUT DE RECHERCHE EN INFORMATIQUE ET SYSTÈMES ALÉATOIRES
Campus de Beaulieu – 35042 Rennes Cedex – France
Tél. : (33) 02 99 84 71 00 – Fax : (33) 02 99 84 71 71
http://www.irisa.fr

Adaptive Observer for MIMO Linear Time Varying Systems

Qinghua Zhang

Thème 4 — Simulation et optimisation
de systèmes complexes
Projet Sigma2

Publication interne n ° 1379 — Janvier 2001 — 29 pages

Abstract: The purpose of adaptive observer is to perform *joint state-parameter estimation* of parameterized state space system. In this paper, we propose a new approach to adaptive observer design for multi-input-multi-output (MIMO) linear time varying (LTV) or *state-affine* systems. It is conceptually simple and computationally efficient. In the case of noise free system with constant unknown parameters, *global exponential convergence* for joint state-parameter estimation is established. In the presence of noises, it is proved that the estimation errors are bounded and converge in the mean to zero if the noises are bounded and have zero means. We also present a unified formulation for some known adaptive observers based on dynamic transformations. This general framework enhances the conceptual simplicity of the proposed approach. Potential applications of the adaptive observer are on-line continuous-time system identification, fault detection and isolation, and adaptive control. Two numerical examples are presented to illustrate the performance of the proposed adaptive observer.

Key-words: adaptive observer, state and parameter estimation, multi-input-multi-output, linear time varying system, state-affine system, continuous-time system, system identification.

Update correcting some equations made in September 2007

Observateur adaptatif pour des systèmes linéaires variables dans le temps multi-entrée-multi-sortie

Résumé : Les observateurs adaptatifs ont pour but d'estimer conjointement les états et les paramètres de systèmes dynamiques. Nous proposons dans ce rapport une nouvelle méthode de conception d'observateur adaptatif pour des systèmes linéaires variables dans le temps multi-entrée-multi-sortie ou pour des systèmes affines en état. C'est une méthode conceptuellement simple et efficace en terme de calcul. Dans le cas où le système considéré est sans bruit et ses paramètres inconnus sont constants, la convergence globale et exponentielle pour l'estimation des états et des paramètres est établie. En présence de bruits, il est démontré que les erreurs d'estimation sont bornées et leur moyennes convergent vers zéro si les bruits sont bornés et de moyennes nulles. Nous présentons aussi une formulation unifiée pour des méthodes existantes pour la conception d'observateurs adaptatifs. A travers ce cadre général, la simplicité conceptuelle de la méthode proposée devient évidente. Des applications potentielles des observateurs adaptatifs sont: l'identification en-ligne de systèmes en temps continu, la détection et le diagnostic de pannes, et la commande adaptative. Deux exemples numériques sont présentés pour illustrer la performance de la méthode proposée.

Mots clés : observateur adaptatif, estimation d'état et de paramètre, multi-entrée-multi-sortie, système linéaire variable dans le temps, système affine en état, système en temps continu, identification.

1 Introduction

State estimation for linear systems has well established solutions with observers and the Kalman filter. Joint estimation of state and some unknown parameters has also several known solutions, with the so-called *adaptive observers* (Kreisselmeier, 1977; Bastin and Gevers, 1988; Marino and Tomei, 1995b; Besançon, 2000). Though the former problem has no particular difficulty for multi-input-multi-output (MIMO) systems, it is not the case for the latter one. As a matter of fact, most published adaptive observers are in single-output form, and their generalization to MIMO systems is not simple. Moreover, because the convergence analysis of adaptive observers is often based on the strictly positive realness of some transfer function, most of them are restricted to linear time invariant (LTI) systems.

In this paper, we propose a new approach to the design of adaptive observers. It is conceptually simple, computationally efficient, and well suited to MIMO linear time varying (LTV) systems.

1.1 Problem statement

We consider state space systems of the form

$$\dot{x}(t) = A(t)x(t) + B(t)u(t) + \Psi(t)\theta \quad \text{(1a)}$$

$$y(t) = C(t)x(t) \quad \text{(1b)}$$

where $x(t) \in \mathbb{R}^n$, $u(t) \in \mathbb{R}^l$, $y(t) \in \mathbb{R}^m$ are respectively the state, input, output of the system, $A(t), B(t), C(t)$ are known time varying matrices of appropriate sizes, $\theta \in \mathbb{R}^p$ is an unknown parameter vector assumed constant unless otherwise specified, $\Psi(t) \in \mathbb{R}^{n\times p}$ is a matrix of known signals. All the matrices $A(t), B(t), C(t), \Psi(t)$ are assumed piecewise continuous and uniformly bounded in time. The problem considered in this paper is the joint estimation of $x(t)$ and θ from measured $u(t), y(t)$ and $\Psi(t)$.

Remark 1 The class of systems considered in this paper includes in fact the so-called *state-affine nonlinear* systems, in the form of

$$\dot{x}(t) = A(t, u, y)x(t) + B(t, u, y)u(t) + \Psi(t, u, y)\theta + \varphi(t, u, y) \quad \text{(2a)}$$

$$y(t) = C(t, u, y)x(t) + D(t, u, y)u(t) \quad \text{(2b)}$$

where the dependence of $A, B, C, D, \Psi, \varphi$ on t, u, y can be nonlinear. As we do not need the time derivatives of $A, B, C, D, \Psi, \varphi$ in the proposed algorithm, their dependence on the known signals u and y can simply be viewed as the dependence on the time t. This class of systems has been considered in (Besançon, 2000) as an example of adaptive observer design for nonlinear systems. Particular cases with constant matrices A, C and single output y have been considered in (Bastin and Gevers, 1988; Marino and Tomei, 1995b) as canonical forms obtained by some nonlinear transformations of nonlinear systems. Note that the extra terms $\varphi(t, u, y)$ and $D(t, u, y)u(t)$ do not increase any difficulty, since we can incorporate them into $B(t, u, y)u(t)$ and $y(t)$, respectively. Therefore, we do not explicitly consider these two terms in this paper for presentation simplicity. □

Remark 2 It is known that nonlinear systems of the form

$$\begin{aligned} \dot{\zeta} &= f(\zeta) + g(\zeta, u) + p(\zeta, u)\theta \\ y &= h(\zeta) \end{aligned}$$

can be transformed into the form of (2) with a coordinate change and output injection if some differential-geometric conditions are satisfied (Marino and Tomei, 1995b; Hammouri and Kinnaert, 1996). These conditions are, however, very restrictive and difficult to check, as they involve multiple Lie-derivatives-brackets and the solution of partial differential equations. □

1.2 Relation with the Kalman filter

A natural idea for joint state-parameter estimation is to consider the extended system by putting the unknown parameters θ into the state. When system (1) is considered, the extended system remains linear, so the Kalman filter can be applied. However, note that even in the case with constant matrices A, B, C, the extended system is *time varying*. It is not easy to guarantee the convergence of the Kalman filter for time varying systems. Application of classical results requires uniform complete observability (Jazwinski, 1970). In practice, it is difficult to check the uniform complete observability of the *extended* system that should take into account some persistent excitation condition. Therefore, the application of the Kalman filter to the extended system is not a trivial problem.

1.3 Motivations

The first motivation for the study of adaptive observers is *on-line system identification*. In this case the main purpose is parameter estimation. Note that, apparently in the considered system (1), the unknown parameters are required to be linear coefficients in front of some measured signals. This requirement is in fact not as restrictive as it seems, since in some situations where it is not apparently satisfied, some transformation may put the system into the required form. For example, consider the single-input-single-output (SISO) system

$$y^{(n)} = a_1 y^{(n-1)} + a_2 y^{(n-2)} + \cdots + a_n y + b_1 u^{(n-1)} + b_2 u^{(n-2)} + \cdots + b_n u$$

with the unknown parameters $\theta = [a_1, \ldots, a_n, b_1, \ldots, b_n]^T$. Arbitrary state space realizations of this system may not be in the form of (1), as unknown parameters appear in front of the derivatives of u and y. However, as pointed out by (Marino and Tomei, 1995b), the following equivalent representation

$$\begin{aligned} \dot{x} &= \begin{bmatrix} 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \\ 0 & 0 & \cdots & 0 \end{bmatrix} x + \begin{bmatrix} y & 0 & \cdots & 0 & u & 0 & \cdots & 0 \\ 0 & y & \cdots & 0 & 0 & u & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & y & 0 & 0 & \cdots & u \end{bmatrix} \theta \\ y &= \begin{bmatrix} 1 & 0 & \cdots & 0 \end{bmatrix} x \end{aligned}$$

fits into the form of (1) (or rather, the form of (2)).

The second motivation for the study of adaptive observers is *adaptive control*. Due to their ability for on-line state and parameter estimation, adaptive observers can be integrated into model-based controllers.

Another motivation is *fault detection and isolation* (FDI). In a model-based approach, faults are typically modeled as parameter changes. Adaptive observers can provide direct and indirect ways for the detection and isolation of parameter changes (Ding and Frank, 1993; Zhang, 2000). The use of continuous-time models, *i.e.*, ordinary differential equations (ODE) is particularly well suited to FDI, as physical models of faulty systems are often in continuous-time.

1.4 Some existing methods

Adaptive observers for linear systems have been studied since 1970's. In (Lüders and Narendra, 1973), an adaptive observer was proposed with a parameter adaptation algorithm suggested by integrating the state estimation error equation. Several years later, three adaptive observers with exponential convergence were proposed in (Kreisselmeier, 1977), each based on the minimization of a particular criterion.

More recent developments of adaptive observers are often based on dynamical transformations putting the original system into some canonical form in which the presence of the unknown parameter θ is simplified to some extent (Bastin and Gevers, 1988; Marino and Tomei, 1995b; Besançon, 2000). Typically, for single output systems, the transformed system has the form

$$\begin{aligned}\dot{z}(t) &= A_o z(t) + Bu + \gamma \xi^T(t)\theta \\ y(t) &= c_o z(t)\end{aligned}$$

where the matrix A_o and the vector c_o are in some special form, $\gamma \in \mathbb{R}^n$ is a constant column vector, $\xi(t) \in \mathbb{R}^p$ is a vector of signals obtained by filtering $u(t), y(t)$. It is important to notice that the parameter vector θ affects the state equation through the scalar product $\xi^T(t)\theta$ and the column vector γ. In Section 4 we will give a unified summary of such methods and compare them with the one proposed in this paper.

