Math. Control Signals Systems (2005) ■: 1–34
DOI 10.1007/s00498-005-0161-8

ORIGINAL ARTICLE

Alessandro Astolfi · Laurent Praly

Global complete observability and output-to-state stability imply the existence of a globally convergent observer

Received: 15 June 2004 / Accepted: 14 April 2005 / Published online: ■■■

Abstract We consider systems which are globally completely observable and output-to-state stable. The former property guarantees the existence of coordinates such that the dynamics can be expressed in observability form. The latter property guarantees the existence of a state norm observer and therefore the possibility of bounding any continuous state functions. Both properties allow to conceptually build an observer from an approximation of an exponentially attractive invariant manifold in the space of the system state and an output driven dynamic extension. The proposed observer provides convergence to zero of the estimation error within the domain of definition of the solutions.

Keywords ■■

1 Introduction

Consider a globally completely observable system whose dynamics can be represented globally by:

A. Astolfi (✉)
Electrical Engineering Department, Imperial College London, Exhibition Road, London SW7 2BT UK
E-mail: a.astolfi@imperial.ac.uk

L. Praly
Centre Automatique et Systèmes, École des Mines de Paris, 35 Rue Saint Honoré, 77305 Fontainebleau, France
E-mail: Laurent.Praly@ensmp.fr

$$\begin{cases} \dot{x}_0 &= x_1, \\ &\vdots \\ \dot{x}_{n-1} &= x_n, \\ \dot{x}_n &= f_n(x_0, \ldots, x_n), \end{cases} \tag{1}$$

where f_n is continuously differentiable. For this system, we wish to establish the existence of a global observer when the only available measurement is:

$$y = x_0. \tag{2}$$

Such a problem has received a lot of attention from a wide variety of view points.

For linear systems, Luenberger [19] showed that an observer can be constructed by simply considering the interconnection of the system for which the state has to be asymptotically reconstructed (the plant), with a linear system of appropriate dimension (the observer), and noting that the interconnected system has an invariant manifold, which can be rendered asymptotically attractive by a proper selection of the dynamics of the observer. The invariant manifold is then used to obtain an approximation of the plant state.

This simple approach can be generalized in several directions. In [6], it has been shown that finite time (and not asymptotic) convergence of the estimation error can be obtained provided two parallel Luenberger observers are used and the two estimates are properly post-processed. In [4] the ideas of Luenberger have been exploited to construct a globally convergent observer for globally completely observable systems imposing some global linear growth conditions on the systems nonlinearities.

Alternatively, for nonlinear globally completely observable systems, the class of so-called high-gain observers (see [8,9] for the general theory and [10,21] for some applications) has been shown to provide asymptotically converging estimates provided the state to be estimated is confined into a known compact set. This approach relies on Lyapunov techniques and does not have a straightforward interpretation in terms of existence of an invariant manifold for the composite system (plant-observer). Nevertheless, it has proved to be an efficient tool to address semi-global stabilization problem, see e.g. [12,15,24].

Finally, in [16] Kreisselmeier and Engell have proposed a different observer design approach for nonlinear globally Lipschitz systems, which is based on the construction of a linear filter, with sufficiently large dimension, and a nonlinear output map, which is the left inverse of a suitably defined observation map. Note that, in this approach, left invertibility of the observation map is ensured by a special selection of the linear filter which processes the available measurements.

For the sake of completeness, let us recall that a very large number of publications have been devoted to a completely different approach from the above and what follows. It exploits the fact that, in one way or another, coordinates can be found so that the dynamics is linear in the unmeasured coordinates [2,11,17,18].

The route followed in this paper to build a global asymptotic observer is related to the classical ideas of Luenberger, and takes its starting point in a contribution[1] of Kazantzis and Kravaris. In [14], they have generalized, to the nonlinear case, Luenberger's early ideas proposed in [19] for linear systems (see also [3, Sect. 7.4, Method II]). However, their analysis is a local one and requires too stringent assumptions aiming at getting an analytic observer. Our intent is to remove these extra assumptions and to deal with the global case. For the latter, we need to add an assumption besides global complete observability, namely output-to-state stability of the system. It must be noted that this restrictive assumption can be relaxed in several directions (see Sect. 6).

Before moving to the technical discussion, we stress a few important points. First, the considered class of systems is very special, i.e. while it is theoretically convenient to deal with uniformly observable systems it is worth noting that several systems arising in applications do not possess this nice property. Second, the paper presents a conceptual result, i.e. it may be extremely difficult to explicitly construct the proposed observer for a given system, even if the system is known to be globally completely observable and output-to-state stable. This is mainly due to the fact that several bounding functions have to be computed. Finally, (numerical) implementation issues are not discussed.

The paper is organized as follows. In Sect. 2 we recall a few definitions, discuss the relation between the proposed approach and the invariant manifold approach of Kazantzis–Kravaris, and formulate the main questions that will be answered in the paper. In Sect. 3 we provide an upper bound for the norm of the state in terms of the state of a norm estimator, and rewrite the system in a new time scale, which is instrumental to deal with systems with unbounded trajectories or with trajectories with finite escape time.

Section 4 provides the (conceptual) observer construction and an in-depth description of its properties. The observer is composed of a series of one-dimensional (linear) filters with gains depending on the state of the norm estimator, and such that certain manifolds are rendered invariant and attractive. The convergence properties of the observer are studied using standard Lyapunov techniques. It must be noted, however, that the Lyapunov function introduced in the paper differs substantially from the Lyapunov functions used in standard (semi-global) high-gain observer design. This latter cannot be used in the present context because we are dealing with a global problem (i.e. we allow unbounded trajectories) and because the observer gains are nonconstant.

In Sect. 5 it is shown that the proposed observer can be implemented in the classical reduced order observer form, and that the gains of such observer can be computed using a simple matrix expression. This also shows that the proposed observer shares the structure of classical high-gain observers, yet it differs substantially in the way gains are assigned and by the introduction of a saturation function acting on scaled nonlinearities. Moreover, it is shown that with a minor modification it is possible to design a full order observer, which again has the same structure of classical high-gain observers.

[1] This contribution has been extended in various ways by Kazantzis and Kravaris themselves but also by Xiao and Krener (see [26] and the references therein). Note, however, that they remain in the same context of looking for a C^∞ observer or at least one admitting a formal power series representation.

Finally, Sect. 6 discusses a way of relaxing the output-to-state-stability assumption, and shows that the weaker property of unboundedness observability is indeed necessary to construct an asymptotic observer for general systems.

2 Preliminaries

To begin with we recall the notion of output-to-state stability.

Assumption 1 ([23]) *The system (1) is output-to-state stable, i.e. there exist C^1 nonnegative functions γ_1, γ_2 and V satisfying:*

$$|x| \le \gamma_2(V(x)), \tag{3}$$

and:

$$\dot{\overbrace{V(x)}} \le -V(x) + \gamma_1(x_0). \tag{4}$$

By Eq. (3) in Assumption 1 and the continuity of f_n, we conclude that there exists a C^1 nondecreasing function γ, lower bounded by 1 say, and satisfying:

$$|x_0| + \cdots + |x_n| + |f_n(x_0, \ldots, x_n)| \le \gamma(V(x)). \tag{5}$$

Later on, specifically in (70), another lower bound for γ will be imposed. With the help of this function, we can define a new time τ as the solution of:[2]

$$\dot{\tau} = \gamma(V(x)), \qquad \tau(0) = 0.$$

Then, by denoting:

$$\mathring{a} = \frac{da}{d\tau} = \frac{\dot{a}}{\gamma(V(x))},$$

the system:

$$\left\{ \begin{array}{rcl} \mathring{x}_0 &=& \dfrac{x_1}{\gamma(V(x))}, \\ &\vdots& \\ \mathring{x}_{n-1} &=& \dfrac{x_n}{\gamma(V(x))}, \\ \mathring{x}_n &=& \dfrac{f_n(x_0, \ldots, x_n)}{\gamma(V(x))} \end{array} \right. \tag{6}$$

is complete. Actually its solutions do not grow faster than $|\tau|$ both forward and backward in the new time τ. As a consequence, for any Hurwitz $p \times p$ matrix A and any p vector B, the function:

$$R(x) = \int_{-\infty}^{0} \exp(-A\tau)\, B\, x_0(\tau)\, d\tau, \tag{7}$$

[2] Note that $\tau(t) \ge t$ for all positive t.

where

$$x = (x_0, \dots, x_n),$$

and $x_0(\tau)$ is the first component of the solution[3] $x(\tau)$ of (6), issued from x, is well defined and continuous (see [5, Théorème 3.149]). Our interest in R comes from the fact that:

$$z = R(x)$$

defines a globally attractive invariant manifold for the system (6) coupled with:

$$\overset{\circ}{z} = A\,z + B\,x_0. \tag{8}$$

Indeed, by computing the limit for h going to 0 of $\frac{R(x(h))-R(x)}{h}$, we can check that R satisfies (8) when evaluated along the solutions of (6). Moreover, integrating (7) by parts, yields:

$$\begin{aligned} R(x) = &-\frac{A^{-(n+1)}}{\gamma(V(x))^n}\left(B\ \gamma(V(x))AB\ \cdots\ \gamma(V(x))^{n-1}A^{n-1}B\right)\begin{pmatrix} x_n \\ x_{n-1} \\ \vdots \\ x_1 \end{pmatrix} \\ &-A^{-1}B\,x_0 - \int_{-\infty}^{0} \exp(-A\tau)\left(\frac{\overset{\circ}{\overline{\gamma(V(x))}}}{\gamma(V(x)}\sum_{i=1}^{n}\frac{iBA^{-(i+1)}x_i}{\gamma(V(x))^i}\right. \\ &\left. - \frac{A^{-(n+1)}Bf_n(x_0(\tau),\dots,x_n(\tau))}{\gamma(V(x))^{n+1}}\right)d\tau. \end{aligned}$$

It follows that, if the pair (A, B) is controllable and $p \geq n$, we may expect that, possibly by modifying γ to make the second line negligible, the map $(x_0, \overline{x}) \mapsto (x_0, R(x_0, \overline{x}))$ is left invertible, with $\overline{x}$ collecting the unmeasured components of x, i.e.

$$\overline{x} = (x_1, \dots, x_n).$$

In such a case there would exist a function S defined on the image of this map, subset of $\mathbb{R} \times \mathbb{R}^n$, and satisfying:

$$S(x_0, R(x_0, \overline{x})) = \overline{x} \qquad \forall\,(x_0, \overline{x}).$$

Furthermore, we may expect that the function S can be extended into a uniformly continuous function $\mathfrak{S}$ defined on $\mathbb{R} \times \mathbb{R}^n$. For instance, as shown in [19] (see also [3, Theorem 7.10]), $\mathfrak{S}$ does exist when f_n is a linear function, the pair (A, B) is controllable, and the spectrum of A and that of the system (1) are separated. The existence of $\mathfrak{S}$ is also established, in [14], locally around the origin assumed to be an equilibrium point of (1), under the assumption that f_n is analytic, the pair (A, B) is controllable and a more restrictive condition of spectral separation.

[3] A less ambiguous and more standard notation would be $X(x, t)$. But, at this point, this would make the following less readable.

