Luenberger observers for non autonomous nonlinear systems

Pauline Bernard, Vincent Andrieu

Luenberger observers for non autonomous nonlinear systems

Pauline Bernard and Vincent Andrieu

***Abstract*—We show how the (initial) Luenberger methodology presented in [1] for linear systems can be used to design causal observers for controlled nonlinear systems. Their implementation relies on the resolution of a time-varying PDE, the solutions of which transform the dynamics into linear asymptotically stable ones. We prove the existence and injectivity (after a certain time) of such transformations, under standard observability assumptions such as differential observability or backward-distinguishability. We show on examples how this PDE can be solved and how the observability assumptions can be checked. Also, we show that similarly to the high gain framework, it is possible to use a time-independent transformation when the system is observable for any input and strongly differentially observable of order the dimension of the system.**

***Index Terms*—observer, Kazantzis-Kravaris, Luenberger**

I. INTRODUCTION

A. Context

Online estimation of the state of a dynamical system is crucial in practice, especially for monitoring or control purposes. However, very few general observer design methods exist for nonlinear time varying systems or for nonlinear systems with dynamics depending on an exogenous input. Some, such as the popular *extended Kalman filters* ([2]) rely on linearization methods, but thus provide only local convergence. Others consist in finding a reversible input-independent change of coordinates, which transforms the dynamics into a more favorable form such as state-affine time-varying forms ([3],[4] among others), for which a Kalman filter can be used, or a triangular form ([5], [6]) for which a high gain observer can be used. But the existence of such a change of coordinates usually requires restrictive assumptions on the system, such as the so-called *uniform observability* for triangular forms, or can be applied to some particular classes of nonlinear systems (see for instance the Immersion & Invariance approach in [7]).

On the other hand, if we allow the transformation to depend on the input, the range of possibilities widens. For instance, the transformation obtained by considering the output and a certain number of its derivatives, transforms the dynamics into the so-called *phase-variable form* ([8], [9]) under the *ACP*(N) *condition* ([9]), which roughly says that the Nth-derivative of the output can be expressed in a "Lipschitz" way in terms of the first $N-1$ ones. In this case, a classical high gain observer can be used. But a drawback of this transformation is that it involves the input's time derivatives (appearing when differentiating the output), which may make this solution unsuitable for practical applications.

P. Bernard was with MINES ParisTech, PSL Research University, Paris, France. e-mail: `pauline.bernard@mines-paristech.fr`.

V. Andrieu is with LAGEP, CNRS, CPE, Université Lyon 1, France. e-mail: `vincent.andrieu@gmail.com`

In a completely independent line of research, some researchers have tried to reproduce Luenberger's initial[1] methodology presented in [1] for linear systems on nonlinear systems. Indeed, initially, Luenberger's method consisted in transforming the system into a linear asymptotically stable one for which a trivial observer (made of a copy of the dynamics) exists. The extension of such a method to autonomous nonlinear systems was proposed and analyzed in a general context by [10], and rediscovered later by [11] where a local analysis close to an equilibrium point was given under conditions then relaxed in [12]. The localness as well as most of the restrictive assumptions were finally by-passed in [13], leading to the so-called *Kazantzis-Kravaris-Luenberger* (KKL) observers.

In this paper, we want to extend the use of those Luenberger observers to non autonomous systems. By non autonomous systems, we mean systems which may be time varying or which may depend on exogeneous signals. Exactly as in the high gain framework two paths are possible when considering exogenous inputs : either we keep the stationary transformation obtained for some constant value of the input (typically $u \equiv 0$) and hope the additional terms due to the presence of the input do not prevent convergence. Or we take a transformation taking into account (implicitly or explicitly) the input.

As far as we know, no result concerning this problem exists in the literature apart from [14], [15] which follows and extends [16]. The idea pursued in [14] belongs to the first path : the transformation is stationary and the input is seen as a disturbance which must be small enough. Although the construction is extended in a cunning fashion to a larger class of inputs, namely those which can be considered as output of a linear generator model with small external input, this approach remains theoretic and restrictive. On the other hand, in [15], the author rather tries to use a time-varying transformation but its injectivity is proved only under the so-called "finite-complexity" assumption, initially introduced in [16] for autonomous systems. Unfortunately, this property is very restrictive and hard to check. Besides, no indication about the dimension of the target form is given and the transformation cannot be computed online because it depends on the whole past trajectory of the output.

That is why, in this paper, we endeavor to give results of existence and injectivity of the transformation under more standard observability assumptions and keeping in mind the practical implementation of this method. Preliminary results presented in [17] showed that any system which can't blow up in finite backward time can be transformed through a time-varying transformation into a Hurwitz asymptotically stable

[1] We write "initial" insofar as what is nowadays usually called "Luenberger observer" differs from what is in [1].

form and the injectivity of this transformation is achieved under a differential observability condition. We complete here this result by showing that injectivity can actually be ensured under a weaker backward observability condition for "almost any" choice of the eigenvalues of the target Hurwitz form with sufficiently large dimension. We also show that it is possible to take a stationary transformation in the case of instantaneously uniformly observable systems whose order of differential observability equals the system's dimension.

B. Problem statement

Consider a general system of the form

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

where x is the state in $\mathbb{R}^{d_x}$, y the output in $\mathbb{R}^{d_y}$, f a continuously differentiable (C^1) function, h a continuous function, and $u : [0,+\infty) \to U \subset \mathbb{R}^{d_u}$ in a set $\mathcal{U} \subset \mathcal{L}^{\infty}_{loc}([0,+\infty))$ of considered inputs[2]. We denote $X(x,t;s;u)$ the value at time s of the (unique) solution to system (1) with input u, initialized at x at time t, and $Y(x,t;s;u)$ the corresponding output function at time s. We consider a subset $\mathcal{X}_0$ of $\mathbb{R}^{d_x}$ containing all the possible initial conditions for the system. We introduce the following assumption :

Assumption 1. *Solutions to System* (1) *initialized in* $\mathcal{X}_0$ *are well defined in positive time and belong to an open set* $\mathcal{X}$*. In other words, for all* u *in* $\mathcal{U}$*, for all* x_0 *in* $\mathcal{X}_0$ *and for all* s *in* $[0,+\infty)$*,* $X(x_0,0;s;u)$ *is well defined and is in* $\mathcal{X}$*.*

In this paper, we want to design an observer for system (1), via the Luenberger-like methodology developed in [1], [10], [11], [13]. We assume the inputs and outputs are known in a causal way, namely the observer can use only their past or current values, i.e., at time t, $u|_{[0,t]}$ and $y|_{[0,t]}$ only[3]. The idea is to transform system (1) into a Hurwitz form[4]

$$\dot{\xi} = A\,\xi + B\,y \tag{2}$$

with A Hurwitz in $\mathbb{R}^{d_\xi \times d_\xi}$, B a matrix in $\mathbb{R}^{d_\xi \times d_y}$, for some strictly positive integer d_ξ, i.e. for each u in $\mathcal{U}$, find a transformation[5] $T : \mathcal{X} \times [0,+\infty) \to \mathbb{R}^{d_\xi}$ such that for any x in $\mathcal{X}$ and any time t in $[0,+\infty)$,

$$\frac{\partial T}{\partial x}(x,t) f(x,u(t)) + \frac{\partial T}{\partial t}(x,t) = A\,T(x,t) + B\,h(x,u(t)) \ . \tag{3}$$

Indeed, implementing the dynamics (2) with any initial condition would then provide an asymptotically converging estimate of $T(X(x_0,0;s;u),s)$. If $T(\cdot,s)$ is besides injective (at least after a certain time), one could deduce an estimate for the system solution $X(x_0,0;s;u)$. More precisely, the following theorem is proved in [17] :

Theorem 1 ([17])**.** *Assume Assumption 1 is satisfied. Consider a strictly positive integer* d_ξ*, a Hurwitz matrix* A *in* $\mathbb{R}^{d_\xi \times d_\xi}$*, and a matrix* B *in* $\mathbb{R}^{d_\xi \times d_y}$*. Suppose that for any input* u *in* $\mathcal{U}$*, there exists a function* $T : \mathcal{X} \times [0,+\infty) \to \mathbb{R}^{d_\xi}$ *such that*

1) *T is a C^1 solution to PDE* (3) *on* $\mathcal{X} \times [0,+\infty)$ *;*
2) *there exists a time $\bar{t} \geq 0$ and a concave $\mathcal{K}^\infty$ function ρ_T such that for all (x_1,x_2) in $\mathcal{X}^2$ and all $t \geq \bar{t}$*

$$|x_1 - x_2| \leq \rho_T(|T(x_1,t) - T(x_2,t)|)$$

i.e. T becomes injective uniformly in time and in space after a certain time $\bar{t}$.

Then, there exists a function $T^* : \mathbb{R}^{d_\xi} \times [\bar{t},+\infty) \to \mathbb{R}^{d_x}$ *such that for any* x_0 *in* $\mathcal{X}_0$*, and any* ξ_0 *in* $\mathbb{R}^{d_\xi}$*, the (unique) solution* $(X(x_0,0;s;u),\Xi(\xi_0,0;s;u,y_{x_0}))$ *to*

$$\begin{array}{rclcl} \dot{x} & = & f(x,u) & , & y = h(x,u) \\ \dot{\xi} & = & A\,\xi + B\,y & , & \hat{x} = T^*(\xi,t) \end{array} \tag{4}$$

verifies

$$\lim_{s \to +\infty} |X(x_0,0;s;u) - \hat{X}(s)| = 0$$

with

$$\hat{X}(s) = T^*(\Xi(\xi_0,0;s;u,y_{x_0}),s) \ .$$

We conclude that it is sufficient to find a solution T to PDE (3) that becomes injective uniformly in time and in space at least after a certain time to obtain an observer for system (1).

In Section II, we show that the existence of the time-varying transformation T itself is achieved under mild assumptions, and that its injectivity can be ensured by observability assumptions, similar to those presented in [13] for autonomous systems. Then, in Section III, we show on practical examples how an explicit expression for such a transformation can be computed. Finally, in Section IV, we prove that, similarly to [5], [6] for a high gain design, in the case of a uniformly observable (see [18, Definition 1]) input-affine system whose drift system is strongly differentially observable of order d_x (see [18, Definition 2]), a Luenberger-type observer can be built with a stationary transformation $T : \mathcal{X} \to \mathbb{R}^{d_\xi}$.

Notations

1) Since h (resp Y) takes values in $\mathbb{R}^{d_y}$, we denote h_i (resp Y_i) its ith-component.
2) For some integer m which we choose later in Assumptions 3, any solution x to system (1) with some C^{m+1} input u, is such that $(x,u,\dot{u},...,u^{(m)})$ is solution to the extended system

$$\begin{array}{rcl} \dot{x} & = & f(x,\nu_0) \\ \dot{\nu}_0 & = & \nu_1 \\ & \vdots & \\ \dot{\nu}_m & = & u^{(m+1)} \end{array}$$

[2]Systems of the type (1) encompass time varying systems in the form

$$\dot{x} = f(x,t) \ , \ y = h(x,t),$$

simply by taking $\mathcal{U} = \{t \mapsto t\}$, $u(t) = t$ and $U = [0,+\infty)$. Following this route, systems in the form $\dot{x} = f(x,t,u)$, $y = h(x,t,u)$ could also be considered.