In (Marino and Tomei, 1995a), some results on the design of adaptive observers with arbitrary exponential convergence rate are presented. The adaptive observer of (Marino and Tomei, 1995b) is revisited by (Marino and Santosuosso, 1999) for robustness issues.

1.5 Main contributions of this paper

The first contribution of this paper is to propose a new approach to the design of global exponential adaptive observers for MIMO linear time varying systems. It is of wider scope than most existing adaptive observers restricted to single output linear systems with constant matrices A and C. Up to our knowledge, the adaptive observer proposed by (Besançon, 2000) is the only existing one covering also MIMO linear time varying systems (among some more general nonlinear systems). Compared to this adaptive observer, our algorithm has a better convergence result, is computationally more efficient and conceptually simpler. See Section 4 for a more detailed comparison.

Another main contribution is to present a unified formulation for some known adaptive observers based on dynamic transformations. Through this general framework, presented in Section 4, the functioning of these adaptive observers and their relation with the one proposed in this paper become clear.

We also provide some robustness analysis of the proposed adaptive observer in the presence of modeling and measurement noises. In particular, the analysis of the convergence in the mean for noise corrupted systems seems to be the first result of this nature for adaptive observers.

The paper is organized as follows. In Section 2 the proposed adaptive observer is presented. Some robustness issues of the adaptive observer in the presence of noises are discussed in Section 3. Section 4 is devoted to a unified overview of some existing methods through the framework of a general dynamic transformation and to the comparison of these methods with our proposed one. Two numerical examples are presented in Section 5. Finally, some concluding remarks are drawn in Section 6.

2 The proposed adaptive observer

Before formally formulating the proposed adaptive observer, we first present some heuristics.

2.1 Some heuristics

First rewrite (1a) as

$$\dot{x}(t) = [A(t) - K(t)C(t)]x(t) + B(t)u(t) + K(t)y(t) + \Psi(t)\theta$$

with some feedback matrix $K(t)$. It can be seen that two different kinds of "exogenous excitations" contribute to the generation of $x(t)$, namely $B(t)u(t) + K(t)y(t)$ and $\Psi(t)\theta$. Accordingly, let us split $x(t)$ into $x(t) = x_u(t) + x_\theta(t)$ with

$$\begin{aligned}\dot{x}_u(t) &= [A(t) - K(t)C(t)]x_u(t) + B(t)u(t) + K(t)y(t)\\ \dot{x}_\theta(t) &= [A(t) - K(t)C(t)]x_\theta(t) + \Psi(t)\theta\end{aligned}$$

It is easy to estimate $x_u(t)$ with the observer

$$\dot{\hat{x}}_u(t) = [A(t) - K(t)C(t)]\hat{x}_u(t) + B(t)u(t) + K(t)y(t) \tag{3}$$

Let us try to estimate $x_\theta(t)$ with

$$\dot{\hat{x}}_\theta(t) = [A(t) - K(t)C(t)]\hat{x}_\theta(t) + \Psi(t)\hat{\theta}(t) + \omega(t) \tag{4}$$

where $\hat{\theta}(t)$ is an estimate of θ computed somehow and the meaning of $\omega(t)$ will be explained later. Assume that there exists a time varying matrix $\Upsilon(t) \in \mathbb{R}^{n\times p}$ such that

$$\hat{x}_\theta(t) = \Upsilon(t)\hat{\theta}(t) \tag{5}$$

Then equation (4) becomes

$$\dot{\Upsilon}(t)\hat{\theta}(t) + \Upsilon(t)\dot{\hat{\theta}}(t) = [A(t) - K(t)C(t)]\Upsilon(t)\hat{\theta}(t) + \Psi(t)\hat{\theta}(t) + \omega(t)$$

If we let $\omega(t) = \Upsilon(t)\dot{\hat{\theta}}(t)$, then equation (5) can be ensured by generating $\Upsilon(t)$ through

$$\dot{\Upsilon}(t) = [A(t) - K(t)C(t)]\Upsilon(t) + \Psi(t)$$

Notice that $\Upsilon(t)$ is generated from the known signal matrix $\Psi(t)$.

Now let us combine $\hat{x}_u(t)$ and $\hat{x}_\theta(t)$ to estimate $x(t)$ by $\hat{x}(t) = \hat{x}_u(t) + \hat{x}_\theta(t)$. Then, according to (3) and (4), $\hat{x}(t)$ satisfies the equation

$$\dot{\hat{x}}(t) = [A(t) - K(t)C(t)]\hat{x}(t) + B(t)u(t) + K(t)y(t) + \Psi(t)\hat{\theta}(t) + \Upsilon(t)\dot{\hat{\theta}}(t) \qquad (6)$$

where we have used the relation $\omega(t) = \Upsilon(t)\dot{\hat{\theta}}(t)$.

In the adaptive observer presented below, the state estimation follows the scheme of (6). The parameter estimation algorithm and the convergence analysis will be detailed in the following.

2.2 Main result

The result presented below essentially states that, if for any known parameter θ an exponential observer can be designed to estimate the state $x(t)$ of system (1), then under some persistent excitation condition, an adaptive observer can be designed to jointly estimate $x(t)$ and θ.

Now we state some assumptions.

Assumption 1 *Assume that the matrix pair $(A(t), C(t))$ in system* (1) *is such that there exists a bounded time-varying matrix $K(t) \in \mathbb{R}^{n\times m}$ so that the system*

$$\dot{\eta}(t) = [A(t) - K(t)C(t)]\eta(t) \qquad (7)$$

is globally exponentially stable.

Assumption 2 *Let $\Upsilon(t) \in \mathbb{R}^{n\times p}$ be a matrix of signals generated by the ODE system*

$$\dot{\Upsilon}(t) = [A(t) - K(t)C(t)]\Upsilon(t) + \Psi(t) \qquad (8)$$

Assume that $\Psi(t)$ is persistently exciting so that there exist two positive constants δ, T and some bounded symmetric positive definite matrix $\Sigma(t) \in \mathbb{R}^{m\times m}$ such that for all t the following inequality holds

$$\int_t^{t+T} \Upsilon^T(\tau)C^T(\tau)\Sigma(\tau)C(\tau)\Upsilon(\tau)d\tau \geq \delta I \qquad (9)$$

Assumption 1 states that, for any given parameter θ, a state observer can be designed for system (1) with the gain matrix $K(t)$. Assumption 2 is a persistent excitation condition, typically required for system identification.

The proposed adaptive observer is stated in the following theorem.

Theorem 1 *Let $\Gamma \in \mathbb{R}^{p\times p}$ be any symmetric positive definite matrix. Under Assumptions 1 and 2 and for constant θ, the ODE system*

$$\dot{\hat{x}}(t) = A(t)\hat{x}(t)+B(t)u(t)+\Psi(t)\hat{\theta}(t)+\left[K(t)+\Upsilon(t)\Gamma\Upsilon^T(t)C^T\Sigma(t)\right][y(t)-C(t)\hat{x}(t)] \quad (10a)$$

$$\dot{\hat{\theta}}(t) = \Gamma\Upsilon^T(t)C^T(t)\Sigma(t)\,[y(t)-C(t)\hat{x}(t)] \quad (10b)$$

is a global exponential adaptive observer for system (1)*, i.e., for any initial conditions $x(t_0), \hat{x}(t_0), \hat{\theta}(t_0)$ and $\forall\theta \in \mathbb{R}^p$, the errors $\hat{x}(t)-x(t)$ and $\hat{\theta}(t)-\theta$ tend to zero exponentially fast when $t \to \infty$.*

Note that substituting (10b) into (10a) will lead to an equation equivalent to (6).

The proof of this theorem requires the following two lemmas.

Lemma 1 *Let $\phi(t) \in \mathbb{R}^{m\times p}$ be a bounded and piecewise continuous matrix and $\Gamma \in \mathbb{R}^{p\times p}$ be any symmetric positive definite matrix. If there exist positive constants T, α, β such that $\forall t$*

$$\alpha I \leq \int_t^{t+T} \phi^T(\tau)\phi(\tau)d\tau \leq \beta I \quad (11)$$

then the system

$$\dot{z}(t) = -\Gamma\phi^T(t)\phi(t)z(t) \quad (12)$$

is globally exponentially stable.

A proof of this lemma is given in Appendix A.

Lemma 2 *If the autonomous linear time varying system*

$$\dot{\zeta}(t) = F(t)\zeta(t)$$

is globally exponentially stable, $u(t)$ is bounded and integrable, and $u(t) \to 0$ when $t \to \infty$, then $z(t)$ driven by $u(t)$ through the ODE system

$$\dot{z}(t) = F(t)z(t) + u(t)$$

is bounded and also converges to zero. If moreover $u(t)$ vanishes exponentially fast, then $z(t)$ also vanishes exponentially fast.

A proof of this lemma can be found in Appendix B.

Now we are ready to prove Theorem 1.

Proof of Theorem 1. For notation convenience, we do not explicitly write the dependence on t of the variables, though the proof is valid for time varying systems.

Substitute (10b) into (10a) to obtain

$$\dot{\hat{x}} = A\hat{x} + Bu + \Psi\hat{\theta} + K(y - C\hat{x}) + \Upsilon\dot{\hat{\theta}}$$

Let $\tilde{x} = \hat{x} - x$, $\tilde{\theta} = \hat{\theta} - \theta$ and notice that $\dot{\theta} = 0$, then

$$\dot{\tilde{x}} = (A - KC)\tilde{x} + \Psi\tilde{\theta} + \Upsilon\dot{\tilde{\theta}} \tag{13}$$

The key step of the proof is to define the following linear combination of $\tilde{x}$ and $\tilde{\theta}$:

$$\eta(t) = \tilde{x}(t) - \Upsilon(t)\tilde{\theta}(t)$$

then we have

$$\begin{aligned}\dot{\eta} &= (A - KC)(\eta + \Upsilon\tilde{\theta}) + \Psi\tilde{\theta} - \dot{\Upsilon}\tilde{\theta} \\ &= (A - KC)\eta + [(A - KC)\Upsilon + \Psi - \dot{\Upsilon}]\tilde{\theta}\end{aligned}$$

Because Υ is generated by (8), we have simply

$$\dot{\eta} = (A - KC)\eta \tag{14}$$

By assumption system (14) is globally exponentially stable, so $\eta \to 0$ with global and exponential convergence.

