

# Design of local observers for autonomous nonlinear systems not in observability canonical form

Daniele Astolfi, Corrado Possieri

▶ **To cite this version:**

Daniele Astolfi, Corrado Possieri. Design of local observers for autonomous nonlinear systems not in observability canonical form. Automatica, 2019, 103, pp.443-449. 10.1016/j.automatica.2019.02.030 . hal-02059938

**HAL Id: hal-02059938**
**https://hal.science/hal-02059938**

Submitted on 7 Mar 2019

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

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

# Design of local observers for autonomous nonlinear systems not in observability canonical form ⋆

Daniele Astolfi [a] and Corrado Possieri [b]

[a] *Univ Lyon, Université Claude Bernard Lyon 1, CNRS, LAGEPP UMR 5007, 43 boulevard du 11 novembre 1918, F-69100, Villeurbanne, France.*

[b] *Dipartimento di Elettronica e Telecomunicazioni, Politecnico di Torino, 10129 Torino, Italy*

**Abstract**

The main objective of this paper is to propose a framework for the design of local asymptotic observers, with arbitrarily fast rate of convergence, for autonomous nonlinear systems that are not in observability canonical form. The proposed methodology does not require the knowledge of the inverse of the observability map. Such a goal is pursued by coupling a high-gain observer with a system that is able to locally dynamically invert the observability map.

## 1 Introduction

The main objective of this paper is to design a state observer for the autonomous nonlinear system

$$\dot{x} = f(x), \qquad y = h(x), \tag{1}$$

where $x \in \mathbb{R}^n$ is the state, $y \in \mathbb{R}$ is the output and $f : \mathbb{R}^n \to \mathbb{R}^n$, $h : \mathbb{R}^n \to \mathbb{R}$ are smooth functions. A classical approach to address such a problem (Gauthier and Kupka, 2001; Besançon, 2007) consists in determining an injective change of coordinates $\Phi : \mathbb{R}^n \to \mathbb{R}^{k+1}$, with $k \geqslant n-1$, that recasts system (1) in the so called *canonical observability form* (see Theorem 4.1 of Gauthier and Kupka, 2001):

$$\begin{aligned} \dot{z}_i &= z_{i+1}, && i = 0, \dots, k-1, \\ \dot{z}_k &= \varphi(z_0, \dots, z_k), \\ y &= z_0. \end{aligned} \tag{2}$$

Thus, letting $\hat{z}$ be an estimate of the state $z = [\, z_0 \ \cdots \ z_k \,]^\top$ of system (2) (obtained, for instance, by using the high-gain observer originally introduced in Esfandiari and Khalil, 1987; Tornambe, 1992), the current state of system (1) can be estimated as $\hat{x} = \Phi^{-1}(\hat{z})$. Note that the knowledge of $\Phi^{-1}(\cdot)$ is usually required also to determine a closed-form expression for the function $\varphi$ in (2). In practice, the computation in closed form of $\Phi^{-1}(\cdot)$ can be a too demanding task, although some attempts have been made by using algebraic geometric tools, e.g. Menini et al. (2016), high-gain approaches, e.g. Nicosia et al. (1992) Menini et al. (2017), discrete-time iterative schemes, e.g., Moraal and Grizzle (1995), or sliding mode tools, e.g. Menini et al. (2018b). An alternative approach, that does not require the knowledge of $\Phi^{-1}(\cdot)$, consists in writing a high-gain observer in the original coordinates (1) and using the inverse of the Jacobian of $\Phi(\cdot)$ in order to guarantee convergence of the estimation error in the coordinates (2). This technique has been investigated when $\Phi$ is a diffeomorphism (i.e., $k = n-1$) in Ciccarella et al. (1993); Maggiore and Passino (2003); Astolfi and Praly (2013), and extended to systems with multiple outputs in Astolfi and Praly (2017). Recently, conditions to use the same technique when $k > n-1$ have been studied in Bernard et al. (2015, 2018) using a coordinate augmentation approach.

The main objective of this work is to propose a simple and flexible framework to design a (local) asymptotic observer for system (1) that does not require the knowledge of the inverse of $\Phi(\cdot)$, that can be applied for any $k \geqslant n-1$, and with a rate of convergence that can be chosen arbitrarily fast. This is obtained by coupling in a precise way two different tools: a dynamical system that estimates the time derivatives of the output of (1) and a dynamical algorithm able to determine the inverse of a

⋆ This work has been partially prepared when the first author was working at Université de Lorraine, CRAN, F-54000 Nancy, France, and supported in part by the "Région Grand-Est" of France.

*Email addresses:* daniele.astolfi@univ-lyon1.fr (Daniele Astolfi), corrado.possieri@polito.it (Corrado Possieri).

nonlinear map. We allow a large flexibility in the design of the two aforementioned tools. For instance, the first system can be designed with any technique able to guarantee practical estimation of the output derivatives, like a standard high-gain observer, see e.g. Khalil and Praly (2014), or any variation of it (see, for instance, the approaches proposed in Astolfi and Marconi, 2015; Teel, 2016; Astolfi et al., 2018). Similarly, the second system does not strictly require the use of the inverse of the Jacobian of $\Phi$. For instance, differently from the approaches proposed in Ciccarella et al. (1993) and Hammouri et al. (2018), we allow here to use also the transpose of the Jacobian of $\Phi$. Finally, in the proposed observation scheme we do not need the knowledge of the function $\varphi(\cdot)$ in (2), which can not be determined, in general, without computing $\Phi^{-1}(\cdot)$.

With respect to the other approaches, we highlight the following differences. Differently from the techniques given in Ciccarella et al. (1993); Maggiore and Passino (2003); Astolfi and Praly (2013, 2017), we do not require the map $\Phi$ to be a diffeomorphism, but we allow the dimension of the system (2) in the canonical observability form to be larger than the dimension of the system (1) in the original coordinates, namely $k \geqslant n-1$. Furthermore, differently from Hammouri et al. (2018), we do not need any strict convexity assumption on the function $\Phi$. We believe also that this approach is in general more simple to apply than the one proposed in Bernard et al. (2015, 2018), where an injective immersion is needed to be transformed into a surjective diffeomorphism. On the other hand, as the dynamical algorithm inverting the map $\Phi$ is in general not global, the convergence properties of the proposed approach are not semi-global (namely the initial conditions of the observer cannot be chosen arbitrarily far from the ones of system (1)), whence only local convergence properties can be established *a priori*. We highlight however, that the design of local asymptotic observer may be of large interest when combined with global practical observers, as shown in Astolfi et al. (2017).

The proposed approach, at present, can be viewed as a true alternative to Extended Kalman Filters (briefly, EKF), since only local convergence is guaranteed (see Bonnabel and Slotine, 2015 and references therein). However, with respect to EKF, a certain number of advantages are ensured: we do not need extra hypotheses which may be hard to verify, such as the boundedness of the solution to a Riccati equation (which is essential to prove convergence of the EFK); we can apply the proposed design in case of injective maps $\Phi$, namely when $k > n-1$; the basin of attraction of the proposed observation scheme can be determined analytically by using the tools given in Menini et al. (2018a); the rate of convergence can arbitrarily assigned; the dimension of the observer is smaller than the one of an EKF ($2n$ versus $n+n^2$ for the case $k = n-1$).

Finally, differently from the observer given in Menini et al. (2017), the observation scheme proposed in this paper guarantees asymptotic (rather than practical) estimation of the state of system (1).

The paper is organized as follows. We introduce some preliminary notations and notions in Section 2. We provide the main results of this paper in Section 3. A numerical example is shown in Section 4. Conclusions and future developments are discussed in Section 5.