[3]Time 0 thus corresponds to the initial time of data recording.

[4]We could have considered a more general Hurwitz form $\dot{\xi} = A\,\xi + B(u,y)$ with B any nonlinear function, but taking $B(u,y) = y$ is sufficient to obtain satisfactory results.

[5]The function T implicitly depends on u in $\mathcal{U}$, so we should write T_u. But we drop this too heavy notation to ease the comprehension. What matters is that the target Hurwitz form (2), namely d_ξ, A and B, be the same for all u in $\mathcal{U}$ and that the dependence on u be causal.

with input $u^{(m+1)}$. We denote $\overline{u}_m = (u, \dot{u}, ..., u^{(m)})$, $\overline{\nu}_m = (\nu_0, ..., \nu_m)$, $\bar{f}$ the extended vector field

$$\bar{f}(x, \overline{\nu}_m, u^{(m+1)}) = \Big(f(x, \nu_0), \nu_1, \ \dots \ , \nu_m, u^{(m+1)}\Big)$$

and $\overline{h}$ the extended measurement function

$$\bar{h}_i(x, \overline{\nu}_m) = h_i(x, \nu_0) \ .$$

Note that while $\overline{\nu}_m$ is an element $\mathbb{R}^{d_u(m+1)}$, $\overline{u}_m$ is a function defined on $[0, +\infty)$ such that $\overline{u}_m(s) = (u(s), \dot{u}(s), ..., u^{(m)}(s))$ is in $\mathbb{R}^{d_u(m+1)}$ for all s in $[0, +\infty)$. We denote $\overline{U}_m$ a subset of $\mathbb{R}^{d_u(m+1)}$ such that for all u in $\mathcal{U}$, $\overline{u}_m([0, +\infty)) \subset \overline{U}_m$. For $1 \leq i \leq d_y$, the successive time derivatives of Y_i are related to the Lie derivatives of $\bar{h}_i$ along the vector fields $\bar{f}$, namely for $j \leq m$

$$\frac{\partial^j Y_i}{\partial s^j}(x, t; s; u) = L_{\bar{f}}^j \bar{h}_i(X(x, t; s; u), \overline{u}_m(s)) \ .$$

II. Time-varying transformation

We make the following assumption :

Assumption 2. *Solutions to System* (1) *initiated from $\mathcal{X}$ do not blow-up in finite backward time, namely for any u in $\mathcal{U}$, any (x, t) in $\mathcal{X} \times [0, +\infty)$ and any s in $[0, t]$, $X(x, t; s; u)$ is defined.*

The existence of a C^1 time-varying solution to PDE (3) is achieved thanks to the following lemma :

Lemma 1 ([17])**.** *Consider a strictly positive number d_ξ, a Hurwitz matrix A in $\mathbb{R}^{d_\xi \times d_\xi}$, a matrix B in $\mathbb{R}^{d_\xi \times d_y}$, and an input u in $\mathcal{U}$. Under Assumption 2, the function T^0 defined on $\mathcal{X} \times [0, +\infty)$ by*[6]

$$T^0(x, t) = \int_0^t e^{A(t-s)} B\, Y(x, t; s; u)\, ds \tag{5}$$

is a C^1 solution to PDE (3) *on $\mathcal{X} \times [0, +\infty)$.*

When Assumption 2 does not hold, it may still be possible to construct a function T^0 solution to PDE (3) on $\mathcal{X} \times [0, +\infty)$. This is the case if there exists a subset $\mathcal{X}'$ of $\mathbb{R}^{d_x}$ such that $\mathtt{cl}(\mathcal{X}) \subset \mathcal{X}'$ from which no solution blows up at infinity in backward time before leaving the set[7] $\mathcal{X}'$. Indeed, any modified dynamics

$$\dot{x} = \chi(x) f(x, u) \ , \tag{6}$$

with a C^∞ function $\chi : \mathbb{R}^{d_x} \to \mathbb{R}$ satisfying

$$\chi(x) = \left\{ \begin{array}{lcl} 1 & , & \text{if } x \in \mathtt{cl}(\mathcal{X}) \\ 0 & , & \text{if } x \notin \mathcal{X}' \end{array} \right.$$

is then backward complete and satisfies Assumption 2, and the PDE associated to system (6) is the same as PDE (3) on $\mathcal{X} \times [0, +\infty)$. In particular, we deduce :

Corollary 1. *Assume $\mathcal{X}$ is bounded. Consider a strictly positive number d_ξ, a Hurwitz matrix A in $\mathbb{R}^{d_\xi \times d_\xi}$, a matrix B in $\mathbb{R}^{d_\xi \times d_y}$, and an input u in $\mathcal{U}$. There exists a C^1 function T^0 solution to PDE* (3) *on $\mathcal{X} \times [0, +\infty)$.*

In the extreme scenario where Assumption 2 does not hold even for system (6), it could also be relaxed by considering an output dependent time rescaling as in [13, Section 2.6] when the system has some (backward) unbounded observability property.

Note that extending directly what is done in [16], [13] would rather lead us to the solution

$$T^\infty(x, t) = \int_{-\infty}^t e^{A(t-s)} B\, Y(x, t; s; u)\, ds \ .$$

The drawback is that some assumptions about the growth of Y have to be made to ensure its continuity, unless Y is bounded backward in time. As for the C^1 property, and even if the solutions are bounded backward in time, it is achieved only if the eigenvalues of A are sufficiently negative. In fact, it is not absolutely needed that the solution be C^1, one could look for continuous solutions to

$$\lim_{\delta \to 0} \frac{T(X(x, t; t+\delta; u), t) - T(x, t)}{\delta} = AT(x, t) + Bh(x, u(t))$$

instead of PDE (3). The major disadvantage of this solution is rather that T^∞ is not easily computable since it depends on the values of u on $(-\infty, t]$. Nevertheless, it may still be useful. For example, that is the solution chosen in [19] for the specific application of a permanent synchronous motor, where it is proved to be injective.

Unlike T^∞, T^0 depends only on the values of the input u on $[0, t]$. Therefore, it is theoretically computable online. However, for each couple (x, t), one would need to integrate backwards the dynamics (1) until time 0, which is quite heavy. If the input u is known in advance (for instance $u(t) = t$) it can also be computed offline. We will see in Section III on practical examples how we can find a solution to PDE (3) in practice, without relying on the expression T^0.

We finally conclude that a C^1 time-varying transformation into a Hurwitz form always exists under the mild Assumption 2, but the core of the problem is to ensure its injectivity.

A. Injectivity with strong differential observability

Assumption 3. *There exists a subset $\mathcal{S}$ of $\mathbb{R}^{d_x}$ such that :*

1) *For any u in $\mathcal{U}$, any x in $\mathcal{X}$ and any time t in $[0, +\infty)$, $X(x, t; s; u)$ is defined in $\mathcal{S}$ for all s in $[0, t]$.*
2) *The quantity*

$$M_f = \sup_{\substack{x \in \mathcal{S} \\ \nu_0 \in U}} \left| \frac{\partial f}{\partial x}(x, \nu_0) \right|$$

is finite.

3) *There exist d_y integers $(m_1, \dots, m_{d_y})$ such that the functions*

$$\begin{aligned} H_i(x, \overline{\nu}_m) = \Big(\bar{h}_i(x, \overline{\nu}_m), L_{\bar{f}} \bar{h}_i(x, \overline{\nu}_m), \\ \dots, L_{\bar{f}}^{m_i - 1} \bar{h}_i(x, \overline{\nu}_m)\Big) \end{aligned} \tag{7}$$

is well defined on $\mathcal{S} \times \mathbb{R}^{d_u(m+1)}$ with $m = \max_i m_i$ and $1 \leq i \leq d_y$ verify :

[6] Following what has been done for the function T, to simplify the presentation, the dependance on u of the function T^0 has been dropped.

[7] This property is named completeness within $\mathcal{X}'$ in [13].

- *for any $\overline{\nu}_m$ in $\overline{U}_m$, $H_i(\cdot,\overline{\nu}_m))$ is Lipschitz on $\mathcal{S}$.*
- *there exists L_H such that the function*

$$H(x,\overline{\nu}_m)=\Big(H_1(x,\overline{\nu}_m),\,\dots\,,H_i(x,\overline{\nu}_m),\\ \dots\,,H_{d_y}(x,\overline{\nu}_m)\Big) \quad (8)$$

verifies for any (x_1,x_2) in $\mathcal{S}^2$ and any $\overline{\nu}_m$ in $\overline{U}_m$

$$|x_1-x_2|\leq L_H|H(x_1,\overline{\nu}_m)-H(x_2,\overline{\nu}_m)|$$

namely H is Lipschitz-injective on $\mathcal{S}$, uniformly with respect to $\overline{\nu}_m$ in $\overline{U}_m$.

4) *For all $1\leq i\leq d_y$, there exists L_i such that for all (x_1,x_2) in $\mathcal{S}^2$ and for all $\overline{\nu}_m$ in $\overline{U}_m$,*

$$|L_{\bar{f}}^{m_i}\bar{h}_i(x_1,\overline{\nu}_m)-L_{\bar{f}}^{m_i}\bar{h}_i(x_2,\overline{\nu}_m)|\leq L_i|x_1-x_2|$$

namely $L_{\bar{f}}^{m_i}\bar{h}_i(\cdot,\overline{\nu}_m)$ is Lipschitz on $\mathcal{S}$, uniformly with respect to $\overline{\nu}_m$ in $\overline{U}_m$.

Theorem 2 ([17]). *Suppose Assumption 3 is satisfied. For all integer $1\leq i\leq d_y$ and m_i defined in Assumption 3, consider a Hurwitz matrix A_i in $\mathbb{R}^{m_i\times m_i}$ and a vector B_i in $\mathbb{R}^{m_i}$ such that the pairs (A_i,B_i) is controllable. Denote $m=\max_i m_i$. There exists a strictly positive real number $\overline{k}$ such that for all $k\geq\overline{k}$ and for all C^{m+1} input u in $\mathcal{U}$, there exists $\bar{t}_{k,u}$ such that any C^1 solution T to PDE* (3) *on $\mathcal{X}\times[0,+\infty)$ with*

- *$d_\xi=\sum_{i=1}^{d_y}m_i$*
- *A in $\mathbb{R}^{d_\xi\times d_\xi}$ and B in $\mathbb{R}^{d_\xi\times d_y}$ defined by*

$$A=\begin{pmatrix} kA_1 & & & & \\ & \ddots & & & \\ & & kA_i & & \\ & & & \ddots & \\ & & & & kA_{d_y}\end{pmatrix}\ ,\ B=\begin{pmatrix} B_1 & & & & \\ & \ddots & & & \\ & & B_i & & \\ & & & \ddots & \\ & & & & B_{d_y}\end{pmatrix}$$

- *$T(\cdot,0)$ Lipschitz on $\mathcal{X}$*

is such that $T(\cdot,t)$ is injective on $\mathcal{X}$ for all $t\geq\bar{t}_{k,u}$, uniformly in time and in space. More precisely, there exists a constant L_k such that for any (x_1,x_2) in $\mathcal{X}^2$, any u in $\mathcal{U}$ and any time $t\geq\bar{t}_{k,u}$

$$|x_1-x_2|\leq L_k\,|T(x_1,t)-T(x_2,t)|\ .$$

Besides, for any $t\geq\bar{t}_{k,u}$, $T(\cdot,t)$ is an injective immersion on $\mathcal{X}$.