Now we should study the behavior of $\tilde{\theta}$. As $\dot{\theta} = 0$, then

$$\begin{aligned}\dot{\tilde{\theta}} &= \Gamma\Upsilon^T C^T \Sigma(y - C\hat{x}) \\ &= -\Gamma\Upsilon^T C^T \Sigma C\tilde{x} \\ &= -\Gamma\Upsilon^T C^T \Sigma C(\eta + \Upsilon\tilde{\theta})\end{aligned} \tag{15}$$

Let us first look at the homogeneous part of system (15), that is

$$\dot{\tilde{\theta}} = -\Gamma\Upsilon^T C^T \Sigma C\Upsilon\tilde{\theta} \tag{16}$$

As Ψ is bounded, Υ generated from the exponentially stable system (8) is also bounded. From the persistent excitation condition (9) and by applying Lemma 1 with $\phi = \Sigma^{\frac{1}{2}}C\Upsilon$, system (16) is globally exponentially stable.

Now from the exponential convergences of η and of system (16), by applying Lemma 2, we obtain the global and exponential convergence to zero of $\tilde{\theta}$ governed by system (15).

Finally, from $\eta \to 0$, $\tilde{\theta} \to 0$ and the fact that Υ is bounded, we conclude $\tilde{x} = \eta + \Upsilon\tilde{\theta} \to 0$ with global and exponential convergence. □

Remark 3 A remarkable point in the proof of this theorem is that the variable $\eta(t)$ is completely decoupled from $\tilde{\theta}(t)$ and $\tilde{x}(t)$. Its stability is guaranteed by simply assuming that $K(t)$ is an observer gain matrix for system (1) when θ is known. This property of $\eta(t)$ does not require the persistent excitation condition. However, when the excitation condition is not satisfied, neither $\tilde{\theta}(t)$ nor $\tilde{x}(t)$ can be guaranteed to converge to zero. In this case, we can only show that $C(t)\tilde{x}(t) \to 0$ (Zhang and Delyon, 2001). It simply means that the prediction error $C(t)\hat{x}(t) - y(t)$ converges to zero. □

Remark 4 Though the total gain matrix in the state estimation equation is $K(t) + \Upsilon(t)\Gamma\Upsilon^T(t)C^T\Sigma$, it is $K(t)$ that mainly plays the role of stabilizing state estimation, or at least stabilizing the transformed dynamics of $\eta(t)$. This is the only requirement on $K(t)$. In principle, $\Sigma(t)$ can be any uniformly bounded positive definite matrix and Γ any constant positive definite matrix. In practice, they are chosen to balance the convergence speeds of state estimation and parameter estimation. □

2.3 Some related results

Obviously, the key point for the design of adaptive observer according to Theorem 1 is how to find the gain matrix $K(t)$. In general this is a difficult task for time varying systems. The Kalman gain matrix can be used for this purpose under some condition as stated in the following.

Corollary 1 *Let $\Phi(t,\tau)$ be the transition matrix associated to the matrix $A(t)$. If the matrix pair $(A(t), C(t))$ is such that there exit $0 < \alpha_o < \beta_o < \infty$, $0 < T < \infty$ and some symmetric positive definite matrix $R(t) \in \mathbb{R}^{m\times m}$, such that for all t*

$$\alpha_o I \leq \int_t^{t+T} \Phi^T(\tau, t+T) C^T(\tau) R^{-1}(\tau) C(\tau) \Phi(\tau, t+T) d\tau \leq \beta_o I \tag{17}$$

then the adaptive observer stated in Theorem 1 can be designed by taking the gain matrix

$$K(t) = P(t) C^T(t) R^{-1}(t) \tag{18a}$$

$$\dot{P}(t) = A(t)P(t) + P(t)A^T(t) - P(t)C^T(t)R^{-1}(t)C(t)P(t) + Q(t) \tag{18b}$$

with some symmetric positive definite matrix $Q(t) \in \mathbb{R}^{n\times n}$ such that

$$\alpha_c I \leq \int_t^{t+T} \Phi(t+T, \tau) Q(\tau) \Phi^T(t+T, \tau) d\tau \leq \beta_c I \tag{19}$$

for some $0 < \alpha_c < \beta_c < \infty$.

This Corollary follows from the fact that, when system (1) with any known θ is uniformly completely observable[1] (as defined by (17)) and uniformly completely controllable regarding the state noise (as defined by (19)), then the Kalman filter is stable (Jazwinski, 1970). This Kalman filter is then an exponential observer of the deterministic system (1) with known θ. Some variants, the so-called Kalman-like observers, can also be used for this purpose (Bornard et al., 1988).

The situation is much simpler when the matrices A, C are constant. In this case, it is well known that the detectability condition is sufficient and necessary for the existence of a stabilizing K matrix. This leads to the following corollary.

Corollary 2 *When the matrix pair (A, C) is constant and satisfies the detectability condition (i.e., the unobservable modes of A are asymptotically stable), then a constant matrix $K \in \mathbb{R}^{n\times m}$ can be chosen to stabilize $A - KC$ and the adaptive observer stated in Theorem 1 can be designed with such a gain matrix.*

The constant gain matrix K can be designed either by pole-placement or through the stationary Riccati equation as in the Kalman filter for LTI systems.

So far in this paper, the unknown parameter θ has been assumed constant. Before closing this section, let us state a result about the tracking ability of the proposed adaptive observer when the unknown parameter is time varying.

[1] In this paper, we follow the definition of uniform complete observability of (Jazwinski, 1970).

Theorem 2 *If in system* (1) *the unknown parameter is time varying and that* $\dot{\theta}(t)$ *is bounded, then, under Assumptions 1 and 2, the adaptive observer* (10) *gives a state estimate* $\hat{x}(t)$ *and a parameter estimate* $\hat{\theta}(t)$ *of system* (1) *with bounded errors.*

Proof of Theorem 2. The proof of this theorem is performed by simply adapting the proof of Theorem 1. As $\dot{\theta} \neq 0$, equation (13) becomes

$$\dot{\tilde{x}} = (A - KC)\tilde{x} + \Psi\tilde{\theta} + \Upsilon\dot{\tilde{\theta}} + \Upsilon\dot{\theta}$$

Consequently, equation (14) becomes

$$\dot{\eta} = (A - KC)\eta + \Upsilon\dot{\theta} \tag{20}$$

Since the autonomous system (14) is exponentially stable and Υ is bounded, η governed by (20) is also bounded.

Now for $\tilde{\theta}$, equation (15) becomes

$$\dot{\tilde{\theta}} = -\Gamma\Upsilon^T C^T \Sigma C(\eta + \Upsilon\tilde{\theta}) - \dot{\theta}$$

Its homogeneous part is still given by equation (16) and is exponentially stable. Therefore $\tilde{\theta}$ driven by the bounded η and $\dot{\theta}$ is also bounded. Finally, $\tilde{x} = \eta + \Upsilon\tilde{\theta}$ is also bounded. □

3 Robustness and convergence in the mean in the presence of noises

In this section we show some properties of the proposed adaptive observer when the considered system is disturbed by noises. For this purpose, consider the noise corrupted system

$$\dot{x}(t) = A(t)x(t) + B(t)u(t) + \Psi(t)\theta(t) + w(t) \tag{21a}$$
$$\dot{\theta}(t) = q(t) \tag{21b}$$
$$y(t) = C(t)x(t) + v(t) \tag{21c}$$

where $w(t) \in \mathbb{R}^n$, $q(t) \in \mathbb{R}^p$ and $v(t) \in \mathbb{R}^m$ are respectively state, parameter and observation noises.

Theorem 3 *If in addition to Assumptions 1 and 2, the noises* $w(t)$*,* $v(t)$ *and* $q(t)$ *in system* (21) *are assumed bounded, then the state and parameter estimation errors of the adaptive observer* (10) *applied to system* (21)*, namely* $\tilde{x}(t) = \hat{x}(t) - x(t)$ *and* $\tilde{\theta}(t) = \hat{\theta}(t) - \theta(t)$*, are also bounded. Moreover, if the noises* $w(t)$*,* $v(t)$ *and* $q(t)$ *have zero means for all* t *and are independent of* $A(t), C(t), K(t), \Sigma(t), \Psi(t)$*, then when* $t \to \infty$*,* $\mathbf{E}\tilde{x}(t)$ *and* $\mathbf{E}\tilde{\theta}(t)$ *converge to zero exponentially fast.*

Proof of Theorem 3. Similarly as in the proof of Theorem 1, define

$$\eta(t) = \tilde{x} - \Upsilon\tilde{\theta}$$

the error system writes

$$\dot{\eta} = (A - KC)\eta - w + Kv + \Upsilon q \tag{22a}$$

$$\dot{\tilde{\theta}} = -\Gamma\Upsilon^T C^T \Sigma C(\eta + \Upsilon\tilde{\theta}) + \Gamma\Upsilon^T C^T \Sigma v - q \tag{22b}$$

Recall that the time varying matrices K, Υ, C, Σ are all uniformly bounded. As the homogeneous part of (22a) is exponentially stable, η is bounded when the noises w, v, q are bounded, according to Lemma 2. In the proof of Theorem 1 we have shown that the homogeneous part of (22b) as shown in (16) is exponentially stable, therefore $\tilde{\theta}$ is also bounded according to Lemma 2.