Since the set $\{(z,x) : z = R(x)\}$ is exponentially attractive, the existence and the uniform continuity of $\mathfrak{S}$ imply that for each solution $(x(\tau), z(\tau))$, one has:

$$\lim_{\tau\to+\infty} (\mathfrak{S}(x_0(\tau), z(\tau)) - \overline{x}(\tau)) = 0.$$

This says that:

$$\begin{cases} \overset{\circ}{z} = A\,z + B\,x_0, \\ \widehat{\overline{x}} = \mathfrak{S}(x_0, z), \end{cases}$$

with the new time τ, or, if $\gamma(V(x))$ were known,

$$\begin{cases} \dot{z} = \gamma(V(x))\,[A\,z + B\,x_0]\,, \\ \widehat{\overline{x}} = \mathfrak{S}(x_0, z), \end{cases}$$

with the initial time t, there is an observer of $\overline{x}$, with $\widehat{\overline{x}}$ converging to $\overline{x}$, as the new time τ goes to infinity. In terms of the initial time t, this says that the convergence occurs in infinite time if there is no finite escape time, or at the time of the escape if there is a finite escape time. This observer, but with $\gamma(V(x)) = 1$, is the one presented by Luenberger in [19] for linear systems and by Kazantzis and Kravaris in [14] locally, for nonlinear systems.

One of the objectives of this paper is to round or solve the problems left in our way in the above presentation. These problems are:

1. How to get an upper-bound of $\gamma(V(x))$ expressed from only the knowledge of x_0?
2. How to modify γ and the system (8) in order to enforce the existence of a uniformly continuous function $\mathfrak{S} : \mathbb{R} \times \mathbb{R}^m \to \mathbb{R}^n$ satisfying:

$$\mathfrak{S}(x_0, R(x_0, \overline{x})) = \overline{x} \qquad \forall\,(x_0, \overline{x})? \tag{9}$$

3. How to get an expression of $\mathfrak{S}$?

The first problem is addressed in Sect. 3. The other two in Sect. 4. This will allow us to exhibit C^1 functions $\mathfrak{f}$ and $\mathfrak{h}$ such that the n-dimensional system:

$$\begin{cases} \dot{\mathcal{X}} = \mathfrak{f}(\mathcal{X}, y), \\ \widehat{\overline{x}} = \mathfrak{h}(\mathcal{X}, y) \end{cases} \tag{10}$$

provides an estimate $\widehat{\overline{x}}$ converging in the new time τ to the actual unmeasured state component $\overline{x}$.

3 An upper-bound for $\gamma(V(x))$

In this section we show how it is possible, exploiting Assumption 1, to obtain an upper-bound for $\gamma(V(x))$. For, we follow the norm-estimator idea proposed in [23] and [13, Lemma 3.1]. Equation (4) in Assumption 1 states that V satisfies:

$$\dot{\overbrace{V(x)}} \leq -V(x) + \gamma_1(x_0).$$

Therefore, let w be a solution of the system:

$$\dot{w} = -w + \gamma_1(x_0) \tag{11}$$

with positive initial condition. For a solution $(x(t), w(t))$ of (1),(11), issued from (x, w), one has:

$$V(x(t)) \leq w(t) + [V(x) - w]_+ \exp(-t) \tag{12}$$

for all t for which this solution exists and with the notation:

$$r_+ = \max\{r, 0\}.$$

If $x(t)$ is right as maximally defined in $[0, t_0)$, then $w(t)$ is defined at least in the same interval. Moreover the following holds.

1. If t_0 is infinite, then, because of the exponential decay in (12), there exists t_v, depending on (x, w), satisfying:

$$V(x(t)) \leq w(t) + 1 \qquad \forall t \in [t_v, +\infty).$$

2. If t_0 is finite, then from inequality (3) in Assumption 1,

$$\lim_{t \to t_0} V(x(t)) = +\infty.$$

This implies the existence of a time t_v, depending on (x, w), satisfying:

$$\max\{V(x) - w, 0\} \exp(-t) \leq \frac{1}{2} V(x(t)) \qquad \forall t \in [t_v, t_0).$$

Hence, by inequality (12), this yields:

$$V(x(t)) \leq 2\, w(t) \qquad \forall t \in [t_v, t_0).$$

From these two cases we conclude that, for each solution, there exists a new time $\tau_v \quad (= \tau(t_v))$ satisfying:

$$V(x(\tau)) \leq w(\tau) + [V(x) - w]_+ \qquad \forall \tau \in [0, \tau_v], \tag{13}$$
$$\leq 2\, w(\tau) + 1 \qquad \forall \tau \in [\tau_v, \infty). \tag{14}$$

Note that τ_v depends on the initial condition (x, w) of the solution.

To simplify the forthcoming notations, we associate, to any nondecreasing function $c : \mathbb{R}_+ \to \mathbb{R}$, a function $c_* : \mathbb{R}_+ \to \mathbb{R}_+$ defined as:

$$c_*(r) = c(r) - c(0).$$

Observe that, if a and b are arbitrary nonnegative real numbers, then either $a-b \geq b$ and we have:

$$\begin{aligned} c(a) - c(0) = c_*((a-b)+b) &\leq c_*(2(a-b))) \\ &\leq c(2b) - c(0) + c_*(2(a-b)), \end{aligned}$$

or $b \geq [a-b]_+$ and we have:

$$c(a) \leq c(2b).$$

This says that we have the inequality

$$c(a) \leq c(2b) + c_*(2[a-b]_+), \quad \forall (a,b) \in \mathbb{R}_+^2. \tag{15}$$

So, in particular we have

$$\gamma(V(x)) \leq \gamma(4w+2) + \gamma_*(2[V(x)-2w-1]_+) \tag{16}$$

and, by condition (5),

$$\frac{|x_0| + \cdots + |x_n| + |f_n(x_0, \ldots, x_n)|}{\gamma(4w+2)} \leq 1 + \gamma_*(2[V(x)-2w-1]_+). \tag{17}$$

Also, with (14), for each solution of (6),(11) (with (11) considered in the new time τ), we obtain the inequality:

$$\begin{aligned} \frac{\gamma(V(x(\tau)))}{\gamma(4w(\tau)+2)} &\leq 1 + \gamma_*(2[V(x)-2w-1]_+) \quad \forall \tau \in [0, \tau_v], \\ &\leq 1 \quad \forall \tau \in [\tau_v, \infty), \end{aligned} \tag{18}$$

It follows that $\gamma(4w+2)$ is a good candidate to replace $\gamma(V(x))$. This leads us to the notation:

$$\widehat{\gamma}(w) \geq \gamma(4w+2) \tag{19}$$

and another new time $\widehat{\tau}$ as a solution of

$$\dot{\widehat{\tau}} = \widehat{\gamma}(w) \quad , \qquad \widehat{\tau}(0) = 0.$$

Note that

$$\overset{\circ}{\widehat{\tau}} = \frac{\widehat{\gamma}(w)}{\gamma(V(x))}.$$

Hence, by condition (18), along any solution, we have

$$\limsup_{\tau \to +\infty} \frac{\widehat{\tau}(\tau)}{\tau} \geq 1.$$

This implies that any boundedness or convergence result established with the time $\widehat{\tau}$ holds also with the time τ. Moreover, denoting:

$$\overset{*}{a} = \frac{da}{d\widehat{\tau}} = \frac{\dot{a}}{\widehat{\gamma}(w)},$$

we conclude that the system:

$$\left\{\begin{array}{rl} \overset{*}{x}_0 & = \frac{x_1}{\widehat{\gamma}(w)}, \\ & \vdots \\ \overset{*}{x}_{n-1} & = \frac{x_n}{\widehat{\gamma}(w)}, \\ \overset{*}{x}_n & = \frac{f_n(x_0,\dots,x_n)}{\widehat{\gamma}(w)}, \\ \overset{*}{w} & = -\frac{w-\gamma_1(x_0)}{\widehat{\gamma}(w)} \end{array}\right. \tag{20}$$

is complete.

4 Existence and construction of $\mathfrak{S}$

In this section we provide an explicit construction of a global observer for system (1), provided it satisfies Assumption 1. We first show how the observer can be recursively constructed using a set of *one-dimensional filters*, and then we study in details its convergence properties.

4.1 Observer design

From the above discussion, it should be clear that if there exists a uniformly continuous function $\mathfrak{S}$ satisfying Eq. (9) and if this function were known, then we would have an observer asymptotically converging in the new time τ. The route followed to prove the existence and to express $\mathfrak{S}$ is actually to modify (8). This modification is built, in what follows, step by step with in particular the objective of getting the function R as a linear map with a triangular representation in the $\overline{x}$ coordinates.

4.1.1 Estimate of x_1

Consider the system:

$$\dot{z}_1 = -a_1\, z_1 + b_1\, x_0 - u_1 \tag{21}$$

where a_1, b_1 and u_1 remain to be defined. In particular, u_1 is an extra term added to (8). This equation gives readily, for any C^1 function r_{10},

$$\begin{aligned} \overset{\cdot}{\overbrace{z_1 - r_{10}x_0 + x_1}} &= -a_1\, z_1 + b_1\, x_0 - u_1 - \dot{r}_{10}\, x_0 - r_{10}\, x_1 + x_2, \\ &= -a_1 \left(z_1 - \frac{b_1 - \dot{r}_{10}}{a_1}\, x_0 + \frac{r_{10}}{a_1}\, x_1 \right) + x_2 - u_1. \end{aligned}$$

Therefore, selecting:

$$u_1 = x_2 \tag{22}$$

and:

$$\frac{b_1 - \dot{r}_{10}}{a_1} = r_{10} = a_1,$$

i.e.:

$$b_1 = \dot{a}_1 + a_1^2, \tag{23}$$

yields:

$$\overset{\cdot}{\overbrace{z_1 - a_1 x_0 + x_1}} = -a_1 \left(z_1 - a_1 x_0 + x_1\right)$$

or equivalently:

$$\overset{*}{\overbrace{z_1 - a_1 x_0 + x_1}} = -\frac{a_1}{\widehat{\gamma}(w)} \left(z_1 - a_1 x_0 + x_1\right)$$

It follows that the set $\{(z_1, x) : z_1 = a_1 x_0 - x_1\}$ is an invariant manifold of (1),(21), which is exponentially attractive in the original time t and the new time $\widehat{\tau}$ provided the ratio $\frac{a_1}{\widehat{\gamma}(w)}$ is lower-bounded away from zero. This leads to propose an estimate for x_1 in the form:

$$\widehat{x}_1 = a_1\, x_0 - z_1 .$$

The only problem with the above discussion is that Eq. (22) does not provide a legitimate choice for u_1, since x_2 is not measured. Therefore, we continue the design leaving u_1 unspecified. However, for uniformity of notations, let:

$$u_1 = v_1,$$

and record that the best choice for v_1 is:

$$v_1 = x_2 .$$

On the other hand, if we restrict a_1 to depend only on w, Eq. (23) can indeed be realized as (see (11)):

$$b_1(w, x_0) = a_1'(w) \left[-w + \gamma_1(x_0)\right] + a_1(w)^2 .$$

In conclusion, a *candidate* estimate for x_1 is obtained from the system:

$$\begin{cases} \dot{z}_1 = -a_1\, z_1 + [\dot{a}_1 + a_1^2]\, x_0 - v_1, \\ \widehat{x}_1 = a_1\, x_0 - z_1 . \end{cases}$$