## 2 Preliminaries

### *2.1 Notation*

Let $\mathbb{R}$, $\mathbb{R}_{\geqslant 0}$, $\mathbb{R}_{>0}$, $\mathbb{Z}$, $\mathbb{Z}_{\geqslant 0}$ and $\mathbb{Z}_{>0}$ be the sets of real, nonnegative real, positive real, integer, natural numbers, and strictly positive natural numbers respectively. Let $|x|$ be the Euclidean norm of $x \in \mathbb{R}^n$, $\|x\|_{[0,t]} := \sup_{s\in[0,t]} |x(s)|$, and $\|x\|_\infty := \sup_{s\geqslant 0} |x(s)|$. Given $r \in \mathbb{R}_{\geqslant 0}$, let $\mathrm{sat}_r(\cdot)$ denote the saturation function, i.e., $\mathrm{sat}_r(x) = \max(\min(x, r), -r)$ for each $x \in \mathbb{R}$. Given $\delta \in \mathbb{R}_{>0}$, let $\mathbb{B}_\delta$ be the closed ball of radius $\delta$. Given $A \in \mathbb{R}^{n\times m}$, $n \geqslant m$, $A^\dagger := (A^\top A)^{-1}A^\top$ is the *Moore-Penrose pseudoinverse* of $A$. Given $\alpha = [\,\alpha_1 \ \cdots \ \alpha_n\,]^\top \in \mathbb{Z}^n_{\geqslant 0}$, let $|\alpha| = \sum_{i=1}^n \alpha_i$, $x^\alpha = \prod_{i=1}^n x_i^{\alpha_i}$, and $\frac{\partial^\alpha}{\partial x^\alpha} = \frac{\partial^{|\alpha|}}{\partial x_1^{\alpha_1}\cdots\partial x_n^{\alpha_n}}$.

### *2.2 Observability of a nonlinear system*

The solution of system (1) from $x(0) = x_0 \in \mathbb{R}^n$ is $x(t) := \phi(t, x_0)$ and is assumed to exist and to be unique for all times $t \in \mathbb{R}$, $t \geqslant 0$. The *observability map of order* $k+1$, $k \in \mathbb{Z}_{\geqslant 0}$, $k \geqslant n-1$, for system (1) is

$$O_k(x) := [\,h(x) \ \cdots \ L_f^k h(x)\,]^\top,$$

with $L_f^{i+1}h(x) = \frac{\partial}{\partial x}(L_f^i h(x))f(x)$, $i = 0, \ldots, k-1$, and $L_f^0 h(x) = h(x)$. The observability map $O_k(\cdot)$ relates the solution $\phi(t, x_0)$ with the time derivatives of the output $y$. Namely, by letting $z(t) = [\,y(t) \ \frac{d}{dt}y(t) \ \cdots \ \frac{d^k}{dt^k}y(t)\,]^\top$, one has

$$z(t) = O_k(\phi(t, x_0)), \quad \forall x_0 \in \mathbb{R}^n, \forall t \in \mathbb{R}_{\geqslant 0}.$$

Note that if $\mathcal{X} \subset \mathbb{R}^n$ is compact and positively invariant for system (1), i.e., $x_0 \in \mathcal{X}$ implies $\phi(t, x_0) \in \mathcal{X}$ for all $t \in \mathbb{R}_{\geqslant 0}$, then $x_0 \in \mathcal{X}$ implies that $z(t)$ is bounded for all $t \in \mathbb{R}_{\geqslant 0}$ due to the continuity of the map $O_k(\cdot)$. Let the observability map $O_k(\cdot)$ of order $k+1$, $k \in \mathbb{Z}_{\geqslant 0}$, for system (1) be given. System (1) is *locally $k$-differentially observable on $\mathcal{X} \subset \mathbb{R}^n$ if $J(x) := \frac{\partial}{\partial x}O_k(x)$ is such that* $\mathrm{rank}(J(x)) = n$, $\forall x \in \mathcal{X}$ (Isidori, 2013). Local $k$-differential observability of system (1) implies the existence of a local inverse $O_k^{-1}(\cdot)$ of $O_k(\cdot)$ in the neighborhood of each $x \in \mathcal{X}$ (Gauthier and Kupka, 2001). As a

consequence, by defining $\Phi(x) = O_k(x)$ for system (1), the function $\varphi$ in (2) is defined as $\varphi(z) = L_f^{k+1} h \circ \Phi^{-1}(z)$.

## 3 Main Results

This section is structured as follows. We detail the assumptions that we need and the design of the proposed observer in Section 3.1. We then discuss the feasibility of the proposed assumptions in Sections 3.2 and 3.3.

### *3.1 Observer design*

The main objective of this section is to show how to design a (local) observer for system (1) based on the time derivatives of the output $y$ without resorting to the inverse of the observability map. We state now the following assumptions.

**Assumption 1.** *There exist a compact set $\mathcal{X} \subset \mathbb{R}^n$, a continuous function $G : \mathcal{X} \to \mathbb{R}^{n\times(k+1)}$ and $\delta, \lambda \in \mathbb{R}_{>0}$, such that the following statements hold.*

1) *The set $\mathcal{X}_\delta := \{x \in \mathcal{X} : x + \tilde{x} \in \mathcal{X}, \forall \tilde{x} \in \mathbb{B}_\delta\}$ is nonempty.*
2) *For all $(x, \tilde{x}) \in \mathcal{X} \times \mathbb{B}_\delta$, we have that*

$$\tilde{x}^\top G(x)(O_k(x + \tilde{x}) - O_k(x)) \geqslant \lambda |\tilde{x}|^2.$$

3) *The set $\mathcal{X}_0 \subseteq \mathcal{X}_\delta$ of admissible initial states of system (1) is such that the solutions to (1) satisfy $\phi(t, x_0) \in \mathcal{X}_\delta$ for any $x \in \mathcal{X}_0$ and for all $t \geqslant 0$.*

As detailed in the subsequent Section 3.2, if system (1) is (locally) $k$-differentially observable in $\mathcal{X}$, then there exist $\lambda \in \mathbb{R}_{>0}$ and a function $G(\cdot)$ that satisfy item 2) of Assumption 1. Then, under such an assumption, it is possible to design a dynamical algorithm able to (locally) invert the mapping $O_k(\cdot)$. In particular, this is given by

$$\dot{\hat{x}} = f(\hat{x}) + \mu G(\hat{x})(O_k(x) - O_k(\hat{x})) \tag{3}$$

with $\hat{x} \in \mathbb{R}^n$ and $\mu \in \mathbb{R}_{>0}$ large enough. Solutions to (3) guarantee that the estimation error $|\hat{x} - x|$ converges exponentially to zero if its initial value is small enough (see Nicosia et al., 1994). In other words, if the mapping $O_k(\cdot)$ is known, algorithm (3) allows to (asymptotically) recover $x$, i.e., in other words, to invert $O_k(\cdot)$. However, since the observability map $O_k(\cdot)$ of system (1) is usually not directly measurable, we cannot implement system (3). As a consequence, in the next assumption, we assume that $O_k(\cdot)$ can be estimated by means of another dynamical system.

**Assumption 2.** *Let Assumption 1 hold and let $z(t) := O_k(\phi(t, x_0))$ and $y(t) := h(\phi(t, x_0))$ for all $t \geqslant 0$. There exist functions $\vartheta_\ell(\cdot), \psi_\ell(\cdot)$ parametrized by $\ell \in \mathbb{R}_{>0}$, and $\underline{\ell} \in \mathbb{R}_{>0}$, such that, for any $\ell \geqslant \underline{\ell}$, system*

$$\dot{\xi} = \vartheta_\ell(\xi, \hat{x}, y), \qquad \hat{z} = \psi_\ell(\xi) \tag{4}$$

*satisfies, for any $t \geqslant 0$,*

$$|\hat{z}(t) - z(t)| \leqslant \min\left\{R, a(\ell)\exp(-b\ell t) + \tfrac{d}{\ell}\|\hat{x} - x\|_{[0,t]}\right\} \tag{5}$$

*for any initial condition $\xi(0) \in \Xi$, where $\Xi \subseteq \mathbb{R}^m$, $m \geqslant k \geqslant n - 1$, is a compact set, $\hat{x}(t) \in \mathbb{R}^n$ is an external input which is defined for all $t \geqslant 0$, $b, d, R \in \mathbb{R}_{>0}$ are constants independent of $\ell$, and $a : \mathbb{R} \to \mathbb{R}_{>0}$ is a $\mathcal{C}^0$ function satisfying*

$$\lim_{\ell\to\infty} a(\ell)\exp(-b\ell) = 0. \tag{6}$$

Assumption 2 concerns the existence of a dynamical system that provides a practical estimate of the observability map $O_k(\cdot)$ of system (1), with an arbitrarily fast convergence rate. Feasibility of this assumption is discussed in Section 3.3.