Under some regularity conditions, we have $\mathbf{E}(\dot{\eta}) = d(\mathbf{E}\eta)/dt$, $\mathbf{E}(\dot{\tilde{\theta}}) = d(\mathbf{E}\tilde{\theta})/dt$. Then using the assumptions that $\mathbf{E}w = 0$, $\mathbf{E}v = 0$, $\mathbf{E}q = 0$ and w, v, q are independent of A, C, K, Σ, Ψ, therefore also independent of Υ, we obtain

$$\frac{d(\mathbf{E}\eta)}{dt} = (A - KC)\mathbf{E}\eta$$

$$\frac{d(\mathbf{E}\tilde{\theta})}{dt} = -\Gamma\Upsilon^T C^T \Sigma C(\mathbf{E}\eta + \Upsilon\mathbf{E}\tilde{\theta})$$

It turns out that the behavior of $\mathbf{E}\eta$ and $\mathbf{E}\tilde{\theta}$ are exactly as that of η and $\tilde{\theta}$ in the deterministic case as shown in (14) and (15). It then follows that $\mathbf{E}\eta \to 0$, $\mathbf{E}\tilde{\theta} \to 0$, and therefore $\mathbf{E}\tilde{x} \to 0$, all with exponential convergence. □

Theorem 3 states that, when the system is corrupted by centered noises, $\mathbf{E}\tilde{\theta}(t) \to 0$, but it does not mean $\tilde{\theta}(t) \to 0$. This result suggests that averaging $\hat{\theta}(t)$ after the transient of the adaptive observer would lead to a better estimate of constant θ.

4 A general formulation of some existing methods

In this section we present a unified overview of some existing methods for adaptive observer design and discuss their relation with the adaptive observer proposed in this paper, all through a general dynamic transformation framework. It is assumed in this section that the unknown parameter θ is constant and the considered system is noise free.

Remark 5 The description of these existing methods given below does not exactly follow their original references, due to our attempt to a unified formulation. □

4.1 General linear dynamic transformation

Once again, let us consider system (1). Most methods for adaptive observer design require a dynamic transformation (sometimes also referred to as filtered transformation) to put the considered system into some special form, for the purpose of *simplifying the presence of the unknown parameters* θ in the transformed system. These transformations, though apparently different from their original presentation, can all be formulated in the following general framework:

$$\dot{\Xi}(t) = F\Xi(t) + G\Psi(t) \tag{23a}$$

$$\Omega(t) = H\Xi(t) \tag{23b}$$

$$z(t) = x(t) - \Omega(t)\theta \tag{23c}$$

where F, G, H are matrices of appropriate sizes defining the dynamic transformation, Ξ and Ω are the state and the output of the transformation, z is the transformed state of the considered system. The sizes of the matrices F, G, H and of $\Xi(t), \Omega(t)$ will be specified in the description of each method. Assume that a transformation with output injection has changed the matrix $A(t)$ of system (1) into $A_o(t)$ satisfying some requirement. Then, the system is further transformed through (23) into

$$\dot{z}(t) = A_o(t)z(t) + B(t)u(t) + [A_o(t)H\Xi(t) + \Psi(t) - HF\Xi(t) - HG\Psi(t)]\theta$$

In the transformed system, the part $A_o(t)z(t) + B(t)u(t)$ remains unchanged, whereas the part involving θ varies depending on the choice of the transformation defined by F, G, H. Appropriate choices of F, G, H can simplify the presence of θ in the transformed system, as shown below with various examples.

In the following we show how the transformations proposed by different authors can be reformulated in this general framework with particular choices of the matrices F, G, H.

4.2 Method A

Let us first look at the method presented in (Marino and Tomei, 1995b). Only single output systems with constant matrices A, C are considered. It is assumed that the pair (A, C) is observable and that some coordinate change with output injection has transformed the considered system into the following form

$$\begin{aligned} \dot{x}(t) &= A_o x(t) + Bu(t) + \Psi(t)\theta + ay(t) && \text{(24a)} \\ y(t) &= c_o x(t) && \text{(24b)} \end{aligned}$$

with

$$A_o = \begin{bmatrix} 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \\ 0 & 0 & \cdots & 0 \end{bmatrix} \qquad c_o = \begin{bmatrix} 1 & 0 & \cdots & 0 \end{bmatrix} \qquad a = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}$$

Note that $ay = ax_1$ could be integrated into the term $A_o x$ so that a becomes the first column of A_o. According to Remark 1 in the introduction, the term ay is treated as $\varphi(y)$ and will be dropped in the following.

The dynamic transformation required by this method corresponds to

$$F = \begin{bmatrix} -\gamma_2 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ -\gamma_{n-1} & 0 & \cdots & 1 \\ -\gamma_n & 0 & \cdots & 0 \end{bmatrix}_{(n-1)\times(n-1)} \qquad G = \begin{bmatrix} -\gamma_2 & 1 & \cdots & \cdots & 0 \\ \vdots & \vdots & \ddots & & \vdots \\ -\gamma_{n-1} & 0 & \cdots & \ddots & 0 \\ -\gamma_n & 0 & \cdots & \cdots & 1 \end{bmatrix}_{(n-1)\times n}$$

$$H = \begin{bmatrix} 0 & \cdots & \cdots & 0 \\ 1 & \cdots & \cdots & 0 \\ \vdots & \ddots & & \vdots \\ 0 & \cdots & \ddots & 0 \\ 0 & \cdots & \cdots & 1 \end{bmatrix}_{n\times(n-1)}$$

where $\gamma_2, \ldots, \gamma_n$ are some chosen values, the sizes of the matrices F, G, H have been indicated. Accordingly, the sizes of $\Xi(t)$ and $\Omega(t)$ are respectively $(n-1)\times p$ and $n \times p$.

After some lengthy computation, system (24) is transformed through (23) into

$$\begin{aligned} \dot{z}(t) &= A_o z(t) + Bu(t) + \gamma\xi^T(t)\theta \\ y(t) &= c_o z(t) \end{aligned}$$

where

$$\gamma = \begin{bmatrix} 1 \\ \gamma_2 \\ \vdots \\ \gamma_n \end{bmatrix} \qquad \xi^T(t) = \Xi_1(t) + \Psi_1(t)$$

with Ξ_1 the first row of Ξ and Ψ_1 the first row of Ψ. Notice the particular feature of this form with a constant vector γ. It means that the scalar excitation $\xi^T(t)\theta$ proportionally affects all the state equations. An adaptive observer is then designed in the form of

$$\begin{aligned} \dot{\hat{z}}(t) &= A_o\hat{z}(t) + Bu(t) + \gamma\xi^T(t)\hat{\theta}(t) + k[y(t) - c_o\hat{z}(t)] \\ \dot{\hat{\theta}}(t) &= \Gamma\xi(t)[y(t) - c_o\hat{z}(t)] \end{aligned}$$

with some gain vector k and a positive definite matrix Γ. In order to ensure the convergence of the adaptive observer, the gain vector k must be chosen such that the transfer function $c_o(sI - A_o + kc_o)^{-1}\gamma$ is *strictly positive real*. See also (Anderson et al., 1986) for convergence issues related to the strictly positive real condition. Because the convergence analysis is based on the strictly positive realness of a transfer function, this method is limited to the case with constant A, C matrices. See (Marino and Tomei, 1995b) for more details.

4.3 Method B

Now let us review the method proposed by (Bastin and Gevers, 1988). Only single output systems with constant matrices A, C are considered. It is assumed that the pair (A, C) is observable and that some coordinate change with output injection has transformed the considered system into the following form

$$\dot{x}(t) = A_o x(t) + Bu(t) + \Psi(t)\theta + ay(t) \tag{25a}$$
$$y(t) = c_o x(t) \tag{25b}$$

with

$$A_o = \begin{bmatrix} 0 & a_{12} \\ 0 & A_{22} \end{bmatrix} \qquad c_o = \begin{bmatrix} 1 & 0 & \cdots & 0 \end{bmatrix} \qquad a = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}$$

where a_{12} is a $1 \times (n-1)$ vector and A_{22} *is an asymptotically stable* $(n-1) \times (n-1)$ *matrix.* As for the previous method, the term $ay = ax_1$ will be dropped in the following.

The dynamic transformation required by this method corresponds to

$$F = A_{22} \qquad G = \begin{bmatrix} 0 & 1 & \cdots & \cdots & 0 \\ \vdots & \vdots & \ddots & & \vdots \\ 0 & 0 & \cdots & \ddots & 0 \\ 0 & 0 & \cdots & \cdots & 1 \end{bmatrix}_{(n-1)\times n} \qquad H = \begin{bmatrix} 0 & \cdots & \cdots & 0 \\ 1 & \cdots & \cdots & 0 \\ \vdots & \ddots & & \vdots \\ 0 & \cdots & \ddots & 0 \\ 0 & \cdots & \cdots & 1 \end{bmatrix}_{n\times(n-1)}$$

where F is $(n-1) \times (n-1)$, the sizes of G, H have been indicated. Accordingly, the sizes of $\Xi(t)$ and $\Omega(t)$ are respectively $(n-1) \times p$ and $n \times p$.

After some computation, system (25) is transformed through (23) into

$$\dot{z}(t) = A_o z(t) + Bu(t) + \gamma\xi^T(t)\theta \qquad (26a)$$

$$y(t) = c_o z(t) \qquad (26b)$$

where

$$\gamma = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \qquad \xi^T(t) = a_{12}\Xi(t) + \Psi_1(t)$$

with Ψ_1 the first row of of Ψ.

Note that, due to the particular value of the vector γ, only the first transformed state equation is affected by the scalar excitation $\xi^T(t)\theta$.

Similarly to the previous case, the adaptive observer is designed in the form of

$$\dot{\hat{z}}(t) = A_o\hat{z}(t) + Bu(t) + \gamma\xi^T(t)\hat{\theta}(t) + k[y(t) - c_o\hat{z}(t)]$$

$$\dot{\hat{\theta}}(t) = \Gamma\xi(t)[y(t) - c_o\hat{z}(t)]$$

As shown in (Bastin and Gevers, 1988), in order to ensure the convergence of the adaptive observer, the transfer function $c_o(sI - A_o + kc_o)^{-1}\gamma$ is also required to be *strictly positive real.* Due to the particular form of A_o, c_o, γ, this requirement is satisfied by choosing a gain vector $k = [k_1, 0 \ldots, 0]^T$ with $k_1 > 0$. Because the convergence analysis is based on the strictly positive realness of a transfer function, this method is limited to the case with constant A, C matrices. See (Bastin and Gevers, 1988) for more details.