4.1.2 Estimate of x_2

With the estimate $\widehat{x}_1$ of x_1 available, we introduce a second system:

$$\dot{z}_2 = -a_2\, z_2 + b_2\, x_0 - u_2 \tag{24}$$

where again a_2, b_2 and u_2 remain to be defined. For any C^1 functions r_{20} and r_{21}, one has:[4]

$$\begin{aligned} &\dot{\overbrace{z_2 - r_{20}x_0 - r_{21}x_1 + x_2}} \\ &= -a_2\, z_2 + b_2\, x_0 - u_2 - \dot{r}_{20}\, x_0 - [r_{20} + \dot{r}_{21}]\, x_1 \\ &\quad -r_{21}\, x_2 + x_3, \\ &= -a_2 \left(z_2 - \frac{b_2 - \dot{r}_{20}}{a_2}\, x_0 + \frac{r_{20}}{a_2}\, x_1 + \frac{r_{21}}{a_2}\, x_2 \right) \\ &\quad -\dot{r}_{21}\, x_1 + x_3 - u_2. \end{aligned} \tag{25}$$

This shows that selecting:

$$u_2 = x_3 - \dot{r}_{21}\, x_1 \tag{26}$$

and:

$$\begin{aligned} \frac{r_{21}}{a_2} &= 1, \\ \frac{r_{20}}{a_2} &= -r_{21}, \\ \frac{b_2 - \dot{r}_{20}}{a_2} &= r_{20}, \end{aligned}$$

i.e.

$$\begin{aligned} r_{21} &= a_2, \\ r_{20} &= -a_2^2, \\ b_2 &= -a_2^3 - 2a_2\dot{a}_2, \end{aligned}$$

the set $\{(z_2, x) : z_2 = -a_2^2 x_0 + a_2 x_1 - x_2\}$ is an invariant manifold of (1),(24). This leads to propose an estimate for x_2 in the form:

$$\widehat{x}_2 = -a_2^2\, x_0 + a_2\, x_1 - z_2.$$

[4] Another way of writing Eq. (25) is:

$$\dot{\overbrace{z_2 - r_{20}x_0 - r_{21}x_1 + x_2}} = -a_2 \left(z_2 - \frac{b_2 - \dot{r}_{20}}{a_2}\, x_0 + \frac{r_{20} + \dot{r}_{21}}{a_2}\, x_1 + \frac{r_{21}}{a_2}\, x_2 \right) + x_3 - u_2.$$

This leads to the choice:

$$r_{21} = a_2, \qquad r_{20} = -\dot{a}_2 - a_2^2, \qquad b_2 = -a_2^3 - 3a_2\dot{a}_2 - \ddot{a}_2.$$

The drawback of this selection is that b_2, and therefore the system (24), involves $\ddot{a}_2$. Indeed, if a_2 is a function of w, then $\dot{a}_2$ depends on (w, x_0) and $\ddot{a}_2$ on (w, x_0, x_1), but x_1 is unknown. It follows that, in systems like (24), we can allow the presence only of a_i and $\dot{a}_i$ and no higher order derivatives whenever a_i is allowed to depend on w.

Unfortunately, such an estimate involves x_1 which is unknown. However, the estimate $\widehat{x}_1$ is available, therefore $\widehat{x}_2$ can be defined as:

$$\widehat{x}_2 = -a_2^2\, x_0 + a_2\, \widehat{x}_1 - z_2.$$

As in the design of the observer for x_1, the problem we are facing is that u_2 defined in Eq. (26) involves x_1 and x_3 which are unknown. However, let:

$$u_2 = v_2 - \dot{r}_{21}\, \widehat{x}_1,$$

and note that the best choice for v_2 is:

$$v_2 = x_3.$$

To sum up, a *candidate* estimate for x_2 is obtained from the system:

$$\begin{cases} \dot{z}_2 = -a_2\, z_2 - [a_2^3 + 2a_2\dot{a}_2]\, x_0 + \dot{a}_2\, \widehat{x}_1 - v_2, \\ \widehat{x}_2 = -a_2^2\, x_0 + a_2\, \widehat{x}_1 - z_2. \end{cases}$$

where (as already remarked), if we choose a_2 as a function of w only, we have:

$$\dot{a}_2 = a_2'(w)\left[-w + \gamma_1(x_0)\right].$$

4.1.3 Estimate of x_i

Proceeding along the same lines outlined above for i ranging from 3 to n, we design an observer for x_i, from the system:

$$\dot{z}_i = -a_i\, z_i + b_i\, x_0 - u_i. \tag{27}$$

For any C^1 functions r_{ij}, we obtain:

$$\begin{aligned} \overbrace{z_i - \sum_{j=0}^{i-1} r_{ij} x_j + x_i}^{\cdot} = {} & -a_i\, z_i + b_i\, x_0 - u_i - \dot{r}_{i0}\, x_0 \\ & - \sum_{j=1}^{i-1} [r_{i(j-1)} + \dot{r}_{ij}]\, x_j - r_{i(i-1)}\, x_i + x_{i+1}, \\ = {} & -a_i \left(z_i - \frac{b_i - \dot{r}_{i0}}{a_i}\, x_0 + \sum_{j=1}^{i-1} \frac{r_{i(j-1)}}{a_i}\, x_j + \frac{r_{i(i-1)}}{a_i}\, x_i \right) \\ & - \sum_{j=1}^{i-1} \dot{r}_{ij}\, x_j + x_{i+1} - u_i, \end{aligned}$$

where, to simplify the notations for $i = n$, we let formally,

$$x_{n+1} = f_n(x_0, \ldots, x_n).$$

Hence, choosing:

$$u_i = x_{i+1} - \sum_{j=1}^{i-1} \dot{r}_{ij}\, x_j$$

and:

$$\begin{aligned} \frac{r_{i(i-1)}}{a_i} &= 1, \\ \frac{r_{i(j-1)}}{a_i} &= -r_{ij} \qquad j \in \{1, \ldots, i-1\}, \\ \frac{b_i - \dot{r}_{i0}}{a_i} &= r_{i0}, \end{aligned}$$

i.e.:

$$\begin{aligned} r_{ij} &= -(-a_i)^{i-j}, \\ b_i &= [i\dot{a}_i + a_i^2]\,(-a_i)^{i-1}, \end{aligned}$$

the set:

$$\left\{ (z_i, x) : z_i = -\sum_{j=0}^{i-1} (-a_i)^{i-j} x_j - x_i \right\}$$

defines an invariant manifold of (1),(27). This motivates an observer for x_i in the form:

$$\left\{ \begin{aligned} \dot{z}_i &= -a_i\, z_i + [i\dot{a}_i + a_i^2]\,(-a_i)^{i-1}\, x_0 + \dot{a}_i \sum_{j=1}^{i-1} (i-j)(-a_i)^{i-j-1} \widehat{x}_j - v_i, \\ \widehat{x}_i &= -(-a_i)^i x_0 - \sum_{j=1}^{i-1} (-a_i)^{i-j} \widehat{x}_j - z_i, \end{aligned} \right. \tag{28}$$

where the best choice for v_i is:

$$v_i = x_{i+1}.$$

4.2 Observer properties

In this section we study the properties of the proposed observer, whose generic expression is given by Eq. (28).

To state our main key technical result, we need to introduce the *manifold error* ε_i, defined as:

$$\varepsilon_i = z_i + (-a_i)^i x_0 + \sum_{j=1}^{i-1} (-a_i)^{i-j} x_j + x_i. \tag{29}$$

It can be used as a coordinate in place of z_i. We have

Theorem 1 *Given the functions V, γ_1, γ_2 and f_n, it is possible to find expressions for the functions a_i's and v_i's such that the overall dynamics (i.e. the dynamics of the system (1), the observer (28) and the norm estimator (11)) admit (x, w, ε) as state and, in the new time $\widehat{\tau}$, the set:*

$$\mathcal{A} = \{(x, w, \varepsilon) : V(x) \leq 2w + 1, \varepsilon = 0\} \tag{30}$$

is globally asymptotically stable.

The proof of this Theorem is given in Sect. 4.2.4. For its presentation, we need preliminaries where in particular the error dynamics are described (Sect. 4.2.1) and the functions a_i's (Sect. 4.2.2) and v_i's (Sect. 4.2.3) are made precise.

4.2.1 Error dynamics

To study the error dynamics, let us first observe that the manifold error ε_i, defined in (29), satisfies the equation:

$$\begin{aligned}
\dot{\varepsilon}_i &= \dot{z}_i + \overset{\cdot}{\overline{(-a_i)^i}} x_0 + \sum_{j=1}^{i-1} \overset{\cdot}{\overline{(-a_i)^{i-j} x_j}} + \dot{x}_i \\
&= -a_i\, z_i + [i \dot{a}_i + a_i^2]\, (-a_i)^{i-1}\, x_0 + \dot{a}_i \sum_{j=1}^{i-1} (i-j)(-a_i)^{i-j-1}\, \widehat{x}_j - v_i \\
&\quad - i \dot{a}_i\, (-a_i)^{i-1}\, x_0 + (-a_i)^i\, x_1 \\
&\quad + \sum_{j=1}^{i-1} \overset{\cdot}{\overline{(-a_i)^{i-j}}} x_j + \sum_{j=1}^{i-1} (-a_i)^{i-j} x_{j+1} \\
&\quad + x_{i+1}, \\
&= -a_i\, z_i - a_i\, (-a_i)^i\, x_0 - \sum_{j=1}^{i-1} \overset{\cdot}{\overline{(-a_i)^{i-j}}}\, \widehat{x}_j - v_i \\
&\quad - a_i\, (-a_i)^{i-1}\, x_1 - a_i \sum_{j=1}^{i-2} (-a_i)^{i-j-1} x_{j+1} - a_i\, x_i \\
&\quad + \sum_{j=1}^{i-1} \overset{\cdot}{\overline{(-a_i)^{i-j}}} x_j \\
&\quad + x_{i+1}, \\
&= -a_i\, \varepsilon_i + \sum_{j=1}^{i-1} \overset{\cdot}{\overline{(-a_i)^{i-j}}}\, [x_j - \widehat{x}_j] + [x_{i+1} - v_i].
\end{aligned} \tag{31}$$

To get a better grip on these dynamics we consider also the estimation error:

$$e_i = x_i - \widehat{x}_i. \tag{32}$$

By Eq. (28), it is related to the manifold error ε by the relation:

$$e_i - \varepsilon_i = -\sum_{j=1}^{i-1} (-a_i)^{i-j} e_j .$$

To continue the analysis we need to express the e_i values in terms of the ε_i values. Let L be the strict lower triangular matrix whose (i, j) entry is $(-a_i)^{i-j}$, i.e.:

$$L = \begin{pmatrix} 0 & \dots & \dots \; \dots & 0 \\ -a_2 & 0 & & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ \vdots & & \ddots \; \ddots & \vdots \\ (-a_n)^{n-1} & (-a_n)^{n-2} & \dots \; -a_n & 0 \end{pmatrix} .$$

With the notations

$$\varepsilon = (\varepsilon_i)_{i=1,\dots,n} , \qquad e = (e_i)_{i=1,\dots,n} ,$$

we get:

$$\varepsilon = (I + L)\, e,$$

and:

$$e = (I + L)^{-1} \varepsilon. \tag{33}$$

Observe now that L is nilpotent, i.e.:

$$L^n = 0,$$

which implies that:

$$(I + L)^{-1} = I + \sum_{i=1}^{n-1} (-L)^i .$$

Therefore, from the expression of the powers of L, we infer that the (i, j) entry of $(I + L)^{-1}$, denoted by ℓ_{ij} in the following:

1. is zero if $j > i$;
2. is 1 if $j = i$;
3. depends only on a_{j+1} to a_i if $j \leq i - 1$.