Note that the additional assumption "$T(\cdot,0)$ Lipschitz on $\mathcal{X}$" is not very restrictive because the solution T can usually be chosen arbitrarily at initial time 0 (see examples in Section III). In particular, the elementary solution T^0 found in Lemma 1 is the zero constant function at time 0 and thus clearly verifies this assumption.

Applying successively Lemma 1, Theorem 2 and Theorem 1, we conclude that under Assumption 3, it is possible to write an observer for system (1) by choosing any (A_i,B_i) controllable and k sufficiently large.

Remark 1. It is important to note that $\overline{k}$ does not dependent on u, thanks to the fact that L_H, M_f and L_i given by Assumptions 3.2, 3.3, 3.4 are the same for all $\overline{\nu}_m$ in $\overline{U}_m$. However, the time $\bar{t}_{k,u}$ after which the solution becomes injective a priori depends on k and u. This is not a problem in practice since we only want to be sure that for k sufficiently large, any solution will become injective after a certain time. If we want this time $\bar{t}_{k,u}$ to be uniform in u, the Lipschitz constants of $H_i(\cdot,\overline{u}_m(0)))$ and of $T(\cdot,0)$ must be the same for all u in $\mathcal{U}$.

Remark 2. If we choose $m=\max_i m_i$ sufficiently large distinct strictly positive real numbers λ_j, and take $A_i=-\mathsf{diag}(\lambda_1,\dots,\lambda_{m_i})$ and $B_i=(1,\dots,1)^\top$, then, the PDEs to solve are simply

$$\frac{\partial T_{\lambda,i}}{\partial x}(x,t)f(x,u(t))+\frac{\partial T_{\lambda,i}}{\partial t}(x,t)=-\lambda\,T_{\lambda,i}(x,t)+h_i(x,u(t)) \quad (9)$$

for each $1\leq i\leq d_y$ and λ in $\{\lambda_1,\dots,\lambda_{m_i}\}$. Then, one take

$$T(x,t)=\Big(T_{\lambda_1,1},\dots,T_{\lambda_{m_1},1},\dots,T_{\lambda_1,d_y},\dots,T_{\lambda_{m_{d_y}},d_y}\Big)\ .$$

Remark 3. Under Assumption 3.3, 3.4, the system could also be transformed via H into a Lipschitz phase-variable form of dimension $d_y\times\max_i m_i\geq\sum_{i=1}^{d_y}m_i$ ([20], [8]) and a high gain observer could be used. If we wanted to use only m_i derivatives for each output and obtain an observer of dimension $\sum_{i=1}^{d_y}m_i$, each $L_{\bar{f}}^{m_i}\bar{h}$ would have to satisfy an additional triangularity assumption. But in any case, the crucial difference with the Luenberger observer presented in this chapter is that unlike H, the transformation used here does not require the computation of the derivatives of the input (see examples in Section III).

In order to check Assumption 3 more easily in practice, we prove the following result :

Lemma 2 ([17]). *Assume that $\mathcal{S}$ is compact and there exist d_y integers $(m_1,\dots,m_{d_y})$ such that $\overline{U}_m$ with $m=\max_i m_i$ is compact and for any $\overline{\nu}_m$ in $\overline{U}_m$, $H(\cdot,\overline{\nu}_m)$ in* (8) *is a well defined injective immersion*[8] *on $\mathcal{S}$. Then, Assumptions 3.2, 3.3, 3.4 are satisfied.*

In other words, since the additional assumption "$T(\cdot,0)$ Lipschitz on $\mathcal{S}$" made in Theorem 2 is automatically verified when $\mathcal{S}$ is compact (T is C^1), the result of Theorem 2 holds under the only assumptions of Lemma 2 if $\mathcal{S}$ satisfies Assumption 3.1.

B. Injectivity with backward distinguishability

In the previous section, we have shown that finding an injective transformation into an Hurwitz form is possible under a strong differential observability property, namely that the function $H(\cdot,\overline{\nu}_m)$ made of each output and a certain number of its derivatives is an injective immersion. We investigate in this section if injectivity is still ensured when we have only a weak differential observability (i.e. $H(\cdot,\overline{\nu}_m)$ is injective but not an immersion) or even only backward-distinguishability as in [13, Theorem 3] for autonomous systems.

[8] $H(\cdot,\overline{\nu}_m)$ is injective on $\mathcal{S}$ and $\frac{\partial H}{\partial x}(x,\overline{\nu}_m)$ is full-rank for any x in $\mathcal{S}$

Theorem 3. *Take u in $\mathcal{U}$. Assume that for this input, system (1) is backward-distinguishable on $\mathcal{X}$ in time $\bar{t}_u$, i.e. for any $t \geq \bar{t}_u$ and any (x_a, x_b) in $\mathcal{X}^2$,*

$$Y(x_a, t; s; u) = Y(x_b, t; s; u)\, \forall s \in [t - \bar{t}_u, t] \implies x_a = x_b \ .$$

Assume also that Assumption 2 holds. Then, there exists a set $\mathcal{R}$ of zero-Lebesgue measure in $\mathbb{C}^{d_x+1}$ such that for any $(\lambda_1, \ldots, \lambda_{d_x+1})$ in $\Omega^{d_x+1} \setminus \mathcal{R}$ with $\Omega = \{\lambda \in \mathbb{C}\,,\, \Re(\lambda) < 0\}$, and any $t \geq \bar{t}_u$, the function T^0 defined in (5) with

- $d_\xi = d_y \times (d_x + 1)$
- *A in $\mathbb{R}^{d_\xi \times d_\xi}$ and B in $\mathbb{R}^{d_\xi \times d_y}$ defined by*

$$A = \begin{pmatrix} \tilde{A} & & & & \\ & \ddots & & & \\ & & \tilde{A} & & \\ & & & \ddots & \\ & & & & \tilde{A} \end{pmatrix} \quad , \quad B = \begin{pmatrix} \tilde{B} & & & & \\ & \ddots & & & \\ & & \tilde{B} & & \\ & & & \ddots & \\ & & & & \tilde{B} \end{pmatrix}$$

and

$$\tilde{A} = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_{d_x+1} \end{pmatrix} \quad , \quad \tilde{B} = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} .$$

is such that $T^0(\cdot, t)$ is injective on $\mathcal{X}$ for $t > \bar{t}_u$.

Note that the assumption of backward-distinguishability in finite time is in particular verified when the system is instantaneously backward-distinguishable, and a fortiori when the map made of the output and its derivatives up to a certain order is injective, namely the system is weakly differentially observable.

Of course, if T^0 has been built with system (6) instead of (1) to satisfy Assumption 2, the assumption of backward distinguishability needed here should hold for system (6), namely the outputs should be distinguishable in backward time before the solutions leave $\mathcal{X}$.

Proof. Let us define for λ in $\mathbb{C}$, the function $T^0_\lambda : \mathcal{X} \times \mathbb{R}^+ \to \mathbb{C}^{d_y}$

$$T^0_\lambda(x, t) = \int_0^t e^{-\lambda(t-s)} Y(x, t; s; u)\, ds \ . \tag{10}$$

Given the structure of A and B, and with a permutations of the components,

$$T^0(x, t) = \left(T^0_{\lambda_1}(x, t), \ldots, T^0_{\lambda_{d_x+1}}(x, t)\right) \ .$$

We need to prove that T^0 is injective for almost all $(\lambda_1, \ldots, \lambda_{d_x+1})$ in Ω^{d_x+1} (in the sense of the Lebesgue measure). For that, we define the function

$$\Delta T(x_a, x_b, t, \lambda) = T^0_\lambda(x_a, t) - T^0_\lambda(x_b, t)$$

on $\Upsilon \times \Omega$ with

$$\Upsilon = \{(x_a, x_b, t) \in \mathcal{X}^2 \times (\bar{t}_u, +\infty) \ : \ x_a \neq x_b\} \ .$$

We are going to use the following lemma whose proof[9] can be found in [13]:

[9]More precisely, the result proved in [13] is for Υ open set of $\mathbb{R}^{2d_x}$ instead of $\mathbb{R}^{2d_x+1}$. But the proof turns out to be still valid with $\mathbb{R}^{2d_x+1}$ because the only constraint is that the dimension of Υ be strictly less than $2(d_\xi + 1)$.

Lemma 3 (Coron's lemma)**.** *Let Ω and Υ be open sets of $\mathbb{C}$ and $\mathbb{R}^{2d_x+1}$ respectively. Let $\Delta T : \Upsilon \times \Omega \to \mathbb{C}^{d_y}$ be a function which is holomorphic in λ for all $\underline{x}$ in Υ and C^1 in $\underline{x}$ for all λ in Ω. If for any $(\underline{x}, \lambda)$ in $\Upsilon \times \Omega$ such that $\Delta T(\underline{x}, \lambda) = 0$, there exists i in $\{1, \ldots, d_y\}$ and $k > 0$ such that $\dfrac{\partial^k \Delta T_i}{\partial \lambda^k}(\underline{x}, \lambda) \neq 0$, then the set*

$$\mathcal{R} = \bigcup_{\underline{x} \in \Upsilon} \Big\{(\lambda_1, \ldots, \lambda_{d_x+1}) \in \Omega^{d_x+1} \ : \ \Delta T(\underline{x}, \lambda_1) = \ldots = \Delta T(\underline{x}, \lambda_{d_x+1}) = 0\Big\}$$

has zero Lebesgue measure in $\mathbb{C}^{d_x+1}$.