4.4 Method C

In (Besançon, 2000), a canonical form for a class of nonlinear systems is proposed for the purpose of adaptive observer design. As an example, the design of adaptive observer for linear time varying (or state-affine) systems in the form of (1) (or of (2)) is considered.

In order to introduce the method of (Besançon, 2000) as applied to system (1), let us first re-examine the canonical form proposed by (Bastin and Gevers, 1988). If we denote

$$z(t) = \begin{bmatrix} z_1(t) \\ \underline{z}(t) \end{bmatrix} \qquad B = \begin{bmatrix} b_1 \\ \underline{B} \end{bmatrix}$$

where z_1 and b_1 are respectively the first row of z and B, then system (26) can be written as

$$\dot{z}_1(t) = a_{12}\underline{z}(t) + b_1 u(t) + \xi^T(t)\theta \quad (27a)$$
$$\dot{\underline{z}}(t) = A_{22}\underline{z}(t) + \underline{B}u(t) \quad (27b)$$
$$y(t) = z_1(t) \quad (27c)$$

A remarkable property of the state equation is that $\underline{z}$ depends only on u. As A_{22} has been designed to be asymptotically stable, $\underline{z}$ can be estimated from u only. Note that z_1 is directly measured by y. Intuitively, in equation (27a) everything is "known" except θ, it is thus possible to estimate θ.

The canonical adaptive observer form of (Besançon, 2000) follows the same idea, but formulated in a more general framework. When it is applied to system (1), the transformation can also be reformulated in the general framework of (23).

Assume that a coordinate change with output injection has put system (1) into the form

$$\dot{x}(t) = A_o(t)x(t) + B(t)u(t) + \Psi(t)\theta \quad (28a)$$
$$y(t) = C_o x(t) \quad (28b)$$

with

$$A_o(t) = \begin{bmatrix} A_{11}(t) & A_{12}(t) \\ A_{21}(t) & A_{22}(t) \end{bmatrix} \qquad C_o = \begin{bmatrix} I_{m\times m} & 0_{m\times \bar{m}} \end{bmatrix}$$

where $\bar{m} = n - m$, $A_{11}(t)$ is a $m \times m$ matrix, $A_{22}(t)$ is a $\bar{m} \times \bar{m}$ *asymptotically stable* matrix[2], $I_{m\times m}$ is the $m \times m$ identity matrix, $0_{m\times\bar{m}}$ is the $m \times \bar{m}$ zero matrix.

Note that for time varying systems, it is not easy to find coordinate changes so that $A_{22}(t)$ is asymptotically stable. Such a coordinate change is proposed in (Besançon, 2000). It involves the inversion of part of the time varying gain matrix of a Kalman like observer. A consequence of this coordinate change is that the implementation of the whole adaptive observer requires the inversion of some time varying matrix.

[2] In this paper, by abuse of terminology, we say that a time varying matrix $A(t)$ is (asymptotically or exponentially) stable when the the autonomous system $\dot{x}(t) = A(t)x(t)$ is (asymptotically or exponentially) stable. Similarly, we talk about the *dynamics of* $A(t)$ in stead of the dynamics of the associated autonomous system.

In order to obtain the adaptive observer form of (Besançon, 2000), let

$$F(t) = A_{22}(t) \qquad G = \begin{bmatrix} 0_{\bar{m}\times m} & I_{\bar{m}\times\bar{m}} \end{bmatrix} \qquad H = \begin{bmatrix} 0_{m\times\bar{m}} \\ I_{\bar{m}\times\bar{m}} \end{bmatrix}$$

where the sizes of F, G and H are respectively $\bar{m} \times \bar{m}$, $\bar{m} \times n$ and $n \times \bar{m}$. Accordingly, the sizes of $\Xi(t)$ and $\Omega(t)$ are respectively $\bar{m} \times p$ and $n \times p$.

Now transform system (28) through (23) into

$$\begin{aligned} \dot{z}(t) &= A_o(t)z(t) + B(t)u(t) + \begin{bmatrix} I_{m\times m} \\ 0_{\bar{m}\times m} \end{bmatrix} [A_{12}(t)\Xi(t) + \Psi_1(t)]\theta \\ y(t) &= C_o z(t) \end{aligned}$$

where Ψ_1 is the first m rows of Ψ.

If we accordingly divide z and B into the first m and the last $\bar{m}$ rows,

$$z(t) = \begin{bmatrix} z_1(t) \\ z_2(t) \end{bmatrix} \qquad B(t) = \begin{bmatrix} B_1(t) \\ B_2(t) \end{bmatrix}$$

then the transformed system writes

$$\dot{z}_1(t) = A_{11}(t)z_1(t) + A_{12}(t)z_2(t) + B_1(t)u(t) + [A_{12}(t)\Xi(t) + \Psi_1(t)]\theta \qquad (29a)$$

$$\dot{z}_2(t) = A_{21}(t)z_1(t) + A_{22}(t)z_2(t) + B_2(t)u(t) \qquad (29b)$$

$$y(t) = z_1(t) \qquad (29c)$$

The adaptive observer proposed by (Besançon, 2000) has the form

$$\begin{aligned} \dot{\hat{z}}_1(t) &= A_{11}(t)\hat{z}_1(t) + A_{12}(t)\hat{z}_2(t) + B_1(t)u(t) + [A_{12}(t)\Xi(t) + \Psi_1(t)]\hat{\theta}(t) + K[y(t) - \hat{z}_1(t)] \\ \dot{\hat{z}}_2(t) &= A_{21}(t)y(t) + A_{22}(t)\hat{z}_2(t) + B_2(t)u(t) \\ \dot{\hat{\theta}}(t) &= \Gamma[A_{12}(t)\Xi(t) + \Psi_1(t)]^T[y(t) - \hat{z}_1(t)] \end{aligned}$$

It is clear from (29b) and (29c) that $z_2(t)$ can be estimated from $u(t)$ and $y(t)$ only, as $A_{22}(t)$ is asymptotically stable. Intuitively, from the estimated $\hat{z}_2$(t) and the other known signals, it is possible to estimate θ from (29a). Accordingly, the proof of the convergence of the adaptive observer presented in (Besançon, 2000) is based on a Lyapounov function of the following form

$$V(\tilde{z}, \tilde{\theta}) = \varepsilon V_1(\tilde{z}_1) + V_2(\tilde{z}_2) + \varepsilon V_3(\tilde{\theta})$$

where V_1, V_2, V_3 are three positive definite functions, $\varepsilon > 0$, $\tilde{z}_1 = \hat{z}_1 - z_1, \tilde{z}_2 = \hat{z}_2 - z_2$, $\tilde{\theta} = \hat{\theta} - \theta$. The negative definiteness of the derivative of $V(\tilde{z}, \tilde{\theta})$ is established for some *sufficiently small* ε. The requirement on this sufficiently small ε intuitively shows that the convergence of $\tilde{z}_2$ precedes those of $\tilde{z}_1$ and $\tilde{\theta}$. Only asymptotic convergence is proved in (Besançon, 2000), not exponential convergence.

Up to our knowledge, this is the only known result on adaptive observer design covering time varying and MIMO systems, prior to the one proposed in this paper. It has the advantage of generality. However, it has the drawback of some kind of convergence singularity related to the presence of ε in the Lyapounov function. Moreover, it requires the inversion of some time varying matrix, and only asymptotic convergence is proved. Clearly, our proposed adaptive observer improves these aspects.

4.5 The proposed adaptive observer

Now let us examine our proposed adaptive observer in the framework of the general dynamic transformation. First perform an appropriate output injection to modify the dynamics of $A(t)$ into $A_o(t) = A(t) - K(t)C(t)$.

Remark that, in contrast to all the above reviewed methods, *no* coordinate change is required here, as *no* particular form of $A_o(t)$ is required in the proposed method. This fact is *especially important for time varying systems* for which time varying coordinate change is computationally expensive.

Our adaptive observer corresponds to the transformation with

$$F(t) = A(t) - K(t)C(t) \qquad G = I_{n\times n} \qquad H = I_{n\times n}$$

where the sizes of F, G and H are all $n \times n$. Accordingly, both $\Xi(t)$ and $\Omega(t)$ have the size $n \times p$.

Then it is easy to check that system (1) is transformed through (23) into

$$\dot{z}(t) = [A(t) - K(t)C(t)]z(t) + B(t)u(t) + K(t)y(t)$$

Remind that, for all the above reviewed methods, the purpose of the dynamic transformation is to simplify the presence of θ in the transformed system. Indeed for the adaptive observer proposed in this paper, this transformation realizes the extreme simplification: *θ is not present at all* in the transformed state equation! This choice should be the most natural one. The other choices all conserve a presence of θ in the transformed state equation, probably by fearing losing the ability to estimate θ. However, as shown in this paper, our choice does allow the estimation of θ, with a simple, general and efficient algorithm.

5 Numerical examples

In this section we present two simulation examples: the proposed adaptive observer is applied to a single link robot arm and to a controlled satellite. The simulations are performed in Simulink with the ODE45 solver.

5.1 A single link robot arm

This example is borrowed from (Marino and Tomei, 1995a). It is a single link robot arm rotating in a vertical plane as illustrated in Figure 1. The equation of motion is

$$I\ddot{q} + \frac{1}{2}mgl\sin q = u$$

where q is the rotation angle, u the input torque, I the moment of inertia, g the gravity constant, m the mass and l the length of the arm.