It follows that we can specify the relation between manifold and estimation errors in:

$$e_i = \varepsilon_i + \sum_{j=1}^{i-1} \ell_{ij}(a_{j+1}, \dots, a_i)\, \varepsilon_j . \tag{34}$$

Also exploiting these compact notations, Eq. (31) reads as:

$$\dot{\varepsilon} = -\mathtt{diag}(a_i)\,\varepsilon + \dot{L}\,e + \mathtt{vect}(x_{i+1} - v_i), \tag{35}$$

$$= -\left(\mathtt{diag}(a_i) - \dot{L}\,(I+L)^{-1}\right)\,\varepsilon + \mathtt{vect}(x_{i+1} - v_i). \tag{36}$$

Here we remark that the (i, j) entry of $\dot{L}(I + L)^{-1}$, denoted by $\dot{a}_i h_{ij}$, has $\dot{a}_i$ in factor and its other factor h_{ij}

1. is zero if $j \geq i$;
2. depends only on a_{j+1} to a_i if $j \leq i - 1$.

These various remarks show that the overall dynamics can be described by the equations:

$$\begin{cases} \overset{*}{x}_0 = \dfrac{x_1}{\widehat{\gamma}(w)}, \\ \quad\vdots \\ \overset{*}{x}_{n-1} = \dfrac{x_n}{\widehat{\gamma}(w)}, \\ \overset{*}{x}_n = \dfrac{f_n(x_0, \ldots, x_n)}{\widehat{\gamma}(w)}, \\ \overset{*}{w} = -\dfrac{w - \gamma_1(x_0)}{\widehat{\gamma}(w)}, \\ \overset{*}{\varepsilon}_i = -\dfrac{a_i}{\widehat{\gamma}(w)}\,\varepsilon_i - \displaystyle\sum_{j=1}^{i-1} \overset{*}{a}_i h_{ij}(a_{j+1}, \ldots, a_i)\,\varepsilon_j + \dfrac{x_{i+1} - v_i}{\widehat{\gamma}(w)}, \end{cases}$$

4.2.2 Choice of the functions a_i's

To motivate the choice of the functions a_i's, and later of the functions v_i's, consider the partial Lyapunov function:

$$U = \frac{1}{2}\sum_{j=1}^{n} (\mu_j \varepsilon_j)^2, \tag{37}$$

where we define:

$$\begin{cases} \mu_n^2 = 4, \\ \mu_j^2 = 4\left(1 + \displaystyle\sum_{i=j+1}^{n} (i-1)\,\ell_{ij}(a_{j+1}, \ldots, a_i)^2\right) \qquad j \in \{1, \ldots, (n-1)\}. \end{cases} \tag{38}$$

This function U is positive definite and radially unbounded in the variables ε_i values. It is worth stressing that each of the functions μ_j in (38) depends on a_{j+1} to a_n only. The motivation for defining these functions μ_j in this way follows from Eq. (34) and the inequality:

$$|e_i| \leq \sqrt{2}\sqrt{\varepsilon_i^2 + (i-1)\sum_{j=1}^{i-1} \ell_{ij}(a_{j+1}, \ldots, a_i)^2\,\varepsilon_j^2}.$$

Indeed this yields the inequalities:

$$\sum_{i=1}^{n} e_i^2 \leq 2 \sum_{i=1}^{n} \left(\varepsilon_i^2 + (i-1) \sum_{j=1}^{i-1} \ell_{ij}(a_{j+1}, \ldots, a_i)^2 \, \varepsilon_j^2 \right), \tag{39}$$

$$\leq 2 \sum_{i=1}^{n} \varepsilon_i^2 + 2 \sum_{i=1}^{n} (i-1) \left(\sum_{j=1}^{i-1} \ell_{ij}(a_{j+1}, \ldots, a_i)^2 \, \varepsilon_j^2 \right), \tag{40}$$

$$\leq 2 \sum_{j=1}^{n} \varepsilon_j^2 + 2 \sum_{j=1}^{n-1} \left(\sum_{i=j+1}^{n} (i-1) \ell_{ij}(a_{j+1}, \ldots, a_i)^2 \, \varepsilon_j^2 \right), \tag{41}$$

$$\leq \sum_{j=1}^{n-1} 2 \left(1 + \sum_{i=j+1}^{n} (i-1) \ell_{ij}(a_{j+1}, \ldots, a_i)^2 \right) \varepsilon_j^2 + 2 \, \varepsilon_n^2. \tag{42}$$

Hence, with (38), we get

$$\sum_{i=1}^{n} e_i^2 \leq \frac{1}{2} \sum_{j=1}^{n} (\mu_j \varepsilon_j)^2 = U. \tag{43}$$

Note now that:

$$\overset{*}{U} = \sum_{i=1}^{n} \left[-\left(\frac{a_i}{\widehat{\gamma}(w)} - \frac{\overset{*}{\mu}_i}{\mu_i} \right) (\mu_i \varepsilon_i)^2 - \mu_i^2 \varepsilon_i \sum_{j=1}^{i-1} \overset{*}{a}_i h_{ij}(a_{j+1}, \ldots, a_i) \varepsilon_j \right.$$
$$\left. + \mu_i^2 \varepsilon_i \frac{x_{i+1} - v_i}{\widehat{\gamma}(w)} \right].$$

At this stage, to deal with the terms $\mu_i^2 \varepsilon_i \dfrac{x_{i+1} - v_i}{\widehat{\gamma}(w)}$ and for analysis purpose only, we introduce parameters λ_i, functions of w, which have to be made precise later on, and we complete the squares as:

$$\mu_i^2 \varepsilon_i \frac{x_{i+1} - v_i}{\widehat{\gamma}(w)} \leq \frac{\mu_i^2 \lambda_i(w)^2}{2 \widehat{\gamma}(w)^2} (\mu_i \varepsilon_i)^2 + \frac{1}{2} \left(\frac{x_{i+1} - v_i}{\lambda_i(w)} \right)^2.$$

Therefore, by adding and subtracting U, we get:

$$
\begin{aligned}
\overset{*}{U} \leq & -U + \frac{1}{2}\sum_{i=1}^{n}\left(\frac{x_{i+1}-v_i}{\lambda_i(w)}\right)^2 \\
& -\sum_{i=1}^{n}\left[\left(\frac{a_i}{\widehat{\gamma}(w)} - \frac{\overset{*}{\mu}_i}{\mu_i} - \frac{\frac{\mu_i^2\lambda_i(w)^2}{2\widehat{\gamma}(w)^2}+1}{2}\right)(\mu_i\varepsilon_i)^2 \right. \\
& \left. +\mu_i^2\varepsilon_i\sum_{j=1}^{i-1}\overset{*}{a}_i h_{ij}(a_{j+1},\ldots,a_i)\,\varepsilon_j\right], \\
\leq & -U + \frac{1}{2}\sum_{i=1}^{n}\left(\frac{x_{i+1}-v_i}{\lambda_i(w)}\right)^2 - \begin{pmatrix}\mu_1\varepsilon_1 & \ldots & \mu_n\varepsilon_n\end{pmatrix} M \begin{pmatrix}\mu_1\varepsilon_1 \\ \vdots \\ \mu_n\varepsilon_n\end{pmatrix}, \qquad (44)
\end{aligned}
$$

where M is the symmetric matrix:

$$
M = \begin{pmatrix}
\frac{a_1}{\widehat{\gamma}(w)} - \frac{\overset{*}{\mu}_1}{\mu_1} - \frac{\frac{\mu_1^2\lambda_1(w)^2}{\widehat{\gamma}(w)^2}+1}{2} & \frac{\mu_2\overset{*}{a}_2 h_{21}(a_2)}{2\mu_1} & \ldots & \frac{\mu_n\overset{*}{a}_n h_{n1}(a_2,\ldots,a_n)}{2\mu_1} \\
\frac{\mu_2\overset{*}{a}_2 h_{21}(a_2)}{2\mu_1} & \frac{a_2}{\widehat{\gamma}(w)} - \frac{\overset{*}{\mu}_2}{\mu_2} - \frac{\frac{\mu_2^2\lambda_2(w)^2}{\widehat{\gamma}(w)^2}+1}{2} & & \\
 & & \ddots & \vdots \\
\vdots & \ddots & \ddots & \frac{\mu_n\overset{*}{a}_n h_{n(n-1)}(a_n)}{2\mu_{n-1}} \\
\frac{\mu_n\overset{*}{a}_n h_{n1}(a_2,\ldots,a_n)}{2\mu_1} & \ldots & \frac{\mu_n\overset{*}{a}_n h_{n(n-1)}(a_n)}{2\mu_{n-1}} & \frac{a_n}{\widehat{\gamma}(w)} - \frac{\overset{*}{\mu}_n}{\mu_n} - \frac{\frac{\mu_n^2\lambda_n(w)^2}{\widehat{\gamma}(w)^2}+1}{2}
\end{pmatrix}
$$

To go on, let us point out some properties of the entries of the matrix M.

1. Recall that μ_j depends only on a_{j+1} to a_n and μ_n is a constant. Hence, given the C^1 functions a_{j+1} to a_n, depending on w only, it is possible to find continuous nondecreasing functions c_{1jw} satisfying:

$$2 \leq \left|\mu_j(a_{j+1}(w),\ldots,a_n(w))\right| \leq c_{1jw}(w). \qquad (45)$$

2. We have:

$$\begin{aligned}
\frac{\overset{*}{\mu}_j}{\mu_j} &= \frac{1}{2}\,\frac{\overset{*}{\widehat{\mu_j^2}}}{\mu_j^2},\\
&= 4\,\frac{\sum_{i=j+1}^{n}\left[(i-1)\,\ell_{ij}(a_{j+1},\dots,a_i)\displaystyle\sum_{k=j+1}^{i}\frac{\partial \ell_{ij}}{\partial a_k}(a_{j+1},\dots,a_i)\overset{*}{a}_k\right]}{\mu_j^2},\\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad (46)\\
&= 4\,\frac{\sum_{i=j+1}^{n}\left[(i-1)\,\ell_{ij}(a_{j+1},\dots,a_i)\displaystyle\sum_{k=j+1}^{i}\frac{\partial \ell_{ij}}{\partial a_k}(a_{j+1},\dots,a_i)a_k'(w)\right]}{\mu_j^2}\\
&\quad\times\frac{-w+\gamma_1(x_0)}{\widehat{\gamma}(w)}. \qquad (47)
\end{aligned}$$

Given the C^1 functions a_{j+1} to a_n of w, the left hand side of (47) is a continuous function of (w, x_0). So we can find continuous nondecreasing functions c_{2jw} and c_{j0} satisfying:

$$\frac{|\overset{*}{\mu}_j|}{\mu_j} \le \frac{c_{2jw}(w)+c_{j0}(|x_0|)}{\widehat{\gamma}(w)}.$$

As a result, using (5), (15) and (19) and the inequality $\widehat{\gamma}\ge 1$, we get:

$$\begin{aligned}
\frac{|\overset{*}{\mu}_j|}{\mu_j} &\le \frac{c_{2jw}(w)+c_{j0}(\gamma(V(x)))}{\widehat{\gamma}(w)},\\
&\le c_{2jw}(w)+c_{j0}(\widehat{\gamma}(w))+\frac{d_{j0*}(2[V(x)-2w-1]_+)}{\widehat{\gamma}(w)}, \qquad (48)
\end{aligned}$$

where the function d_{j0*} is the composition $c_{j0}\circ\gamma$. Without loss of generality, it is possible to choose the functions c_{j0} such that d_{j0*} is decreasing with j, i.e.:

$$d_{j0*}(r) \ge d_{(j+1)0*}(r).$$

3. Similarly, if necessary by increasing the functions c_{2jw} and c_{j0}, we can obtain, for the functions $\frac{\mu_i\overset{*}{a}_i h_{ij}(a_{j+1},\dots,a_i)}{2\mu_j}$, with $i\in\{j+1,\dots,n\}$,

$$\begin{aligned}
&\left(\frac{\mu_i\overset{*}{a}_i h_{ij}(a_{j+1},\dots,a_i)}{2\mu_j}\right)^2\\
&\quad\le c_{2jw}(w)+c_{j0}(\widehat{\gamma}(w))+\frac{d_{j0*}(2[V(x)-2w-1]_+)}{\widehat{\gamma}(w)}. \qquad (49)
\end{aligned}$$

4. Finally, we stress that the functions c_{1jw}, c_{2jw} and c_{j0} and d_{j0*} are given once the functions a_{j+1} to a_n are fixed. Also we can define the functions c_{2nw}, c_{n0} and d_{n0*} as being the zero constant.

