Tutorial article

Observer design for continuous-time dynamical systems

Pauline Bernard [a,*], Vincent Andrieu [b], Daniele Astolfi [b]

[a] *Centre Automatique et Systèmes, Mines Paris, Université PSL, Paris, France*
[b] *LAGEPP CNRS, Université Claude Bernard Lyon 1, Lyon, France*

ARTICLE INFO

ABSTRACT

We review the main design techniques of state observer design for continuous-time dynamical systems, namely algorithms which reconstruct online the full information of a dynamical process on the basis of partially measured data. Starting from necessary conditions for the existence of such asymptotic observers, we classify the available methods depending on the detectability/observability assumptions they require. We show how each class of observer relies on transforming the system dynamics in a particular normal form which allows the design of an observer, and how each observability condition guarantees the invertibility of its associated transformation and the convergence of the observer. Finally, some implementation aspects and open problems are briefly discussed.

1. Introduction

In many applications, knowing the current state of a dynamical system is crucial either to build a controller or to obtain real time information on the system for decision-making or monitoring, see, e.g., Kang, Krener, Xiao, and Xu (2013), Praly (2015), Radke and Gao (2006) and references therein. A common way of addressing this problem is to place some sensors on the physical system in order to have access to such information. In practice, however, not all the state variables are directly measurable: either because we want to reduce the number of sensors to reduce costs, due to physical constraints, or simply because such a sensor does not exist. In this case, an estimation algorithm is thus needed to process the incomplete and imperfect information provided by the sensors and thereby reconstruct a reliable estimate of the whole system state. The number and quality of sensors being often limited in practice due to cost and physical constraints, such estimation algorithms play a decisive role in a lot of applications. See, as a few examples, the problems of estimating the state of the charge of a battery (Blondel, Postoyan, Raël, Benjamin, & Desprez, 2018), the position of sensorless permanent magnet synchronous machines (Bernard & Praly, 2020; Malaizé, Praly, & Henwood, 2012) or the size of the crystals in a crystallization process (Brivadis et al., 2020).

In this paper, we restrict our attention to systems modeled by finite-dimensional continuous-time dynamics and look for an estimation algorithm under the form of a dynamical system, denoted as *asymptotic observer*, that takes as input the sensor measurements (and possible known control actions) and produces an estimate which asymptotically converges to the plant state. This framework is detailed in Section 2.

Of course, such an algorithm can exist only if the measurements somehow contain enough information to determine uniquely (asymptotically) the state of the system, namely if the system is *detectable*. Then, some additional requirements (finite-time convergence, stability of the estimation error, tunability, etc.) typically lead to a very wide range of detectability/observability properties, which all represent necessary conditions for the existence of certain classes of estimation algorithms. These are reviewed in Section 3.

Then, it comes the question of actually designing such an algorithm. As we shall see, this problem has been solved in a general and constructive way for systems modeled by linear dynamics, but only for particular classes in the nonlinear context. Therefore, observer design usually consists in looking for a change of coordinates that transforms the dynamics into one of these *normal forms* where an observer design is available. Existence and invertibility of such a transformation typically involve the observability conditions that characterize each "family" of designs. The goal of this paper is therefore to gather a large part of the huge amount of contributions scattered throughout the literature and classify them according to the "type" of observability properties they

---

* Corresponding author.

E-mail address: pauline.bernard@minesparis.psl.eu (P. Bernard).

Table 1
Main families of observers with their associated detectability/observability properties and convergence conditions.

Observability property	Transformation T	Normal form		Observer	Conditions for convergence
Differential detectability (Def 3.2)	Identity	Given coordinates		Metric based observer (35)	Convexity property (Prop. 4.2)
Linearizability by output injection (Section 5.4)	No systematic design	State affine form (46)		Kalman or Kalman-like (49)	Uniform complete observability (47)
Differential observability of order m (Section 6.1)	Output and its $m-1$ Lie derivatives (Section 6.1)	Triangular form	General (Definition 6.2)	High-gain	Global Lipschitz and conditions on jacobians (Prop. 6.1)
				Low-power high-gain	
			(A,C) constant in the form (55)	Extended Kalman high-gain	Global Lipschitz and C^1 (Prop. 6.3)
				Homogeneous high-gain	Homogeneous incremental bound of degree τ (Prop. 6.4)
				Interconnection of mixed homogeneous 2nd order	Affine incremental bound (Prop. 6.7)
			$A(y,t)$ and C constant	Kalman-like high-gain	Instantaneous complete observability and global Lipschitz bound (Prop. 6.2)
				Bi-homogeneous high-gain with dynamic scaling	non-global Lipschitz incremental bound (Prop. 6.5)
Backward distinguishability	Solution to PDE (98) or (100)	Hurwitz form Eq. (96)		Nonlinear Luenberger (or KKL) (97)	–

require. This is done throughout Sections 4 to 7 . A recapitulative table is given in Table 1.

Finally, even when such a design has been achieved, some issues may need to be addressed before an actual implementation can be carried out. These include inverting the change of coordinates, tuning the observer to ensure some desired performances, taking into account unknown parameters, disturbances, measurement noise, state constraints, or the sampling of the output, etc. Such issues and some of the corresponding solutions available in the literature are briefly reviewed in Section 8.

Notation. $\mathbb{R}$ (resp. $\mathbb{N}$) denotes the set of real numbers (resp. integers) and $\mathbb{R}_{\geq 0} := [0, +\infty)$, $\mathbb{N}_{>0} := \mathbb{N} \setminus \{0\}$. A map $\rho : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ is a $\mathcal{K}$-map if $\rho(0) = 0$ and ρ is continuous and increasing, and a $\mathcal{K}^\infty$-map if it is also unbounded. A map $\beta : \mathbb{R}_{\geq 0} \times \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ is a $\mathcal{KL}$-map if for all $t \in \mathbb{R}_{\geq 0}$, $r \mapsto \beta(r, t)$ is a class-$\mathcal{K}$ map and for all $r \in \mathbb{R}_{\geq 0}$, $t \mapsto \beta(r, t)$ is decreasing with $\lim_{t \to \infty} \beta(r, t) = 0$. For a matrix M, $\text{He}(M) := M + M^\top$, and for a positive definite matrix $P \in \mathbb{R}^{n \times n}$, $\| \cdot \|_P$ denotes the norm associated to P, namely $\|x\|_P := \sqrt{x^\top P x}$ for $x \in \mathbb{R}^n$. We denote with $\otimes$ the Kronecker product.

2. Problem statement

2.1. Observation problem

We consider a general plant described by finite-dimensional continuous-time nonlinear dynamics of the form

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

with state $x \in \mathbb{R}^{n_x}$, output (or measurement) $y \in \mathbb{R}^{n_y}$ and maps $f : \mathbb{R}^{n_x} \times \mathbb{R} \to \mathbb{R}^{n_x}$, $h : \mathbb{R}^{n_x} \times \mathbb{R} \to \mathbb{R}^{n_y}$ sufficiently regular, with $n_x, n_y \in \mathbb{N}_{>0}$. This paper aims at reviewing existing methods available in the literature to estimate online the full state $x(t)$ of a solution to system (1) at time t, based on the knowledge of the plant model described by the functions f, h, and of the output y up to time t. Generally in

applications, we are interested in estimating only physically relevant solutions which are known to remain in a certain set $\mathcal{X}$, often taken compact and modeling known physical bounds on the states. Hence, the following assumption holds throughout the paper.

Standing Assumption. *We consider sets $\mathcal{X}_0 \subseteq \mathcal{X} \subseteq \mathbb{R}^{n_x}$ such that any solution of (1) initialized in $\mathcal{X}_0$ is defined on $[0, +\infty)$ and remains in $\mathcal{X}$ at all positive times.*

The set $\mathcal{X}_0$ does not need to be known for design, but will be used in the theorem statements to refer to the plant solutions that need to be estimated. On the other hand, the knowledge of $\mathcal{X}$ may actively be used to design the estimation algorithms. Note that we restrict our attention to solutions defined on $[0, +\infty)$ because we will mainly review asymptotic estimation methods. Otherwise, the estimate needs to converge in finite time, before solutions cease to exist.

The system (1) encompasses a large variety of systems among which *autonomous systems*

$$\dot{x} = f(x), \qquad y = h(x), \tag{2}$$

and *systems with (known) input*

$$\dot{x} = f(x,u), \qquad y = h(x,u), \tag{3}$$

with $u : \mathbb{R} \to \mathbb{R}^{n_u}$, although the treatise of (3) may present subtle differences compared to (1) in terms of causality and uniformity with respect to inputs.

2.2. Approaches for real-time state estimation

A first naive approach to solve the desired estimation problem would be to simulate (1) simultaneously for a set of initial conditions and progressively removing from the set those producing an output trajectory "too far" from the witnessed $y(t)$ (with the notion of "far" to be defined). If the output y determines a unique possible solution

asymptotically, then the set of possible initial conditions shrinks to only one possible asymptotically. Otherwise, a certain distribution of possible initial conditions remains. However, the trade off between amount of computations and estimation precision is typically not reachable, and the open-loop numerical integration of (1) is often not reliable in the presence of disturbances and model uncertainties. This path has nevertheless aroused a lot of research:

- either through stochastic approaches, considering random processes in the plant model (1), and following the probability distribution of the possible values of the state (see, e.g., Jazwinski, 1970);
- or in a deterministic way, considering the presence of bounded disturbances in the dynamics (1), and producing a "set-valued observer" (see, e.g., Lin, Zhai, & Antsaklis, 2003), "interval observer" (see, e.g., Efimov, Raïssi, Chebotarev, & Zolghadri, 2013; Gouzé, Rapaport, & Hady-Sadok, 2000; Khan, Xie, Bo, & Liu, 2021; Mazenc & Bernard, 2011), or "norm observers" (see, e.g., Sontag & Wang, 1997), guaranteed to contain the plant state.

Another natural approach is to proceed by a minimization approach (see, e.g., Zimmer, 1994), namely to solve online at each time

$$\hat{x}(t) = \underset{\hat{x}_f}{\arg\min} \int_0^t \left| Y(\hat{x}_f, t; \tau) - y(\tau) \right|^2 d\tau$$

or rather with finite memory

$$\hat{x}(t) = \underset{\hat{x}_f}{\arg\min} \int_{t-\bar{t}}^t \left| Y(\hat{x}_f, t; \tau) - y(\tau) \right|^2 d\tau,$$

where $Y(\hat{x}_f, t; \tau)$ denotes the output at time τ of the solution to (1) going through $\hat{x}_f$ at time t. Some methods have been developed to solve this optimization problem online, in spite of its non-convexity and the presence of local minima (see, e.g., Alamir, 2007 for a survey of existing algorithms). Such algorithms are often denoted as *finite-horizon observers* and the theory is usually developed in the discrete-time context.

In this paper, the path we choose to follow is rather to look for a dynamical system, called an *observer*, using the current value of the output $y(t)$ and whose state is guaranteed to provide (at least asymptotically) enough information to asymptotically reconstruct the state of (1).

2.3. Asymptotic observers

An asymptotic observer takes the form

$$\dot{\hat{z}} = \mathcal{F}(\hat{z}, y, t), \qquad \hat{x} = \mathcal{T}(\hat{z}, y, t), \tag{4}$$

with state $\hat{z} \in \mathbb{R}^{n_z}$ for $n_z \in \mathbb{N}_{>0}$ and maps $\mathcal{F} : \mathbb{R}^{n_z} \times \mathbb{R}^{n_y} \times \mathbb{R} \to \mathbb{R}^{n_z}$, $\mathcal{T} : \mathbb{R}^{n_z} \times \mathbb{R}^{n_y} \times \mathbb{R} \to \mathbb{R}^{n_x}$ chosen such that the following holds.

Definition 2.1. The system (4) is an *asymptotic observer* for the system (1) if there exists $\mathcal{Z}_0 \subset \mathbb{R}^{n_z}$ such that for any solution $t \mapsto x(t)$ to (1) defined on $[0, +\infty)$ with $x(0) \in \mathcal{X}_0$, any[1] solution $t \mapsto \hat{z}(t)$ to (4) with $\hat{z}(0) \in \mathcal{Z}_0$ and input $y(t) = h(x, t)$, is defined on $[0, +\infty)$ and verifies

$$\lim_{t \to +\infty} \left| \hat{x}(t) - x(t) \right| = 0 \tag{5}$$

with $\hat{x}(t) = \mathcal{T}(\hat{z}(t), y(t), t)$.

In other words, $\hat{x}(t)$ is an estimate of the current plant state and the error made with this estimation asymptotically converges to 0 as time goes to infinity.