In light of Assumptions 1 and 2, we combine systems (3) and (4) in order to obtain a (local) observer,denoted *high-gain observer with dynamical inversion* (HGOWDI), for system (1) as follows

$$\dot{\xi} = \vartheta_\ell(\xi, \hat{x}, y), \tag{7a}$$
$$\dot{\hat{x}} = f(\hat{x}) + \mu G(\hat{x})(\psi_\ell(\xi) - O_k(\hat{x})), \tag{7b}$$

where $(\xi, \hat{x}) \in \mathbb{R}^m \times \mathbb{R}^n$ is the state and $\mu \in \mathbb{R}_{>0}$ is a parameter to be chosen large enough. The following statement holds.

**Theorem 1.** *Let Assumptions 1 and 2 hold, and consider system (1) and observer (7). For any $c \in [0, 1)$, there exists $\mu^\star \in \mathbb{R}_{>0}$, and for any $\mu \geqslant \mu^\star$ there exists $\ell^\star \in \mathbb{R}_{>0}$, such that for all $\ell > \ell^\star$, the set*

$$\mathcal{A} := \big\{(x, \hat{x}, \xi) \in \mathcal{X}_\delta \times \mathcal{X} \times \Xi \ : \ \hat{x} - x = 0,\ \psi_\ell(\xi) - O_k(x) = 0\big\} \tag{8}$$

*is exponentially stable for any $\xi(0) \in \Xi$ and $\hat{x}(0) \in \mathcal{X}$ satisfying $|x(0) - \hat{x}(0)| \leqslant c\delta$. Furthermore, it is possible to make the errors $|\hat{x} - x|$ and $|\psi_\ell(\xi) - O_k(x)|$ arbitrarily small in an arbitrarily small amount of time by letting the gains $\mu$ and $\ell$ be sufficiently large.*

*Proof.* Define the error $\tilde{x} := x - \hat{x}$, whose dynamics are given by

$$\begin{aligned}\dot{\tilde{x}} &= f(x) - f(\hat{x}) - \mu G(\hat{x})(\hat{z} - O_k(\hat{x}))\\ &= f(x) - f(\hat{x}) - \mu G(\hat{x})(z - O_k(\hat{x}) + \hat{z} - z)\\ &= f(\hat{x} + \tilde{x}) - f(\hat{x})\\ &\qquad - \mu G(\hat{x})(O_k(\hat{x} + \tilde{x}) - O_k(\hat{x}) + \hat{z} - z),\end{aligned} \tag{9}$$

where, by definition, $z = O_k(x) = O_k(\hat{x} + \tilde{x})$ and $\hat{z} = \psi_\ell(\xi)$. Recall that $|\tilde{x}(0)| \leqslant c\delta$. Define [1]

$$T_m := \inf\{t \in \mathbb{R}_{\geqslant 0} : \tilde{x} \notin \mathbb{B}_\delta\},$$

[1] We use the convention $\inf \emptyset = +\infty$ as customary when dealing with the set of extended real numbers $\mathbb{R} \cup \{\pm\infty\}$.

In light of item 3) of Assumption 1 and by the absolute continuity of the solutions to systems (1) and (7), we have that $T_m > 0$ and $(x(t), \hat{x}(t), \tilde{x}(t)) \in \mathcal{X}_\delta \times \mathcal{X} \times \mathbb{B}_\delta$ for all $t \in [0, T_m)$. Recall that $f$ is smooth, $\mathcal{X}$ is compact, and $G$ is continuous. Therefore, there exist $L, \overline{G} \in \mathbb{R}_{>0}$ such that $|f(x) - f(\hat{x})| \leqslant L|\tilde{x}|$ for all $x, \hat{x} \in \mathcal{X}$ and $\sup_{x\in\mathcal{X}} |G(x)| \leqslant \overline{G}$. Moreover, by substituting $x$ by $\hat{x}$ in item 2) of Assumption 2, we obtain $\tilde{x}^\top G(\hat{x})(O_k(\hat{x} + \tilde{x}) - O_k(\hat{x})) \geqslant \lambda|\tilde{x}|^2$. As a consequence, the derivative of the function $\tilde{x} \mapsto |\tilde{x}|^2$ along the solutions to system (9) satisfies

$$\begin{aligned}\frac{d}{dt}|\tilde{x}|^2 &= 2\tilde{x}^\top \Big( f(\hat{x} + \tilde{x}) - f(\hat{x}) \\ &\qquad - \mu G(\hat{x})(O_k(\hat{x} + \tilde{x}) - O_k(\hat{x}) + \hat{z} - z)\Big) \\ &\leqslant 2L|\tilde{x}|^2 - 2\mu\tilde{x}^\top G(\hat{x})(O_k(\hat{x} + \tilde{x}) - O_k(\hat{x})) \\ &\qquad - 2\mu\tilde{x}^\top G(\hat{x})(\hat{z} - z) \\ &\leqslant -2(\mu\lambda - L)|\tilde{x}|^2 + 2\mu\overline{G}|\tilde{x}||\hat{z} - z|, \end{aligned} \tag{10}$$

for all $t \in [0, T_m)$. Thus, by integrating (10) and by recalling $|\tilde{x}(0)| \leqslant c\,\delta$, we obtain

$$\begin{aligned}|\tilde{x}(t)| \leqslant \exp(-(\lambda\mu - L)t)\, c\,\delta + \\ \mu\overline{G} \textstyle\int_0^t \exp(-(\lambda\mu - L)(t - s))|\tilde{z}(s)|ds, \end{aligned} \tag{11}$$

for all $t \in [0, T_m)$, where we denoted $\tilde{z} := \hat{z} - z$. Fix any $\mu^\star > \frac{L}{\lambda}$ and let $\varrho \in \mathbb{R}_{\geqslant 0}$ be such that

$$\tfrac{\mu\overline{G}}{\lambda\mu - L} \leqslant \varrho \qquad \forall\, \mu \geqslant \mu^\star\,. \tag{12}$$

Such a $\varrho$ exists since the term on the left is strictly decreasing for $\mu \in [\mu^\star, \infty)$. Thus, fix $\mu \geqslant \mu^\star$. By using (11) and the previous inequality, we obtain

$$\begin{aligned}|\tilde{x}(t)| \leqslant \exp(-(\lambda\mu - L)t)c\delta \\ + \varrho(1 - \exp(-(\lambda\mu - L)t)) \sup_{s\in[0,t]} |\tilde{z}(s)|, \end{aligned} \tag{13}$$

for all $t \in [0, T_m)$. Recall now that $|\tilde{z}(t)| \leqslant R$ for all $t \in \mathbb{R}_{\geqslant 0}$ in view of (5). Therefore, let $T \in \mathbb{R}_{>0}$ satisfy

$$\varrho(1 - \exp(-(\lambda\mu - L)T))R \leqslant \tfrac{(1-c)\delta}{2}\,. \tag{14}$$