With all the above properties and inequalities at hand, consider the sequence of matrices:

$$M_n = \frac{a_n}{\widehat{\gamma}(w)} - \frac{\frac{\mu_n^2 \lambda_n(w)^2}{\widehat{\gamma}(w)^2}+3}{2},$$

$$M_{n-1} = \left(\begin{array}{c|c} \frac{a_{n-1}}{\widehat{\gamma}(w)} - \frac{\overset{*}{\mu}_{n-1}}{\mu_{n-1}} - \frac{\frac{\mu_{n-1}^2 \lambda_{n-1}(w)^2}{\widehat{\gamma}(w)^2}+3}{2} & \frac{\mu_n \overset{*}{a}_n h_{n(n-1)}(a_n)}{2\mu_{n-1}} \\ \hline \frac{\mu_n \overset{*}{a}_n h_{n(n-1)}(a_n)}{2\mu_{n-1}} & 1 + M_n \end{array}\right), \tag{50}$$

$$M_j = \left(\begin{array}{c|ccc} \frac{a_j}{\widehat{\gamma}(w)} - \frac{\overset{*}{\mu}_j}{\mu_j} - \frac{\frac{\mu_j^2 \lambda_j(w)^2}{\widehat{\gamma}(w)^2}+3}{2} & \frac{\mu_{j+1} \overset{*}{a}_{j+1} h_{(j+1)j}(a_{j+1})}{2\mu_j} & \dots & \frac{\mu_n \overset{*}{a}_n h_{nj}(a_{j+1},\dots,a_n)}{2\mu_j} \\ \hline \frac{\mu_{j+1} \overset{*}{a}_{j+1} h_{(j+1)j}(a_{j+1})}{2\mu_j} & & & \\ \vdots & & I + M_{j+1} & \\ \frac{\mu_n \overset{*}{a}_n h_{nj}(a_{j+1},\dots,a_n)}{2\mu_j} & & & \end{array}\right). \tag{51}$$

They are such that:

$$M = I + M_1 .$$

Assume for the time being that for each j, the functions a_{j+1} to a_n can be chosen such that:

$$M_{j+1} > -[n-j]\, \frac{d_{(j+1)0*}(2[V(x) - 2w - 1]_+)}{\widehat{\gamma}(w)}\, I . \tag{52}$$

Then, in particular for $j = 0$, this yields:

$$M = I + M_1 > \left[1 - n \frac{d_{10*}(2[V(x) - 2w - 1]_+)}{\widehat{\gamma}(w)}\right] I .$$

But by introducing this result in Eq. (44) we get the key inequality:

$$\overset{*}{U} \le \left[-3 + 2n \frac{d_{10*}(2[V(x) - 2w - 1]_+)}{\widehat{\gamma}(w)}\right] U + \frac{1}{2} \sum_{i=1}^{n} \left(\frac{x_{i+1} - v_i}{\lambda_i(w)}\right)^2 . \tag{53}$$

So, for this inequality (53) to hold, it is sufficient to establish (52). In order to do so, we remark that if for some real number m one has:

$$M_{j+1} + m\, I > 0, \tag{54}$$

then the following inequality holds :

$$\begin{aligned} &\left(\frac{\mu_{j+1} \overset{*}{a}_{j+1} h_{(j+1)j}(a_{j+1})}{2\mu_j} \dots \frac{\mu_n \overset{*}{a}_n h_{nj}(a_{j+1},\dots,a_n)}{2\mu_j}\right) ((1+m)I + M_{j+1})^{-1} \\ &\times \begin{pmatrix} \frac{\mu_{j+1} \overset{*}{a}_{j+1} h_{(j+1)j}(a_{j+1})}{2\mu_j} \\ \vdots \\ \frac{\mu_n \overset{*}{a}_n h_{nj}(a_{j+1},\dots,a_n)}{2\mu_j} \end{pmatrix} \le \sum_{i=j+1}^{n} \left(\frac{\mu_i \overset{*}{a}_i h_{ij}(a_{j+1},\dots,a_i)}{2\mu_j}\right)^2 . \end{aligned} \tag{55}$$

As a consequence, from the definition (51), we see that, if m satisfies (54) and we have:

$$m + \frac{a_j}{\widehat{\gamma}(w)} - \frac{\overset{*}{\mu}_j}{\mu_j} - \frac{\frac{\mu_j^2 \lambda_j(w)^2}{\widehat{\gamma}(w)^2} + 3}{2} > \sum_{i=j+1}^{n} \left(\frac{\mu_i \overset{*}{a}_i h_{ij}(a_{j+1}, \dots, a_i)}{2\mu_j} \right)^2, \tag{56}$$

then, with the Schur formula, it follows that:

$$M_j + m\, I > 0. \tag{57}$$

So let us now establish (52) by induction.

- For $j = n-1$, since d_{n0*} is zero, we need to prove that M_n is strictly positive. From (50), this is obtained by picking a_n as a C^1 function satisfying:

$$a_n(w) > \left(\frac{\frac{\mu_n^2 \lambda_n(w)^2}{\widehat{\gamma}(w)^2} + 3}{2} \right) \widehat{\gamma}(w). \tag{58}$$

- Assume that the functions a_{j+1} to a_n have been chosen such that (52) holds. Since d_{j0*} decreases with j, it follows that, by selecting m as:

$$\begin{aligned} m &= [n-j+1]\, \frac{d_{j0*}(2[V(x) - 2w - 1]_+)}{\widehat{\gamma}(w)}, \\ &\geq [n-j]\, \frac{d_{(j+1)0*}(2[V(x) - 2w - 1]_+)}{\widehat{\gamma}(w)}, \end{aligned} \tag{59}$$

  (54) is satisfied.
  About (56), we remark that the inequalities (45), (48) and (49), give:

$$\begin{aligned} &\frac{|\overset{*}{\mu}_j|}{\mu_j} + \frac{\frac{\mu_j^2 \lambda_j(w)^2}{\widehat{\gamma}(w)^2} + 3}{2} + \sum_{i=j+1}^{n} \left(\frac{\mu_i \overset{*}{a}_i h_{ij}(a_{j+1} w), \dots, a_i(w))}{2\mu_j} \right)^2 \\ &\leq \frac{\frac{c_{1jw}(w)^2 \lambda_j(w)^2}{\widehat{\gamma}(w)^2} + 3}{2} + [n-j+1][c_{2jw}(w) + c_{j0}(\widehat{\gamma}(w))] \\ &\quad + [n-j+1] \frac{d_{j0*}(2[V(x) - 2w - 1]_+)}{\widehat{\gamma}(w)}. \end{aligned} \tag{60}$$

  So let us choose a_j as a C^1 function satisfying:

$$a_j(w) > \left(\frac{\frac{c_{1jw}(w)^2 \lambda_j(w)^2}{\widehat{\gamma}(w)^2} + 3}{2} + [n-j+1][c_{2jw}(w) + c_{j0}(\widehat{\gamma}(w))] \right) \widehat{\gamma}(w). \tag{61}$$

  This is always possible since the functions c_{1jw}, c_{2jw} and c_{j0} are obtained from the functions a_{j+1} to a_n only. With (59) and (60), this choice for a_j implies the inequality (56). Thus, we have established that (54) and (56) hold with m given by (59). It follows from the remark above that (57) holds. This gives

$$M_j + [n - j + 1]\,\frac{d_{j0*}(2[V(x) - 2w - 1]_+)}{\widehat{\gamma}(w)}\,I > 0.$$

This is nothing but (52) for $j - 1$.

In conclusion, by choosing the a_i's as C^1 functions satisfying the constraints (58) and (61), we have the inequality (53).

4.2.3 *Choice of the v_i's and the λ_i's*

We now concentrate on the last term in the right hand side of (53), i.e. on the choice of the v_i values and the λ_i values. As discussed in Sect. 4.1, the best choice for v_i would be:

$$v_i = x_{i+1} \quad \forall i \in \{1, \ldots, n-1\}, \tag{62}$$

$$v_n = f_n(x_0, x_1, \ldots, x_n). \tag{63}$$

However, this selection cannot be implemented since the x_i values are unknown. Instead we implement, for i ranging from 1 to $n - 1$,

$$v_i = \widehat{x}_{i+1}.$$

This yields :

$$\frac{1}{2}\sum_{i=1}^{n-1}\left(\frac{x_{i+1} - v_i}{\lambda_i(w)}\right)^2 = \frac{1}{2}\sum_{i=1}^{n-1}\left(\frac{e_i}{\lambda_i(w)}\right)^2.$$

Then, from inequalities (43) and (53), it is apparent that an appropriate choice for the λ_i values, for i ranging from 1 to $n - 1$, is:

$$\lambda_i(w) = 1 \quad \forall i \in \{1, \ldots, n-1\}.$$

This selection yields simply:

$$\frac{1}{2}\sum_{i=1}^{n-1}\left(\frac{x_{i+1} - v_i}{\lambda_i(w)}\right)^2 \leq \frac{1}{2}\,U,$$

hence inequality (53) becomes:

$$\overset{*}{U} \leq \left[-\frac{5}{2} + 2n\,\frac{d_{10*}(2[V(x) - 2w - 1]_+)}{\widehat{\gamma}(w)}\right] U + \frac{1}{2}\left(\frac{f_n(x_0, x_1, \ldots, x_n) - v_n}{\lambda_n(w)}\right)^2. \tag{64}$$

In a similar way, one may be tempted to choose:

$$v_n = f_n(x_0, \widehat{x}_1, \ldots, \widehat{x}_n). \tag{65}$$

However, contrary to the case of $v_i = x_{i+1}$, where the Lipschitz constant is trivially bounded by 1, the use of Eq. (65) would require a bound for the Lipschitz constant of f_n on a noncompact set, which is not available. To round this problem recall first that by condition (5), the function $\frac{f_n(x_0,x_1,\ldots,x_n)}{\gamma(V(x))}$ is bounded by 1. Mimicking this property for $\frac{v_n}{\widehat{\gamma}(w)}$, we impose that v_n satisfy

$$\frac{v_n}{\widehat{\gamma}(w)} \leq 1. \tag{66}$$

Otherwise, for the time being, v_n remains free. By inequalities (5), (17) and (19), this gives:

$$\left(\frac{f_n(x_0,x_1,\ldots,x_n)-v_n}{\widehat{\gamma}(w)}\right)^2 \leq 2\left[1+\left(\frac{f_n(x_0,x_1,\ldots,x_n)}{\widehat{\gamma}(w)}\right)^2\right], \tag{67}$$

$$\leq 2\left[1+\left(1+\frac{\gamma_*(2[V(x)-2w-1]_+)}{\widehat{\gamma}(w)}\right)^2\right]. \tag{68}$$

Hence, selecting:

$$\lambda_n(s) = 2\widehat{\gamma}(w),$$

Equation (64) gives

$$\overset{*}{U} \leq \left[-\frac{5}{2}+2n\frac{d_{10*}(2[V(x)-2w-1]_+)}{\widehat{\gamma}(w)}\right]U + \frac{1}{4}\left[1+\left(1+\frac{\gamma_*(2[V(x)-2w-1]_+)}{\widehat{\gamma}(w)}\right)^2\right]. \tag{69}$$

From this, the actual selection of v_n will follow by studying the solution after the time $\widehat{\tau}(t_v)$. Indeed, by Eq. (14), after the time $\widehat{\tau}(t_v)$, $2[V(x(\widehat{\tau}))-2w(\widehat{\tau})-1]_+$ is zero. Therefore, after the time $\widehat{\tau}(t_v)$, (69) is nothing but simply:

$$\overset{*}{U} \leq -\frac{5}{2}U+\frac{1}{2}.$$

It follows that, for any solution, there exists a finite time $\widehat{\tau}$ after which, U is smaller than 1. Moreover, by Eq. (43), the same holds for all the e_i values, i.e. the $\widehat{x_i}$ deviate from x_i by a distance less than 1. As a consequence, only a bound of the Lipschitz constant of f_n for a compact set around x is needed. However, because of (3), such a bound can be expressed in terms of $V(x)$ and therefore of w.