In the next paragraph and in the rest of the paper, we will see that $\hat{z}$ typically estimates a certain function of the state $T(x, t)$, which is

injective with respect to x, and with $\mathcal{T}$ representing its left-inverse. In the favorable case where $\mathcal{T}$ is the identity or a mere projection from $\mathbb{R}^{n_z}$ to $\mathbb{R}^{n_x}$, namely $\hat{x}$ can directly be read from n_x components of $\hat{z}$, we say that the observer is *in the given coordinates*.

The time-dependence in the mappings $\mathcal{F}(\cdot, \cdot, t)$ and $\mathcal{T}(\cdot, \cdot, t)$ allows to take into account the time dependence of the functions f and h in (1). It may be removed for an autonomous system (2). In case of a system with input (3), we will see that this time dependency may appear through the value $u(t)$ at time t, or a certain number of time derivatives of $t \mapsto u(t)$ at time t, or more generally model a dependence on the whole past trajectory of u.

The output dependence in the mapping $\mathcal{T}(\cdot, y, \cdot)$ enables to cover the case where the knowledge of the output is used (explicitly or implicitly) to build the estimate $\hat{x}$ from the observer state $\hat{z}$. This sometimes allows to reduce the observer dimension, thus obtaining a *reduced-order observer*, see, e.g., Chen (1984). An example is the so-called *Immersion & Invariance* approach ("I&I") developed in Astolfi and Ortega (2003), Karagiannis and Astolfi (2005). A drawback of those approaches is that the estimate $\hat{x}$ then depends directly on y and is therefore more affected by measurement noise, while it is typically filtered through $\mathcal{F}$ in $\hat{z}$. We remark also that, under some mild conditions, the equivalence between the existence of a full-state observer and a reduced-order observer is well established, see, e.g., Sanfelice and Praly (2015), Shim and Praly (2003). In other words, when one obtains a full-state observer, then a reduced-order observer may be derived, and *vice-versa*.

Finally, we highlight that this article considers full-state observers in which the whole state x needs to be reconstructed from $\hat{z}$. When only a certain part/function of the state needs to be estimated, one may consider the so-called *functional observers*, see, e.g., Jennings, Fernando, and Trinh (2011), Kravaris (2016).

2.4. Desired observer properties

Depending on the specifications, different additional properties can be imposed on the observer.

Exponential observer: If system (4) is an asymptotic observer and Eq. (5) is replaced by

$$\left| \hat{x}(t) - x(t) \right| \le c\left(\hat{z}(0), x(0) \right) \exp(-\lambda t) \tag{6}$$

where c is a continuous positive function which estimates the overshoot of the observer and λ is the convergence rate. Note that $c\left(\hat{z}(0), x(0) \right)$ may not always be written as a function of $|\hat{x}(0) - x(0)|$, namely there is not necessarily stability/uniformity with respect to the initial estimation error.

Uniformly asymptotically stable observer: If system (4) is an asymptotic observer and Eq. (5) is replaced by

$$\left| \hat{x}(t) - x(t) \right| \le \beta\left(|\hat{x}(0) - x(0)|, t \right) \tag{7}$$

where β is a $\mathcal{KL}$-map. The observer is said to be **uniformly exponentially stable** if the class $\mathcal{KL}$-map is $\beta(r, t) = ar \exp(-\lambda t)$ for some positive real numbers a and λ. Besides the attractivity described in (5), property (7) requires stability of the estimation error, namely the fact that for any $\varepsilon > 0$, there exists $\delta > 0$ such that

$$\left| \hat{x}(0) - x(0) \right| \le \delta \implies \left| \hat{x}(t) - x(t) \right| \le \varepsilon \quad \forall t \ge 0,$$

namely the estimation error remains arbitrarily small during the transient if the initial error is sufficiently small. Actually, (7) also requires the attractivity and stability to be uniform for any initial condition in $(x(0), \hat{z}(0)) \in \mathcal{X}_0 \times \mathcal{Z}_0$. This property is quite strong and may be relaxed in different ways depending on the context. In any case, remark that the stability imposes that the 0-error set $x = \hat{x}$ is invariant by the dynamics. We will see that such stability properties may sometimes hold in certain well-chosen coordinates, and not in the initial x-coordinates.

---

[1] We say "any solution" because $\mathcal{F}$ may only be continuous, or even set-valued, and may admit several solutions. This is not a problem as long as any such solution verifies the required convergence property.

Finite time observer: If system (4) is an asymptotic observer and Eq. (5) is replaced by

$$\left| \hat{x}(t) - x(t) \right| = 0, \qquad \forall t \geq t_c(\hat{z}(0), x(0)), \tag{8}$$

where t_c is the convergence time depending on the initialization of the system and observer in $\mathcal{X}_0 \times \mathcal{Z}_0$.

Uniform finite time observer: If system (4) is a finite time observer and t_c is the convergence time uniform for all $(x(0), \hat{z}(0)) \in \mathcal{X}_0 \times \mathcal{Z}_0$.

Tunable observers: If we know a family of observers able to provide an arbitrarily small estimation error in an arbitrarily short time. More precisely, it implies that there exists a set of asymptotic observers such that for all $\epsilon > 0$ and for all $t_d > 0$, one of those observers (in this set) $(\mathcal{F}, \mathcal{T})$ ensures that any solution $t \mapsto (x(t), \hat{z}(t))$ to (1), (4) defined on $[0, +\infty)$ with $(x(0), \hat{z}(0)) \in \mathcal{X}_0 \times \mathcal{Z}_0$ verifies

$$\left| \hat{x}(t) - x(t) \right| \leq \epsilon, \qquad \forall t \geq t_d, \tag{9}$$

Robust observer: If system (4) is an asymptotic observer and admits an asymptotic gain in presence of disturbances on the plant dynamics (1). More precisely, there exist $\bar{v} \in \mathbb{R}_{\geq 0} \cup \{+\infty\}$ and a $\mathcal{K}$-map ρ such that, for any measurable disturbance $v = (v_x, v_y) : \mathbb{R} \to \mathbb{R}^{n_x} \times \mathbb{R}^{n_y}$ such that $|v(t)| \leq \bar{v}$ for almost all $t \geq 0$, and for any solution $t \mapsto x(t)$ to

$$\dot{x} = f(x, t) + v_x, \qquad y = h(x, t) + v_y,$$

defined on $[0, +\infty)$ and with $x(0) \in \mathcal{X}_0$, the corresponding solution $t \mapsto \hat{z}(t)$ to (4), with $\hat{z}(0) \in \mathcal{Z}_0$ and input y, is defined on $[0, +\infty)$ and verifies

$$\limsup_{t \to +\infty} \left| \hat{x}(t) - x(t) \right| \leq \rho \left(\limsup_{t \to +\infty} |v(t)| \right). \tag{10}$$

Again, more or less restrictive robustness properties may be stated depending on the context, such as non-uniformity of ρ with respect to initial conditions, or input-to-state-stability (ISS) conditions (see, e.g., Astolfi, Jungers, & Zaccarian, 2021; Shim & Liberzon, 2016; Sontag, 2008). Note that a uniformly asymptotically stable observer is a robust observer.

All in all, the role of an observer is to estimate in real time the plant state based on the knowledge of its output. This means that this signal somehow contains enough information to determine uniquely the whole state of the system. This brings us to the notions of observability or detectability, which characterize the necessary conditions that the plant must verify in order for an observer to exist. We will see in Section 3 that the strength of those conditions depend on the convergence/stability/robustness properties that are required from the observer.

2.5. Observers for linear autonomous systems

Before treating the observer problem for general nonlinear systems, we start by recalling how the problem is solved for linear autonomous systems of the form

$$\dot{x} = Ax + Bu, \qquad y = Cx \tag{11}$$

as a way of introducing the challenges appearing with nonlinearities and time dependence.

First, a necessary condition for the existence of an observer is that the output y should contain enough information to determine uniquely the trajectory at least asymptotically. This means that any pair of solutions $t \mapsto x_a(t)$ and $t \mapsto x_b(t)$ giving the same output $y = Cx_a = Cx_b$, i.e., indistinguishable from the output, should at least converge

to each other asymptotically. This property is called *detectability* and its necessity extends to the nonlinear context as will be detailed in Section 3. If there does not even exist such indistinguishable pairs of solutions, namely each pair of distinct solutions $t \mapsto x_a(t)$ and $t \mapsto x_b(t)$ produce two different outputs $y_a = Cx_a \neq y_b = Cx_b$, we speak of *observability*.

A peculiarity of linearity is that for any indistinguishable pair of solutions, the difference $\delta_x = x_a - x_b$ is solution to

$$\dot{\delta}_x = A\delta_x, \qquad C\delta_x = 0 \tag{12}$$

meaning that observability/detectability properties can be characterized considering only trajectories with output constantly equal to 0 and independently from the input u. This is no longer true for nonlinear systems where incremental properties need to be considered and some inputs may destroy observability. The class of *uniformly observable* systems, i.e., observable for any input, will be studied in Section 6 with the *triangular forms*.

Detectability of (11) thus means that any solution to (12) asymptotically converges to 0. This is equivalent to the so-called Hautus test

$$\text{rank}\left(\begin{bmatrix} \lambda \,\text{Id} - A \\ C \end{bmatrix} \right) = n_x \qquad \forall \lambda \in \mathbb{C} \ : \ \Re(\lambda) \geq 0 \tag{13}$$

which means that any eigenvector of A corresponding to a nonnegative eigenvalue cannot be in the kernel of C. This necessary condition is actually also sufficient to design an asymptotic observer for (11). Indeed, one can show that there exists $L \in \mathbb{R}^{n_x \times n_y}$ such that $A - LC$ is Hurwitz and an asymptotic observer is therefore given by

$$\dot{\hat{x}} = A\hat{x} + Bu + L(y - C\hat{x}), \tag{14}$$

whose estimation error follows the Hurwitz dynamics

$$\frac{d}{dt}(\hat{x} - x) = (A - LC)(\hat{x} - x)$$

and asymptotically converges to 0 for any initial conditions $x(0)$ and $\hat{x}(0)$. More precisely, the observer is uniformly exponentially stable, with a decreasing rate λ related to the eigenvalues of $A - LC$. The fact that a global observer of dimension n_x can be designed under a mere detectability condition and in the form (14) actually relies on very stringent metric constraints that happen to be verified by linear systems as will be seen in Section 4. On the other hand, when such constraints are not verified, nonlinear observers typically require a prior change of coordinates, transforming the dynamics into specific forms possibly of larger dimension $n_z \geq n_x$ where an observer can be written. Besides, the existence/invertibility of such a transformation typically relies on stronger observability conditions and is not always globally valid.

Beyond detectability, system (11) is actually observable if it does not admit any indistinguishable pairs, i.e., if the only solution to (12) is the constant solution $\delta_x = 0$. One can show in this case that the eigenvalues of $A - LC$ can be arbitrarily assigned and the observer is thus tunable. Actually, for linear systems, observability is equivalent to *instantaneous observability* or even *differential observability*, namely the fact that the knowledge of the output and its derivatives at a single time is enough to determine the state. Indeed, having $C\delta_x(t) = 0$ on an arbitrarily small interval $[t_0, t_0 + \epsilon)$ is equivalent to having it for all times, and to having $\delta_x(t_0)$ in the kernel of the observability matrix

$$\mathcal{O} := \text{col}[C, CA, \ldots, CA^{n_x - 1}]. \tag{15}$$

Hence observability for linear systems is equivalent to $\mathcal{O}$ being full rank, which is also equivalent to the Hautus test (13) being valid for any $\lambda \in \mathbb{C}$ (see Chen, 1984, Theorem 6.O.1). For nonlinear systems, observability, instantaneous observability and differential observability are no longer equivalent but we will see in Section 3 that the link between tunability of the observer and instantaneous observability remains. Besides, the property of differential observability can be extended and exploited in the so-called high-gain observer designs presented in Section 6.

By solving (12), another equivalent way of seeing observability is to say that for any t_0 and any $\epsilon > 0$,

$$C \exp(At)\delta_x(t_0) = 0 \quad \forall t \in [t_0, t_0 + \epsilon] \quad \implies \quad \delta_x(t_0) = 0,$$

and therefore, integrating $\|C \exp(At)\delta_x(t_0)\|$ in time,

$$\delta_x(t_0)^\top \left[\int_{t_0}^{t_0+\epsilon} \exp(At)^\top C^\top C \exp(At)dt \right] \delta_x(t_0) = 0$$
$$\implies \quad \delta_x(t_0) = 0.$$

Observability is thus equivalent to the invertibility of the so-called *observability Gramian* (see Chen, 1984, Theorem 6.O.1). The advantage of such a characterization is that it extends to time-varying linear systems, i.e., with $A(t)$ and $C(t)$, as long as $\exp(At)$ is replaced by the corresponding transition matrix. This will be detailed in Section 5.