Such a $T$ exists since the term on the left is 0 when $T = 0$, and increasing for $T > 0$. As a consequence, we obtain that if $|\tilde{x}(0)| \leqslant c\delta$, then $|\tilde{x}(t)| \leqslant \delta$ for all $t \in [0, T]$ (i.e., $T_m \geqslant T$). With such a $T$, let $\ell_1 \geqslant \underline{\ell}$ be such that

$$a(\ell)\exp\left(-b\ell\tfrac{T}{2}\right) + \tfrac{d}{\ell}\delta \leqslant \min\{R, \tfrac{(1-c)\delta}{2\varrho}\}$$

for any $\ell \geqslant \ell_1$. Such an $\ell_1$ does always exist in view of (6). With such choice, we obtain that $|\tilde{z}(t)| \leqslant \frac{(1-c)\delta}{2\varrho}$ for all $t \in [\frac{T}{2}, T_m)$. By using again (11) and inequality (14), we obtain

$$\begin{aligned}|\tilde{x}(t)| &\leqslant \exp(-(\lambda\mu - L)t)c\delta \\ &\qquad + \mu\overline{G}\textstyle\int_0^{\frac{T}{2}} \exp(-(\lambda\mu - L)(\frac{T}{2} - s))|\tilde{z}(s)|ds \\ &\qquad + \mu\overline{G}\textstyle\int_{\frac{T}{2}}^t \exp(-(\lambda\mu - L)(t - s))|\tilde{z}(s)|ds \\ &\leqslant \exp(-(\lambda\mu - L)t)c\delta + \tfrac{(1-c)\delta}{2} \\ &\qquad + \varrho\sup_{s\in[\frac{T}{2},t)}|\tilde{z}(s)|, \end{aligned} \tag{15}$$

for all $t \in [0, T_m)$. Hence, assume that $T_m < \infty$. By (15), this implies that there exists $t \in [\frac{T}{2}, T_m)$ such that $|\tilde{z}(t)| > \frac{(1-c)\delta}{2\varrho}$, that is in contradiction with $|\tilde{z}(t)| \leqslant \frac{(1-c)\delta}{2\varrho}$, for all $t \in [\frac{T}{2}, T_m)$. As a consequence, we have that $T_m = +\infty$ and therefore it results that $|\tilde{x}(t)| \leqslant \delta$ for all $t \geqslant 0$ and $|\tilde{z}(t)| \leqslant \frac{(1-c)\delta}{2\varrho}$ for all $t \in [\frac{T}{2}, \infty)$. Thus, by (5) and (13) , there exist functions $\varpi_1, \varpi_2$ of class $\mathcal{KL}$ (parametrized by $\ell$) such that, if $x(0) \in \mathcal{X}$, $|\tilde{x}(0)| < \delta$, and $\xi(0) \in \Xi$, then,

$$\begin{aligned}|\tilde{z}(t)| &\leqslant \max\left\{\varpi_1(|\tilde{z}(0)|, t), \tfrac{d}{\ell}\|\tilde{x}\|_\infty\right\}, \\ |\tilde{x}(t)| &\leqslant \max\left\{\varpi_2(|\tilde{x}(0)|, t), \varrho\|\tilde{z}\|_\infty\right\}, \end{aligned}$$

for all $t \in \mathbb{R}_{\geqslant 0}$. Thus, if we further impose $\ell > \ell^\star$, with $\ell^\star = \max\{\ell_1, \ell_2\}$ and $\ell_2 = d\varrho$, the small gain condition given in Teel (1996) is verified. Thus, the set $\{\tilde{z} = 0, \tilde{x} = 0\}$ is locally exponentially stable with basin of attraction including $\{|\tilde{z}| \leqslant \frac{(1-c)\delta}{2\varrho},\ |\tilde{x}| \leqslant \delta\}$. This concludes the first part of the proof.

In order to prove the second part, we need now to show that for any $\epsilon \in \mathbb{R}_{>0}$ and any $T_\epsilon \in \mathbb{R}_{>0}$, there exists $\ell, \mu \in \mathbb{R}_{>0}$ such that $|\tilde{x}(t)| + |\tilde{z}(t)| \leqslant \epsilon$ for all $t \geqslant T_\epsilon$. Hence, first select $\mu^\diamond \in \mathbb{R}_{>0}$ such that

$$\exp\left(-(\lambda\mu - L)T_\epsilon\right)\delta \leqslant \tfrac{\epsilon}{4}, \tag{16a}$$

for any $\mu \geqslant \mu^\diamond$. Fix any $\mu > \max\{\mu^\star, \mu^\diamond\}$, where $\mu^\star$ is defined as in the first part of the proof, and pick any $\overline{T} \in (0, T_\epsilon]$ such that

$$\varrho(1 - \exp(-(\lambda\mu - L)\overline{T}))R \leqslant \tfrac{\epsilon}{4}. \tag{16b}$$

Finally, let $\ell^\diamond \in \mathbb{R}_{>0}$ be such that

$$a(\ell)\exp\left(-b\ell\overline{T}\right) + \tfrac{d}{\ell}\delta \leqslant \min\left\{\tfrac{\epsilon}{4\varrho}, \tfrac{\epsilon}{4}\right\}, \tag{16c}$$

for all $\ell \geqslant \ell^\diamond$. Fix any $\ell > \max\{\ell^\star, \ell^\diamond\}$, where $\ell^\star$ is defined as in the first part of the proof. As a consequence, by recalling (5), the fact that $|\tilde{x}(t)| \leqslant \delta$ for all $t \geqslant 0$ and (16), we obtain $|\tilde{z}(t)| \leqslant R$ for $t \in [0, \overline{T}]$ and $|\tilde{z}(t)| \leqslant \frac{\epsilon}{4}$ for all $t \geqslant \overline{T}$. Furthermore, by using the bounds given in

(11), and the definition of $\overline{T}$ in (16) and of $\varrho$ in (12), we obtain

$$\begin{aligned}|\tilde{x}(t)| &\leqslant \exp(-(\lambda\mu - L)t)\delta \\ &\quad + \mu\overline{G}\int_0^{\overline{T}} \exp(-(\lambda\mu - L)(\overline{T} - s))|\tilde{z}(s)|ds \\ &\quad + \mu\overline{G}\int_{\overline{T}}^{t} \exp(-(\lambda\mu - L)(t - s))|\tilde{z}(s)|ds \\ &\leqslant \tfrac{\epsilon}{4} + \tfrac{\epsilon}{4} + \varrho \sup_{s\in[\overline{T},t)}|\tilde{z}(s)| \leqslant \tfrac{3}{4}\epsilon,\end{aligned}$$

for all $t \geqslant T_\varepsilon$. Hence, for the selected values of the high-gain parameters $\ell$ and $\mu$, we have that $|\tilde{x}(t)| + |\tilde{z}(t)| \leqslant \epsilon$ for all $t \geqslant T_\epsilon$. This concludes the proof. ■

### *3.2 Feasibility of Assumption 1*

Consider the following statement that provides verifiable conditions ensuring that items 1) and 2) of Assumption 1 hold.