This leads to introduce a new constraint on the function γ. Not only do we want γ to satisfy condition (5), but also to satisfy:

$$\sup_{\sum_{i=1}^n \eta_i^2 \leq 1}\left\{\frac{|f_n(x_0,x_1+\eta_1,\ldots,x_n+\eta_n)-f_n(x_0,x_1,\ldots,x_n)|}{\sqrt{\sum_{i=1}^n \eta_i^2}}\right\} \leq \gamma(V(x)). \tag{70}$$

This is possible since f_n is continuously differentiable, V satisfies condition (3) and the set

$$\left\{(\eta_1, \cdots, \eta_n) \mid \sum_{i=1}^{n} \eta_i^2 \le 1\right\}$$

is compact.

Finally, the following implications hold:

$$U \le 1 \Rightarrow \sum_{i=1}^{n}(x_i - \widehat{x}_i)^2 = \sum_{i=1}^{n} e_i^2 \le U \le 1, \tag{71}$$

$$\Rightarrow \frac{|f_n(x_0, \widehat{x}_1, \dots, \widehat{x}_n) - f_n(x_0, x_1, \dots, x_n)|}{\gamma(V(x))} \le \sqrt{U}. \tag{72}$$

Therefore, recalling Eq. (66), this motivates the selection:

$$v_n = \widehat{\gamma}(w)\, \mathtt{sat}\left(\frac{f_n(x_0, \widehat{x}_1, \dots, \widehat{x}_n)}{\widehat{\gamma}(w)}\right),$$

with sat the standard saturation function. As a result:

1. when $U \le 1$, we have nine cases[5] to consider to get an upper-bound on $|f_n(x_0, x_1, \dots, x_n) - v_n|$. By using (5) and exploiting the fact that the saturation function is "pushing" v_n in the interval $[-\widehat{\gamma}(w), \widehat{\gamma}(w))]$, an upper-bound is given in the following table:

$(\lvert f_n(x_0, x_1, \dots, x_n) - v_n\rvert \quad \le \quad)$	$\widehat{f_n} < -\widehat{\gamma}$	$-\widehat{\gamma} \le \widehat{f_n} \le \widehat{\gamma}$	$\widehat{\gamma} < \widehat{f_n}$
$\widehat{\gamma} < f_n$	$\gamma\sqrt{U}$	$\gamma\sqrt{U}$	$\gamma - \widehat{\gamma}$
$-\widehat{\gamma} \le f_n \le \widehat{\gamma}$	$\gamma\sqrt{U}$	$\gamma\sqrt{U}$	$\gamma\sqrt{U}$
$f_n < -\widehat{\gamma}$	$\gamma - \widehat{\gamma}$	$\gamma\sqrt{U}$	$\gamma\sqrt{U}$

where we have used the compact notations:

$$f_n = f_n(x_0, x_1, \dots, x_n), \quad \widehat{f_n} = f_n(x_0, \widehat{x}_1, \dots, \widehat{x}_n),$$
$$\gamma = \gamma(V(x)), \quad \widehat{\gamma} = \widehat{\gamma}(w)$$

Using (16), we get in any case:

$$\frac{|f_n(x_0, x_1, \dots, x_n) - v_n|}{\widehat{\gamma}(w)} \le \left(1 + \frac{\gamma_*(2[V(x) - 2w - 1]_+)}{\widehat{\gamma}(w)}\right)\sqrt{U} + \frac{\gamma_*(2[V(x) - 2w - 1]_+)}{\widehat{\gamma}(w)}. \tag{73}$$

2. when $1 \le U$, with (66), we have:

$$\begin{aligned}\frac{|f_n(x_0, x_1, \dots, x_n) - v_n|}{\widehat{\gamma}(w)} &\le \left(1 + \frac{\gamma_*(2[V(x) - 2w - 1]_+)}{\widehat{\gamma}(w)}\right) + 1, \\ &\le 2\sqrt{U} + \frac{\gamma_*(2[V(x) - 2w - 1]_+)}{\widehat{\gamma}(w)}.\end{aligned} \tag{74}$$

[5] Actually by symmetry there are only six cases to consider.

Replacing these bounds (73) and (74) in Eq. (64) yields:

$$\begin{aligned}\overset{*}{U} &\leq \left[-\frac{5}{2} + 2n\frac{d_{10*}(2[V(x)-2w-1]_+)}{\widehat{\gamma}(w)}\right] U \\ &\quad + \frac{1}{4}\left(2 + \frac{\gamma_*(2[V(x)-2w-1]_+)}{\widehat{\gamma}(w)}\right)^2 U + \frac{1}{4}\frac{\gamma_*(2[V(x)-2w-1]_+)^2}{\widehat{\gamma}(w)^2}, \\ &\leq -\frac{3}{2}U + \left[2n\frac{d_{10*}(2[V(x)-2w-1]_+)}{\widehat{\gamma}(w)} + \frac{\gamma_*(2[V(x)-2w-1]_+)}{\widehat{\gamma}(w)}\right. \\ &\quad \left. + \frac{1}{4}\frac{\gamma_*(2[V(x)-2w-1]_+)^2}{\widehat{\gamma}(w)^2}\right] U + \frac{1}{4}\frac{\gamma_*(2[V(x)-2w-1]_+)^2}{\widehat{\gamma}(w)^2}.\end{aligned}$$

Let:

$$\begin{aligned}k_1(r) &= 2nd_{10*}(2r) + \gamma_*(2r) + \frac{1}{4}\gamma_*(2r)^2, \\ k_2(r) &= \frac{1}{4}\gamma_*(2r)^2.\end{aligned}$$

These functions are continuous, take nonnegative values and are zero at zero. Recalling that $\widehat{\gamma} \geq 1$, we conclude that:

$$\overset{*}{U} \leq \left[-\frac{3}{2} + \frac{k_1([V(x)-2w-1]_+)}{\widehat{\gamma}(w)}\right] U + \frac{k_2([V(x)-2w-1]_+)}{\widehat{\gamma}(w)}, \quad (75)$$

and, after the time $\widehat{\tau}(t_v)$:

$$\overset{*}{U} \leq -\frac{3}{2}U.$$

Remark 1 The analysis carried out shows that the proposed observer has two modes of operation. In the first mode, the observer does not need any information on the function f_n, but only bounds on the norm of the state to be observed. In the second mode, the function f_n is used to achieve asymptotic convergence of the estimation error to zero. This implies that, if one is merely interested in a practical observer, i.e. in an observer yielding asymptotically a bounded state estimation error, then it is sufficient to select $v_n = 0$.

4.2.4 Proof of Theorem 1

With the functions a_i values and v_i values as defined above and the inequality (75) established, we are now ready to prove Theorem 1.

Consider the function:

$$\mathfrak{L}(x, w, \varepsilon) = \log(1+U) + \int_0^{[V(x)-2w-1]_+} [r + k_1(r) + k_2(r)]\, dr.$$

and the set $\mathcal{A}$ defined in Eq. (30). Let $d((x, w, \varepsilon), \mathcal{A})$ be the distance of a point (x, w, ε) to $\mathcal{A}$, i.e.:

$$d((x, w, \varepsilon), \mathcal{A}) = \inf_{((x_a, w_a, \varepsilon_a) \in \mathcal{A}} \{|\varepsilon - \varepsilon_a| + |w - w_a| + |x - x_a|\}.$$

The function $\mathfrak{L}$ is positive definite with respect to $\mathcal{A}$. Indeed, the following implications hold:

$$\begin{aligned} \mathfrak{L}(x, w, \varepsilon) = 0 \quad &\Longrightarrow \quad U = [V(x) - 2w - 1]_+ = 0, &(76)\\ &\Longrightarrow \quad \{|\varepsilon| = 0, V(x) \leq 2w + 1\}. &(77) \end{aligned}$$

Moreover, for any point (x, w, ε), the point $(x, \frac{[V(x)-1]_+}{2}, 0)$ is in $\mathcal{A}$. Also, if such a point is not in $\mathcal{A}$, we have:

$$\left| w - \frac{[V(x)-1]_+}{2} \right| = \frac{V(x)-1}{2} - w.$$

This implies:

$$0 < d((x, w, \varepsilon), \mathcal{A}) \leq |\varepsilon| + \left[\frac{V(x)-1}{2} - w \right]_+ .$$

Therefore, with (37) and (38), for any nonnegative real number ℓ, one has:

$$\begin{aligned} \mathfrak{L} \leq \ell \Longrightarrow &\{U \leq \exp(\ell) - 1, 2[V(x) - 2w - 1]_+^2 \leq \ell\},\\ &\Longrightarrow d((x, w, \varepsilon), \mathcal{A}) \leq |\varepsilon| + \left[\frac{V(x)-1}{2} - w \right]_+ \leq \frac{\sqrt{\exp(\ell) - 1}}{\sqrt{2}} + \sqrt{\frac{\ell}{8}}. \end{aligned}$$

This implies that the function $\mathfrak{L}$ is radially unbounded with respect to the (noncompact) set $\mathcal{A}$.