Let $x_1 = q, x_2 = \dot{q}, y = q, \theta_1 = mgl/(2I), \theta_2 = 1/I$, then the state space model is

$$\begin{aligned}\dot{x}_1 &= x_2\\ \dot{x}_2 &= -\theta_1 \sin y + \theta_2 u\\ y &= x_1\end{aligned}$$

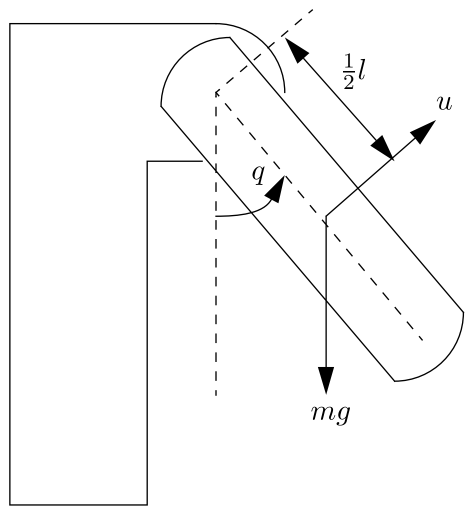

Figure 1: Single link robot arm

which fits into the form of (1) despite the apparent nonlinearity $\sin y$. The proposed adaptive observer is used to estimate the states x_1, x_2 and the parameters θ_1, θ_2.

The same simulation parameters as in (Marino and Tomei, 1995a) are used: $m = 1, l = 1, I = 0.5$. The input signal is $u(t) = 5(\sin 2t + \cos 20t)$. The initial conditions are $x(0) = [1, 1]^T, \hat{x}(0) = [0, 0]^T, \hat{\theta}(0) = [5, 1]^T$. The only difference from (Marino and Tomei, 1995a) is that in our simulation the output *y is corrupted by a Gaussian noise* whose standard deviation is 0.2.

The parameters of the adaptive observer are $K = [1, 1]^T, \Sigma = 10, \Gamma = \mathrm{diag}([2.1, 2])$.

In Figures 2, 3, 4 and 5 are respectively plotted the input-output signals, the simulated state variables, the state estimation errors and the parameter estimates. It can be noticed that after about 3 time units, the convergences of both state and parameter estimation are practically established. Due to the noises added to the output $y(t)$, the estimation errors randomly oscillate around zero instead of tending to zero.

When our simulation result is compared to that of (Marino and Tomei, 1995a), it should be noted that *our simulated output signal is noise corrupted*, whereas it is noise free in (Marino and Tomei, 1995a).

5.2 A controlled satellite

This example comes from (Brockett, 1970). Assume that the nominal orbit of the considered satellite is circular with the radius normalized to 1. The equations of motion of the satellite is linearized around the nominal orbit. Let ω be the nominal angular velocity of the satellite around the earth, then the linearized model writes (see (Brockett, 1970) for

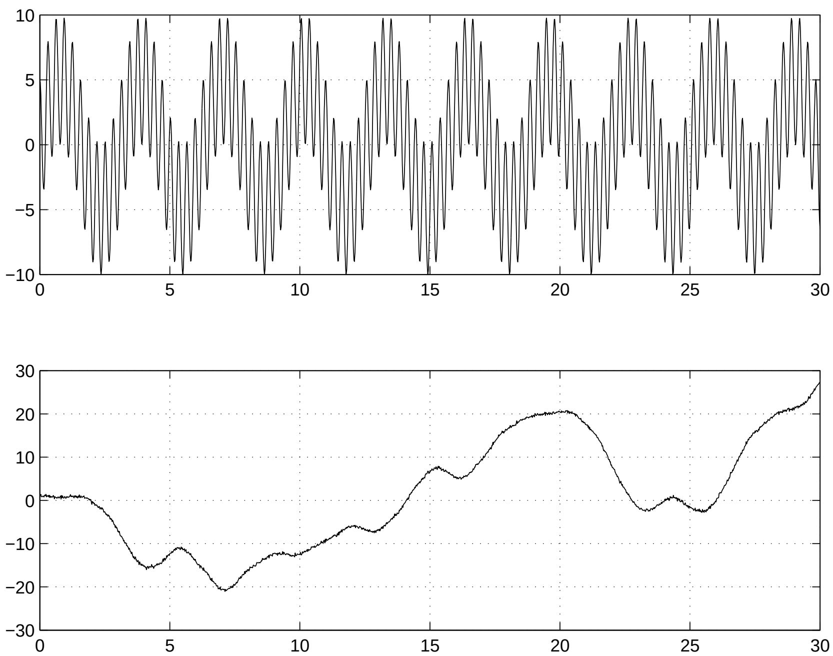

Figure 2: Single link robot arm: simulated input $u(t)$ (top) and output $y(t)$ (bottom).

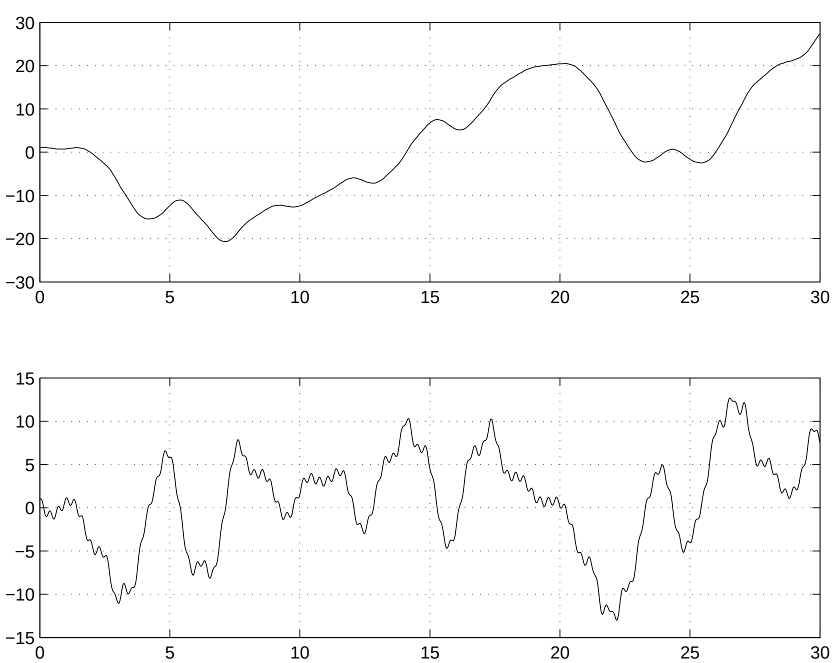

Figure 3: Single link robot arm: simulated states $x_1(t)$ (top) and $x_2(t)$ (bottom).

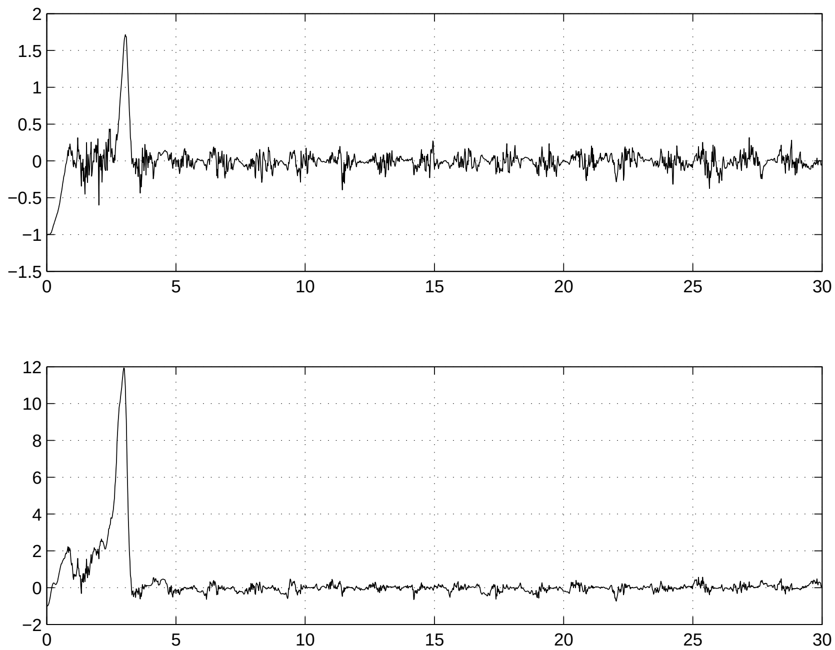

Figure 4: Single link robot arm: state estimation errors $\tilde{x}_1(t)$ (top) and $\tilde{x}_2(t)$ (bottom).

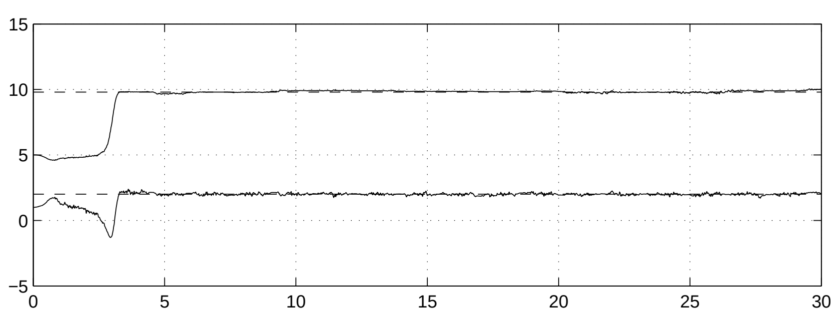

Figure 5: Single link robot arm: parameter estimates $\hat{\theta}_1(t)$ (upper) and $\hat{\theta}_2(t)$ (lower). The true parameter values $\theta_1 = 9.8$ and $\theta_2 = 2$ are shown by the dashed lines.