Finally, it is interesting to know that D. Luenberger initially obtained in Luenberger (1964) the observer (14) for observable autonomous linear systems by linearly transforming (11) into an Hurwitz form with output injection and showing the invertibility of this transformation thanks to observability. Unlike the direct design of (14), this method can actually be extended to nonlinear systems leading to the so-called *nonlinear Luenberger* or *Kazantzis–Kravaris/Luenberger* (KKL) observers presented in Section 7.

3. Necessary conditions and a general sufficient condition for observer design

3.1. Necessary conditions for asymptotic observers

3.1.1. Detectability notions

As explained in Section 2.5, a *detectability* property is necessary for the existence of an observer. See also Chen (1984), Luenberger (1964) for linear systems. This property is defined for general systems as follows.

Definition 3.1. The system (1) is *asymptotically detectable* if any pair of solutions $t \mapsto x_a(t)$ and $t \mapsto x_b(t)$ to (1) initialized in $\mathcal{X}_0$ and defined on $[0, +\infty)$ such that

$$h(x_a(t), t) = h(x_b(t), t), \qquad \forall t \geq 0, \tag{16}$$

verifies

$$\lim_{t \to \infty} \left| x_a(t) - x_b(t) \right| = 0. \tag{17}$$

The property of asymptotic detectability says that when two solutions are not distinguishable from the output, they necessarily converge to each other asymptotically. This ensures one obtains a "good" asymptotic estimate no matter which initial condition we pick. The following result, which can be found for instance in Andrieu, Besançon, and Serres (2013), shows that *asymptotic detectability* is necessary for the existence of an asymptotic observer.

Lemma 3.1. *If there exists an asymptotic observer for the system* (1)*, then the system* (1) *is asymptotically detectable.*

Proof. This result follows from the fact that a solution $\hat{z}$ of (4) with input $h(x_b, t)$ is also a solution with input $h(x_a, t)$, and $\hat{x}(t) = \mathcal{T}(\hat{z}(t), h(x_a(t), t), t) = \mathcal{T}(\hat{z}(t), h(x_b(t), t), t)$ must thus converge both to x_a and x_b, which implies (17) by triangular inequality. □

We could see in the same way that a finite time observer requires a finite time detectability, and that an asymptotically stable observer requires an asymptotically stable detectability, in the sense of the distance between two solutions with same output. On the other hand, a so-called *differential detectability* has been introduced in Sanfelice and Praly (2011a) (see also Tsinias, 1989 for a constant metric case) and can be given (in its stronger version) as the following.

Definition 3.2. The system (1) is *differentially detectable* if there exists three positive real numbers $0 < \underline{p} \leq \bar{p}$ and q and a C^1 matrix function $P : \mathbb{R}^{n_x} \mapsto \mathbb{R}^{n_x \times n_x}$ such that

$$\underline{p}\, I \leq P(x) \leq \bar{p}\, I, \qquad \forall x \in \mathbb{R}^{n_x}, \tag{18a}$$

and[2]

$$\delta_x^\top L_f P(x,t)\delta_x \leq -q\delta_x^\top P(x)\delta_x,$$
$$\forall (x, \delta_x, t) \in \mathcal{X} \times \mathbb{R}^{n_x} \times \mathbb{R}_{\geq 0} \quad \text{such that} \quad \frac{\partial h}{\partial x}(x,t)\delta_x = 0. \tag{18b}$$

As studied in Sanfelice and Praly (2011a), this property establishes that the vector field f is contracting for a Riemannian distance associated to the quadratic form $\delta_x \mapsto \delta_x P(x)\delta_x$ in the direction tangent to the level set of h. Picking $\delta_x \mapsto \delta_x^\top P(x)\delta_x$ as a Lyapunov function, it establishes that complete solutions to

$$\dot{x} = f(x, t), \qquad \dot{\delta}_x = \frac{\partial f}{\partial x}(x, t)\, \delta_x, \tag{19}$$

initialized in $\mathcal{X}_0 \times \mathbb{R}^{n_x}$ and verifying

$$\frac{\partial h}{\partial x}(x(t), t)\, \delta_x(t) = 0, \qquad \forall t \geq 0, \tag{20}$$

satisfy

$$\lim_{t \to +\infty} |\delta_x(t)| = 0. \tag{21}$$

It is shown in Andrieu et al. (2013), Andrieu, Jayawardhana, and Praly (2016) that this property is linked to the existence of an observer giving exponential stability of some error set. Actually, for autonomous systems (2) with $\mathcal{X} = \mathcal{X}_0 = \mathbb{R}^{n_x}$, differential detectability is necessary for the existence of an exponentially stable observer of the form

$$\dot{\hat{z}} = \mathcal{F}(\hat{z}, y, t) = f(\hat{z}) + k(\hat{z}, y), \qquad \hat{x} = \mathcal{T}(\hat{z}, y, t) = \hat{z}, \tag{22}$$

made of a copy of the plant dynamics and an extra correction term.

Lemma 3.2 (*Andrieu, Jayawardhana, & Praly, 2016*). *If there exists a uniformly exponentially stable observer for the system* (2) *in the form* (22) *with* $\mathcal{Z}_0 = \mathbb{R}^{n_x}$ *and with* (f, h, k) *having bounded first and second order derivatives then the system* (1) *is differentially detectable.*

3.1.2. Observability notions

As recalled in Section 2.5 for linear systems, the ability to design an observer with arbitrary pole placement imposes an instantaneous observability property. Various notions of observability can be defined for nonlinear systems. In its weakest version, "observability" means that (16) must actually imply $x_a = x_b$ rather than asymptotic convergence in (17), and "instantaneous observability" means that having (16) on an arbitrarily small interval is enough to imply the equality.

Definition 3.3. The system (1) is *instantaneously observable* if for all $t_d > 0$, any pair of solutions $t \mapsto x_a(t)$ and $t \mapsto x_b(t)$ to (1) initialized in $\mathcal{X}_0$ and defined on $[0, +\infty)$ such that

$$h(x_a(t), t) = h(x_b(t), t), \qquad \forall 0 \leq t < t_d, \tag{23}$$

verifies $x_a(\cdot) = x_b(\cdot)$.

The following result can be found in Andrieu et al. (2013) for autonomous systems but extends readily to time-varying systems.

---

[2] Given a C^1 matrix function $P : \mathbb{R}^{n_x} \mapsto \mathbb{R}^{n_x \times n_x}$, and a C^1 vector field $f : \mathbb{R}^{n_x} \times \mathbb{R} \to \mathbb{R}^{n_x}$, $L_f P$ denotes the Lie derivative in the direction of f of the quadratic form P, i.e.

$$L_f P(x, t) = \lim_{h \to 0} \frac{(I + h\frac{\partial f}{\partial x}(x, t))^\top P(x + hf(x, t + h))(I + h\frac{\partial f}{\partial x}(x, t)) - P(x)}{h},$$

with coordinates $(L_f P(x, t))_{i,j} = \sum_k \left[2P_{ik}\frac{\partial f_k}{\partial x_j}(x, t) + \frac{\partial P_{ij}}{\partial x_k}(x)f_k(x, t) \right].$

Lemma 3.3. *If there exists a tunable observer for the system* (1), *then the system* (1) *is instantaneously observable.*

In the particular case where f and h are analytical, the output y is an analytical function of time and the notions of observability and instantaneous observability are actually equivalent because two analytical functions which are equal on some interval are necessarily equal on their maximal interval of definition. Besides, for any initial condition, there exists t_d such that

$$y(t) = \sum_{k=0}^{+\infty} \frac{y^{(k)}(0)}{k!} t^k, \qquad \forall t \in [0, t_d),$$

and observability is thus closely related to the notion of differential observability defined in Section 6, which roughly says that the state is uniquely determined by the value of the output and of its derivatives (up to a certain order).

Actually, in observer design, one is more interested in estimating the current state $x(t)$, based on the past measurements, than estimating its initial condition. Therefore, notions of *backward distinguishability* (sometimes also called *constructibility* or *determinability*) typically appear as in Section 7, requiring that the current state be uniquely determined by the past outputs. Both are equivalent of course when solutions are unique, but not generally.

3.1.3. Structural property on the observer

To summarize, the following implication can be obtained:

Observer	$\Rightarrow$ Detectability.
Unif. exp. stable observer	$\Rightarrow$ Differential detectability.
Tunable observer	$\Rightarrow$ Instantaneous observability.

In this survey, we will thus classify the observer designs available in the literature depending on the type of observability properties they require of the plant.

In order to justify the common spirit in which they are developed, it is interesting to recall another necessary condition from Andrieu et al. (2013). Indeed, for an autonomous system (2), if there exists an asymptotic observer (4) (without time-dependence) for (2) and a compact subset of $\mathbb{R}^{n_x} \times \mathbb{R}^{n_z}$ which is invariant by the dynamics $(f, \mathcal{F})$, then there exist compact subsets C_x of $\mathbb{R}^{n_x}$ and C_z of $\mathbb{R}^{n_z}$, and a closed set-valued map T defined on C_x such that the set

$$\mathcal{E} = \left\{ (x, \hat{z}) \in C_x \times C_z \; : \; \hat{z} \in T(x) \right\} \tag{24}$$

is invariant, attractive, and verifies

$$\forall (x, \hat{z}) \in \mathcal{E}, \quad \mathcal{T}(\hat{z}, h(x)) = x. \tag{25}$$

In other words, the pair made of the plant state x (following the dynamics f) and the observer state $\hat{z}$ (following the dynamics $\mathcal{F}$) necessarily converges to the graph of some set-valued map T, and $\mathcal{T}$ is a left-inverse of this mapping. Note that this injectivity is of a peculiar kind since it is conditional to the knowledge of the output, namely "$x \mapsto T(x)$ is injective knowing $h(x)$". This result justifies the usual methodology of observer design for autonomous systems which consists in transforming, via a function T, the system into a form for which an observer is available, then designing the observer in those new coordinates z (i.e., find $\mathcal{F}$), and finally deducing an estimate in the original coordinates via inversion of T (i.e., find $\mathcal{T}$). Actually, it follows that some observer properties such as stability or exponential convergence are obtained in the z-coordinates, namely on the error $\hat{z} - T(x)$, but are not necessarily preserved in the x-coordinates through the left-inversion of T, unless T satisfies appropriate extra conditions (uniform injectivity, diffeomorphism, etc.). However, this is generally acceptable since the crucial properties such as attractivity and robust asymptotic gain are typically preserved in the x-coordinates.

Note that in most approaches, the observer is designed from a (maybe time-varying) single-valued map T because it is simpler to manipulate than a set-valued map. In the designs based on some

contraction property presented in Section 4, T is simply taken as the identity. For the high-gain observers/differentiators presented in Section 6, $T(x, t)$ is obtained from the successive Lie derivatives of the output map h. In the case of the KKL observers presented in Section 7, T is selected to transform the plant dynamics (1) into a particular Hurwitz form, namely as a solution to a PDE

$$\frac{\partial T}{\partial x}(x, t) f(x, t) + \frac{\partial T}{\partial t}(x, t) = \Lambda T(x, t) + \Gamma h(x, t)$$

where Λ is a Hurwitz matrix.

3.2. Sufficient condition for observer design

Inspired by the previous section, we introduce the following sufficient condition on which all the designs presented in this paper are based on. See Bernard (2019, Theorem 1.1) for its proof.

Theorem 3.1. *Assume there exist an integer n_z, a C^1 map $T : \mathbb{R}^{n_x} \times \mathbb{R} \to \mathbb{R}^{n_z}$, and continuous maps $F : \mathbb{R}^{n_z} \times \mathbb{R}^{n_y} \times \mathbb{R} \to \mathbb{R}^{n_z}$, $H : \mathbb{R}^{n_z} \times \mathbb{R} \to \mathbb{R}^{n_y}$ and $\mathcal{F} : \mathbb{R}^{n_z} \times \mathbb{R}^{n_y} \times \mathbb{R} \to \mathbb{R}^{n_z}$ such that the following holds.*

(a) *T transforms the plant dynamics* (1) *into[3]*

$$\dot{z} = F(z, H(z, t), t), \qquad y = H(z, t), \tag{26}$$

i.e. for all x in $\mathcal{X}$ and all $t \in [0, +\infty)$,

$$\begin{aligned} \frac{\partial T}{\partial x}(x, t) f(x, t) + \frac{\partial T}{\partial t}(x, t) &= F(T(x, t), h(x, t), t), \\ h(x, t) &= H(T(x, t), t). \end{aligned} \tag{27}$$

(b) *$x \mapsto T(x, t)$ becomes uniformly injective after a certain time $\bar{t} \geq 0$, namely there exists a concave $\mathcal{K}$ function ρ such that for all (x_a, x_b) in $\mathcal{X} \times \mathcal{X}$ and all $t \geq \bar{t}$,*

$$|x_a - x_b| \leq \rho\left(|T(x_a, t) - T(x_b, t)| \right). \tag{28}$$

(c) *The system*

$$\dot{\hat{z}} = \mathcal{F}(\hat{z}, y, t) \tag{29}$$

is an asymptotic observer for (26).