**Proposition 1.** *Suppose there exist a compact set $\overline{\mathcal{X}} \subset \mathbb{R}^n$ and $\delta^\star \in \mathbb{R}_{>0}$ such that system (1) is locally $k$–differentially observable on $\overline{\mathcal{X}} + \mathbb{B}_{\delta^\star}$. Then, there exist $\mathcal{X} \subset \mathbb{R}^n$, $G : \mathcal{X} \to \mathbb{R}^{n\times(k+1)}$, $\delta \in \mathbb{R}_{>0}$, and $\lambda \in \mathbb{R}_{>0}$ such that Assumption 1 holds.*

*Proof.* First of all, $k$–differentially observability implies $\operatorname{rank}(J(x)) = n$ for all $x \in \overline{\mathcal{X}} + \mathbb{B}_{\delta^\star}$. Therefore, there exists a continuous bounded function $G : \overline{\mathcal{X}} + \mathbb{B}_{\delta^\star} \to \mathbb{R}^{n\times(k+1)}$ and $\bar{\lambda} \in \mathbb{R}_{\geqslant 0}$ such that $G(x)J(x) > \bar{\lambda}I$ for all $x \in \overline{\mathcal{X}} + \mathbb{B}_{\delta^\star}$ (see the subsequent Remark 1 for two possible choices of $G$ satisfying such properties). Furthermore, since the map $O_k(\cdot)$ is smooth, by using the multivariate version of Taylor's theorem (Königsberger, 2013), one has that, for all $x \in \mathcal{X}$ and $\tilde{x} \in \mathbb{B}_{\delta^\star}$,

$$O_k(x + \tilde{x}) = O_k(x) + J(x)\tilde{x} + \sum_{|\alpha|=2} \Upsilon_\alpha(x + \tilde{x})\tilde{x}^\alpha,$$

where, for all $\alpha \in \mathbb{Z}^n$ such that $|\alpha| = 2$,

$$\Upsilon_\alpha(x + \tilde{x}) := \frac{2}{\alpha!}\int_0^1 (1 - t)\frac{\partial^2 O_k}{\partial \tilde{x}^\alpha}(x + t\tilde{x})dt$$

In particular, due to the smoothness of $O_k(\cdot)$ and to the compactness of $\mathcal{X}$ and of $\mathbb{B}_{\delta^\star}$, one has that there exists $\overline{H} \in \mathbb{R}_{>0}$ such that, for all $(x, \tilde{x}) \in \mathcal{X} \times \mathbb{B}_{\delta^\star}$

$$|\Upsilon_\alpha(x + \tilde{x})| \leqslant \max_{|\beta|=2,\, x\in\mathcal{X},\, \tilde{x}\in\mathbb{B}_{\delta^\star}} \left|\frac{\partial^2}{\partial\tilde{x}^\beta} O_k(x + \tilde{x})\right| \leqslant \overline{H}.$$

This implies that, for all $(x, \tilde{x}) \in \mathcal{X} \times \mathbb{B}_{\delta^\star}$,

$$\begin{aligned}\left|\sum_{|\alpha|=2} \Upsilon_\alpha(x + \tilde{x})\tilde{x}^\alpha\right| &\leqslant \sum_{|\alpha|=2} |\Upsilon_\alpha(x + \tilde{x})\,|\,|\tilde{x}^\alpha| \\ &\leqslant \overline{H} \sum_{|\alpha|=2} |\tilde{x}^\alpha| \leqslant n^2\, \overline{H}\, |\tilde{x}|^2.\end{aligned}$$

Furthermore, by the boundedness of $G(\cdot)$, there exists $\overline{G} \in \mathbb{R}_{>0}$ such that $|G(x)| \leqslant \overline{G}$ for all $x \in \mathcal{X}$. Therefore, we have that, for all $(x, \tilde{x}) \in \mathcal{X} \times \mathbb{B}_{\delta^\star}$,

$$\tilde{x}^\top G(x)(O_k(x + \tilde{x}) - O_k(x)) \geqslant \bar{\lambda}|\tilde{x}|^2 - n^2\overline{G}\overline{H}|\tilde{x}|^3.$$

Thus, item 2) of Assumption 1 holds with $\delta < \min\left\{\frac{\bar{\lambda}}{2n^2\overline{G}\overline{H}}, \delta^\star\right\}$, $\mathcal{X} = \overline{\mathcal{X}} + \mathbb{B}_\delta$, $\lambda = \frac{\bar{\lambda}}{2}$ and the function $G$ previously defined. Finally, item 1) of Assumption 1 is verified since the set $\mathcal{X}_\delta$ coincides with $\overline{\mathcal{X}}$ which is non empty by assumption. ■

In view of Proposition 1, we conclude that Assumption 1 allows to take into consideration a large class of autonomous systems, namely all $k$-differentially observable systems whose trajectories evolve in some given compact set. This is the case of systems possessing an equilibrium, a limit cycle, or chaotic (bounded) behaviours.

The following remark details how to choose the function $G$ so that the assumptions of Proposition 1 are met.

*Remark* 1. By Menini et al. (2017), if the observability map $O_k : \mathcal{X} \to \mathbb{R}^{k+1}$ is an immersion, whence $\operatorname{rank}(J(x)) = n$ for all $x \in \mathcal{X}$ and the set $\mathcal{X}$ is compact, the following choices for $G : \mathcal{X} \to \mathbb{R}^{n\times(k+1)}$ can be made in order to meet the assumptions of Proposition 1:

1) $G(x) = J^\top(x)$, obtaining a gradient-like algorithm for the inversion of the observability map;
2) $G(x) = J^\dagger(x)$, obtaining a Newton-like algorithm for the inversion of the observability map.

In the first case, since $\operatorname{rank}(J(x)) = n$ for all $x \in \mathcal{X}$, one has that $vJ^\top(x)J(x)v = |J(x)v|^2 > 0$, for all $(x, v) \in \mathcal{X} \times \mathbb{R}^n$ such that $|v| = 1$. Therefore, since the eigenvalues of a matrix are continuous functions of its entries, by letting $\lambda_{\min}(x)$ be the smallest eigenvalue of $J(x)^\top J(x)$, if $\mathcal{X}$ is compact, then there exists $\lambda = \min_{x\in\mathcal{X}} \lambda_{\min}(x)$, $\lambda \in \mathbb{R}_{>0}$, such that $J^\top(x)J(x) > \lambda I$. In the second case, since $\operatorname{rank}(J(x)) = n$ for all $x \in \mathcal{X}$, one obtains $J^\dagger(x)J(x) = I$ for all $x \in \mathcal{X}$. Boundedness of $J^\dagger(x)$ follows again by the compactness of $\mathcal{X}$. In fact, since $J^\top(x)J(x) > \lambda I$ for all $x \in \mathcal{X}$, $J(x)^\top J(x)$ is invertible for each $x \in \mathcal{X}$, whence $J^\dagger(x)$ is continuous in $\mathcal{X}$ and its boundedness follows by the extreme value theorem. Note that if both $f$ and $h$ in (1) are known elementary functions, the matrix $J$ can be easily computed through any software able to carry out symbolic differentiation. Therefore, the main advantage of choosing $G(x) = J^\top(x)$ relies on the fact that, in such a case, a symbolic expression for the map $G$ is usually available, even though the parameter $\lambda$ in item 2) of Assumption 1 may be rather small. On the other hand, if one chooses $G(x) = J^\dagger(x)$, the parameter $\lambda$ in item 2) of Assumption 1 is always equal to 1, but determining a symbolic expression of $J^\dagger(x)$ may be rather challenging, especially if $n$ and $k$ are large integers. □

The following remark provides further details about the parameter $\delta$ that makes Assumption 1 be fulfilled.

*Remark* 2. In the proof of Proposition 1, it has been assumed that the parameter $\delta \in \mathbb{R}_{>0}$ is sufficiently small in order to guarantee that the Taylor approximation up to the second order of the observability map is sufficiently close to its actual value. However, in practical applications, such a parameter need not be small and is essentially related to the positive definiteness of $\tilde{x}^\top G(\hat{x})(O_k(\hat{x}+\tilde{x}) - O_k(\hat{x}))$ for $(\hat{x}, \tilde{x}) \in \mathcal{X} \times \mathbb{B}_\delta$. Numerical techniques able to determine the greatest value of $\delta$ that makes such an expression be positive definite with respect to $\tilde{x}$ are given in Ważewski (1948); Zampieri (1993). On the other hand, in Menini et al. (2018a), an exact technique, based on algebraic geometry tools, has been proposed to estimate the largest value of $\delta$ that satisfies the above requirements. □

### *3.3 Feasibility of Assumption 2*

The main objective of this section is to provide an observer for the time derivatives of the output of system (1) satisfying Assumption 2. In particular, we will show that a classical high-gain observer (see Khalil and Praly, 2014 and references therein), or some of its recent variations (see, e.g., Astolfi and Marconi, 2015; Teel, 2016; Astolfi et al., 2018), can be used for this purpose, although we highlight that other design methods can be used to obtain (4).