In our case, ΔT is clearly holomorphic in λ and C^1 in $\underline{x}$. Since for every $\underline{x}$ in Υ, $\lambda \mapsto \Delta T(\underline{x}, \lambda)$ is holomorphic on the connex set $\mathbb{C}$, its zeros are isolated and admit a finite multiplicity, unless it is identically zero on $\mathbb{C}$. Let us prove that $\lambda \mapsto \Delta T(\underline{x}, \lambda)$ can't be identically zero on $\mathbb{C}$. If it was the case, we would have in particular for any ω in $\mathbb{R}$

$$\int_{-\infty}^{+\infty} e^{-i\omega\tau} g(\tau)\, d\tau = 0$$

with g the function

$$g(\tau) = \begin{cases} Y(x_a, t; t - \tau; u) - Y(x_b, t; t - \tau; u) & \text{, if } \tau \in [0, t] \\ 0 & \text{, otherwise} \end{cases}$$

which is in $\mathcal{L}^2$. Thus, the Fourier transform of g would be identically zero and we deduce that necessarily we would have

$$Y(x_a, t; t - \tau; u) - Y(x_b, t; t - \tau; u) = 0$$

for almost all τ in $[0, t]$ and thus for all τ in $[0, t]$ by continuity. Since $t \geq \bar{t}_u$, it would follow from the backward-distinguishability that $x_a = x_b$ but this is impossible because (x_a, x_b, t) is in Υ. We conclude that $\lambda \mapsto \Delta T(\underline{x}, \lambda)$ is not identically zero on $\mathbb{C}$ and the assumptions of the lemma are satisfied. Thus, $\mathcal{R}$ has zero measure and for all $(\lambda_1, \ldots, \lambda_{d_x+1})$ in $\mathbb{C}^{d_x+1} \setminus \mathcal{R}$, T^0 is injective on $\mathcal{X}$, by definition of $\mathcal{R}$. $\square$

Remark 4. The function T proposed by Theorem 3 takes complex values. To remain in the real frame, one should consider the transformation made of its real and imaginary parts, and instead of implementing for each i in $\{1, \ldots, d_y\}$ and each lambda

$$\dot{\hat{\xi}}_{\lambda,i} = -\lambda \hat{\xi}_\lambda + y_i$$

in $\mathbb{C}$, one should implement

$$\dot{\hat{\xi}}_{\lambda,i} = \begin{pmatrix} -\Re(\lambda) & -\Im(\lambda) \\ \Im(\lambda) & -\Re(\lambda) \end{pmatrix} \hat{\xi}_{\lambda,i} + \begin{pmatrix} y_i \\ 0 \end{pmatrix}$$

in $\mathbb{R}$. Thus, the dimension of the observer is $2 \times d_y \times (d_x + 1)$ in terms of real variables.

Remark 5. It should be noted that Theorem 3 gives for each u in $\mathcal{U}$ a set $\mathcal{R}_u$ of zero measure in which not to choose the λ_i, but unfortunately, there is no guarantee that $\bigcup\limits_{u \in \mathcal{U}} \mathcal{R}_u$ is also of zero-Lebesgue measure.

Remark 6. Unlike Theorem 2 which proved the injectivity of any solution T to PDE (3), Theorem 3 proves only the injectivity of T^0. Note though that as shown at the beginning of the proof of Theorem 2 (see [17]) by the "variation of constants" formula, any solution T verifies

$$T(x,t) = e^{At}\, T(X(x,t;0;u),0) + T^0(x,t)$$

with A Hurwitz, and thus tends to the injective function T^0. We can thus expect T to become injective after a certain time, under some appropriate uniformity assumptions. In particular, this is the case if for all x in $\mathcal{X}$ and all $t \geq 0$, $X(x,t;0;u)$ is in a set $\mathcal{S}$ for which there exists a class $\mathcal{K}$ function ρ such that for all (x_a, x_b) in $\mathcal{S}\times\mathcal{S}$,

$$\begin{aligned} |T(x_a,0) - T(x_b,0)| &\leq \rho(|x_a - x_b|) \ , \\ \left|T^0(x_a,0) - T^0(x_b,0)\right| &\leq \rho(|x_a - x_b|) \ , \end{aligned}$$

and a positive real number ℓ such that for all $t \geq 0$,

$$\left|T^0(x_a,t) - T^0(x_b,t)\right| \geq \ell\, \rho(|x_a - x_b|).$$

In that case,

$$|T(x_a,t) - T(x_b,t)|| \geq \left(\ell - 2\left|e^{At}\right|\right) \rho(|x_a - x_b|).$$

Hence, injectivity on $\mathcal{X}$ is obtained after a certain time. The first part is about uniform (in space) continuity, which is satisfied as soon as $\mathcal{S}$ is bounded, and the second part is about injectivity (in space and in time). In fact, a way of ensuring the injectivity is to take, if possible, a solution T with the boundary condition

$$T(x,0) = 0 \quad \forall x \in \mathcal{S} \ , \tag{11}$$

because in that case, necessarily, $T = T^0$.

It is also interesting to remark that in the case where T is initialized along (11), and the observer state is initialized at $\xi_0 = 0$, we have for all time s

$$\Xi(\xi_0, 0; s; u, y_{x_0}) = T(X(x_0,0;s;u), s)$$

i.e. finite-time convergence is achieved as soon as $T(\cdot,t)$ becomes injective.

We conclude from this section that as soon as no blow-up in finite-time is possible, there always exists a time-varying solution to PDE (3) which is injective under standard observability assumptions. It follows that the only remaining problem to address is the computation of such a solution without relying on the expression (5). This is done in the following section through practical examples.

III. Examples

A. Linear dynamics with polynomial output

Consider a system of the form[10]

$$\dot{x} = A(u,y)\, x + B(u,y) \quad , \quad y = C(u) P_d(x) \tag{12}$$

with $P_d : \mathbb{R}^{d_x} \to \mathbb{R}^{k_d}$ a vector containing the k_d possible monomials of x with degree inferior to d, $A : \mathbb{R}^{d_u} \times \mathbb{R}^{d_y} \to \mathbb{R}^{d_x \times d_x}$, $B : \mathbb{R}^{d_u} \times \mathbb{R}^{d_y} \to \mathbb{R}^{d_x}$ and $C : \mathbb{R}^{d_u} \to \mathbb{R}^{d_y \times k_d}$ matrices depending on u. For any i in $\{1, \cdots, d_y\}$, a transformation $T_{\lambda,i}$ of the form

$$T_{\lambda,i}(x,t) = M_{\lambda,i}(t) P_d(x)$$

with $M_{\lambda,i} : \mathbb{R} \to \mathbb{R}^{1\times k_d}$, verifies

$$\begin{aligned} &\frac{\partial T_{\lambda,i}}{\partial x}(x,t) f(x,u) + \frac{\partial T_{\lambda,i}}{\partial t}(x,t) \\ &\quad = M_{\lambda,i}(t) \frac{\partial P_d}{\partial x}(x) \Big(A(u,y)x + B(u,y) \Big) + \dot{M}_{\lambda,i}(t) P_d(x) \ . \end{aligned}$$

But there exists a matrix of coefficients $D : \mathbb{R}^{d_u} \times \mathbb{R}^{d_y} \to \mathbb{R}^{k_d}$ such that

$$\frac{\partial P_d}{\partial x}(x) \Big(A(u,y)x + B(u,y) \Big) = D(u,y) P_d(x)$$

so that we get

$$\begin{aligned} &\frac{\partial T_{\lambda,i}}{\partial x}(x,t) f(x,u) + \frac{\partial T_{\lambda,i}}{\partial t}(x,t) \\ &\qquad\qquad = \Big(M_{\lambda,i}(t) D(u,y) + \dot{M}_{\lambda,i}(t) \Big) P_d(x) \ . \end{aligned}$$

It follows that by choosing the coefficients $M_{\lambda,i}$ as solutions of the filters

$$\dot{M}_{\lambda,i}(t) + \lambda\, M_{\lambda,i}(t) = -M_{\lambda,i}(t) D(u(t),y) + C_i(u(t)) \ ,$$

$T_{\lambda,i}$ is solution to the PDE (9) with $\mathcal{X} = \mathbb{R}^{d_x}$.

A practical example of this kind of systems is a Permanent Magnet Synchronous Motor (PMSM), which can be modeled by

$$\dot{x} = u(t) - Ri(t) \quad , \quad y = |x - Li(t)|^2 - \Phi^2 = 0 \tag{13}$$

where x is in $\mathbb{R}^2$, the voltages u and currents i are time varying exogenous signals taking value in $U = \mathbb{R}^2$, the resistance R, impedance L and flux Φ are known scalar parameters and the measurement y is constantly zero. Applying the method presented above and removing the unnecessary terms, we find that we can choose T_λ of the form ($d_y = 1$)

$$T_\lambda(x,t) = |x|^2 + a_\lambda(t)^\top x + b_\lambda(t)$$

with the dynamics of a_λ and b_λ given by

$$\begin{aligned} \dot{a}_\lambda &= -\lambda\, a_\lambda - 2(u(t) - Ri(t)) + 2Li(t) \\ \dot{b}_\lambda &= -\lambda\, b_\lambda - a_\lambda^\top (u(t) - Ri(t)) + L^2 |i(t)|^2 - \Phi^2 \end{aligned} \tag{14}$$

Once this solution has been found, an observability analysis must be carried out to know the number of eigenvalues λ which are necessary to ensure the injectivity of the transformation. This is developed in [17].

Note that for this particular system, a classical gradient observer of smaller dimension exists ([21], [22]). The Luenberger observer that we would obtain here offers the advantage of depending only on filtered versions of u and i, which can be useful in presence of significant noise. On the other hand, no high gain design would have been possible for this system without computing the derivatives of i, which is not desirable in practice.

[10]This is an abuse of notation to highlight the fact that A and B are functions of known signals which can thus be used in the observer. In truth, the dynamics are given by $\dot{x} = A(u, C(u)P_d(x))\, x + B(u, C(u)P_d(x))$.

B. A time-varying transformation for an autonomous system

It was observed in [23, Section 8.4] that it is sometimes useful to allow the transformation to be time-varying even for an autonomous system. Only results concerning stationary transformations were available at the time, so that the framework of dynamic extensions had to be used. This is no longer necessary thanks to Theorems 2 and 3. Indeed, consider for instance the system

$$\left\{\begin{array}{rcl}\dot{x}_1 & = & x_2^3\\ \dot{x}_2 & = & -x_1\end{array}\right. \quad , \quad y = x_1 \tag{15}$$

which admits bounded trajectories (the quantity $x_1^2 + x_2^4$ is constant). This system is weakly differentially observable of order 2 on $\mathbb{R}^2$ since $x \mapsto \mathbf{H}_2(x) = (x_1, x_2^3)$ is injective on $\mathbb{R}^2$. It is thus a fortiori instantaneously backward-distinguishable and [13, Theorem 3] holds. Applying Luenberger's methodology to this system would thus bring us to look for a stationary transformation T_λ into

$$\dot{\xi}_\lambda = -\lambda\,\xi_\lambda + x_1\ , \tag{16}$$

for which a possible solution is

$$T_\lambda(x) = \int_{-\infty}^{0} e^{\lambda\,\tau} Y(x;\tau) d\tau\ .$$

Although the injectivity of $T = (T_{\lambda_1}, T_{\lambda_2}, T_{\lambda_3})$ is satisfied for a generic choice of $(\lambda_1, \lambda_2, \lambda_3)$ in $\{\lambda \in \mathbb{C}\ :\ \Re(\lambda) > 0\}^3$ according to [13, Theorem 3], it is difficult to compute T numerically and as far as we are concerned, we are not able to find an explicit expression.