Then, there exists a map $\mathcal{T} : \mathbb{R}^{n_z} \times \mathbb{R} \to \mathbb{R}^{n_x}$ such that for all $t \geq \bar{t}$, $z \mapsto \mathcal{T}(z, t)$ is uniformly continuous[4] and such that

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

Besides, (4) *defined with the maps $(\mathcal{F}, \mathcal{T})$ is an asymptotic observer for system* (1).

This result formalizes the design methodology presented in the previous section which consists in finding a (maybe time varying) uniformly injective change of coordinates $z = T(x, t)$ transforming (1) into a normal form (26), then designing an observer (29) for (26), and finally find a left-inverse of $x \mapsto T(x, t)$ to recover an estimate in the initial x-coordinates.

Remark 3.1. Without the assumption of concavity of ρ in (28), it is still possible to show that $x \mapsto T(x, t)$ admits a continuous left-inverse $\mathcal{T}$ defined on $\mathbb{R}^{n_z}$. But, as shown in Shim and Liberzon (2016, Example 4), continuity of $\mathcal{T}$ is not enough to deduce the convergence of $\hat{x}$ from that of $\hat{z}$: uniform continuity is necessary. Note that if $\mathcal{X}$ is bounded, the concavity of ρ is no longer a constraint, since a concave upper-approximation can always be obtained by saturation of ρ (see McShane, 1934 for more details).

---

[3] The expression of the dynamics under the form $F(z, H(z, t), t)$ can appear strange and abusive at this point because it is highly non unique and we should rather write $F(z)$. However, we will see how specific structures of dynamics $F(z, y, t)$ allow the design of an observer (29).

[4] A function γ is uniformly continuous if and only if $\lim_{n \to +\infty} |x_n - x'_n| = 0$ implies $\lim_{n \to +\infty} |\gamma(x_n) - \gamma(x'_n)| = 0$.

Remark 3.2. By imposing stronger constraints on the mapping T and on the observer (29) for (26), the properties that have been listed previously may be stated for the observer (4) obtained for system (1).

- If the observer (29) for (26) is *uniformly asymptotically stable* then the observer (4) for (1) is *robust*.
- If the injectivity property (28) is strengthened into

$$|x_a - x_b| \leq c|T(x_a, t) - T(x_b, t)|$$

for some positive real number c and if (29) is an exponential observer for (26) then the observer (4) is an *exponential* observer for (1).

- If a uniform continuity condition on T is added to the injectivity property (28), namely

$$\underline{\rho}\left(|T(x_a, t) - T(x_b, t)|\right) \leq |x_a - x_b| \leq \overline{\rho}\left(|T(x_a, t) - T(x_b, t)|\right)$$

for some concave $\mathcal{K}$-maps $(\underline{\rho}, \overline{\rho})$ and (29) is a stable (resp. uniformly asymptotically stable) observer for (26) then the observer (4) is a *stable* (resp. uniformly asymptotically stable) observer for (1) when the observer initialization set verifies $\mathcal{Z}_0 \subset T(\mathcal{X})$.

In this review paper, we attempt to make a list of the normal forms (26) available in the literature and their associated observer (29). We classify them depending of the kind of observability properties they require and the corresponding class of systems (1) that can be transformed into each form. All this is summed up in Table 1.

4. Observers from detectability conditions

In this section, observer designs based on the differential detectability condition are considered. Such observers are typically asymptotically stable and written directly in the given coordinates in the form (22), where the correction term k is such that $k(x, h(x, t), t) = 0$. We thus rewrite it directly in a more conventional way

$$\dot{\hat{x}} = f(\hat{x}, t) + k(\hat{x}, y, t), \tag{30}$$

or sometimes

$$\dot{\hat{x}} = F(\hat{x}, y, t) + k(\hat{x}, y, t), \tag{31}$$

when $F(x, h(x), t) = f(x, t)$ and we employ an output injection directly in observer dynamics. The whole question is now to find means of designing the correction term k to ensure the observer convergence. This is the purpose of this section.

4.1. Finsler-like relaxation of differential detectability

To understand how the differential detectability property may be employed to design an observer, it is enlightening to consider the case in which system (1) is linear, i.e., is in the form

$$\dot{x} = Ax, \quad y = Cx,$$

with A and C of appropriate dimension. The differential detectability property (see Definition 3.2) implies (when considering constant metric without loss of generality) the existence of a positive definite matrix P in $\mathbb{R}^{n_x \times n_x}$ and a positive real number q such that, for any positive definite R in $\mathbb{R}^{n_y \times n_y}$,

$$\delta_x^\top [PA + A^\top P]\delta_x \leq -q\,\delta_x^\top P\delta_x ,$$

$$\forall \delta_x \in \mathbb{R}^{n_x} \text{ such that } \delta_x^\top C^\top RC\delta_x = 0 . \tag{32}$$

By Finsler's lemma, given R, this implication is equivalent to the existence of a positive real number k such that

$$PA + A^\top P - kC^\top RC < 0 . \tag{33}$$

In other words, the matrix $A - kP^{-1}C^\top RC$ is stable and consequently

$$\dot{\hat{x}} = A\hat{x} + kP^{-1}C^\top R(y - C\hat{x})$$

defines an asymptotic observer. We recover the well-known Luenberger observer for detectable linear systems. In this section, our aim is to follow the same route to obtain converging observers in the nonlinear context.

4.2. A local observer assuming strong differential detectability

For nonlinear systems, the differential detectability property also implies the existence of an observer. However, without further assumptions, the obtained observer is local. Indeed, going from the differential detectability property, i.e., the implication (18b), towards a property of the form (33) is possible but with a parameter which is state and time dependent. More precisely, assuming (18b), it is possible to show the existence of a continuous function $\rho : \mathbb{R}^{n_x} \times \mathbb{R}_{\geq 0} \mapsto \mathbb{R}_{\geq 0}$ such that

$$L_f P(x, t) \leq \rho(x, t)\frac{\partial h}{\partial x}(x, t)^\top \frac{\partial h}{\partial x}(x, t) - qP(x) ,$$

$$\forall (x, t) \in \mathcal{X} \times \mathbb{R}_{\geq 0} . \tag{34}$$

When particularizing the result of Sanfelice and Praly (2015) to the case in which there is only one output, the following result may be obtained for strongly differentially detectable autonomous system in the form (2).

Theorem 4.1. *Consider an autonomous system in the form (2) with $n_y = 1$. Assume that the couple (f, h) is differentially detectable with $\mathcal{X} = \mathcal{X}_0 = \mathbb{R}^{n_x}$ and assume that the function ρ in (34) is time-independent and there exists κ such that $\kappa \geq \rho(x)$ for all x. Assume moreover that h has bounded first and second order derivatives. Then, (30) with*

$$k(\hat{x}, y) = \kappa P(\hat{x})^{-1}\frac{\partial h}{\partial x}(\hat{x})^\top (y - h(\hat{x})) , \tag{35}$$

is a local observer in the sense that there exists c_0 such that for all solutions defined on $[0, +\infty)$, if $|\hat{x}(0) - x(0)| \leq c_0$ then (5) is satisfied.

Differential detectability thus only provides a local observer, i.e., the initial error needs to be sufficiently small to ensure the observer convergence. In other words, it is not an asymptotic observer in the sense of Definition 2.1.

Remark 4.1. Note also that the result in Sanfelice and Praly (2015) is more general since a multidimensional output is considered. Moreover, the case in which $\mathcal{X}$ is only a subset of $\mathbb{R}^{n_x}$ is considered, requiring a convexity property on $\mathcal{X}$. Moreover, the upper bounds on P and first and second derivatives of h are replaced by a boundedness assumption on the Riemannian hessian matrix of h.

4.3. Toward regional observer assuming convexity of the output map

In Sanfelice and Praly (2011a), a further assumption is made allowing to obtain a semi-global observer. In order to introduce this one, notice that a C^1 function P which satisfies (18a) defines a complete Riemannian metric on $\mathbb{R}^{n_x}$ for which we can define geodesics[5] (straight line for Euclidean metrics).

Definition 4.1. *Given y in $\mathbb{R}$, the set $\{x \in \mathbb{R}^{n_x} : h(x) = y\}$ is said to be totally geodesic if any geodesic γ satisfying*

$$h(\gamma(0)) = y, \qquad \frac{dh\circ\gamma}{ds}(0) = 0,$$

satisfies $h(\gamma(s)) = y$ for all s in the maximal time domain of definition of γ.

---

[5] When P is C^2, geodesics are curves in $\mathbb{R}^{n_x}$ γ defined on an open time domain of $\mathbb{R}$ and solution to the geodesic equation which in coordinates reads

$$\frac{d}{ds}\left(\frac{d\gamma}{ds}(s)^\top P(\gamma(s))\right) = \frac{1}{2}\frac{\partial}{\partial x}\left(\frac{d\gamma}{ds}(s)^\top P(x)\frac{d\gamma}{ds}(s)\right)\Bigg|_{x=\gamma(s)} . \tag{36}$$

This assumption is related to a geodesic convexity property of these level sets (see Sanfelice & Praly, 2011a). It is satisfied if for instance the output map is linear and P is constant (see next subsection). With this assumption, the following result is obtained in Sanfelice and Praly (2011a).

Theorem 4.2. *Consider an autonomous system in the form* (2) *with* $n_y = 1$. *Assume that* $\mathcal{X}$ *is bounded. Assume moreover that the pair* (f, h) *is differentially detectable and that the level sets of* h *are totally geodesic* (4). *Then, picking* $\mathcal{Z}_0 = \mathcal{X}_0$, *there exists a positive real number* κ *such that* (30) *with* (35) *defines a uniformly asymptotically stable observer.*

Remark 4.2. Again, the result in Sanfelice and Praly (2011a) is more general since a multidimensional output is considered and $\mathcal{X}$ is *closed weakly geodesically* and not necessarily bounded. However the result is semi-global in terms of the initial error: κ has to be taken sufficiently large depending on the initial estimation error.

4.4. The case of an Euclidean metric describing detectability

A case which has been deeply studied in the literature is the case in which the output map is linear

$$h(x) = Cx \,, \tag{37}$$

for some matrix C of appropriate dimension. In that case, to apply the former results, it suffices to find a constant matrix P in $\mathbb{R}^{n_x \times n_x}$ such that the system is differentially detectable with respect to this matrix P, i.e. such that

$$\delta_x^\top \left[P \frac{\partial f}{\partial x}(x, t) + \frac{\partial f}{\partial x}(x, t)^\top P \right] \delta_x \le -q \delta_x^\top P \delta_x \,,$$

$$\forall (x, \delta_x) \in \mathbb{R}^{n_x} \times \mathbb{R}^{n_x} \text{ such that } \delta_x^\top C^\top C \delta_x = 0 \,. \tag{38}$$

Indeed, in that case, since P is constant (Euclidean case), the level sets of $y = Cx$ are totally geodesic. Consequently, the former theorem applies and an observer may be obtained for a bounded $\mathcal{X}$. This result was previously published in the autonomous case in Praly (2001). Assuming furthermore a global Lipschitz property (uniformly with respect to time) on f, it is possible to show that there exists κ such that

$$P \frac{\partial f}{\partial x}(x, t) + \frac{\partial f}{\partial x}(x, t)^\top P - \kappa C^\top C \le -q P \,, \tag{39}$$

and an observer may be obtained for $\mathcal{X} = \mathbb{R}^{n_x}$ as shown in Tsinias (1989). The observer obtained is in the form (30) with

$$k(\hat{x}, y, t) = \kappa P^{-1} C^\top (y - C \hat{x}) \,. \tag{40}$$

Hence, one of the means of designing an observer for nonlinear systems with linear measured output is to construct P such that (39) holds.

Note that in this case and as shown in Besançon (2000), a reduced-order observer may be constructed (see also Sanfelice & Praly, 2015 for a Riemannian version of this result).

4.5. Finding an Euclidean metric to obtain a contraction

With (38), it can be noticed that (30) with (40) satisfies

$$\text{He} \left\{ P \left[\frac{\partial f}{\partial x}(x, t) + \frac{\partial k}{\partial x}(x, y, t) \right] \right\} \le -q P \,, \qquad \forall (y, x, t) \,. \tag{41}$$

We recognize here the Lohmiller–Slotine condition given for instance in Lohmiller and Slotine (1998). It establishes that (30) defines an exponential contraction in the sense that any two trajectories of the observer are exponentially converging to each other. Actually, contraction analysis is a common approach to design the correction term k as summarized in the following theorem in which a time-varying metric is allowed.