As a result, to complete the proof of the theorem we have to show that $\overset{*}{\mathfrak{L}}$ is negative definite with respect to $\mathcal{A}$. With (4) and (11), and since r, $k_1(r)$ and $k_2(r)$ are zero when r is zero, we get:

$$\begin{aligned} &\overbrace{\int_0^{[V(x)-2w-1]_+} [r + k_1(r) + k_2(r)]\, dr}^{*}\\ &= \big[[V(x) - 2w - 1]_+ + k_1([V(x) - 2w - 1]_+) + k_2([V(x) - 2w - 1]_+)\big]\\ &\quad \frac{-V(x) + 2w - \gamma_1(x_0)}{\widehat{\gamma}(w)}. \end{aligned}$$

Together with (75), this yields :

$$
\begin{aligned}
\overset{*}{\mathfrak{L}} \leq & \left[-\frac{1}{4}+\frac{k_1([V(x)-2w-1]_+)}{\widehat{\gamma}(w)}\right]\frac{U}{1+U}+\frac{k_2([V(x)-2w-1]_+)}{\widehat{\gamma}(w)}\frac{1}{1+U}\\
&+\big[[V(x)-2w-1]_+ + k_1([V(x)-2w-1]_+)\\
&\quad +k_2([V(x)-2w-1]_+)\big]\,\frac{-V(x)+2w-\gamma_1(x_0)}{\widehat{\gamma}(w)},\\
\leq & -\frac{1}{4}\frac{U}{1+U}+\frac{k_1([V(x)-2w-1]_+)+k_2([V(x)-2w-1]_+)}{\widehat{\gamma}(w)}\\
&-\frac{[V(x)-2w-1]_+ + k_1([V(x)-2w-1]_+)+k_2([V(x)-2w-1]_+)}{\widehat{\gamma}(w)}\\
&\times\big[1+\gamma_1(x_0)+[V(x)-2w-1]_+\big],\\
\leq & -\frac{1}{4}\frac{U}{1+U}-(1+\gamma_1(x_0))\,\frac{[V(x)-2w-1]_+^2}{\widehat{\gamma}(w)}.
\end{aligned}
$$

This establishes that $\overset{*}{\mathfrak{L}}$ is negative definite with respect to $\mathcal{A}$, and completes the proof of Theorem 1. □

4.2.5 Summary

The results established so far can be summarized as follows. For system (1) with output (2), and under the stated assumptions, we propose the observer:

$$
\begin{cases}
\dot{z}_i = -a_i\, z_i + [i\dot{a}_i + a_i^2]\,(-a_i)^{i-1}\, x_0 + \dot{a}_i \displaystyle\sum_{j=1}^{i-1}(i-j)(-a_i)^{i-j-1}\widehat{x}_j - v_i,\\
\widehat{x}_i = -(-a_i)^i x_0 - \displaystyle\sum_{j=1}^{i-1}(-a_i)^{i-j}\widehat{x}_j - z_i,\\
\dot{w} = -w + \gamma_1(x_0)
\end{cases}
\tag{78}
$$

where:

$$
\begin{aligned}
v_i &= \widehat{x}_{i+1} \quad \forall i \in \{1,\ldots,n-1\},\\
v_n &= \widehat{\gamma}(w)\,\mathtt{sat}\left(\frac{f_n(x_0,\widehat{x}_1,\ldots,\widehat{x}_n)}{\widehat{\gamma}(w)}\right).
\end{aligned}
$$

Recall that the functions a_i's are only function of w, therefore the notation $\dot{a}_i$ means simply:

$$
\dot{a}_i = \frac{da_i}{dw}(w)\left[-w+\gamma_1(x_0)\right].
$$

Finally, rewriting the observer (78) in the compact form (10) (with $y = x_0$) and, to be precise, denoting with $X(x,t)$ the solutions of the system (1) and with $(X(x,t), \mathcal{X}((x,\mathcal{x}),t))$ the solutions of system (1)–(10), the following result has been established.

Theorem 2 *For the globally completely observable system (1), under Assumption 1, we can construct*[6] *functions a_i values and $\widehat{\gamma}$, the a_i values being C^1, such that for each solution $X(x,t)$ of (1), right maximally defined on $[0,T)$, with $T \leq +\infty$, and for each initial condition $\mathcal{X}$, the associated solution $(X(x,t), \mathcal{X}((x,\mathcal{X}),t))$ of (1)–(10), is defined also on $[0,T)$ and satisfies:*

$$\lim_{t\to T} |X(x,t) - \mathfrak{h}(\mathcal{X}((x,\mathcal{X}),t), X_0(x,t))| = 0.$$

This result states that we have designed an observer providing an estimate of the state of the system which converges to the actual value in infinite time, if there is no finite escape time of this actual value, and at the time of the escape if there is a finite escape time.

5 Observer implementation

5.1 Implementation as a reduced order observer

The observer dynamics described by Eq. (78) are somewhat involved, and may be difficult or computationally expensive to implement.

It is now shown that, with a proper change of coordinates, it is possible to implement the observer in a simpler form. To begin with consider the dynamics of the estimation error e. By Eq. (32) we have, for i ranging from 1 to $n-1$,

$$x_{i+1} - v_i = x_{i+1} - \hat{x}_{i+1} = e_{i+1}.$$

[6] To be precise, the function a_i values are any C^1 functions satisfying:

$$a_j(w) > \left(\frac{\frac{c_{1jw}(w)^2}{\widehat{\gamma}(w)^2} + 3}{2} + [n-j+1][c_{2jw}(w) + c_{j0}(\widehat{\gamma}(w))] \right) \widehat{\gamma}(w),$$

for j smaller than n, and

$$a_n(w) > \frac{11}{2}\,\widehat{\gamma}(w),$$

$$a_n(w) > 2\,\widehat{\gamma}(w),$$

where the functions $\widehat{\gamma}$, c_{1jw}, c_{2jw}, c_{j0} can be expressed from the problem data γ_1 and γ_2 and from the function γ, chosen to satisfy:

$$|x_0| + \cdots + |x_n| + |f_n(x_0,\ldots,x_n)| \leq \gamma(V(x))$$

and

$$\sup_{|\eta_i| \leq 1} \left\{ \frac{|f_n(x_0, x_1+\eta_1, \ldots, x_n+\eta_n) - f_n(x_0, x_1, \ldots, x_n)|}{\sqrt{\sum_{i=1}^{n} \eta_i^2}} \right\} \leq \gamma(V(x)).$$

Hence, defining

$$S = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ \vdots & & & 0 & 1 \\ 0 & \dots & \dots & \dots & 0 \end{pmatrix} \quad , \qquad E_n = \begin{pmatrix} 0 \\ \vdots \\ \vdots \\ 0 \\ 1 \end{pmatrix},$$

Eq. (35) rewrites as:

$$\dot{\varepsilon} = -\mathtt{diag}(a_i)\,\varepsilon + \left(\dot{L} + S\right) e + E_n\,\Delta f_n,$$

where:

$$\Delta f_n = f_n(x_0, \dots, x_n) - v_n.$$

Exploiting Eq. (33) we obtain:

$$\begin{aligned} \dot{e} &= -(I+L)^{-1}\dot{L}(I+L)^{-1}\varepsilon + (I+L)^{-1}\dot{\varepsilon}, \\ &= (I+L)^{-1}\left[-\mathtt{diag}(a_i)\,(I+L) + S\right] e + (I+L)^{-1} E_n\,\Delta f_n, \\ &= (I+L)^{-1}\left[-\mathtt{diag}(a_i)\,(I+L) + S\right] e + E_n\,\Delta f_n. \end{aligned}$$

Observe now that we have the identity:

$$-\mathtt{diag}(a_i)\,(I+L) = \left(\begin{array}{c|cccc} -a_1 & 0 & \dots & \dots & 0 \\ a_2^2 & -a_2 & 0 & & \vdots \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ \vdots & \vdots & & \ddots & 0 \\ (-a_n)^n & (-a_n)^{n-1} & (-a_n)^{n-2} & \dots & -a_n \end{array}\right)$$

where we recognize the matrix L without its last column in the right block of the right hand term. It follows that:

$$(I+L)^{-1}\left[-\mathtt{diag}(a_i)\,(I+L) + S\right] = \begin{pmatrix} k_1 & 1 & 0 & \dots & 0 \\ \vdots & 0 & \ddots & \ddots & \vdots \\ k_i & & \ddots & \ddots & 0 \\ \vdots & & & 0 & 1 \\ k_n & \dots & \dots & \dots & 0 \end{pmatrix},$$

where the k_i's are the components of the vector k defined as:

$$\begin{pmatrix} -a_1 \\ a_2^2 \\ \vdots \\ (-a_n)^n \end{pmatrix} = (I+L)\,k. \tag{79}$$

Note that each k_i is a function of w. As a result the dynamics of the estimation error e is simply:

$$\dot{e} = \begin{pmatrix} k_1 & 1 & 0 & \dots & 0 \\ \vdots & 0 & \ddots & \ddots & \vdots \\ k_i & & \ddots & \ddots & 0 \\ \vdots & & & 0 & 1 \\ k_n & \dots & \dots & \dots & 0 \end{pmatrix} e + \begin{pmatrix} 0 \\ \vdots \\ \vdots \\ 0 \\ \Delta f_n \end{pmatrix}. \tag{80}$$

Consider now the coordinates:

$$\mathcal{X}_i = x_i + k_i\, x_0,$$

and note that:

$$\begin{aligned} \dot{\mathcal{X}}_i &= x_{i+1} + \dot{k}_i\, x_0 + k_i\, x_1, & (81)\\ &= \mathcal{X}_{i+1} + k_i\, \mathcal{X}_1 + \left(\dot{k}_i - k_{i+1} - k_i k_1\right) x_0 & (82) \end{aligned}$$

Similarly, setting:

$$\widehat{\mathcal{X}}_i = \widehat{x}_i + k_i\, x_0 = x_i + k_i\, x_0 - e_i \quad , \qquad \widehat{\mathcal{X}}_{n+1} = v_n,$$

yields, by Eq. (80),

$$\dot{\widehat{\mathcal{X}}}_i = \widehat{\mathcal{X}}_{i+1} + k_i \widehat{x}_1 + \left(\dot{k}_i - k_{i+1} - k_i k_1\right) x_0.$$

This implies that the estimate $\widehat{x}$ of x can be obtained as:

$$\dot{\widehat{\mathcal{X}}} = \begin{pmatrix} k_1 & 1 & 0 & \dots & 0 \\ \vdots & 0 & \ddots & \ddots & \vdots \\ k_i & & \ddots & \ddots & 0 \\ \vdots & & & 0 & 1 \\ k_n & \dots & \dots & \dots & 0 \end{pmatrix} \widehat{\mathcal{X}} + \begin{pmatrix} 0 \\ \vdots \\ \vdots \\ 0 \\ v_n \end{pmatrix} + \begin{pmatrix} \left(\dot{k}_1 - k_2 - k_1^2\right) \\ \vdots \\ \vdots \\ \left(\dot{k}_{n-1} - k_n - k_{n-1}k_1\right) \\ \left(\dot{k}_n - k_n k_1\right) \end{pmatrix} x_0, \tag{83}$$

$$\widehat{x} = \widehat{\mathcal{X}} - k\, x_0. \tag{84}$$

This establishes that the observer we propose can be implemented as a classical reduced order observer. This observer is intrinsically high-gain. However a main departure from usual high-gain observers as those studied in [9] for instance is that we have multiple high-gains. we thus do not have a single gain κ such that the ratios of $\frac{k_{i+1}}{\kappa k_i}$ remain bounded as κ goes to infinity. Following a completely new route, our ratios are designed in such a way that we can bound the estimation error in terms of the partial Lyapunov function U and the Lyapunov function $\mathcal{L}$, even when the norm of the state of system (1) goes to infinity. Such a property cannot be obtained in the classical framework of high-gain observer as discussed in [9].

5.2 A full order observer

The arguments above showing that the observer we propose can be implemented as a classical reduced order observer can be used inversely to propose a full order version of our observer.