First, let $\mathcal{X}$ be the set claimed in Assumption 1 and define

$$r_i := \sup_{x\in\mathcal{X}} |L_f^i h(x)|, \quad i = 0, \dots, k+1. \tag{17}$$

The observer (4) satisfying Assumption 2 can be designed as

$$\begin{aligned} \dot{\xi}_i &= \xi_{i+1} + \gamma_i \ell^{i+1}(y - \xi_0), \quad i = 0, \dots, k-1, \\ \dot{\xi}_k &= \mathrm{sat}_{r_{k+1}}(L_f^{k+1} h(\hat{x})) + \gamma_k \ell^{k+1}(y - \xi_0), \\ \hat{z} &= [\, \mathrm{sat}_{r_0}(\xi_0) \ \cdots \ \mathrm{sat}_{r_k}(\xi_k)\,]^\top, \end{aligned} \tag{18}$$

where $\xi = (\xi_1, \dots, \xi_n)^\top \in \mathbb{R}^n$ is the state, $\gamma_0, \dots, \gamma_k \in \mathbb{R}_{>0}$ are coefficients selected such that the polynomial $s^{k+1} + \sum_{i=0}^{k} \gamma_i s^{k-i}$ is Hurwitz, and $\ell \in \mathbb{R}_{>0}$ is a high-gain parameter to be chosen large enough. Note that the observer (18) has been obtained from the classical high-gain observer (see Khalil and Praly, 2014 and references therein), by substituting $\varphi(\hat{z}) := L_f^{k+1} h \circ O_k^{-1}(\hat{z})$ with $\mathrm{sat}_{r_{k+1}}(L_f^{k+1} h(\hat{x}))$ and by saturating its output. The following proposition states that the observer (18) satisfies Assumption 2.

**Proposition 2.** *Let Assumption 1 hold. Then, Assumption 2 holds with system* (4) *defined as* (18).

*Proof.* Note that $L_f^{k+1} h(\cdot)$ is Lipschitz on $\mathcal{X}$ and that $O_{k+1}(\mathcal{X})$ is bounded due to the smoothness of $f$ and $h$ and to the compactness of $\mathcal{X}$. Since the output of system (18) is saturated and $z$ ranges in the compact set $O_k(\mathcal{X})$, we have that, for all $t \in \mathbb{R}_{\geqslant 0}$,

$$|z(t) - \hat{z}(t)| \leqslant \sum_{i=0}^{k} |z_i(t) - \hat{z}_i(t)| \leqslant R, \tag{19}$$

with $R = 2\sum_{i=0}^{k} r_i$. Define $\varepsilon_i = \ell^{-i}(z_i - \xi_i)$, $i = 0, \dots, k$, and let $\varepsilon = [\, \varepsilon_0 \ \cdots \ \varepsilon_k \,]^\top$, whose dynamics are

$$\dot{\varepsilon} = \ell A_k \varepsilon + \ell^{-k} B u_{k-1}, \tag{20}$$

where $A_k$ is the companion matrix of the Hurwitz polynomial $s^{k+1} + \sum_{i=1}^{k+1} \gamma_i s^{k+1-i}$, $B = [\, 0 \ \cdots \ 0 \ 1 \,]^\top$ and $u_{k-1} = \mathrm{sat}_{r_{k+1}}(L_f^{k+1} h(\hat{x})) - L_f^{k+1} h(x)$. Since $L_f^{k+1} h(\cdot)$ is Lipschitz and bounded on $\mathcal{X}$, in light of the saturation function, there exists $L \in \mathbb{R}_{>0}$ such that

$$|\mathrm{sat}_{r_{k+1}}(L_f^{k+1} h(\hat{x})) - L_f^{k+1} h(x)| \leqslant L|\hat{x} - x|$$

for all $x \in \mathcal{X}$ and $\hat{x} \in \mathbb{R}^n$. Therefore, we have that $\|u_{k-1}\|_{[0,t]} \leqslant L\|\hat{x} - x\|_{[0,t]}$, and hence, by (20) and since $A_k$ is Hurwitz, there exist $c_1, c_2 \in \mathbb{R}_{>0}$ such that

$$\begin{aligned} |\varepsilon(t)| &\leqslant |\exp(\ell A_k t)||\varepsilon(0)| \\ &\quad + \ell^{-k} \textstyle\int_0^t |\exp(\ell A_k (t-\tau))||B u_{k-1}(\tau)| d\tau \\ &\leqslant c_1 \exp(-c_2 \ell t)|\varepsilon(0)| + \tfrac{c_1}{c_2 \ell^{k+1}} L \|\hat{x} - x\|_{[0,t]}. \end{aligned}$$

Recall that for any $\ell \geqslant 1$, we have $|\hat{z} - z| \leqslant |\xi - z| \leqslant \ell^k |\varepsilon|$ and $|\varepsilon(0)| \leqslant |z(0) - \xi(0)|$. As a consequence, by using previous inequalities and (19), we obtain that the conditions of Assumption 2 are satisfied with $\underline{\ell} = 1$, $b = c_2$, $d = \frac{L c_1}{c_2}$, and $a(\ell) = 2 c_1 c_3 \ell^k$, where $c_3 \in \mathbb{R}_{>0}$ is such that $O_{k+1}(\mathcal{X}) \subseteq \mathbb{B}_{c_3}$ and $\Xi \subseteq \mathbb{B}_{c_3}$. ■

A result similar to Proposition 2 can be given also for the variations of the high-gain observer proposed in Astolfi and Marconi (2015); Teel (2016); Astolfi et al. (2018). For instance, the low-power peaking-free high-gain observer given in Astolfi et al. (2018) can be modified as

$$\dot{\hat{z}}_i = \eta_i + \alpha_i \ell e_i, \qquad i = 0, \dots, k-1, \tag{21a}$$
$$\dot{\hat{z}}_k = \mathrm{sat}_{r_{k+1}}(L_f^{k+1} h(\hat{x})) + \alpha_k \ell e_k, \tag{21b}$$
$$\dot{\eta}_i = \mathrm{sat}_{r_{i+2}}(\eta_{i+1}) + \beta_i \ell^2 e_i, \ i = 0, \dots, k-2, \tag{21c}$$
$$\dot{\eta}_{k-1} = \mathrm{sat}_{r_{k+1}}(L_f^{k+1} h(\hat{x})) + \beta_{k-1} \ell^2 e_{k-1}, \tag{21d}$$
$$\hat{z} = [\, \hat{z}_0 \ \cdots \ \hat{z}_k \,]^\top \tag{21e}$$

where $\xi = (\hat{z}, \eta) \in \mathbb{R}^k \times \mathbb{R}^{k-1}$ is the state of the observer, $\alpha_0, \cdots, \alpha_k$ and $\beta_0, \cdots, \beta_{k-1}$ are coefficients satisfying the low-power strong stability requirement (see

Definition 2 of Astolfi et al., 2018), $\ell \in \mathbb{R}_{>0}$, $\ell \geqslant 1$, is the high-gain parameter, and

$$e_0 = y - \hat{z}_0, \tag{21f}$$
$$e_i = \mathrm{sat}_{r_i}(\eta_{i-1}) - \hat{z}_i, \qquad i = 1, \dots, k. \tag{21g}$$