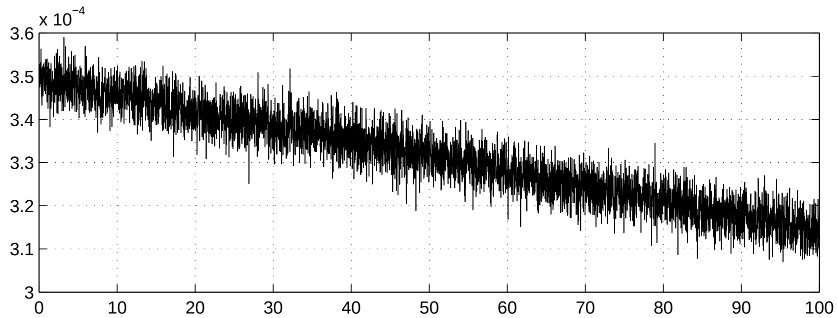

Figure 6: Controlled satellite: the signal $\omega(t)$.

the details)

$$\begin{bmatrix} \dot{x}_1(t) \\ \dot{x}_2(t) \\ \dot{x}_3(t) \\ \dot{x}_4(t) \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 3\omega^2 & 0 & 0 & 2\omega \\ 0 & 0 & 0 & 1 \\ 0 & -2\omega & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \\ x_4(t) \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ \theta_1 & 0 \\ 0 & 0 \\ 0 & \theta_2 \end{bmatrix} \begin{bmatrix} u_1(t) \\ u_2(t) \end{bmatrix}$$

$$\begin{bmatrix} y_1(t) \\ y_2(t) \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \\ x_4(t) \end{bmatrix}$$

where the states $x_1 = r - 1$ (r is the orbit radius, 1 is the normalized nominal orbit radius), $x_2 = \dot{r}$, $x_3 = \varphi - \omega t$ (φ is the rotation angle), $x_4 = \dot{\varphi} - \omega$, u_1 and u_2 are the radial and tangential thrusts with the coefficients θ_1 and θ_2, the outputs y_1 and y_2 are related to distance and angle observations.

In order to simulate a time varying linear system, the value of ω varies linearly from 3.49×10^{-4} to 3.14×10^{-4} over 100 seconds and is disturbed by a Gaussian noise whose standard deviation is 3.49×10^{-6}. Such a signal is shown in Figure 6. The true parameter θ_1 switches between 1 and 0.75, whereas θ_2 between 1.5 and 1.25. The two inputs are square impulses and shown in Figure 7. Each of the two outputs is corrupted by a Gaussian noise whose standard deviation is 0.01.

The initial values are $x(0) = [1, 0, 0, \omega(0)]^T$, $\hat{x}(0) = [0.9, 0, 0, 0.9\omega(0)]^T$, $\hat{\theta}(0) = [0.5, 0.5]^T$.

The adaptive observer parameters are $\Sigma = \text{diag}([1, 1])$, $\Gamma = 5.0 \times 10^2 \text{diag}([2, 2.4])$. $K(t)$ is computed as the Kalman gain (18) with $Q = 2.0 \times 10^{-3} \text{diag}([1, 1])$, $R = 1.0 \times 10^{-4} \text{diag}([1, 1])$.

In Figures 7, 8, 9 and 10 are respectively plotted the input-output signals, the simulated state variables, the state estimation errors and the parameter estimates. It can be observed that, for each parameter change, the convergence of the parameter estimation errors are re-established after a transient less than 10 seconds. Due to the noises added to the output $y(t)$, the estimation errors randomly oscillate around zero instead of tending to zero.

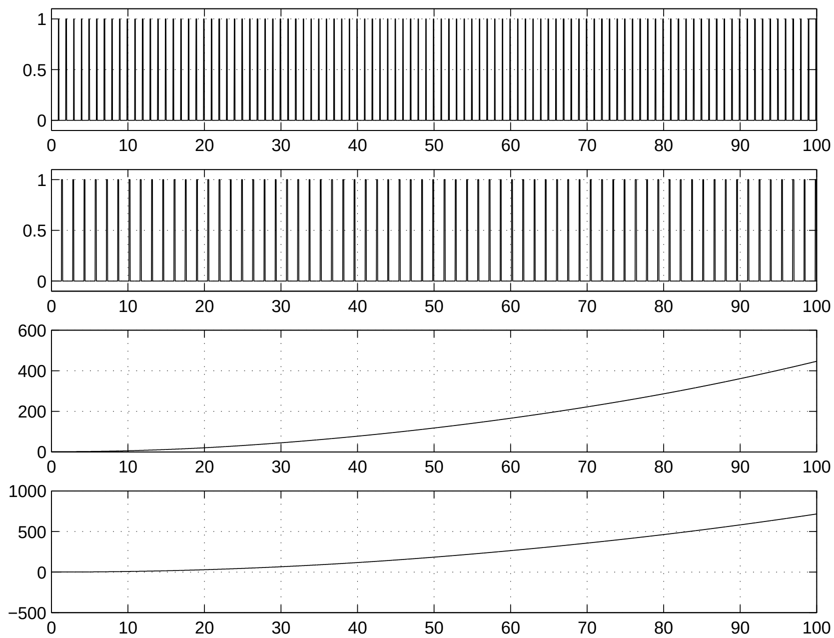

Figure 7: Controlled satellite: simulated input $u_1(t), u_2(t)$ (the two upper plots) and output $y_1(t), y_2(t)$ (the two lower plots).

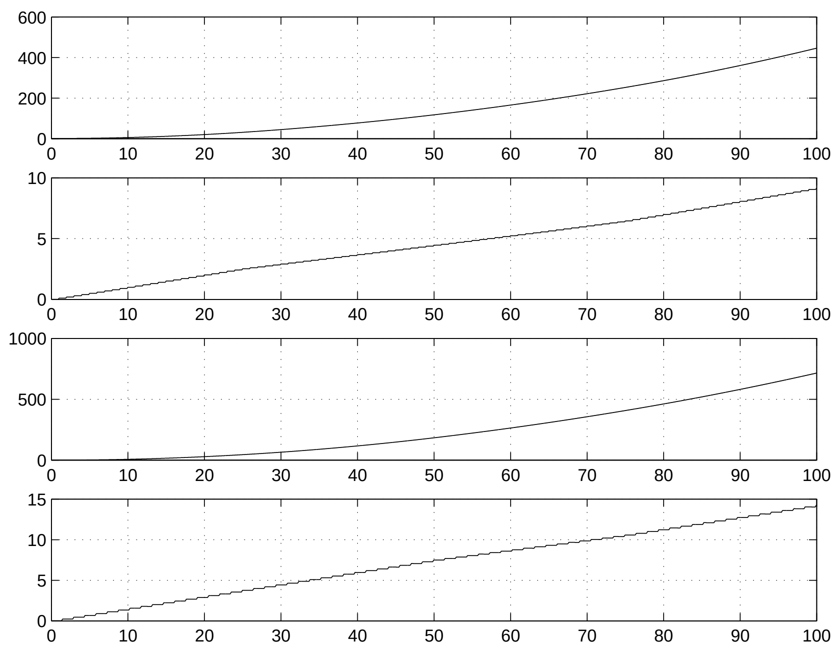

Figure 8: Controlled satellite: simulated states $x_1(t), x_2(t), x_3(t), x_4(t)$.

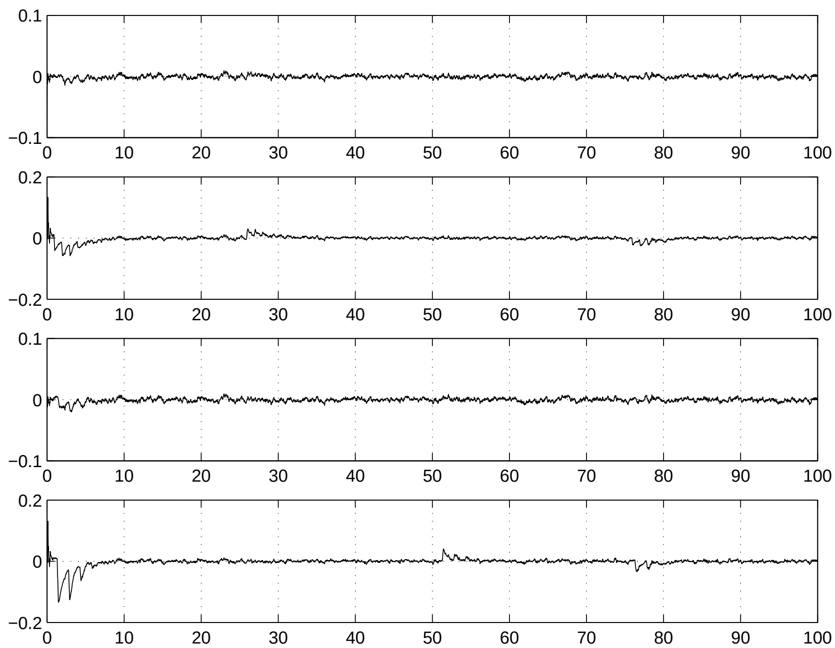

Figure 9: Controlled satellite: state estimation errors $\tilde{x}_1(t), \tilde{x}_2(t), \tilde{x}_3(t), \tilde{x}_4(t)$.

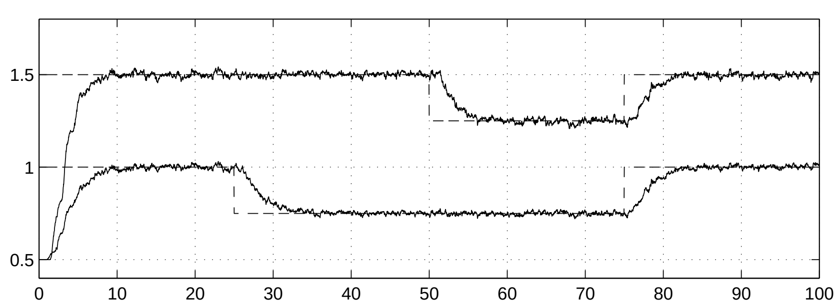

Figure 10: Controlled satellite: parameter estimates $\hat{\theta}_1(t)$ (lower) and $\hat{\theta}_2(t)$ (upper). The true parameter values are shown by the dashed lines.

6 Conclusion

In this paper we have proposed a new approach to global exponential adaptive observer design for MIMO linear time varying systems. In addition to its generality, it is conceptually simple and computationally efficient. The robustness of the proposed adaptive observer and its convergence in the mean in the presence of noises have been established. We have also presented a unified formulation for some known adaptive observers based on dynamic transformations. Through this general framework, the functioning of these adaptive observers and their relation with the one proposed in this paper become clear. Two Numerical examples have been presented to illustrate the performance of the proposed adaptive observer.