Theorem 4.3. *Consider a (possibly time-varying or controlled) system in the form* (1) *which satisfies*

$$f(x, t) = F(x, h(x), t) \,, \tag{42a}$$

for some smooth function F. *Suppose there exist* C^1 *functions* $P : \mathbb{R} \mapsto \mathbb{R}^{n_x \times n_x}$ *and* $k : \mathbb{R}^{n_x} \times \mathbb{R}^{n_y} \times \mathbb{R} \mapsto \mathbb{R}^{n_x}$, *and positive real numbers* $(q, \underline{p}, \overline{p})$ *such that*

$$k(x, h(x), t) = 0 \,, \qquad \forall (x, t) \in \mathbb{R}^{n_x} \times \mathbb{R} \,, \tag{42b}$$

$$\dot{P}(t) + \text{He} \left\{ P(t) \left[\frac{\partial F}{\partial \hat{x}}(\hat{x}, y, t) + \frac{\partial k}{\partial \hat{x}}(\hat{x}, y, t) \right] \right\} \le -q P(t) \,,$$
$$\forall (y, \hat{x}, t) \in \mathbb{R}^{n_y} \times \mathbb{R}^{n_x} \times \mathbb{R} \,, \tag{42c}$$

$$\underline{p} \, I \le P(t) \le \overline{p} \, I \,, \qquad \forall t \in \mathbb{R} \,. \tag{42d}$$

Then (31) *defines a uniformly exponentially stable observer.*

Proof. Let $V(t) = \frac{1}{2}(x(t) - \hat{x}(t))^\top P(t)(x(t) - \hat{x}(t))$. Its time derivative is computed as

$$\dot{V}(t) = \frac{1}{2}(x - \hat{x})^\top \dot{P}(t)(x - \hat{x})$$
$$+ (x - \hat{x})^\top P(t) \left[F(x, y, t) - F(\hat{x}, y, t) - k(\hat{x}, y, t) \right] \,,$$

where the time-dependence of the variables $x, \hat{x}$ and y has been omitted for compactness. Furthermore, with (42b), we have for all (x, t) in $\mathbb{R}^{n_x} \times \mathbb{R}$

$$F(x, y, t) - F(\hat{x}, y, t) - k(\hat{x}, y, t)$$
$$= \left(\int_0^1 \frac{\partial F}{\partial x}(\hat{x} + s(x - \hat{x}), y, t) + \frac{\partial k}{\partial x}(x + s(\hat{x} - x), y, t) ds \right)$$
$$\times (x - \hat{x}) \,,$$

for $y = h(x)$. This gives, with (42c),

$$\dot{V}(t) \le -q V(t) \,,$$

from which, using (42d), the conclusion holds. $\quad\square$

Note that a similar proof of the former theorem can be found in Manchester (2018). Theorem 4.3 is the core of most of the publications related to observers, as shown in the forthcoming examples.

4.5.1. Designs based on LMI

In Arcak and Kokotović (2001), is considered the case in which the dynamical system (1) is in the form

$$f(x) = Ax + G\phi(\zeta, y, t) \,, \quad h(x) = Cx \,, \quad \zeta = Mx \,, \tag{43}$$

where (A, G, M, C) are matrices of appropriate dimensions and ϕ is a smooth function with ζ in $\mathbb{R}^{d_\zeta}$. The observers considered in Arcak and Kokotović (2001) take the form

$$\dot{\hat{x}} = A\hat{x} + K(y - C\hat{x}) + G\phi(\hat{\zeta}, y, t) \quad , \quad \hat{\zeta} = M\hat{x} + E(y - C\hat{x}) \,, \tag{44}$$

where E, K are matrices in $\mathbb{R}^{n_\zeta \times n_y}$ and $\mathbb{R}^{n_x \times n_y}$ respectively. We recognize an observer in the form (31). A sufficient condition to guarantee the convergence of the observer is to select the matrices E and K to ensure that the observer defines a uniform (with respect to (y, t)) contraction in the sense of Theorem 4.3. In other words, the goal is to find P, K, E and q satisfying

$$\text{He} \left\{ P \left[A - KC + G \frac{\partial \phi}{\partial \zeta}(\zeta, y, t)(M - EC) \right] \right\} \le -q P \,,$$
$$P > 0 \,, \ q > 0 \,, \tag{45}$$

for all (ζ, y, t). This is an infinite dimensional matrix inequality since it has to be satisfied for all (ζ, t, y). But imposing some constraints on φ allows to rephrase this infinite dimensional matrix inequality into a finite dimensional LMI.

For instance, in Rajamani (1998), under the globally Lipschitz condition

$$\left| G \frac{\partial \phi}{\partial \zeta}(\zeta, y, t) M \right| \le \overline{b} \,, \qquad \forall (\zeta, y, t) \,,$$

for some positive real number $\bar{b}$, a sufficient condition to obtain inequality (45) is to find $P > 0$ and R such that

$$\begin{bmatrix} A^\top P + PA - R^\top C - C^\top R + I & P \\ P & -\frac{1}{\bar{b}}I \end{bmatrix} \leq 0$$

with the gain of the observer (44) selected as $K = P^{-1}R^\top$ and $E = 0$. In this case, the resulting observer is uniformly exponentially stable. Based on different generalized globally Lipschitz conditions, other LMI-based designs can be obtained. We refer, for instance, to Hu (2006), Ibrir (2009), Raghavan and Hedrick (1994), Rajamani (1998), Thau (1973), Vijayaraghavan (2014), Zemouche, Boutayeb, and Bara (2008), Zemouche, Zhang, Mazenc, and Rajamani (2018). An early formulation can also be found in Kou, Elliott, and Tarn (1975).

Also, in Arcak and Kokotović (2001) is considered the case in which each component ϕ_i of the map ϕ satisfies

$$(\zeta_a - \zeta_b)\big[\phi_i(\zeta_a, y, t) - \phi_i(\zeta_b, y, t)\big] \geq 0, \qquad \forall (\zeta_a, \zeta_b, y, t).$$

Then, as shown in Arcak and Kokotović (2001) (see also Fan & Arcak, 2003), the inequality (45) is verified if there exist $P > 0$, E and $q > 0$ satisfying the following LMI

$$\begin{bmatrix} (A - KC)^\top P + P(A - KC) + qI & PG + (M - EC)^\top W \\ G^\top P + W(M - EC) & 0 \end{bmatrix} \leq 0,$$

for some (fixed) diagonal matrix $W > 0$. In such a case, the observer (44) with K, E satisfying the previous inequality is uniformly exponentially stable for system (1) with f, h of the form (43).

4.5.2. Structured nonlinearities

Finally, the conditions of Theorem 4.3 can be satisfied by imposing some structures on the nonlinearities. For instance, some approaches aim at finding coordinates in which φ depends only on the measured output. This is the so-called *linearization by output injection*, see Section 5.4.

A different approach is based on assuming that the nonlinearities satisfy an upper triangular structure, resulting in the so-called *high-gain observer* approach, see Section 6.

5. Observers from observability Gramian

In this section, we consider the class of state-affine normal forms

$$\dot{z} = A(y, t)z + \varphi(y, t), \qquad y = C(t)z, \tag{46}$$

where $A : \mathbb{R}^{n_y} \times \mathbb{R} \to \mathbb{R}^{n_z \times n_z}$, $\varphi : \mathbb{R}^{n_y} \times \mathbb{R} \to \mathbb{R}^{n_z}$ and $C : \mathbb{R} \to \mathbb{R}^{n_y \times n_z}$. Note that (46) is not really affine in z, since y depends on z, but it is written so in order to highlight the dependency on the known signal y that is available for observer design. That is why the process of transforming a system into (46) is called in the literature *linearization by output injection*.

5.1. Observability Gramian

As observed in Section 2.5, when analyzing the observability of a system of the form (46), one naturally comes across the notion of *observability Gramian*. See, e.g., Chen (1984).

Definition 5.1. For a given signal $t \mapsto y(t)$, the *observability Gramian* associated to (46) on an interval $[t_0, t_1] \subset [0, +\infty)$ is the positive symmetric matrix defined by

$$\mathcal{G}_y(t_0, t_1) = \int_{t_0}^{t_1} \Psi_y(\tau, t_0)^\top C(\tau)^\top C(\tau) \Psi_y(\tau, t_0) d\tau,$$

where Ψ_y is the *transition matrix* associated to the linear dynamics $\dot{\chi} = A(y, t)\chi$, namely the unique solution to

$$\frac{\partial \Psi_y}{\partial \tau}(\tau, t) = A(y(\tau), \tau)\Psi_y(\tau, t),$$
$$\Psi_y(t, t) = I.$$

Indeed, consider a pair of solutions z_a and z_b to system (46) having same output y for all $t \in [0, \bar{t}]$ for some $\bar{t} > 0$. Then, their difference $\delta_z = z_a - z_b$ verifies

$$\dot{\delta}_z = A(y, t)\delta_z, \qquad C(t)\delta_z = 0, \qquad \forall t \in [0, \bar{t}],$$

and thus

$$\delta_z(t) = \Psi_y(t, 0)\delta_z(0) = \Psi_y(t, 0)(z_a(0) - z_b(0)),$$
$$\|C(t)\delta_z(t)\|^2 = 0 \qquad \forall t \in [0, \bar{t}].$$

It follows that

$$(z_a(0) - z_b(0))^\top \mathcal{G}_y(0, \bar{t})(z_a(0) - z_b(0)) = 0.$$

This implies that $z_a(0) = z_b(0)$ and thus $z_a = z_b$, namely (46) is observable in time $\bar{t}$, if and only if $\mathcal{G}_y(0, \bar{t})$ is invertible. In other words, the invertibility of $\mathcal{G}_y(0, \bar{t})$ for each output trajectory $t \mapsto y(t)$ characterizes the observability of (46) in time $\bar{t}$.

When A is independent from y, so is the Gramian, and any observability property is valid for any initial condition. Otherwise, one must take care to require uniformity with respect to any possible output signal y.

Putting aside the dependency on y, several observability properties can be defined depending on the "amplitude", "uniformity", "persistence", etc of the Gramian invertibility. For instance, Kalman's well-known *uniform complete observability* Kalman and Bucy (1961) requires the existence of $\alpha_1, \alpha_2 > 0$ and $\bar{t} > 0$ such that

$$\alpha_1 I \leq \mathcal{G}(t - \bar{t}, t) \leq \alpha_2 I \qquad \forall t. \tag{47}$$

The existence of α_2 being guaranteed as long as A and C are bounded, this property is mainly about α_1, namely the uniformity of the Gramian invertibility over intervals of length $\bar{t}$. See Theorem 5.1 below. If the time dependency of A and C comes through an input u, then such observability properties depend on this particular input. For instance, an input is said to be *regularly persistent* if it guarantees the existence of α_1 at least after a certain time. See Bornard, Celle-Couenne, and Gilles (1995) for more details.

5.2. Kalman or Kalman-like observers

The most famous observer used for state-affine systems is *Kalman's and Bucy's observer* presented in Kalman and Bucy (1961) in a stochastic context, for linear time-varying systems with independent white-noise disturbances impacting the initial condition, the dynamics and the output. Their approach was to find an optimal estimate $\hat{z}(t)$ of $z(t)$ minimizing at each time t the conditional expectation $E\left(|\hat{z}(t) - z(t)|^2 \mid y_{[0,t]}\right)$ assuming known the covariance of each noise process.

On the other hand, parallel to the Kalman school, *Kalman-like observers* were introduced as an optimal solution to a deterministic optimization problem (Bornard, Couenne, & Celle, 1989; Hammouri & De Leon Morales, 1990). More precisely, at each time t, the estimate $\hat{z}(t)$ is chosen as $\hat{z}(t) = \Psi_y(t, 0)\hat{z}^O$, where

$$\hat{z}^O = \arg\min e^{-\lambda t}\|\hat{z}_0 - \hat{z}^O\|_{\Pi_0^{-1}}$$
$$+ \int_0^t e^{-\lambda(t-\tau)}\|y(\tau) - C(\tau)\Psi_y(\tau, 0)\hat{z}^O\|_{R^{-1}}, \tag{48}$$

for some positive matrices Π_0 and R.