**Proposition 3.** *Let Assumption 1 hold. Then, Assumption 2 holds with system* (4) *defined as* (21)*.*

*Proof.* By Proposition 3 of Astolfi et al. (2018), the trajectories of system (21) are bounded whenever $\mathrm{sat}_{r_{k+1}}(L_f^{k+1}h(\hat{x})) - L_f^{k+1}h(x)$ is bounded and for each $T \in \mathbb{R}_{>0}$, there exists $\ell^\star \in \mathbb{R}_{>0}$, $\ell^\star \geqslant 1$, such that, for each $\ell > \ell^\star$ and for all $t > T$,

$$\mathrm{sat}_{r_{i+1}}(\eta_i(t)) = \eta_i(t), \quad i = 0, \dots, k-1.$$

Therefore, letting $T$ and $\ell^\star$ be as above, if $\ell > \ell^\star$, then, by letting $\varepsilon_i = [\, \frac{1}{\ell^i}(\hat{z}_i - z_i) \;\; \frac{1}{\ell^{i+1}}(\eta_i - z_{i+1}) \,]^\top$ and $\varepsilon_{e,k} = [\, \varepsilon_0 \; \cdots \; \varepsilon_k \,]^\top$, the dynamics of $\varepsilon_{e,k}$ are, for $t \geqslant T$,

$$\dot{\varepsilon}_{e,k} = \ell M_k \varepsilon_{e,k} - \ell^{-k} \bar{B} u_{k-1},$$

where $M_k$ is a Hurwitz matrix, $\bar{B} = [\, 0 \; \cdots \; 0 \; 1 \; 1 \,]^\top$ and $u_{k-1} = \mathrm{sat}_{r_{k+1}}(L_f^{k+1}h(\hat{x})) - L_f^{k+1}h(x)$. Thus, the proof follows by the same reasoning used for Proposition 2. ■

Note that, in practical applications (as the one reported in the forthcoming section), the design (21) may be preferred to (18), due to the fact that the "low-power design" (21) has better sensitivity properties with respect to high-frequency noise, see Astolfi et al. (2018).

## 4 Numerical example

Consider the system (Thomas, 1999)

$$\begin{aligned} \dot{x}_1 &= \sin(x_2), \quad \dot{x}_2 = \sin(x_3), \quad \dot{x}_3 = \sin(x_1), \\ y &= x_1, \end{aligned} \tag{22}$$

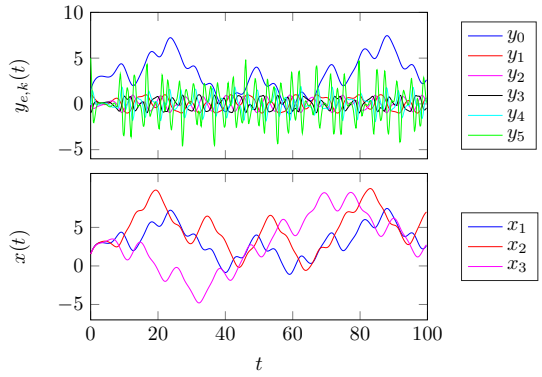
that can be viewed as the dynamics of an undamped particle moving in a three-dimensional lattice of forces. These dynamics exhibits anomalous diffusion and has been described as a form of deterministic Brownian motion (Rowlands and Sprott, 2008), thus making the state estimation process particularly challenging. It is worth noticing that, although system (22) is globally Lipschitz, the approach based on the solution of LMIs given for instance in Zemouche and Boutayeb (2013) does not allow the design of an observer for such a system. As a matter of fact, the LMI given in (35) of Zemouche and Boutayeb (2013) is unfeasible for the considered system. Furthermore, note that the linearization of system (22) about the point $[\, \frac{\pi}{2} \;\; \frac{\pi}{2} \;\; \frac{\pi}{2} \,]^\top$ is not observable. Hence, even if one allows for time-varying gains, the design of an observer based on the linearization of system (22) may be particularly challenging. Note also that system (22) is not globally observable since $L_f^j h(x) = L_f^j h(x + [\, 0 \;\; 2k_1\pi \;\; 2k_2\pi \,]^\top)$ for all $j \in \mathbb{Z}_{>0}$ and $k_1, k_2 \in \mathbb{Z}$.

As it has already been noticed in Menini et al. (2017), the observability map $O_5(x)$ of system (22) is locally invertible in the neighborhood of each $x \in \mathbb{R}^3$, but not globally. However, as shown in the following, the observer (7), with (7a) selected as in (21), can be used to determine an asymptotic estimate of the state of system (22), provided $\hat{x}(0)$ is sufficiently close to $x(0)$.

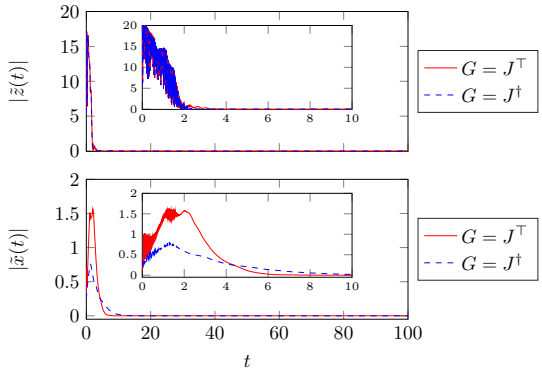
A numerical simulation has been carried out to test the observer (7b), (21) for both the choices suggested in Remark 1 (i.e., $G(x) = J^\top(x)$ and $G(x) = J^\dagger(x)$). For comparative purposes, in all the simulations the following parameters have been assumed: $x_0 = [\, 1.5808 \;\; 1.5408 \;\; 1.5508 \,]^\top$, $\hat{x}(0) = [\, 1.6808 \;\; 1.3408 \;\; 1.4508 \,]^\top$, $\hat{z}_i(0) = 0$, $i = 0, \dots, 6$, $\eta_i(0) = 0$, $i = 0, \dots, 5$, $\alpha_0 = 1$, $\alpha_1 = 2$, $\alpha_2 = 3$, $\alpha_3 = 4$, $\alpha_4 = 5$, $\alpha_5 = 1$, $\beta_0 = 1$, $\beta_1 = 1.3628$, $\beta_2 = 0.2498$, $\beta_3 = 0.3331$, $\beta_4 = 0.2651$ (that have been designed by using the algorithm given in Appendix A.1 of Astolfi et al. (2018)), $r_i = 10$, $i = 0, \dots, 6$, $\ell = 200$ and $\mu = 0.9$. Figure 1 depicts the results of the simulation.

Several properties of the proposed observers can be noticed from such Figure 1. First of all, we remark that in light of the low-power peaking-free high-gain observer design (21), the estimates $\hat{z}$ are not affected by the peaking phenomenon that is common in classical high-gain observers (see Khalil and Praly, 2014 for further details). From Figure 1(b), it can be noticed that the choices $G(x) = J^\top(x)$ and $G(x) = J^\dagger(x)$ perform similarly, leading to estimation errors for $z(t)$ and $x(t)$ that go to 0 exponentially.

It is worth noticing that, for both the choices $G(x) = J^\top(x)$ and $G(x) = J^\dagger(x)$, the basin of attraction of the proposed observer can be estimated by using the tools given in Gorni and Zampieri (1994); Menini et al. (2018a). Namely, in view of Assumption 1, it suffices to determine the largest $\delta \in \mathbb{R}_{>0}$ such that $\tilde{x}^\top G(x)(O_k(x + \tilde{x}) - O_k(x)) - \lambda|\tilde{x}|^2 \geqslant 0$ for some $\lambda \in \mathbb{R}_{>0}$ and for all $\tilde{x} \in \mathbb{B}_\delta$. It turns out that, for $\lambda = \frac{1}{2}$, the choices $G(x) = J^\top(x)$ and $G(x) = J^\dagger(x)$ lead to a maximal allowed norm of the initial error equal to 0.0759 and 1.0808, respectively. Therefore, the selection $G(x) = J^\dagger(x)$ leads to a larger admissible initial error, even though, with such a choice, the observer given in (7b), (21) has a more




(a) Time derivatives of $y$ and solution $x$.




(b) Norm of $\tilde{z} = z - \hat{z}$ and $\tilde{x} = x - \hat{x}$.