Actually, the only important points in our design is that the estimation error satisfies (80) with the gain k_i given by (79). So let us introduce an estimation error also for x_0 and correspondingly a gain k_0 so that (80) and (79) still hold. This implies that a new function a_0 also has to be introduced. From our analysis, we know that it is sufficient to select it as a C^1 function satisfying

$$a_0(w) > \left(\frac{\frac{c_{10w}(w)^2}{\widehat{\gamma}(w)^2} + 3}{2} + [n+1][c_{20w}(w) + c_{00}(\widehat{\gamma}(w))] \right) \widehat{\gamma}(w).$$

where the functions c_{10w}, c_{20w} and c_{00} are obtained from the functions a_1 to a_n already selected.

With this data, the observer is:

$$\left\{ \begin{array}{ll} \dot{\widehat{x}}_0 & = k_0\,(x_0 - \widehat{x}_0) + \widehat{x}_1, \\ & \vdots \\ \dot{\widehat{x}}_{n-1} & = k_{n-1}\,(x_0 - \widehat{x}_0) + \widehat{x}_n, \\ \dot{\widehat{x}}_n & = k_n\,(x_0 - \widehat{x}_0) + \widehat{\gamma}(w)\,\mathtt{sat}\left(\frac{f_n(x_0, \widehat{x}_1, \ldots, \widehat{x}_n)}{\widehat{\gamma}(w)}\right), \\ \dot{w} & = -w + \gamma_1(x_0). \end{array} \right. \tag{85}$$

The interest of this extension is that now the derivatives of the k_i's are not needed, as opposed to (83). Also Theorem 2 holds in this case. Namely, we have the following theorem.

Theorem 3 *For the globally completely observable system (1), under Assumption 1, we can construct functions k_i's and $\widehat{\gamma}$, such that for each solution $X(x,t)$ of (1), right maximally defined on $[0, T)$, with $T \leq +\infty$, and for each initial condition $(\widehat{x}, w)$, the associated solution $(X(x,t), \widehat{X}((\widehat{x}, w), t), W(w,t))$ of (1)–(85) is defined also on $[0, T)$ and satisfies:*

$$\lim_{t \to T} |X(x,t) - \widehat{X}((\widehat{x}, w), t)| = 0.$$

6 On the necessity of unboundedness observability

Compared to previous results on observers and in particular to those on high gain observers, the main new point here is the exploitation of output-to-state stability as described by Assumption 1. The role of this assumption is to allow the construction of a state norm estimator for the system (1) with output (2). However, it is by far not necessary, and can be replaced by any (weaker) assumption which allows to

construct a dynamic state norm estimator. Precisely, it is clear from our analysis that the only property we need is as described in what follows.
Rewrite the system (1) with (2) in the compact form:

$$\dot{x} = F(x) \quad , \qquad y = H(x). \tag{86}$$

We need the knowledge of continuous functions $\overline{\gamma}_1$ and $\overline{\gamma}_2$ such that, for any initial condition (x, w), if $X(x,t)$ is a solution of (86) right maximally defined on $[0, T)$, then the solution $(X(x,t), W((x,w),t))$ of (86) augmented with:

$$\dot{w} = \overline{\gamma}_1(w, y)$$

is defined also on $[0, T)$ and there exists a time t_v in $[0, T)$ such that we have:

$$|X(x,t)| \leq \overline{\gamma}_2(W((w,x),t)) \qquad \forall t \in [t_v, T).$$

This last property motivates the name of state norm estimator. Clearly the existence of such an estimator implies that the solution $X(x,t)$ cannot escape in finite time if its observation $y(t) = H(X(x,t))$ does not. This is the unboundedness observability as exhibited in [20]. But conversely, we know from [1, Theorem 1] that unboundedness observability implies the existence of a state norm estimator. Moreover, the following holds.

Proposition 1 *For the system (86) with F and H locally Lipschitz functions, if there exist locally Lipschitz functions $\mathfrak{f}$ and $\mathfrak{h}$ such that the observer (10) gives an estimation error $x - \widehat{\overline{x}}$ converging to zero within the domain of existence of the solution, then there exists a C^1 and proper function V and of a nonnegative function γ_1 satisfying:*

$$\dot{\overbrace{V(x)}} \leq 1 + \gamma_1(h(x)).$$

Proof In view of [1, Theorem 1], it is sufficient to establish the following claim
Claim If there exists an initial condition x_* and a real number Y such that the corresponding solution $X(x_*, t)$ of (86) has a finite escape time T and we have:

$$|H(X(x_*, t))| \leq Y \qquad \forall t \in [0, T).$$

then, for any local Lipschitz functions $\mathfrak{f}$ and $\mathfrak{h}$ and any initial condition $\mathcal{X}_+$, we can always find an initial condition x_+ such that the corresponding solution $X(x_+, t)$ of (86) has a finite escape time t_+ and the corresponding solution $(X(x_+, t), \mathfrak{h}(\mathcal{X}((x_+, \mathcal{X}_+), t)))$ of (10)–(86), defined on $[0, t_+)$, satisfies:

$$\lim_{t \to t_+} |X(x_+, t) - \widehat{x}(t)| = +\infty, \tag{87}$$

with

$$\widehat{x}(t) = \mathfrak{h}((\mathcal{X}((x_+, \mathcal{X}_+), t)), H(X(x_+, t))). \tag{88}$$

To prove this claim, let $\mathfrak{f}$ and $\mathfrak{h}$ be any given local Lipschitz functions and $\mathcal{X}_+$ any given initial condition. Let also t_+ and x_+ be defined as:

$$\begin{aligned} t_+ &= \min\left\{\frac{T}{2}, \frac{1}{2 \sup\limits_{|\mathcal{X}-\mathcal{X}_+|\leq 1, |y|\leq Y} \{\mathfrak{f}(\mathcal{X}, y)\}}\right\}, \\ x_+ &= X(x_*, T - t_+). \end{aligned}$$

It follows that the solution $X(x_+, t)$ of (86) is right maximally defined on $[0, t_+)$. This allows us to define a time function as:

$$\begin{aligned} y(t) &= H(X(x_+, t)) \; \forall\, t \in [0, t_+), \\ &= H(x_+) \qquad \forall\, t \in [t_+, \infty). \end{aligned}$$

Consider now the solution $\mathcal{X}_y(\mathcal{X}_+, t)$ of the time-dependent ordinary differential equation:

$$\dot{\mathcal{X}} = \mathfrak{f}(\mathcal{X}, y(t)),$$

issued from $\mathcal{X}_+$. From the definition of t_+, this solution is defined at least on $[0, 2t_+]$ and satisfies:

$$|\mathcal{X}_y(\mathcal{X}_+, t) - \mathcal{X}_+| \leq 1 \qquad \forall t \in [0, 2t_+].$$

Moreover, when restricted to $[0, t_+)$, it is the $\mathcal{X}$-component of the solution of (10)–(86) issued from $(x_+, \mathcal{X}_+)$. It follows that $\widehat{x}(t)$, defined in (88), remains for all $t \in [0, t_+)$ in the compact set

$$\{\widehat{x} : \widehat{x} = \mathfrak{h}(\mathcal{X}, y), |\mathcal{X} - \mathcal{X}_+| \leq 1, |y| \leq Y\}.$$

As a result, Eq. (87) holds, since:

$$\lim_{t\to t_+} |X(x_+, t)| = \lim_{t\to T} |X(x_*, t)| = +\infty.$$

□

7 Conclusions

In this paper the problem of global observer design for general globally completely observable nonlinear systems has been studied and solved. It has been shown that global complete observability and output-to-state stability allow to explicitly construct a global observer, i.e. an observer which provides convergence to zero of the estimation error within the domain of definition of the solutions. This observer is analyzed using classical Lyapunov techniques, and it is shown that it could be implemented using a standard reduced order observer form. We have also shown that is has a full order counterpart.

While output-to-state stability is not necessary to construct the proposed observer, i.e. this property could be replaced by any weaker property which allows to build a dynamic state norm estimator, it is shown that unboundedness observability, which in principle allows to construct a dynamic norm estimator, is necessary for the design of a time-invariant observer.

References

1. Angeli D, Sontag E (1999) Forward completeness, unboundedness observability, and their Lyapunov characterizations. Syst Control Lett 38:209–217
2. Besançon G (1998) State affine systems and observer-based control. In: Proceedings of IFAC nonlinear control systems designs symposium, Enschede, The Netherlands, July 1998, pp 399–404
3. Chen C-T (1984) Linear systems theory and design. Holt (Rinehart and Winston), New York
4. Ciccarella G, Dalla Mora M, Germani A (1993) A Luenberger-like observer for nonlinear systems. Int J Control 57(3):537–556
5. Deheuvels P (1980) L'intégrale. Presses universitaires de France, Paris
6. Engel R, Kreisselmeier G (2002) A continuous-time observer which converges in finite time. IEEE Trans Automat Control 47(7):1202–1204
7. Filippov A (1988) Differential equations with discontinuous right hand sides. Mathematics and its applications, Kluwer, Dordrecht
8. Gauthier J-P, Kupka I (1994) Observability and observers for nonlinear systems. SIAM J Control Optim 32(4) 975–994
9. Gauthier J-P, Kupka I (2001) Deterministic observation theory and applications. Cambridge University Press, Cambridge
10. Gauthier J-P, Hammouri H, Kupka I (1992) A simple observer for nonlinear systems, application to bioreactor. IEEE Trans Automat Control 37:875–880
11. Hammouri H, de Leon Morales J (1991) On systems equivalence and observer synthesis. In: Proceedings of joint conference, Genoa, Italy, pp 341–347
12. Isidori A (1999) Nonlinear control systems II. Springer, Berlin Heidelberg New York
13. Jiang Z-P, Praly L (1998) Design of robust adaptive controllers for nonlinear systems with dynamic uncertainties. Automatica 34(7):835–840
14. Kazantzis N, Kravaris C (1998) Nonlinear observer design using Lyapunov's auxiliary theorem. Systems & Control Lett 34:241–247
15. Khalil HK, Esfandiari F (1993) Semiglobal stabilization of a class of nonlinear systems using output feedback. IEEE Trans Automat Control 38:1412–1415
16. Kreisselmeier G, Engel R (2003) Nonlinear observers for autonomous Lipschitz continuous systems. IEEE Trans Automat Control 48(3):451–464
17. Krener A, Isidori A (1983) Linearization by output injection and nonlinear observers. Syst Control Lett 3:47–52
18. Krener A, Respondek W (1985) Nonlinear observers with linearizable error dynamics. SIAM J Control Optim 23(2):197–216
19. Luenberger D (1964) Observing the state of a linear system. IEEE Trans Mil Electron MIL-8:74–80
20. Mazenc F, Praly L, Dayawansa WP (1994) Global stabilization by output feedback: Examples and Counter-Examples. Syst Control Lett 23:119–125
21. Nicosia S, Tornambè A (1989) High-gain observers in the state and parameter estimation of robots having elastic joints. Syst Control Lett 13:331–337
22. Sontag E, Wang Y (1996) New characterizations of input-to-state stability. IEEE Trans Automat Control 41(9)1283–1294
23. Sontag E, Wang Y (1997) Output-to-state stability and detectability of nonlinear systems. Syst Control Lett 29:279–290
24. Teel A, Praly L (1994) Global stabilizability and observability imply semi-global stabilizability by output feedback. Syst Control Lett 22:313–325
25. Teel A, Praly L (2000) A smooth Lyapunov function from a class-KL estimate involving two positive semi-definite functions. ESAIM: Control Optim Calc Var 5:313–367
26. Xiao M-Q, Krener AJ (2002) Design of reduced-order observers of nonlinear systems through change of coordinates. In: Proceedings of the 41st IEEE conference on decision and control