Instead, it may be easier to look for a time-varying transformation and apply either Theorem 2 or 3. According to Lemma 1, such a transformation exists whatever the chosen set $\mathcal{X}$ of interest and given the structure of the dynamics, one can try to look for it in the form

$$T_\lambda(x,t) = a_\lambda(t)x_2^3 + b_\lambda(t)x_2^2 + c_\lambda(t)x_2 + d_\lambda(t)x_1 + e_\lambda(t)\ . \tag{17}$$

It verifies the dynamics (16) if for instance

$$\begin{array}{rcl}\dot{a}_\lambda(t) & = & -\lambda\,a_\lambda(t) + d_\lambda(t)\\ \dot{b}_\lambda(t) & = & -\lambda\,b_\lambda(t) + 3a_\lambda(t)y\\ \dot{c}_\lambda(t) & = & -\lambda\,c_\lambda(t) + 2b_\lambda(t)y\\ \dot{d}_\lambda(t) & = & -\lambda\,d_\lambda(t) + 1\\ \dot{e}_\lambda(t) & = & -\lambda\,e_\lambda(t) + c_\lambda(t)y\end{array}$$

Using Remark 6 and applying Theorem 3, we know that, by initializing the filters a_λ, b_λ, c_λ, d_λ and e_λ at 0 at time 0, $x \mapsto (T_{\lambda_1}(x,t), T_{\lambda_2}(x,t), T_{\lambda_3}(x,t))$ is injective on $\mathbb{R}^2$ for $t > 0$ and for a generic choice of $(\lambda_1, \lambda_2, \lambda_3)$ in $\{\lambda \in \mathbb{C}\ :\ \Re(\lambda) > 0\}^3$.

To reduce the dimension of the filters, we can take $d_\lambda(t) = \frac{1}{\lambda}$ and $a_\lambda(t) = \frac{1}{\lambda^2}$. In that case Theorem 3 cannot be properly applied because T_λ is not T_λ^0. However, we have found at least in simulations that injectivity is preserved after a certain time as shown in Figure 1.

Note that since the system is strongly differentially observable of order 4 on $\mathcal{S} = \{(x_1, x_2) \in \mathbb{R}^2\ :\ x_1^2 + x_2^2 \neq 0\}$, i.e. $\mathbf{H}_4$ is an injective immersion on $\mathcal{S}$, Theorem 2 in combination with Lemma 2 says that, for any positive real number $L > 1$, by choosing 4 sufficiently large real strictly positive numbers λ_i, and for any initial conditions for the filters, $x \mapsto (T_{\lambda_1}(x,t), T_{\lambda_2}(x,t), T_{\lambda_3}(x,t), T_{\lambda_4}(x,t))$ becomes injective on $\mathcal{X} = \left\{(x_1, x_2) \in \mathbb{R}^2\ :\ \frac{1}{L} < x_1^2 + x_2^4 < L\right\}$ after some time.

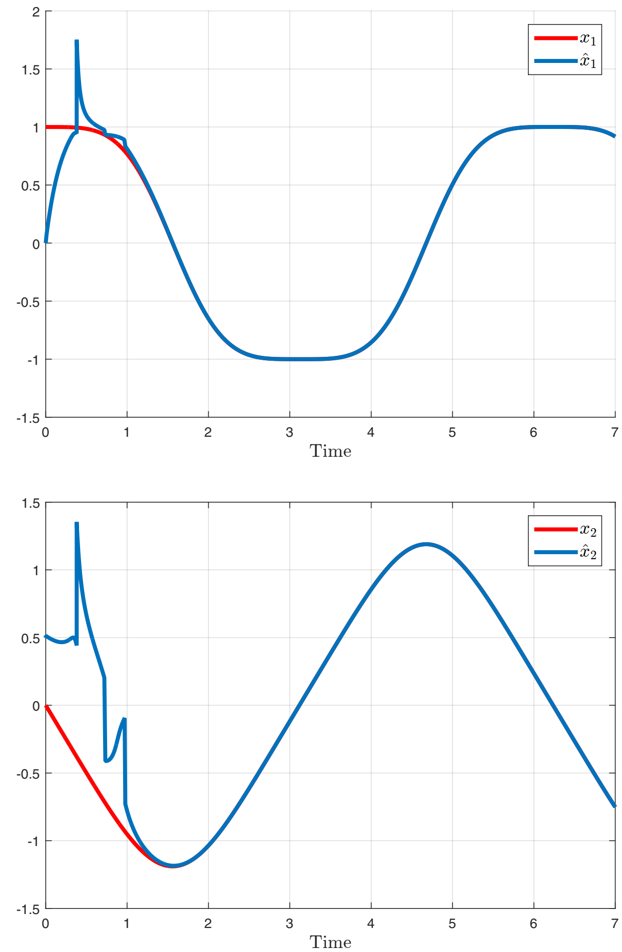

Fig. 1. Nonlinear Luenberger observer for system (15) : dynamics (16) and transformations (17) (with $d_\lambda(t) = \frac{1}{\lambda}$ and $a_\lambda(t) = \frac{1}{\lambda^2}$) for $\lambda_1 = 5$, $\lambda_2 = 6$, $\lambda_3 = 7$. The transformation is inverted by first linearly combining the $T_{\lambda_i} - \xi_{\lambda_i}$ to make x_1 disappear (because the d_{λ_i} are all nonzero), and then searching numerically the common roots of the obtained two polynomials of order 3 in x_2.

Remark 7. In those examples, the time dependence of T comes through filters a_λ, ..., e_λ, that take into account the input and output signals. Therefore, T could also be seen as a stationary transformation $T(x, a_\lambda, \dots, e_\lambda)$ if the filters' states were added to the system and observer states.

IV. Stationary transformation

We have just seen that a time-varying transformation could be used for an autonomous system. We investigate here the converse, i.e. if a stationary transformation can be used for time-varying systems. Consider a control-affine single-output system

$$\dot{x} = f(x) + g(x)u \quad , \quad y = h(x) \in \mathbb{R} \ . \tag{18}$$

We will use the following two notions of observability :

Definition 1 (Differential observability of the drift system)**.** *The drift system of system (18) is said weakly differentially observable of order m on an open subset $\mathcal{S}$ of $\mathbb{R}^{d_x}$ if the function*

$$\mathbf{H}_m = (h(x), L_f h(x), \dots, L_f^{m-1} h(x)$$

is injective on $\mathcal{S}$. If it is also an immersion, we say strongly differentially observable of order m.

Definition 2 (Instantaneous uniform observability)**.** *System (18) is instantaneously uniformly observable on an open subset $\mathcal{S}$ of $\mathbb{R}^{d_x}$ if, for any pair (x_a, x_b) in $\mathcal{S}^2$ with $x_a \neq x_b$, any strictly positive number $\bar{t}$, and any function u defined on $[0, \bar{t})$, there exists a time $t < \bar{t}$ such that $h(X_u(x_a, t)) \neq h(X_u(x_b, t))$ and $(X_u(x_a, s), X_u(x_b, s)) \in \mathcal{S}^2$ for all $s \leq t$.*

In the high gain framework, we know from [5], [6] that when system (18) is uniformly instantaneously observable and its drift dynamics are differentially observable of order d_x, it is possible to keep the stationary transformation associated to the drift autonomous system, because the additional terms resulting from the presence of inputs are triangular and do not prevent the convergence of the observer. It turns out that, inspired from [13, Theorem 5], an equivalent result exists in the Luenberger framework.

Theorem 4. *Let $\lambda_1, \dots, \lambda_{d_x}$ be any distinct strictly positive real numbers, A the Hurwitz matrix $\mathsf{diag}(-\lambda_1, \dots, -\lambda_{d_x})$ in $\mathbb{R}^{d_x \times d_x}$, B the vector $(1, ..., 1)^\top$ in $\mathbb{R}^{d_x}$ and $\mathcal{S}$ an open subset of $\mathbb{R}^{d_x}$ containing $\mathcal{X}_0$. Assume that system (18) is uniformly instantaneously observable on $\mathcal{S}$ and its drift system is strongly differentially observable of order d_x on $\mathcal{S}$. Then, for any positive real number $\bar{u}$, any bounded open subset $\mathcal{X}$ of $\mathbb{R}^{d_x}$ such that*

- $\mathtt{cl}(\mathcal{X}) \subset \mathcal{S}$,
- *for any u in $\mathcal{U}$, for all t in $[0, +\infty)$ and for all x_0 in $\mathcal{X}_0$, $|u(t)| \leq \bar{u}$ and $X(x_0, 0; t; u)$ is in $\mathcal{X}$,*

there exists a strictly positive number $\bar{k}$ such that for any $k > \bar{k}$:

- *there exists a function $T : \mathbb{R}^{d_x} \to \mathbb{R}^{d_x}$ which is a diffeomorphism on $\mathtt{cl}(\mathcal{X})$ and is solution to the PDE associated to the drift dynamics*

$$\frac{\partial T}{\partial x}(x) f(x) = k\, A\, T(x) + B\, h(x) \quad \forall x \in \mathcal{X} \ . \tag{19}$$

- *there exists a Lipschitz function $\bar{\varphi}$ defined on $\mathbb{R}^{d_x}$ verifying*

$$\bar{\varphi}(T(x)) = \frac{\partial T}{\partial x}(x) g(x) \qquad \forall x \in \mathcal{X} \ , \tag{20}$$

and such that, for any function $\mathcal{T} : \mathbb{R}^{d_x} \to \mathbb{R}^{d_x}$ verifying

$$\mathcal{T}(T(x)) = x \qquad \forall x \in \mathcal{X} \ ,$$

the system

$$\dot{\xi} = k\, A\, \xi + B\, y + \bar{\varphi}(\xi)\, u \quad , \quad \hat{x} = \mathcal{T}(\xi) \tag{21}$$

is an observer for system (18) *initialized in* $\mathcal{X}_0$.

Proof. See Appendix. □

Even though Theorem 4 is not constructive in its statement it has to be mentioned that the function involved in the observer definition can be given explicitly. For instance, following [13] the function $T : \mathbb{R}^{d_x} \to \mathbb{R}^{d_x}$ solution to (19) is defined by

$$T(x) = \int_{-\infty}^{0} e^{-kA\tau} B\, h(\check{X}(x, \tau))\, d\tau$$

where $\check{X}$ is the flow of a modified version of the vector field f (see the proof in the Appendix for more details). Similarly to the function T^∞, this mapping is not easily computable. Note however that as shown in [24] some numerical approximation can be considered.

Also, the function $\bar{\varphi}$ is defined on the open set $T(\mathcal{X})$ by (20). If the trajectories of the observer state $\hat{\xi}$ remain in this set, there is no need to extend its domain of definition to the whole $\mathbb{R}^{d_x}$. Otherwise, the only constraint is that the global Lipschitz constant $\mathfrak{a}$ of the extension be such that $k \min |\lambda_i| > \mathfrak{a}\, \bar{u}$, to ensure the convergence of the observer. In the proof below, it is proved that such extensions exist for k sufficiently large (this is not trivial because $\mathfrak{a}$ could a priori depend on k).

Otherwise, instead of extending $\bar{\varphi}$ outside $T(\mathcal{X})$, one could take

$$\bar{\varphi}(\xi) = \frac{\partial T}{\partial x}(\mathcal{T}(\xi)) g(\mathcal{T}(\xi))$$

but the way $\mathcal{T}$ is defined outside $T(\mathcal{X})$ must be such that :

$$\exists \alpha > 0 \ : \ \forall k \geq \bar{k}\, , \forall \hat{\xi} \in \mathbb{R}^{d_x}\, , \forall x \in \mathcal{X}\, ,$$
$$|T(x) - T(\mathcal{T}(\hat{\xi}))| \leq \alpha |T(x) - \hat{\xi}| \ .$$

The constraint here is that α must be independent from k. For instance, the function

$$\mathcal{T}(\xi) = \mathsf{Argmin}_{x \in \mathtt{cl}(\mathcal{X})} |T(x) - \xi|$$

clearly works since

$$|T(x) - T(\mathcal{T}(\hat{\xi}))| \leq |T(x) - \hat{\xi}| + \underbrace{|\hat{\xi} - T(\mathcal{T}(\hat{\xi}))|}_{\leq |\hat{\xi} - T(x)|} \ .$$

Another more regular candidate is the McShane extension

$$\mathcal{T}(\xi) = \min_{x \in \mathtt{cl}(\mathcal{X})} \{x + |T(x) - \xi|\}$$

which also verifies the requirement.