In both cases, the miracle of linearity makes it possible to produce the optimal trajectory $\hat{z}(t)$ as the solution to an observer of the form

$$\dot{\hat{z}} = A(y, t)\hat{z} + \varphi(y, t) + \Pi C(t)^\top R^{-1}(t)(y - C(t)\hat{z}), \tag{49a}$$
$$\dot{\Pi} = A(y, t)\Pi + \Pi A(y, t)^\top - \Pi C(t)^\top R^{-1}(t)C(t)\Pi$$
$$+ \lambda \Pi + D(t)Q(t)D(t)^\top, \tag{49b}$$

or, equivalently, with $P^{-1} = \Pi$,

$$\dot{\hat{z}} = A(y, t)\hat{z} + \varphi(y, t) + P^{-1}C(t)^\top R^{-1}(t)(y - C(t)\hat{z}), \tag{50a}$$

$$\dot{P} = -PA(y,t) - A(y,t)^\top P + C(t)^\top R^{-1}(t)C(t)$$
$$- \lambda P - PD(t)Q(t)D(t)^\top P, \qquad (50b)$$

where

- in a Kalman design, $\lambda \geq 0$, $\Pi(0) = P^{-1}(0)$ (resp. $Q(t)$, resp. $R(t)$) is the covariance matrix of the initial condition (resp. of the model noise, resp. of the output noise), and $D(t)$ is the matrix describing how the model noise enters the dynamics at time t;
- in a Kalman-like design, $Q = 0$, $\lambda > 0$ models a (stabilizing) forgetting factor, $\Pi(0)^{-1} = P(0)$ (resp R^{-1}) describes the weight of the initial error and the output error in the optimized cost (48), and thus the confidence in our initial guess and the output comparatively.

The main advantage of the Kalman-like design with $Q = 0$ is that it makes the dynamics of $P = \Pi^{-1}$ linear and explicitly solvable, with the expression of P directly related to the observability (or more precisely determinability/constructibility) Gramian. This allows to derive lower- and upper-bounds on P and to follow a Lyapunov analysis with the Lyapunov function $V = (\hat{z} - z)^\top P(\hat{z} - z)$ under a regular persistence assumption.

Theorem 5.1 (*Kalman-like Design*). *Assume that there exist $\bar{a} \geq 0$, $\alpha > 0$, $\bar{t} > 0$ and $t_0 \geq \bar{t}$ such that for any trajectory $t \mapsto z(t)$ of (46) defined on $[0, +\infty)$ with output y, $t \mapsto A(y(t), t)$ is bounded by $\bar{a}$ and $\mathcal{G}_y(t - \bar{t}, t) \geq \alpha I$. Then, for any $\lambda > 2\bar{a}$ and any positive definite matrices $\Pi(0)$ and R, (49) with $Q = 0$ is an asymptotic observer for (46), with exponential convergence of the estimation error $\hat{z} - z$ after time t_0.*

For a Kalman design, Kalman and Bucy (1961) similarly requires boundedness of A and uniform complete observability of the pair (A, C), but additionally requires uniform complete controllability of the pair (A, D) and R, Q being positive definite with uniform upper- and lower-bounds.

As discussed in Section 2.5, when the matrices A and C are constant, the invertibility of the observability Gramian is equivalent to the observability of the pair (A, C), namely the fact that the matrix $\mathcal{O} := \text{col}[C, CA, \ldots, CA^{n_x - 1}]$ is full-rank (Chen, 1984). In that case, one can implement the observer dynamics (49a) with a constant gain $K = \Pi_\infty C^\top R^{-1}$, where Π_∞ is the asymptotic equilibrium of (49b), which makes $A - KC$ Hurwitz. Note that this linear observer with constant gain is also often called "Luenberger observer", but we choose to keep this name for observers designed along Luenberger's initial method (Luenberger, 1964) described in Section 7.

On the other hand, when the time-dependence in A and C comes from an input u, this input may be actively chosen to ensure the Gramian condition and maximize observability, leading to the literature of *active sensing* (Franquiz, Muñoz, Udrea, & Balas, 2018; Hinson, Binder, & Morgansen, 2013; Hinson & Morgansen, 2013; Rubio Scola, Besançon, & Georges, 2018), to name a few. The position of the sensors itself may also be optimized, as in the literature of *sensor positioning* (van den Berg, Hoefsloot, Boelens, & Smilde, 2000).

5.3. Extended Kalman filter

The appeal of the Kalman filter mainly lies in its robustness and simplicity, since it is made of a copy of the dynamics and a correction gain obtained from a Riccati differential equation, whose parameters can be linked to physical quantities such as the noise covariance. That is why a very popular method in the industry (especially in the discrete-time formulation) consists in applying the Kalman filter also to nonlinear systems (1), by using the linearizations of f and h along the estimate $\hat{x}$ in the Riccati equation. Such an approach is commonly denoted as *Extended Kalman Filter* (EKF) and takes the form

$$\dot{\hat{x}} = f(\hat{x}, t) + \Pi C(\hat{x}, t)^\top R^{-1}(y - h(\hat{x}, t)), \qquad (51a)$$

$$\dot{\Pi} = A(\hat{x}, t)\Pi + \Pi A(\hat{x}, t)^\top + Q + \lambda \Pi - \Pi C(\hat{x}, t)^\top R^{-1}(t)C(\hat{x}, t)\Pi, \qquad (51b)$$

initialized at $\Pi(0) = \Pi(0)^\top > 0$, with A, C defined as

$$A(x, t) = \frac{\partial f}{\partial x}(x, t), \qquad C(x, t) = \frac{\partial h}{\partial x}(x, t).$$

and Q some positive definite matrix.

Unfortunately, apart from particular uniformly observable triangular structures (Gauthier & Kupka, 2001; Krener, 2003, see Theorem 6.3 below), only local convergence of the estimation error $\hat{x} - x$ is ensured (i.e. only for initial conditions satisfying $|\hat{x}(0) - x(0)| \leq c_0$ with $c_0 > 0$ being small enough), under the additional ad-hoc assumption that the solution Π of the Riccati Eq. (51b) is uniformly bounded in time, namely there exist $\underline{p}, \bar{p} > 0$ such that

$$\underline{p}I \leq \Pi(t) \leq \bar{p}I, \qquad \forall t \geq 0. \qquad (52)$$

See Reif, Sonnemann, and Unbehauen (1998) for $\lambda > 0$ and Bonnabel and Slotine (2015) for $\lambda = 0$. However, since the trajectory of Π depends on that of $\hat{x}$ itself, the assumption (52) cannot be in general checked "a priori", introducing therefore a loop in the stability analysis (the same issue appears in the discrete-time context Boutayeb, Rafaralahy, & Darouach, 1997; Rapp & Nyman, 2004; Song & Grizzle, 1992). An exception is Hammouri, Vu, and Yahoui (2013), where an infinitesimal observability assumption is made on the plant, and not along the estimate. But convergence is still only local.

Recent contributions though, suggest that for controlled systems, the input could be actively chosen online to maximize the Gramian associated to $(A(\hat{x}, t), C(\hat{x}, t))$, namely sufficiently excite the observability of the linearization along the known estimate trajectory $\hat{x}$, in order to guarantee (52) holds and obtain semi-global convergence (Bernard, Mimmo, & Marconi, 2021; Salaris, Cognetti, Spica, & Giordano, 2019).

5.4. Linearization by output injection

Once the community knew how to build an observer for the form (46), researchers attempted to characterize the class of nonlinear systems that could be transformed into such a form.

5.4.1. Constant observable pair (A, C)

The problem of (locally) transforming a nonlinear system into a linear one of the form (46) with a constant observable pair (A, C) via a diffeomorphism (with $n_z = n_x$) was first investigated in Krener and Isidori (1983) for autonomous systems (2) and Krener and Respondek (1985) for systems with inputs (3). Bossane, Rakotopara, and Gauthier (1989) then gave conditions for the existence of a local (and global) immersion with $n_z \geq n_x$ in the particular case of control affine systems. A vast literature followed on the subject, either developing algebraic algorithms to check the existence of a transformation or tools to explicitly find the transformation.

Actually, as noticed in Jouan (2003), a system (1) can be transformed into (46) with constant observable pair (A, C) if and only if it can be transformed into

$$\begin{cases} \dot{z}_1 &= z_2 + \varphi_1(z_1, t) \\ &\vdots \\ \dot{z}_i &= z_{i+1} + \varphi_i(z_1, t) \qquad y = z_1, \\ &\vdots \\ \dot{z}_m &= \varphi_m(z_1, t) \end{cases}$$

if $n_y = 1$ or blocks of this form otherwise. This is therefore more restrictive than the triangular forms of Section 6 where each φ_i is allowed to depend on $z_1, \ldots, z_i$. In particular, for autonomous systems, Back and Seo (2004), Jouan (2003) show that this is equivalent to finding for each output, a change of output $\psi : \mathbb{R} \to \mathbb{R}$ and maps φ_i solution to the *characteristic equation*

$$L_f^{n_z} \tilde{h}_k = L_f^{n_z - 1} \varphi_1 \circ \tilde{h} + L_f^{n_z - 2} \varphi_2 \circ \tilde{h} + \cdots + L_f \varphi_{n_z - 1} \circ \tilde{h} + \varphi_{n_z} \circ \tilde{h}$$

with $\tilde{h} = \psi \circ h$, which does not always admit solutions.

Along the history of linearization, we may also mention some generalizations such as (Keller, 1987), where the function φ is allowed

to depend on the derivatives of the input and later on the derivatives of the output in Plestan and Glumineau (1997), or Guay (2002), Respondek, Pogromski, and Nijmeijer (2004) where it is proposed to use an output-depending time-scale transformation. But generally the existence of the transformation is difficult to check and involves tedious symbolic computations, which do not always give the transformation itself. Even when they do, its validity is often local and its injectivity on $\mathcal{X}$ not guaranteed.

5.4.2. Time-varying pair (A, C)

In parallel, Fliess and Kupka (1983) studied the linearization problem by allowing A and C in (46) to depend on time/the input. This led to the restrictive *finiteness criterion of the observation space*, which roughly says that the linear space containing the successive derivatives of the output along any vector field of the type $f(\cdot, u)$ is finite. Later, Besançon and Bornard (1997), Hammouri and Kinnaert (1996) also allowed A to depend on the output y to broaden the class of concerned systems. But those systems remain difficult to characterize because there are often many possible parametrizations via the output.

6. Observers from differential observability

6.1. Differential observability and normal forms

Consider an autonomous single-output system of the form (2), i.e.,

$$\dot{x} = f(x), \qquad y = h(x), \tag{2}$$

$x \in \mathbb{R}^{n_x}$, $y \in \mathbb{R}$, in which the maps f, h are supposed to be sufficiently differentiable. We have the following definition of *differential observability*, which is based on the analysis of the output y and its time-derivatives $\dot{y}, \ddot{y}, \dots, y^{(k)}$, $k \geq n - 1$,

Definition 6.1. The system (2) is said to be:

- *weakly differentially observable* if there exists $n_z \in \mathbb{N}$ such that the mapping $T : \mathbb{R}^{n_x} \to \mathbb{R}^{n_z}$,

$$T(x) = \begin{pmatrix} h(x) \\ L_f h(x) \\ \vdots \\ L_f^{n_z - 1} h(x) \end{pmatrix} \tag{53}$$

is injective on $\mathcal{X}$;

- *strongly differentially observable* if the previous condition holds, and moreover the mapping T is full rank for all $x \in \mathcal{X}$.

Remark 6.1. For a linear system, the mapping T defined in (53) with $n_z = n_x$ is given by $T(x) = \mathcal{O}x$ where $\mathcal{O}$ is the observability matrix defined in (15). Weak and strong differential observability therefore coincide with the standard rank test for the matrix $\mathcal{O}$.

Under weak differential observability, if T defined in (53) is uniformly injective on $\mathcal{X}$ in the sense of Theorem 3.1, for instance if $\mathcal{X}$ is compact, then T admits a uniformly continuous left-inverse $\mathcal{T}$ defined on $\mathbb{R}^{n_z}$. Then, a direct consequence is that system (2) can be transformed, using $z = T(x)$, into a system of the form

$$\dot{z} = Az + \varphi(z), \qquad y = Cz, \tag{54}$$

with $z \in \mathbb{R}^{n_z}$, $\varphi(z) = \left(0, \dots, 0, \varphi_{n_z}(z)\right)^\top$,

$$A = \begin{pmatrix} 0 & 1 & & 0 \\ \vdots & & \ddots & \\ & & & 1 \\ 0 & & \cdots & 0 \end{pmatrix}, \qquad C = \begin{pmatrix} 1 & 0 & \cdots & 0 \end{pmatrix}, \tag{55}$$

and the continuous function $\varphi_{n_z} : \mathbb{R}^{n_z} \to \mathbb{R}$ is defined as $L_f^{n_z} h \circ \mathcal{T}$. The triplet (A, φ, C) is often said to be in *prime form*. A system of the form (54) is also said to be in *canonical observability form*, or *phase-variable*

form or simply normal form. See, e.g., Gauthier and Kupka (2001). One of the peculiarities of such a normal form is its lower-triangular structure and in particular its chain of integrators in which each i-th component of the state z represents the $(i-1)$-th time derivative of the measured output y. As we shall see in Section 6.2, the interest of such lower-triangular forms relies on the systematic design of *differentiators* able to reconstruct the state z with an arbitrarily fast exponential rate of convergence (in this case, we talk about *high-gain observers*, see Section 6.2) or finite-time convergence (following *sliding-mode* or *homogeneous* approaches, see Section 6.3).