As the proposed adaptive observer is mainly based on a deterministic approach, when it is applied a noise corrupted system, the parameter estimation error does not tend to zero, but oscillates around zero. This problem can be remedied by averaging the parameter estimates after the transient, as suggested by the convergence in the mean of the adaptive observer. Another solution would be to extend the results into a fully stochastic framework.

Some related research topics are: the study of adaptive observers in a fully stochastic framework, extensions to nonlinear systems, applications to adaptive control and to fault detection and isolation.

A Proof of Lemma 1

This lemma has been proved for the case with a scalar Γ in (Anderson et al., 1986). Our proof for the case with a symmetric positive definite matrix Γ is adapted from that of (Anderson et al., 1986).

Define the Lyapounov function candidate

$$V(t) = \frac{1}{2} z^T(t)\Gamma^{-1} z(t)$$

then

$$\dot{V}(t) = -z^T(t)\phi^T(t)\phi(t)z(t) \leq 0$$

The autonomous system (12) is thus stable. It requires more efforts to prove its exponential stability.

Let $\Phi(\tau, t)$ be the transition matrix of system (12), then for all τ and t,

$$z(\tau) = \Phi(\tau, t)z(t)$$

$$\begin{aligned} V(t+T)-V(t) &= \int_t^{t+T} \dot{V}(\tau)d\tau \\ &= -\int_t^{t+T} z^T(\tau)\phi^T(\tau)\phi(\tau)z(\tau)d\tau \\ &= -\int_t^{t+T} z^T(t+T)\Phi^T(\tau,t+T)\phi^T(\tau)\phi(\tau)\Phi(\tau,t+T)z(t+T)d\tau \\ &= -z^T(t+T)\left[\int_t^{t+T}\Phi^T(\tau,t+T)\phi^T(\tau)\phi(\tau)\Phi(\tau,t+T)d\tau\right]z(t+T) \end{aligned}$$

According to Lemma 3 stated below and the condition (11), there exists a constant $\alpha_2 > 0$ such that

$$\alpha_2 I \leq \int_t^{t+T} \Phi^T(\tau,t+T)\phi^T(\tau)\phi(\tau)\Phi(\tau,t+T)d\tau$$

therefore

$$V(t+T)-V(t) \leq -\alpha_2\lambda z^T(t+T)\Gamma^{-1}z(t+T)$$

where $\lambda > 0$ is the smallest eigenvalue of Γ. It follows from the definition of $V(t)$ that

$$V(t+T) \leq \frac{1}{1+2\alpha_2\lambda}V(t)$$

For all $\tau \geq T$, denote by $\left[\frac{\tau}{T}\right]$ the integer part of $\frac{\tau}{T}$. Then

$$\begin{aligned} V(t+\tau) &\leq V\left(t+\left[\frac{\tau}{T}\right]T\right) \\ &\leq \left(\frac{1}{1+2\alpha_2\lambda}\right)^{\left[\frac{\tau}{T}\right]} V(t) \\ &\leq \left(\frac{1}{1+2\alpha_2\lambda}\right)^{\frac{\tau}{T}-1} V(t) \end{aligned}$$

Thus $V(t)$ is exponentially decreasing, that implies that $z(t)$ converges to zero exponentially fast. □

Lemma 3 *Let $\phi(t) \in \mathbb{R}^{m\times p}$ be bounded and piecewise continuous. If there exist positive constants $0 < T < \infty$, $0 < \alpha_1 < \beta_1 < \infty$, such that $\forall t$*

$$\alpha_1 I \leq \int_t^{t+T} \phi^T(\tau)\phi(\tau)d\tau \leq \beta_1 I \tag{30}$$

then there exist positive constants $0 < \alpha_2 < \beta_2 < \infty$ such that

$$\alpha_2 I \leq \int_t^{t+T} \Phi^T(\tau,t+T)\phi^T(\tau)\phi(\tau)\Phi(\tau,t+T)d\tau \leq \beta_2 I \tag{31}$$

where $\Phi(\tau,t)$ is the transition matrix of the system

$$\dot{x}(t) = -\Gamma\phi^T(t)\phi(t)x(t)$$

Proof of Lemma 3. The proof sketch presented here follows (Anderson et al., 1986).

Condition (30) is in fact the uniform complete observability condition of the system

$$\dot{x}(t) = 0$$
$$y(t) = \phi(t)x(t)$$

as the transition matrix of the null system is simply the identity matrix. Adding the output feedback $-\Gamma\phi^T(t)y(t) = -\Gamma\phi^T(t)\phi(t)x(t)$ into the state equation conserves the uniform complete observability of the system, that is to say, the system

$$\dot{x}(t) = -\Gamma\phi^T(t)\phi(t)x(t)$$
$$y(t) = \phi(t)x(t)$$

is also uniformly completely observable. It then follows from the definition of uniform complete observability that there exist constants $0 < \alpha_2 < \beta_2 < \infty$ such the inequalities (31) are satisfied. □

B Proof of Lemma 2

The first part of the lemma, that is, $z(t)$ is bounded and converges to zero if $u(t)$ is bounded and tends to zero, has been proved in (Brockett, 1970). Now we prove that if $u(t)$ vanishes exponentially, then $z(t)$ vanishes also exponentially. Let $\Phi(\tau, t)$ be the transition matrix of the autonomous system, then

$$z(t) = \Phi(t, t_0)z(t_0) + \int_{t_0}^{t} \Phi(t, \tau)u(\tau)d\tau$$

$$\|z(t)\| \leq \|\Phi(t, t_0)z(t_0)\| + \int_{t_0}^{t} \|\Phi(t, \tau)\|\|u(\tau)\|d\tau$$

By assumption $u(t)$ vanishes exponentially, therefore $\|u(t)\| \leq C_1 e^{-\lambda_1(t-t_0)}$ for some constants $C_1 > 0$ and $\lambda_1 > 0$. Moreover, the exponential stability of the autonomous system implies that $\|\Phi(t, \tau)\| \leq C_2 e^{-\lambda_2(t-\tau)}$ for some constants $C_2 > 0$ and $\lambda_2 > 0$. Therefore,

$$\begin{aligned}\|z(t)\| &\leq C_2 e^{-\lambda_2(t-t_0)}\|z(t_0)\| + \int_{t_0}^{t} C_2 e^{-\lambda_2(t-\tau)} C_1 e^{-\lambda_1(\tau-t_0)} d\tau \\ &\leq C_2 e^{-\lambda_2(t-t_0)}\|z(t_0)\| + C_1 C_2 e^{-\lambda_2(t-t_0)} \int_{t_0}^{t} e^{(\lambda_2-\lambda_1)(\tau-t_0)} d\tau\end{aligned}$$

If $\lambda_1 \neq \lambda_2$, then

$$\|z(t)\| \leq C_2 e^{-\lambda_2(t-t_0)}\|z(t_0)\| + \frac{C_1 C_2}{\lambda_2 - \lambda_1}\left(e^{-\lambda_1(t-t_0)} - e^{-\lambda_2(t-t_0)}\right)$$

If $\lambda_1 = \lambda_2$, then

$$\|z(t)\| \leq C_2 e^{-\lambda_2(t-t_0)}\|z(t_0)\| + C_1 C_2 e^{-\lambda_2(t-t_0)}(t - t_0)$$

In both cases, $\|z(t)\|$ tends to zero exponentially fast. □

Bibliography

Anderson, B. D., , Bitmead, R. R., Johnson, C. R. J., Kokotovic, P. V., Kosut, R. L., Mareels, I. M., Praly, L., and Riedle, B. D. (1986). *Stability of adaptive systems: passivity and averaging analysis.* signal processing, optimization, and control. MIT Press, Cambridge, Massachusetts.

Bastin, G. and Gevers, M. (1988). Stable adaptive observers for nonlinear time varying systems. *IEEE Trans. on Automatic Control*, 33(7):650–658.

Besançon, G. (2000). Remarks on nonlinear adaptive observer design. *Systems and Control Letters*, 41(4):271–280.

Bornard, G., Couenne, N., and Celle, F. (1988). Regularly persistent observer for bilinear systems. In *New trends in nonlinear control theory*, volume 122 of *Lecture Notes in Control and Info. Sciences*, pages 130–140. Springer-Verlag.

Brockett, R. W. (1970). *Finite dimensional linear systems.* J. Wiley and sons, New York.

Ding, X. and Frank, P. (1993). An adaptive observer-based fault detection scheme for nonlinear dynamic systems. In *Proc. IFAC World Congress'93*, volume 8, pages 307–311, Sydney.

Hammouri, H. and Kinnaert, M. (1996). A new formulation for time-varying linearization up to output injection. *Systems and Control Letters*, 28:151–157.

Jazwinski, A. H. (1970). *Stochastic Processes and Filtering Theory*, volume 64 of *Mathematics in Science and Engineering.* Academic Press, New York.

Kreisselmeier, G. (1977). Adaptive observers with exponential rate of convergence. *IEEE Trans. on Automatic Control*, 22(1):2–8.

Lüders, G. and Narendra, K. S. (1973). An adaptive observer and identifier for a linear system. *IEEE Trans. on Automatic Control*, 18:496–499.

Marino, R. and Santosuosso, G. L. (1999). Robust adaptive observers for nonlinear systems with bounded disturbances. In *38th IEEE Conference on Decision and Control (CDC'99)*, Phoenix, Arizona.

Marino, R. and Tomei, P. (1995a). Adaptive observers with arbitrary exponential rate of convergence for nonlinear systems. *IEEE Trans. on Automatic Control*, 40(7):1300–1304.

Marino, R. and Tomei, P. (1995b). *Nonlinear control design.* Information and system sciences. Prentice Hall, London, New York.

Zhang, Q. (2000). A new residual generation and evaluation method for detection and isolation of faults in nonlinear systems. *International Journal of Adaptive Control and Signal Processing*, 14:759–773.

Zhang, Q. and Delyon, B. (2001). A new approach to adaptive observer design for MIMO systems. In *ACC'2001*, pages 1545–1550, Arlington.