Example 1. *Consider the bioreactor model used in [6]*

$$\left\{ \begin{array}{rcl} \dot{x}_1 & = & \mu(x_1, x_2)\, x_1 - u x_1 \\ \dot{x}_2 & = & -a_3\, \mu(x_1, x_2)\, x_1 - u x_2 + u a_4 \end{array} \right. \quad , \quad y = x_1 \tag{22}$$

where x_1 (resp x_2) is the concentration of the microorganisms (resp the substrate) in a tank of constant volume, u is bounded positive input and the growth rate is given by the "Contois" model

$$\mu(x_1, x_2) = \frac{a_1 x_2}{a_2 x_1 + x_2}$$

with a_i positive constants. This system is uniformly instantaneously observable on the set

$$\mathcal{S} = \{(x_1, x_2) \in \mathbb{R}^2 \ : \quad x_1 > 0 \ , \ x_2 > 0\}$$

which is invariant by the dynamics (22)*. Besides, it is straightforward to check that the drift system is strongly differentially observable of order 2 on $\mathcal{S}$. Note finally that the input u being bounded, the trajectories are bounded and all the assumptions of Theorem 4 are satisfied.*

Let us look for a transformation T solution to the PDE (19) *associated to the drift system. We first note that the quantity $z = a_3 x_1 + x_2$ is constant along the drift dynamics and to facilitate the computations, we look for T as a function of (x_1, z) instead of (x_1, x_2), namely we solve :*

$$\frac{\partial T_\lambda}{\partial x_1}(x_1, z)\, \mu(x_1, z - a_3 x_1)\, x_1 = -k\lambda\, T_\lambda(x_1, z) + x_1$$

Integrating with respect to x_1, we find that a possible solution is :

$$T_\lambda(x_1, z) = \frac{1}{a_1} \int_0^{x_1} \left(\frac{z - a_3 x_1}{z - a_3 s} \right)^{\frac{k\lambda a_2}{a_1 a_3}} \left(\frac{s}{x_1} \right)^{\frac{k\lambda}{a_1}} \left(\frac{a_2}{z - a_3 s} + 1 \right) ds$$

By taking

$$T(x_1, x_2) = \Big(T_{\lambda_1}(x_1, a_3 x_1 + x_2), T_{\lambda_2}(x_1, a_3 x_1 + x_2) \Big)$$

with λ_1 and λ_2 two distinct strictly positive numbers, we thus obtain a solution to PDE (19) *on $\mathcal{S}$.*

For k sufficiently large, we know from Theorem 4 that there exists at least one solution of (19) *which is a diffeomorphism and we assume the same property holds for this particular solution. Assuming also that ξ remains in $T(\mathcal{X})$ as in [6], the observer writes*

$$\dot{\xi} = kA\,\xi + B\,y + \frac{dT}{dx}(\hat{x})g(\hat{x})u, \quad \hat{x} = T^{-1}(\xi)$$

which may be realized in the x-coordinates as

$$\dot{\hat{x}} = f(\hat{x}) + g(\hat{x})u + \left(\frac{dT}{dx}(\hat{x}) \right)^{-1} (y - \hat{x}_1) \ . \tag{23}$$

The results of a simulation with the same system parameters as in [6] are presented on Figure 2.

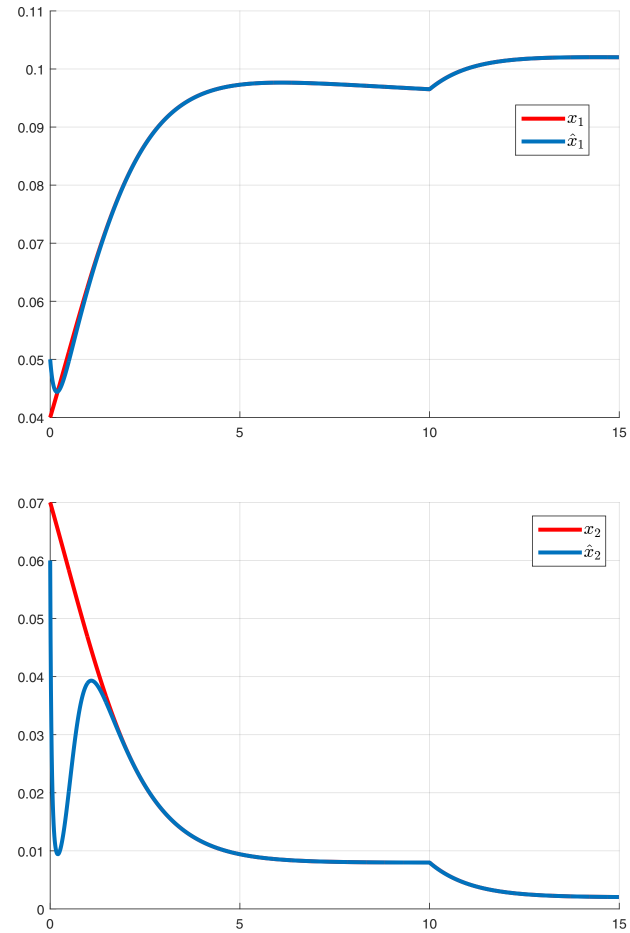

Fig. 2. Nonlinear Luenberger observer (23) for system (22) with $k = 3$, $\lambda_1 = 1$, $\lambda_2 = 2$.

V. Conclusion

We have shown how a Luenberger methodology can be applied to nonlinear controlled systems. It is based on the resolution of a time-varying PDE, the solutions of which exist under very mild assumptions, transform the system into a linear asymptotically stable one, and become injective after a certain time if

- either the function made of the output and a certain number of its derivatives is Lipschitz-injective : this is verified when the system is strongly differentially observable and the trajectories are bounded.
- or the system is backward-distinguishable (uniformly in time), but in this case, injectivity is ensured for "almost all" choice of a diagonal complex matrix A (of sufficiently large dimension) in the sense of the Lebesgue measure in $\mathbb{C}$.

Although solutions to the PDE are guaranteed to exist, they may be difficult to compute. We have shown on practical examples how this can be done by a priori guessing their "structure". The advantage with respect to a more straightforward high gain design is however that the transformation does not depend on the derivatives of the input which thus need not be computed.

Also, it is interesting to remember that exactly as in the high gain paradigm, for uniformly instantaneously observable control-affine systems, we may use the stationary transformation associated to the autonomous drift system when it is strongly differentially observable of order d_x. The result does not stand for higher orders of differential observability, since it relies on the existence of Lipschitz functions $\mathfrak{g}_i$ such that $\mathfrak{g}_i(\mathbf{H}_i(x)) = L_g L_f^{i-1}(x)$, and it is shown in [18] that the Lipschitzness is lost when the drift system is differentially observable of higher order.

A perspective of this work could be to study the impact of the noise in a Luenberger design and in particular see if it is possible to optimize the choice of the eigenvalues of the Hurwitz matrix A in order to limit its effect.

Appendix A
Proof of Theorem 4

Let $\mathcal{X}'$ and $\mathcal{X}''$ be open sets such that

$$\mathtt{cl}(\mathcal{X}) \subset \mathcal{X}' \subset \mathtt{cl}(\mathcal{X}') \subset \mathcal{X}'' \subset \mathtt{cl}(\mathcal{X}'') \subset \mathcal{S} \ .$$

Consider the function $T : \mathbb{R}^{d_x} \to \mathbb{R}^{d_x}$ defined by

$$T(x) = \int_{-\infty}^{0} e^{-kA\tau} B\, h(\check{X}(x,\tau))\, d\tau$$

where $\check{X}(x,\tau)$ denotes the value at time τ of the solution initialized at x at time 0 of the modified autonomous drift system

$$\dot{x} = \chi(x) f(x) \ ,$$

for any C^∞ function $\chi : \mathbb{R}^{d_x} \to \mathbb{R}$ satisfying

$$\chi(x) = \begin{cases} 1 & , \quad \text{if } x \in \mathtt{cl}(\mathcal{X}') \\ 0 & , \quad \text{if } x \notin \mathcal{X}'' \end{cases}$$

According to [25, Proposition 3.3], there exists $\overline{k}_0$ such that for all $k \geq \overline{k}_0$, T is C^1 and verifies PDE (19). Now let us prove that it is injective on $\mathtt{cl}(\mathcal{X}')$ for k sufficiently large[11]. The drift system being strongly differentially observable of order d_x, the function

$$\mathbf{H}_{d_x}(x) = (h(x), L_f h(x), \dots, L_f^{d_x}(x))$$

is an injective immersion on $\mathtt{cl}(\mathcal{X}')$ and by [25, Lemma 3.2], there exists $L_H > 0$ such that for all $(x_a, x_b)^2$ in $\mathtt{cl}(\mathcal{X}')^2$,

$$|\mathbf{H}_{d_x}(x_a) - \mathbf{H}_{d_x}(x_b)| \geq L_H |x_a - x_b| \ .$$

Besides, since $\chi f = f$ on $\mathtt{cl}(\mathcal{X}')$, after several integrations by parts (by integrating $e^{-kA\tau}$ and differentiating $h(X(x,\tau))$ with respect to time, and using that A is Hurwitz and $\mathcal{X}$ bounded for the limit $\tau \to -\infty$), we obtain for all x in $\mathtt{cl}(\mathcal{X}')$

$$T(x) = A^{-d_x} \mathscr{C} \left(-K \mathbf{H}_{d_x}(x) + \frac{1}{k^{d_x}} R(x) \right) \qquad (24)$$

where $K = \mathrm{diag}\left(\frac{1}{k}, ..., \frac{1}{k^{d_x}}\right)$, $\mathscr{C}$ is the invertible controllability matrix

$$\mathscr{C} = [A^{d_x - 1} B \ ... \ AB \ B] \ ,$$

and R the remainder

$$R(x) = \mathscr{C}^{-1} \int_{-\infty}^{0} e^{-kA\tau} B\, L_f^{d_x}(\check{X}(x,\tau))\, d\tau \ .$$

This latter integral makes sense on $\mathtt{cl}(\mathcal{X}')$ because :

- A being diagonal and denoting $a = \min_i |\lambda_i| > 0$, for all $\tau \in (-\infty, 0]$,
$$\left|e^{-kA\tau}\right| \leq e^{ka\tau} \ .$$
- By definition of the function χ, for all x in $\mathtt{cl}(\mathcal{X}')$, $\check{X}(x,\tau)$ is in $\mathtt{cl}(\mathcal{X}')$ for all τ, i.e. $\tau \mapsto L_f^{d_x}(\check{X}(x,\tau))$ is bounded.

[11] This proof is similar to that of [13, Theorem 4].