While for autonomous systems, Definition 6.1 provides sufficient and necessary conditions to obtain a normal form, the extension to time-varying systems (1), resp. controlled systems (3), is more involved because the presence of the time t, resp. input u, may destroy the observability properties of the system. Indeed, starting from Gauthier and Bornard (1981), a lot of work has been done in order to establish the existence of state transformations, dependent or independent from the input t, resp. u, to obtain generalized lower-triangular forms for which high-gain (or homogeneous) approaches may be applied. See, e.g., Bernard, Praly, Andrieu, and Hammouri (2017), Fliess (1990), Proychev and Mishkov (1993), Rudolph and Zeitz (1994), Zeitz (1984) and references therein. For instance, one can allow the state transformation (53) to be t dependent (resp. $u, \dot{u}, \ddot{u}, \dots$ dependent), so that to obtain an injective map $T_t(x)$ (resp. $T_u(x)$), uniform in time (resp. in u and its derivatives) transforming the time varying system (1) (resp. controlled system (3)) into a system of the form (54) in which the function φ is now t (resp. $u, \dot{u}, \ddot{u}, \dots$) dependent. See, e.g., the *phase-variable form* in Zeitz (1984) or Gauthier and Kupka (2001, Section 2.2). A similar equivalent notion is also the *complete uniform observability* introduced in Teel and Praly (1994).

This paper being focused on the design of observers, we state the following definition that generalizes the form (54) for time-varying systems (1).

Definition 6.2. The system (1) admits a triangular normal form if there exists a mapping $T : \mathbb{R}^{n_x} \times \mathbb{R} \to \mathbb{R}^{n_z}$, (uniformly) injective in $\mathcal{X}$, transforming system (1) in the sense of Theorem 3.1 into a system of the form (26) with an output map depending only on z_1 and time, i.e., of the form

$$y = H(z_1, t) \tag{56a}$$

and where the dynamics $F = (F_1, \dots, F_{n_z})$ have a lower triangular structure, with each F_i depending only on $z_1, \dots, z_{i+1}$, i.e., with

$$\frac{\partial F_i}{\partial z_j}(z, y, t) = 0, \qquad j > i + 1, \qquad \forall (z, y, t), \tag{56b}$$

for all $i = 1, \dots, n_z$.

Note that for controlled systems (3), it suffices to replace t in (56) with u (and possibly its time-derivatives $\dot{u}, \ddot{u}, \dots$), see, e.g., Gauthier and Kupka (2001, Section 3.2).

For input-affine systems of the form

$$\dot{x} = f(x) + g(x)u, \qquad y = h(x), \tag{57}$$

with scalar input $u \in \mathbb{R}$ and output $y \in \mathbb{R}$, necessary and sufficient conditions to admit an autonomous transformation into a triangular normal form (without output injection) have been given in (Gauthier & Bornard, 1981, Section 4) or Theorem 4.1 in Gauthier and Kupka (2001, Section 3.4). In particular, if (57) is observable for any input (uniform observability) and the pair (f, h) is strongly differentially observable of order n_x, then T defined in (53) transforms (57) into a triangular form verifying (56) with $H(z_1, t) = z_1$ and with maps F_i that are input-affine and Lipschitz. The possibility of extending this result to any uniformly observable system in the form (57) with higher order of differential observability and less regularity of F has been studied in Bernard, Praly, Andrieu, and Hammouri (2017) but is still not fully characterized.

Another special case of (56) is when the system can be written in the form

$$F(z, y, t) = A(y, t)z + \varphi(y, z, t)$$
$$y = C(t)z + D(t)$$
(58)

where the functions A, φ have the structure

$$A(y, t) = \begin{pmatrix} 0 & a_1(y, t) & & 0 \\ \vdots & & \ddots & \\ & & & a_{n_z-1}(y, t) \\ 0 & & \cdots & 0 \end{pmatrix},$$

$$\varphi(z, t) = \begin{pmatrix} \varphi_1(y, t) \\ \varphi_2(y, z_2, t) \\ \vdots \\ \varphi_{n_z}(y, z_2, \dots, z_{n_z}, t) \end{pmatrix}.$$

See, for instance, Besancon and Ticlea (2007).

We finally remark that extensions to the multiple-output case allow to obtain block-triangular forms in which each block has a triangular form satisfying (56b). For compactness, we refer, for example, to Besancon and Ticlea (2007), Fliess (1990), Gauthier and Bornard (1981), Proychev and Mishkov (1993), Rudolph and Zeitz (1994), Zeitz (1984) and we limit the exposition of Section 6 to the single-output case.

6.2. High-gain observers

6.2.1. General construction

The use of high-gain observers for state-estimation of nonlinear systems first appeared at the end of the 80's in a different number of contemporary works (Deza, Busvelle, Gauthier, & Rakotopara, 1992; Emelyanov, Korovin, Nikitin, & Nikitina, 1989; Esfandiari & Khalil, 1987; Tornambe, 1989). One of the main features of high-gain observers is their "tunability property" which is typically controlled by one single scalar parameter. Such a fundamental property allowed this class of observers to address the problem of output feedback stabilization (see also Section 8.7). We refer to Khalil and Praly (2013) and references therein for an extensive bibliography on the topic.

The high-gain observer construction is based on the fact that for the triangular normal form, it is possible to follow the strategy described in Section 4.5 and Theorem 4.3 to construct an observer. More precisely, it is possible to construct a metric $P : \mathbb{R} \mapsto \mathbb{R}^{n_z \times n_z}$ and a smooth correction term k such that

$$k(z, H(z), t) = 0, \qquad \forall (z, t) \in \mathbb{R}^{n_z} \times \mathbb{R},$$
(59a)

$$\dot{P}(t) + \mathrm{He} \left\{ P(t) \left[\frac{\partial F}{\partial z}(z, y, t) + \frac{\partial k}{\partial z}(z, y, t) \right] \right\} \leq -qI,$$
$$\forall (y, z, t) \in \mathbb{R}^{n_y} \times \mathbb{R}^{n_z} \times \mathbb{R},$$
(59b)

$$\underline{p} I \leq P(t) \leq \bar{p} I, \qquad \forall t \in \mathbb{R}.$$
(59c)

With Theorem 4.3, it implies that

$$\dot{\hat{z}} = F(\hat{z}, y, t) + k(\hat{z}, y, t),$$
(60a)

is a uniformly exponentially stable observer in the z coordinates. Hence, with

$$\hat{x} = \mathcal{T}(\hat{z}, t)$$
(60b)

where $\mathcal{T}$ is a uniformly continuous left inverse of T, (60) defines an asymptotic observer for the given coordinates from Theorem 3.1.

Various results have been obtained to construct P and k such that (59) holds. They are all based on the same two-step design. First of all, assuming F differentiable, let us denote

$$a_i(z, y, t) = \frac{\partial F_i}{\partial z_{i+1}}(z, y, t), \qquad 1 \leq i \leq n_z,$$

$$b_{ij}(z, y, t) = \frac{\partial F_i}{\partial z_j}(z, y, t), \qquad 1 \leq j \leq i,$$
(61)

$$c_1(z, t) = \frac{\partial H}{\partial z_1}(z_1, t).$$

We have the decomposition

$$\frac{\partial F}{\partial z}(z, y, t) = A(z, y, t) + B(z, y, t), \quad \frac{\partial H}{\partial z}(z_1, t) = C(z_1, t)$$
(62)

where (forgetting the dependence in (z, y))

$$A(t) = \begin{pmatrix} 0 & a_1(t) & & 0 \\ \vdots & & \ddots & \\ & & & a_{n_z-1}(t) \\ 0 & & \cdots & 0 \end{pmatrix},$$

$$B(t) = \begin{pmatrix} b_{11}(t) & 0 & & 0 \\ b_{21}(t) & b_{22}(t) & 0 & 0 \\ \vdots & & \ddots & \\ b_{n_z 1}(t) & & \cdots & b_{n_z n_z}(t) \end{pmatrix},$$
(63)

$$C(t) = \begin{pmatrix} c_1(t) & 0 & \cdots & 0 \end{pmatrix}.$$

The two design steps for high-gain observer construction can be described as follows.

Step 1: *Construct a preliminary correction term k_0 only for A.* The problem reduces to finding P and k_0 such that (59a) and (59c) hold and also

$$\dot{P}(t) + \mathrm{He} \left\{ P(t) \left[A(t) + \frac{\partial k_0}{\partial z}(z, y, t) \right] \right\} \leq -qI, \quad \forall (y, z, t).$$
(64)

For instance, if A is constant (as in the prime form or the uniformly observable context), this step can be simply solved by selecting P constant and k_0 linear via pole placement.

Step 2: *Increase the robustness via high-gain scaling.* If one compare (64) and (59b), the difference comes from the B term which is lower triangular according to (63). As it will be shown in the following, a typical global Lipschitz property allows to modify k_0 (and also P) to cope with this extra term. Note that if $\mathcal{X}$ is bounded and if the system admits a locally Lipschitz triangular normal form, it is possible to extend it to a globally Lipschitz triangular normal form outside of $T(\mathcal{X})$. In practice, this extension is obtained by saturation.

Two main types of methods have been developed, both of them following this two-step design: the case in which the metric P is constant (also called *Luenberger-like high-gain observers*) and the case in which P is time varying (also called *Kalman-like high-gain observer* or *Extended-Kalman-like high-gain observer*).

6.2.2. With a constant metric P

For step 1: The case of a constant metric is the most popular one (see Deza et al., 1992; Emelyanov et al., 1989; Esfandiari & Khalil, 1987; Gauthier, Hammouri, & Othman, 1992; Tornambe, 1989). When the A matrix in (63) is constant, namely takes the form (55), it suffices to select $K = (k_1, \dots, k_{n_z})^\top$, with k_i coefficients of a stable polynomial, so that the matrix

$$A - KC = \begin{pmatrix} -k_1 & 1 & & 0 \\ \vdots & & \ddots & \\ & & & 1 \\ -k_{n_z} & & \cdots & 0 \end{pmatrix}$$

is Hurwitz, and P solution of the Lyapunov equation $\mathrm{He}\{P(A - KC)\} \leq -qI$. In its most general form, namely for non-constant matrices A, the following lemma due to W. Dayawansa can be found in Gauthier and Kupka (2001, Lemma 2.1 p. 96) and in Isidori (2017, Section 7.4).

Lemma 6.1 (*Gauthier & Kupka, 2001*). *Consider A and C defined in (63). Suppose there exist $0 < \underline{a} < \bar{a}$ and $0 < \underline{c} < \bar{c}$ such that*

$$\underline{a} < |a_i(t)| < \bar{a}, \qquad \underline{c} < |c_1(t)| < \bar{c}, \qquad \forall t.$$
(65)

Then, there exists a vector K in $\mathbb{R}^{n_z}$ and a positive definite matrix P in $\mathbb{R}^{n_z \times n_z}$, both depending only on $(\underline{a}, \bar{a}, \underline{c}, \bar{c})$, such that

$$\mathrm{He}\{P[A(t) - KC(t)]\} \leq -qI, \qquad \forall t.$$
(66)

For step 2: Employing a high-gain scaling defined with

$$\mathcal{D}_\ell := \mathrm{diag}(\ell, \ell^2, \dots, \ell^{n_z}), \tag{67}$$

where ℓ is the so-called high-gain parameter,[6] it is possible to amplify the robustness of the correction terms to cope with Lipschitz nonlinearities. Based on this approach, the following theorem is obtained, see, e.g., Gauthier and Kupka (2001).

Theorem 6.1. *Assume that the system* (1) *admits a triangular normal form in the sense of* Definition 6.2, *with the maps F and H satisfying for all (z, y, t) in $\mathbb{R}^{n_z} \times \mathbb{R}^{n_y} \times \mathbb{R}$*

$$\underline{a} < \left| \frac{\partial F_i}{\partial z_{i+1}}(z, y, t) \right| < \overline{a}, \quad \underline{c} < \left| \frac{\partial H}{\partial z_1}(z, y, t) \right| < \overline{c}, \tag{68a}$$

$$\left| \frac{\partial F_i}{\partial z_j}(z, y, t) \right| < \overline{b}, \qquad j \le i, \tag{68b}$$

for some positive real numbers $(\underline{a}, \overline{a}, \underline{c}, \overline{c}, \overline{b})$. Then, with K in $\mathbb{R}^{n_z}$ obtained from Lemma 6.1, there exists $\ell^* \ge 1$ such that for all $\ell > \ell^*$, the system (60) *is a robust tunable asymptotic observer with*

$$k(\hat{z}, y, t) = \mathcal{D}_\ell K[y - H(\hat{z}_1, t)] \tag{69}$$

and $\mathcal{D}_\ell$ defined as in (67).