Fig. 1. Results of the numerical simulation.

cumbersome expression due to the presence of pseudo-inverse of the matrix $J$. Note that, by Theorem 1, the gains $\mu$ and $\ell$ do not affect the basin of attraction of the proposed observer, provided they are chosen sufficiently large in order to guarantee its asymptotic stability.

Finally, it is worth stressing again that the observer given in (7b), (21) does not require neither the existence of a global inverse nor a closed-form expression for a local inverse $O_k^{-1}(\cdot)$. Furthermore, such an observer is genuinely nonlinear, thus allowing to estimate the state of the system although its linearization is not observable.

## 5 Conclusions

In this paper, it has been shown that the knowledge of the inverse of the observability map is not a strict requirement to design local state observers for autonomous nonlinear systems that are not in observability canonical form. Namely, it has been shown that, under some mild regularity assumptions, a local state observer can be designed by coupling high-gain observers with a system that dynamically inverts the observability map. Although only systems with scalar output have been taken into account, results wholly similar to the ones given in this paper hold if the output is not scalar, see e.g. Gauthier and Bornard (1981).

As shown in Section 3.3, the proposed framework is very flexible since it does not require the use of a particular high-gain observer, but it can be employed with any tool able to estimate the time derivatives of the output with arbitrarily fast decay rate (thus including the sliding mode techniques given in Shtessel et al., 2014).

Future works will attempt:to extend the proposed technique to systems with inputs (Bernard et al., 2017); to enlarge the domain of attraction of the dynamical inverse (7b) through a suitable choice of the function $G$ and by requiring strong $k$-differential observability of the system rather than just local $k$-differential observability; to characterize the effect of measurement noise on the obtained estimate (Astolfi et al., 2018). Finally, it is worth stressing that, since the state of the proposed observer converges exponentially to the state of system (1), by applying classical Lyapunov arguments, it is possible to show that the convergence is robust with respect to small perturbation, such as model uncertainties, and to small measurement noise.

## References


Astolfi, D. and Marconi, L. (2015). A high-gain nonlinear observer with limited gain power, *IEEE Trans. Automat. Contr.* **60**(11): 3059–3064.

Astolfi, D., Marconi, L., Praly, L. and Teel, A. R. (2018). Low-power peaking-free high-gain observers for nonlinear systems, *Automatica* **98**: 169–179.

Astolfi, D., Postoyan, R. and Nešić, D. (2017). Uniting local and global observers for the state estimation of nonlinear continuous-time systems, *56th IEEE Conf. Decis. Control.*, pp. 3039–3044.

Astolfi, D. and Praly, L. (2013). Output feedback stabilization for siso nonlinear systems with an observer in the original coordinates, *52nd IEEE Conf. Decis. Control.*, pp. 5927–5932.

Astolfi, D. and Praly, L. (2017). Integral action in output feedback for multi-input multi-output nonlinear systems, *IEEE Trans. Automat. Contr.* **62**(4): 1559–1574.

Bernard, P., Andrieu, V. and Praly, L. (2018). Expressing an observer in preferred coordinates by transforming an injective immersion into a surjective diffeomorphism, *SIAM J. Control Optim.* **56**(3): 2327–2352.

Bernard, P., Praly, L. and Andrieu, V. (2015). Tools for observers based on coordinate augmentation, *54th IEEE Conf. Decis. Control.*, pp. 6324–6329.

Bernard, P., Praly, L., Andrieu, V. and Hammouri, H. (2017). On the triangular canonical form for uniformly observable controlled systems, *Automatica* **85**: 293–300.

Besançon, G. (2007). *Nonlinear observers and applications*, Vol. 363, Springer.

Bonnabel, S. and Slotine, J.-J. (2015). A contraction

theory-based analysis of the stability of the deterministic extended Kalman filter, *IEEE Trans. Automat. Contr.* **60**(2): 565–569.

Ciccarella, G., Dalla Mora, M. and Germani, A. (1993). A luenberger-like observer for nonlinear systems, *Int. J. Control* **57**(3): 537–556.

Esfandiari, F. and Khalil, H. K. (1987). Observer-based design of uncertain systems: recovering state feedback robustness under matching conditions, *Proc. Allerton Conf*, pp. 97–106.

Gauthier, J. and Bornard, G. (1981). Observability for any *u(t)* of a class of nonlinear systems, *IEEE Trans. Automat. Contr.* **26**(4): 922–926.

Gauthier, J.-P. and Kupka, I. (2001). *Deterministic observation theory and applications*, Cambr. Univ. Press.

Gorni, G. and Zampieri, G. (1994). Injectivity onto a star-shaped set for local homeomorphisms in n-space, *Ann. Polonici Math.* **59**(2): 171–196.

Hammouri, H., Ahmed, F. and Othman, S. (2018). Observer design based on immersion technics and canonical form, *Syst. Control Lett.* **114**: 19–26.

Isidori, A. (2013). *Nonlinear control systems*, Springer.

Khalil, H. K. and Praly, L. (2014). High-gain observers in nonlinear feedback control, *Int. J. Robust Nonlin. Control* **24**(6): 993–1015.

Königsberger, K. (2013). *Analysis 2*, Springer.

Maggiore, M. and Passino, K. M. (2003). A separation principle for a class of non-uco systems, *IEEE Trans. Automat. Contr.* **48**(7): 1122–1133.

Menini, L., Possieri, C. and Tornambe, A. (2016). Switching signal estimator design for a class of elementary systems, *IEEE Trans. Automat. Contr.* **61**(5): 1362–1367.

Menini, L., Possieri, C. and Tornambe, A. (2017). A "practical" observer for nonlinear systems, *56th IEEE Conf. Decis. Control.*, pp. 6720–6725.

Menini, L., Possieri, C. and Tornambe, A. (2018a). Estimation of the basin of attraction of a practical high-gain observer, *European Control Conference*, pp. 1238–1243.

Menini, L., Possieri, C. and Tornambe, A. (2018b). A Newton–like algorithm to compute the inverse of a nonlinear map that converges in finite time, *Automatica* **89**: 411–414.

Moraal, P. and Grizzle, J. (1995). Observer design for nonlinear systems with discrete-time measurements, *IEEE Trans. Automat. Contr.* **40**(3): 395–404.

Nicosia, S., Tornambe, A. and Valigi, P. (1992). Robust inversion of nonlinear maps, *J. Math. Syst. Estim. Control* **2**(1): 45–69.

Nicosia, S., Tornambe, A. and Valigi, P. (1994). Nonlinear map inversion via state observers, *Circuits Syst. Signal Proces.* **13**(5): 571–589.

Rowlands, G. and Sprott, J. (2008). A simple diffusion model showing anomalous scaling, *Physics of Plasmas* **15**(8).

Shtessel, Y., Edwards, C., Fridman, L. and Levant, A. (2014). *Sliding mode control and observation*, Springer Sci. Bus. Media, New York.

Teel, A. R. (1996). A nonlinear small gain theorem for the analysis of control systems with saturation, *IEEE Trans. Automat. Contr.* **41**(9): 1256–1270.

Teel, A. R. (2016). Further variants of the Astolfi/Marconi high-gain observer, *American Control Conference*, pp. 993–998.

Thomas, R. (1999). Deterministic chaos seen in terms of feedback circuits: Analysis, synthesis, "labyrinth chaos", *Int. J. Bifurcation Chaos* **9**(10): 1889–1905.

Tornambe, A. (1992). High-gain observers for non-linear systems, *Int. J. Syst. Sci.* **23**(9): 1475–1489.

Ważewski, T. (1948). Sur l'évaluation du domaine d'existence des fonctions implicites réelles ou complexes, *Ann. Soc. Po. Math.*

Zampieri, G. (1993). Finding domains of invertibility for smooth functions by means of attraction basins, *J. Differ. Equ.* **104**(1): 11–19.

Zemouche, A. and Boutayeb, M. (2013). On LMI conditions to design observers for Lipschitz nonlinear systems, *Automatica* **49**(2): 585–591.