So now taking (x_a, x_b) in $\mathtt{cl}(\mathcal{X}')^2$, and considering the difference $|T(x_a) - T(x_b)|$, from (24), we obtain

$$|T(x_a) - T(x_b)| \geq \frac{|A^{-d_x}\mathscr{C}|}{k^{d_x}} \Big(|\mathbf{H}_{d_x}(x_a) - \mathbf{H}_{d_x}(x_b)| - |R(x_a) - R(x_b)| \Big) \ ,$$

and if R is Lipschitz with Lipschitz constant L_R, we get

$$|T(x_a) - T(x_b)| \geq \frac{|A^{-d_x}\mathscr{C}|}{k^{d_x}} (L_H - L_R) |x_a - x_b| \ .$$

In order to deduce the injectivity of T, we also need $L_R < L_H$ and we are going to prove that this is true for k sufficiently large. To compute L_R, let us find a bound of $\left|\frac{\partial R}{\partial x}(x)\right|$. By defining

$$c_0 = \max_{x \in \mathtt{cl}(\mathcal{X}')} \left| B \frac{\partial L_f^{d_x} h}{\partial x}(x) \right| \quad , \quad \rho_1 = \max_{x \in \mathtt{cl}(\mathcal{X}')} \left| \frac{\partial f}{\partial x}(x) \right| \ ,$$

we have for all τ in $(-\infty, 0]$ and all x in $\mathtt{cl}(\mathcal{X}')$, $\left| B \frac{\partial L_f^{d_x} h}{\partial x}(\check{X}(x,\tau)) \right| \leq c_0$ and[12]

$$\left| \frac{\partial \check{X}}{\partial x}(x,\tau) \right| \leq e^{-\rho_1 \tau} \ . \qquad (25)$$

We conclude that for $k > \frac{\rho_1}{a}$, R is C^1 and there exists a positive constant c_1 such that for all x in $\mathtt{cl}(\mathcal{X}')$,

$$\begin{aligned} \left| \frac{\partial R}{\partial x}(x) \right| &\leq |\mathscr{C}^{-1}| \int_{-\infty}^{0} \left| e^{-kA\tau} \right| \left| B \frac{\partial L_f^{d_x} h}{\partial x}(\check{X}(x,\tau)) \right| \left| \frac{\partial \check{X}}{\partial x}(x,\tau) \right| d\tau \\ &\leq \frac{c_1}{ka - \rho_1}. \end{aligned}$$

We finally obtain

$$|T(x_a) - T(x_b)| \geq L_T |x_a - x_b| \quad \forall (x_a, x_b) \in \mathtt{cl}(\mathcal{X}')^2 \quad (26)$$

where

$$L_T = \frac{|A^{-d_x}\mathscr{C}|}{k^{d_x}} \left(L_H - \frac{c_1}{ka - \rho_1} \right) \ ,$$

and T is injective on $\mathtt{cl}(\mathcal{X}')$ if $k \geq \overline{k}_1$ with

$$\overline{k}_1 = \max \left\{ \overline{k}_0, \frac{c_1 + \rho_1 L_H}{a L_H} \right\} \ .$$

Moreover, taking x in $\mathcal{X}'$, any v in $\mathbb{R}^m$ and h sufficiently small for $x + hv$ to be in $\mathcal{X}'$, it follows from (26) that

$$\left| \frac{T(x+hv) - T(x)}{h} \right| \geq L_T |v| \ ,$$

and making h tend to zero, we get

$$\left| \frac{\partial T}{\partial x}(x) v \right| \geq L_T |v|$$

and T is full-rank on $\mathcal{X}'$. So T is a diffeomorphism on $\mathcal{X}'$ for $k \geq \overline{k}_1$.

[12] Because $\psi(\tau) = \frac{\partial \check{X}}{\partial x}(x,\tau)$ follows the ODE $\frac{d\psi}{d\tau}(\tau) = \frac{\partial f}{\partial x}(\check{X}(x,\tau))\psi(\tau)$, and $\psi(0) = I$.

Now, let us show that system (21) is an observer for system (18). Suppose for the time being that we have shown that there exists a strictly positive number $\mathfrak{a}$ such that for any $k \geq \overline{k}_1$, there exists a function $\overline{\varphi}$ such that (20) holds and

$$|\overline{\varphi}(\hat{\xi}) - \overline{\varphi}(\xi)| \leq \mathfrak{a}\,|\hat{\xi} - \xi| \qquad \forall (\hat{\xi}, \xi) \in (\mathbb{R}^{d_x})^2 \ . \tag{27}$$

Take u in $\mathcal{U}$, x_0 in $\mathcal{X}_0$ $\hat{\xi}_0$ in $\mathbb{R}^{d_x}$, and consider the solution $X(x_0;t;u)$ of system (18) and any corresponding solution $\hat{\Xi}(\hat{\xi}_0;t;u,y_{x_0})$ of system (21). Since $X(x_0;t;u)$ remains in $\mathcal{X}$ by assumption, the error $e(t) = \hat{\Xi}(\hat{\xi}_0;t;u,y_{x_0}) - T(X(x_0;t;u))$ verifies

$$\dot{e} = kA\ e + \left(\overline{\varphi}(\hat{\Xi}(\hat{\xi}_0;t;u,y_{x_0})) - \overline{\varphi}(T(X(x_0;t;u))\right)u$$

and thus

$$\dot{\overbrace{e^\top e}} \leq -2(ka - \mathfrak{a}\,\overline{u})\,e^\top e \ .$$

Defining $\overline{k}_2 = \max\{\overline{k}_1, \frac{\mathfrak{a}\overline{u}}{a}\}$, we conclude that e asymptotically converges to 0 if $k \geq \overline{k}_2$. Note that for this conclusion to hold, it is crucial to have $\mathfrak{a}$ independent from k. Now, consider an open set $\tilde{\mathcal{X}}$ such that $\mathtt{cl}(\mathcal{X}) \subset \tilde{\mathcal{X}} \subset \mathtt{cl}(\tilde{\mathcal{X}}) \subset \mathcal{X}'$. Since $T(X(x_0;t;u))$ remains in $T(\mathcal{X})$ and $\mathtt{cl}(T(\mathcal{X})) = T(\mathtt{cl}(\mathcal{X}))$ is contained in the open set $T(\tilde{\mathcal{X}})$, there exists a time $\overline{t}$ such that for all $t \geq \overline{t}$, $\hat{\Xi}(\hat{\xi}_0;t;u,y_{x_0})$ is in $T(\tilde{\mathcal{X}})$. $\mathcal{T} = T^{-1}$ is C^1 on the compact set $\mathtt{cl}(T(\tilde{\mathcal{X}}))$ and thus Lipschitz on that set. It follows that $\hat{X}((x_0,\hat{\xi}_0);t;u) = \mathcal{T}(\hat{\Xi}(\hat{\xi}_0;t;u,y_{x_0}))$ converges to $X(x_0;t;u)$.

It remains to show the existence of the functions $\overline{\varphi}$. Since system (18) is uniformly instantaneously observable and its drift system is strongly differentially observable of order d_x on $\mathcal{S}$, we know since [5] that for all i in $\{1,\dots,d_x\}$, there exists a Lipschitz function $\mathfrak{g}_i$ such that

$$L_g L_f^{i-1} h(x) \ = \ \mathfrak{g}_i(h(x),\dots,L_f^{i-1}(x)) \quad \forall x \in \mathtt{cl}(\mathcal{X}) \ . \tag{28}$$

Consider the function

$$\begin{aligned}\varphi(x) &= \frac{\partial T}{\partial x}(x) g(x) \\ &= A^{-d_x}\mathscr{C}\left(-\underbrace{K\frac{\partial \mathbf{H}_{d_x}}{\partial x}(x)g(x)}_{\varphi_H(x)} + \underbrace{\frac{1}{k^{d_x}}\frac{\partial R}{\partial x}(x)g(x)}_{\varphi_R(x)}\right) \ .\end{aligned}$$

Let us first study φ_H. Notice that the ith-component of φ_H is $\varphi_{H,i} = \dfrac{1}{k^i}L_g L_f^{i-1}h(x)$ and according to (28), there exists L_i such that for all $(x,\hat{x})$ in $\mathtt{cl}(\mathcal{X})^2$

$$|\varphi_{H,i}(\hat{x}) - \varphi_{H,i}(x)| \leq L_i \sum_{j=1}^{i}\left|\frac{1}{k^j}\left(L_f^{j-1}(\hat{x}) - L_f^{j-1}(x)\right)\right|$$

and thus L such that for all $(x,\hat{x})$ in $\mathtt{cl}(\mathcal{X})^2$

$$|\varphi_H(\hat{x}) - \varphi_H(x)| \leq L|K\mathbf{H}_{d_x}(\hat{x}) - K\mathbf{H}_{d_x}(x)| \ .$$

But using (24), we get

$$\begin{aligned}|K\mathbf{H}_{d_x}(\hat{x}) - K\mathbf{H}_{d_x}(x)| \leq &|A^{d_x}\mathscr{C}^{-1}||T(\hat{x}) - T(x)| \\ &+ \frac{1}{k^{d_x}}|R(\hat{x}) - R(x)| \ .\end{aligned}$$

We have seen that for all $(x,\hat{x})$ in $\mathtt{cl}(\mathcal{X})^2$

$$|R(\hat{x}) - R(x)| \leq \frac{c_1}{ka-\rho_1}|\hat{x} - x|$$

and according to (26),

$$\frac{1}{k^{d_x}}|R(\hat{x}) - R(x)| \leq \frac{\frac{c_1}{ka-\rho_1}}{L_H - \frac{c_1}{ka-\rho_1}}|A^{d_x}\mathscr{C}^{-1}||T(\hat{x}) - T(x)| \ .$$

We finally obtain, for any $(x,\hat{x})$ in $\mathtt{cl}(\mathcal{X})^2$ and for any $k \geq \overline{k}_1$,

$$\begin{aligned}&|\varphi_H(\hat{x}) - \varphi_H(x)| \\ &\quad \leq L|A^{d_x}\mathscr{C}^{-1}|\left(1 + \frac{\frac{c_1}{ka-\rho_1}}{L_H - \frac{c_1}{ka-\rho_1}}\right)|T(\hat{x}) - T(x)| \\ &\quad \leq L|A^{d_x}\mathscr{C}^{-1}|\left(1 + \frac{c_1}{L_H(\overline{k}_1 a - \rho_1)}\right)|T(\hat{x}) - T(x)| \ .\end{aligned}$$

Let us now study the term $\varphi_R(x)$. For $(x,\hat{x})$ in $\mathtt{cl}(\mathcal{X})^2$,

$$\begin{aligned}\varphi_R(\hat{x}) - \varphi_R(x) = &\frac{1}{k^{d_x}}\mathscr{C}^{-1}\int_{-\infty}^{0} e^{-kA\tau}B \\ &\times (D_1(x,\hat{x},\tau) + D_2(x,\hat{x},\tau) + D_3(x,\hat{x},\tau))d\tau\end{aligned}$$

where

$$\begin{aligned}&D_1(x,\hat{x},\tau) \\ &= \left(\frac{\partial L_f^{d_x}h}{\partial x}(\check{X}(x,\tau)) - \frac{\partial L_f^{d_x}h}{\partial x}(\check{X}(\hat{x},\tau))\right)\frac{\partial \check{X}}{\partial x}(\hat{x},\tau)g(\hat{x}) \\ &D_2(x,\hat{x},\tau) \\ &= \frac{\partial L_f^{d_x}h}{\partial x}(\check{X}(x,\tau))\left(\frac{\partial \check{X}}{\partial x}(\hat{x},\tau) - \frac{\partial \check{X}}{\partial x}(x,\tau)\right)g(\hat{x}) \\ &D_3(x,\hat{x},\tau) = \frac{\partial L_f^{d_x}h}{\partial x}(\check{X}(x,\tau))\frac{\partial \check{X}}{\partial x}(x,\tau)\left(g(\hat{x}) - g(x)\right)\end{aligned}$$

Assuming that $L_f^{d_x}h$ is C^2 and g is C^1, it follows from (25) and the fact that $\check{X}(x,\tau)$ is in the compact set $\mathtt{cl}(\mathcal{X}')$ for all τ in $(-\infty,0]$, that for all $(x,\hat{x})$ in $\mathtt{cl}(\mathcal{X})^2$ and for all τ in $(-\infty,0]$,

$$\begin{aligned}|D_1(x,\hat{x},\tau)| &\leq c_2 e^{-2\rho_1\tau}|x - \hat{x}| \\ |D_3(x,\hat{x},\tau)| &\leq c_3 e^{-\rho_1\tau}|x - \hat{x}| \ .\end{aligned}$$

As for D_2, posing $\varphi(\tau) = \frac{\partial \check{X}}{\partial x}(\hat{x},\tau) - \frac{\partial \check{X}}{\partial x}(x,\tau)$, and differentiating φ with respect to time, we get

$$\begin{aligned}&\varphi(0) = 0 \\ &\varphi'(\tau) = \frac{\partial f}{\partial x}(\check{X}(\hat{x},\tau))\varphi(\tau) \\ &\qquad + \left(\frac{\partial f}{\partial x}(\check{X}(\hat{x},\tau)) - \frac{\partial f}{\partial x}(\check{X}(x,\tau))\right)\frac{\partial \check{X}}{\partial x}(x,\tau) \ .\end{aligned} \tag{29}$$

Since for all τ in $(-\infty,0]$ and for all $(x,\hat{x})$ in $\mathtt{cl}(\mathcal{X})^2$,

$$\left|\frac{\partial f}{\partial x}(\check{X}(\hat{x},\tau))\right| \leq \rho_1 \quad , \quad \left|\frac{\partial \check{X}}{\partial x}(x,\tau)\right| \leq e^{-\rho_1\tau}$$

and

$$\left|\frac{\partial f}{\partial x}(\check{X}(\hat{x},\tau)) - \frac{\partial f}{\partial x}(\check{X}(x,\tau))\right| \leq c_4 e^{-\rho_1\tau}|x - \hat{x}| \ ,$$

we obtain by solving (29) in negative time and taking the norm

$$|D_2(\hat{x},x,\tau)| \leq \left(c_5 e^{-\rho_1\tau} + c_6 e^{-2\rho_1\tau}\right)|x-\hat{x}| \leq c_7\, e^{-2\rho_1\tau}|x-\hat{x}|$$

for all τ in $(-\infty,0]$ and all $(x,\hat{x})$ in $\mathtt{cl}(\mathcal{X})^2$. Therefore, for all $k \geq \overline{k}_1$,

$$\begin{aligned} |\varphi_R(\hat{x}) - \varphi_R(x)| &\leq \frac{1}{k^{d_x}} \frac{c_8}{ka-\rho_1} |x-\hat{x}| \\ &\leq \frac{\frac{c_9}{ka-\rho_1}}{L_H - \frac{c_1}{ka-\rho_1}} |T(x) - T(\hat{x})| \\ &\leq \frac{c_9}{L_H(\overline{k}_1 a - \rho_1)} |T(x) - T(\hat{x})| \ . \end{aligned}$$

Finally, there exists a constant $\mathfrak{a}$ such that for all $k \geq \overline{k}_1$, and for all $(x,\hat{x})$ in $\mathtt{cl}(\mathcal{X})^2$,

$$|\varphi(\hat{x}) - \varphi(x)| \leq \mathfrak{a}\, |T(\hat{x}) - T(x)| \ . \tag{30}$$

Consider now the function

$$\overline{\varphi}(\xi) = \varphi(T^{-1}(\xi))$$

defined on $T(\mathcal{X}')$. According to (30), $\overline{\varphi}$ is Lipschitz on $T(\mathcal{X}')$, and with Kirszbraun-Valentine Theorem [26], [27], it admits a Lipschitz extension on $\mathbb{R}^{d_x}$ with same Lipschitz constant $\mathfrak{a}$, i.e. such that (20) and (27) hold. This concludes the proof.

References

[1] D. Luenberger, "Observing the state of a linear system," *IEEE Transactions on Military Electronics*, vol. 8, pp. 74–80, 1964.

[2] A. H. Jazwinski, *Stochastic processes and filtering theory*. Academic Press, 1970.

[3] D. Bossane, D. Rakotopara, and J. P. Gauthier, "Local and global immersion into linear systems up to output injection," *IEEE Conference on Decision and Control*, pp. 2000–2004, 1989.

[4] P. Jouan, "Immersion of nonlinear systems into linear systems modulo output injection," *SIAM Journal on Control and Optimization*, vol. 41, no. 6, pp. 1756–1778, 2003.

[5] J.-P. Gauthier and G. Bornard, "Observability for any u(t) of a class of nonlinear systems," *IEEE Transactions on Automatic Control*, vol. 26, pp. 922 – 926, 1981.

[6] J.-P. Gauthier, H. Hammouri, and S. Othman, "A simple observer for nonlinear systems application to bioreactors," *IEEE Transactions on Automatic Control*, vol. 37, no. 6, pp. 875–880, 1992.

[7] A. Astolfi, R. Ortega, and A. Venkatraman, "A globally exponential convergent immersion and invariance speed observer for mechanical systems with non-holonomic constraints," *Automatica*, vol. 46, no. 1, pp. 182–189, 2010.

[8] M. Zeitz, "Observability canonical (phase-variable) form for nonlinear time-variable systems," *International Journal of Systems Science*, vol. 15, no. 9, pp. 949–958, 1984.

[9] J.-P. Gauthier and I. Kupka, *Deterministic observation theory and applications*. Cambridge University Press, 2001.

[10] A. Shoshitaishvili, "On control branching systems with degenerate linearization," *IFAC Symposium on Nonlinear Control Systems*, pp. 495–500, 1992.

[11] N. Kazantzis and C. Kravaris, "Nonlinear observer design using Lyapunov's auxiliary theorem," *Systems and Control Letters*, vol. 34, pp. 241–247, 1998.

[12] A. Krener and M. Xiao, "Nonlinear observer design in the Siegel domain," *SIAM Journal on Control and Optimization*, vol. 41, no. 3, pp. 932–953, 2003.

[13] V. Andrieu and L. Praly, "On the existence of a Kazantzis–Kravaris / Luenberger observer," *SIAM Journal on Control and Optimization*, vol. 45, no. 2, pp. 432–456, 2006.

[14] R. Engel, "Exponential observers for nonlinear systems with inputs," Universität of Kassel, Department of Electrical Engineering, Tech. Rep., 2005.

[15] ——, "Nonlinear observers for Lipschitz continuous systems with inputs," *International Journal of Control*, vol. 80, no. 4, pp. 495–508, 2007.

[16] G. Kreisselmeier and R. Engel, "Nonlinear observers for autonomous lipshitz continuous systems," *IEEE Transactions on Automatic Control*, vol. 48, no. 3, pp. 451–464, 2003.

[17] P. Bernard, "Luenberger observers for nonlinear controlled systems," *IEEE Conference on Decision and Control*, 2017.

[18] P. Bernard, L. Praly, V. Andrieu, and H. Hammouri, "On the triangular canonical form for uniformly observable controlled systems," *Automatica*, vol. 85, pp. 293–300, 2017.

[19] F. Poulain, L. Praly, and R. Ortega, "An observer for permanent magnet synchronous motors with currents and voltages as only measurements," *IEEE Conference on Decision and Control*, 2008.

[20] A. Tornambe, "Use of asymptotic observers having high gains in the state and parameter estimation," *IEEE Conference on Decision and Control*, vol. 2, pp. 1791–1794, 1989.

[21] J. Lee, J. Hong, K. Nam, R. Ortega, L. Praly, and A. Astolfi, "Sensorless control of surface-mount permanent-magnet synchronous motors based on a nonlinear observer," *IEEE Transactions on Power Electronics*, vol. 25, no. 2, pp. 290–297, 2010.

[22] J. Malaizé, L. Praly, and N. Henwood, "Globally convergent nonlinear observer for the sensorless control of surface-mount permanent magnet synchronous machines," *IEEE Conference on Decision and Control*, 2012.

[23] V. Andrieu, "Bouclage de sortie et observateur," Ph.D. dissertation, École Nationale Supérieure des Mines de Paris, 2005.

[24] L. Marconi, L. Praly, and A. Isidori, "Output stabilization via nonlinear Luenberger observers," *SIAM Journal of Control and Optimization*, vol. 45, no. 6, pp. 2277–2298, 2007.

[25] V. Andrieu, "Convergence speed of nonlinear Luenberger observers," *SIAM Journal on Control and Optimization*, vol. 52, no. 5, pp. 2831–2856, 2014.

[26] M. D. Kirszbraun, "Über die zusammenziehende und Lipschitzsche transformationen," *Fundamenta Mathematicae*, vol. 22, pp. 77–108, 1934.

[27] F. A. Valentine, "A Lipschitz condition preserving extension for a vector function," *American Journal of Mathematics*, vol. 67, no. 1, pp. 83–93, 1945.

Pauline Bernard graduated in Applied Mathematics from MINES ParisTech in 2014. She joined the Control and Systems Center of MINES ParisTech as a PhD student to work on observer design for nonlinear systems with Laurent Praly and Vincent Andrieu. In 2017, she obtained her PhD in Mathematics and Control from PSL Research University. In 2018, she visited the Hybrid Systems Lab in the Department of Computer Engineering of University California Santa Cruz, USA, and she is now a post-doc at the Department of Electronics, Computer Science and Systems of the University of Bologna, Italy.

Vincent Andrieu graduated in applied mathematics from INSA de Rouen, France, in 2001. After working in ONERA (French aerospace research company), he obtained a PhD degree from Ecole des Mines de Paris in 2005. In 2006, he had a research appointment at the Control and Power Group, Dept. EEE, Imperial College London. In 2008, he joined the CNRS-LAAS lab in Toulouse, France, as a CNRS-chargé de recherche. Since 2010, he has been working in LAGEP-CNRS, Université de Lyon 1, France. In 2014, he joined the functional analysis group from Bergische Universität Wuppertal in Germany, for two sabbatical years. His main research interests are in the feedback stabilization of controlled dynamical nonlinear systems and state estimation problems. He is also interested in practical application of these theoretical problems, and especially in the field of aeronautics and chemical engineering.