Eq. (68b) implies that the mapping F_i is globally Lipschitz with respect to the variable z. The main feature of observer (60a) with (69) is that the high-gain parameter $\ell \ge 1$ is chosen large enough compared to the Lipschitz bound (68b) (see, e.g., Khalil & Praly, 2013, Section 3). The following bound then holds for the estimation error in the z coordinates

$$|z(t) - \hat{z}(t)| \le c\ell^{n_z - 1} \exp(-\ell \lambda t)|z(0) - \hat{z}(0)|, \qquad \forall t, \tag{70}$$

for some constants c, λ independent from ℓ. From this inequality, we obtain that the high-gain observer (60a) with (69) is uniformly exponentially stable in the z coordinates. It is moreover robust and a tunable observer in the given coordinates (see Section 2.4 and Remark 3.2) since its convergence rate can be arbitrarily increased by augmenting the high-gain parameter ℓ. Note, however, that a typical drawback of such a construction is the fact that the overshoot increases in $\ell^{n_z - 1}$. This fact is usually referred to as "peaking phenomenon".

Combining Lipschitz conditions, an LMI-design mixing the aforementioned paradigm and the results in Section 4.5 can also be obtained in order to obtain less restrictive conditions by following (Zemouche et al., 2018).

When the mapping F is no longer Lipschitz, the high-gain observer no longer ensures asymptotic convergence, but practical convergence may be guaranteed under some conditions on the Hölder powers or boundedness of each nonlinearity, see, e.g., Bernard, Praly, and Andrieu (2017). Otherwise, more general homogeneous correction terms can be used, as detailed in Section 6.3.

6.2.3. Kalman-like high-gain observers

For step 1: Another approach to design a correction term for A is to follow the approach obtained from the observability Gramian given in Section 5. In order to do so, one needs to consider the particular case in which A and C depend only on known signals (i.e., y and t or u but not z) and can be employed to update dynamically P. Moreover, the pair (A, C) needs to be observable in the sense that the observability Gramian $\mathcal{G}_y$ of the system (see Definition 5.1)

$$\dot{\chi} = A(y, t)\chi, \qquad y = C(t)\chi,$$

needs to be lower bounded in a particular manner. We have the following result, see, e.g., (Besançon, 1999).

Lemma 6.2. *Assume that there exists a positive real number $\overline{a}$ such that $|A(y, t)| \le \overline{a}$ for all y, t. Assume moreover that there exists ℓ^* that for all $\ell > \ell^*$ and all $t \ge \frac{1}{\ell}$*

$$\mathcal{G}_y\left(t - \frac{1}{\ell}, t\right) \ge c\ell \mathcal{D}_\ell^{-2},$$

with $\mathcal{D}_\ell$ defined as in (67). *Then, for any $\lambda \ge 2\overline{a}$, there exist $\underline{p}$ and $\overline{p}$ such that for all $\ell > \ell^*$ the solution to*

$$\dot{P}(t) = \ell\left(-\lambda P(t) - A(y, t)^\top P(t) - P(t)A(y, t) + C(t)^\top C(t)\right) \tag{71}$$

initialized at $P(0) = P(0)^\top > 0$ satisfies (59c).

The Gramian condition may be referred to as *instantaneous uniform complete observability*, since it requires lower boundedness of the Gramian in arbitrarily short time.

For step 2: Again, for the second step, assuming a global Lipschitz property, it is possible to cope with the nonlinearities.

Theorem 6.2. *Assume that the system* (1) *admits a triangular normal form with a vector field F such that A and C given in* (62), (63) *are independent of z,* (68b) *holds and the assumptions of Lemma 6.2 hold, then there exists ℓ^* such that for all $\ell > \ell^*$, the system* (60a), (60b), (71) *is a robust tunable asymptotic observer with*

$$k(\hat{z}, y, t) = \mathcal{D}_\ell P(t)^{-1} C(t)^\top (y - C(t)\hat{z}) \tag{72}$$

and $\mathcal{D}_\ell$ defined as in (67).

As highlighted in Theorem 6.2, the gain of the high-gain observer (60a), (60b), is obtained from the solution of a differential Lyapunov matrix Eq. (71) similar to the one of (50). In the case in which A, C in (71) are constant, we recall also the works (Chitour, 2002; Ibrir & Diop, 1999) and the early work (Gauthier et al., 1992) in which the asymptotic solution to (71), $P_\infty = \lim_{t \to \infty} P(t)$ is employed.

In Gauthier and Kupka (2001), another version of the Kalman-like high-gain observer is given in which the matrix function B is also employed in the update law of P and where A and C are assumed to be constant. We briefly recall the main result.

Theorem 6.3. *Assume that the system* (1) *admits a triangular normal form with a vector field F such that A and C given in* (62), (63) *are constant and of the form* (55) *and* (68b) *holds then there exists ℓ^* such that for all $\ell > \ell^*$, the system* (60a), (60b) *is a robust tunable asymptotic observer with*

$$k(\hat{z}, y, t) = P(t)^{-1} C^\top (y - C\hat{z}) \tag{73}$$

where

$$\dot{P}(t) = -\left(A^\top + B(\hat{z}, t)\right)^\top P(t) - P(t)\left(A^\top + B(\hat{z}, t)\right) + C^\top C - P(t)\mathcal{D}_\ell^2 P(t), \tag{74}$$

initialized at $P(0) = P(0)^\top > 0$, and $\mathcal{D}_\ell$ defined as in (67).

In that case, the design of the gain (73) of the high-gain observer (60a), (60b) is obtained from the solution of a differential Riccati matrix Eq. (74) which is exactly the same as the Extended Kalman Filter. For this reason, this observer is named the *Extended Kalman like High-gain Observer*.

6.2.4. Adapting the high-gain parameter and dynamics scaling

The idea of dynamically updating the high-gain parameter has been investigated by many researchers. The motivation for this is twofold: for qualitative purpose to reduce the size of the gain (see for instance Farza, Oueder, Abdennour, & M'Saad, 2011), or to allow unknown or time-dependent Lipschitz constants.

---

[6] In the literature, the high-gain parameter is sometimes denoted $\ell = \frac{1}{\varepsilon}$ with ε taken small enough, see, e.g., Khalil and Praly (2013).

- For instance, Lipschitz bounds of the form

$$\left|\frac{\partial F_i}{\partial z_j}(z, y, t)\right| < \bar{b}(y, t), \qquad j \le i, \tag{75}$$

where now $\bar{b}(y, t)$ is a continuous function have been considered in Praly (2003) or Krishnamurthy, Khorrami, and Chandra (2003). In that case the prototypical high-gain parameter updating law takes the form

$$\dot{\ell} = \ell(c_1(c_2 - \ell) + c_3\bar{b}(y, t)), \qquad \ell(0) > c_2,$$

for some positive real numbers (c_1, c_2, c_3). The idea behind such a design is that ℓ dynamically adapts to the Lipschitz bound (75) with a Riccati differential equation.

- In Alessandri and Rossi (2015), Bullinger and Allgower (1997), Lei, Wei, and Lin (2005) the Lipschitz bound is unknown. The idea is then to employ a strictly increasing high-gain parameter. Note that in this case, the obtained observer may not be robust.
- Adaptation has also been employed in combination with the Kalman-like high-gain observer given in Theorem 6.3, see, e.g. Boizot, Busvelle, and Gauthier (2010).

6.3. Homogeneous correction terms

6.3.1. Homogeneity to allow non locally Lipschitz normal form

High-gain observers may also be designed to cope with non-Lipschitz triangular forms. Indeed, it is possible to relax this regularity condition if we employ homogeneous observers (also known as sliding-mode observers in some cases). These type of observers already have an old history which can be traced back to the work of Arie Levant (Levant, 2001, 2003) and has been followed by many other researchers (see, e.g., Andrieu, Praly, & Astolfi, 2008, 2009a, 2009b; Bernard, Praly, & Andrieu, 2017; Levant, 2005; Perruquetti, Floquet, & Moulay, 2008; Qian, 2005a; Qian & Lin, 2006; Yang & Lin, 2004).

The homogeneous correction terms of degree $\mathfrak{r}$ in $[-1, \frac{1}{n_z - 1})$ denoted $k_{\text{Hom}} = (k_{\text{Hom},1}, \ldots, k_{\text{Hom},n_z})$ take the form

$$k_{\text{Hom},i}(s) = \mu_i \, \text{sign}(s) \, |s|^{\frac{r_{i+1}}{r_1}} , \qquad r_i = 1 - \mathfrak{r}(n_z - i). \tag{76}$$

where μ_i are gains to be tuned and sign is a set-valued map defined as

$$\text{sign}(s) = \begin{cases} \{1\} & \text{if } s > 0, \\ [-1, 1] & \text{if } s = 0, \\ \{-1\} & \text{if } s < 0, \end{cases} \tag{77}$$

which is upper semi-continuous with nonempty, compact and convex values. In (76), $r = (r_1, \ldots, r_{n_z+1})$ is a vector in $\mathbb{R}^{n_z+1}$, called weight vector. For instance, for $\mathfrak{r} = -1$ and $n_z = 2$, we obtain

$$k_{\text{Hom},1}(s) = \mu_1 \, \text{sign}(s)|s|^{\frac{1}{2}}, \quad k_{\text{Hom},2}(s) = \mu_2 \, \text{sign}(s), \tag{78}$$

recovering the gains of the well-known two-order sliding-mode observer, see, e.g., Andrieu, Astolfi, and Bernard (2021), Levant (2001) and references therein. For the case $\mathfrak{r} = -1$, the observer dynamics with the correction term (76) must be understood as a differential inclusion. When $\mathfrak{r} = 0$, we recover a linear correction terms as the one given in Lemma 6.1.

The construction strategy in the homogeneous case follows the same two steps design as the one previously given. However, since we may not have differentiable nonlinearities in the triangular form, we do not employ the differential framework. So in the following, we simply assume the system (1) admits a continuous triangular normal form in the form (58) with constant matrices C and A.

Step 1 Homogeneous: When employing homogeneous correction terms, the following result has been obtained for a constant matrix A and C in Cruz-Zavala and Moreno (2019) and Bernard, Praly, and Andrieu (2017, Lemma 1) (see also Levant, 2001 for the case $\mathfrak{r} = -1$)

Lemma 6.3. *Consider A and C defined in (63) and assume A and C are constant and of the form (55). Then, there exist real numbers $(\mu_1, \ldots, \mu_{n_z})$ in $\mathbb{R}^{n_z}$, such that the origin of*

$$\dot{e} = Ae - k_{\text{Hom}}(Ce), \qquad e \in \mathbb{R}^{n_z} \tag{79}$$

is globally asymptotically stable.

The system (79) being homogeneous and its origin being asymptotically stable, when $\mathfrak{r} < 0$, trajectories are converging to the origin in finite time.

Step 2 Homogeneous: Employing the robustness property of homogeneous systems, the following result have been obtained in Bernard, Praly, and Andrieu (2017).

Theorem 6.4. *Assume that the system (1) admits a continuous triangular normal form as in (58) with constant vectors A and C of the form (55) and with the mappings φ which satisfy for all i in $\{1, \ldots, n_z\}$, for all (y, z_a, z_b) in $\mathbb{R} \times \mathbb{R}^{n_z} \times \mathbb{R}^{n_z}$ and $t > 0$,*

$$\left|\varphi_i(y, z_{2a}, \ldots, z_{ia}, t) - \varphi_i(y, z_{2b}, \ldots, z_{ib}, t)\right| \le \bar{b} \sum_{j=2}^{i} |z_{ja} - z_{jb}|^{\alpha_{ij}} , \tag{80}$$

with

$$\frac{1 + (n_z - i - 1)}{1 + (n_z - j)} \le \alpha_{ij} < 1, \qquad j \le i \le n_z. \tag{81}$$

Then, there exists $\mathfrak{r} < 0$ and $\ell^* > 0$ such that for all $\ell > \ell^*$, the system (60a), (60b), (76) is a robust finite-time observer with

$$k(\hat{z}, y, t) = \mathcal{D}_\ell k_{\text{Hom}}(y - C\hat{z}) \tag{82}$$

where $\mathcal{D}_\ell$ is defined in (67) and k_{Hom} in Lemma 6.3.

The homogeneous degree $\mathfrak{r}$ is typically chosen depending on how far the α_{ij} are from their lower-bound in (81) (see Bernard, Praly, & Andrieu, 2017). In the limit case where the lower-bounds of (81) are actually equalities, we take $\mathfrak{r} = -1$ and no restriction is imposed on φ_{n_z} besides boundedness. We recognize the well-known robustness property of the sliding mode observer.

Note also that the use of more involved correction terms may give uniform finite time converging observer (see Andrieu, Praly, & Astolfi, 2009c; Lopez-Ramirez, Polyakov, Efimov, & Perruquetti, 2018).

6.3.2. Homogeneity to allow non-globally Lipschitz normal form

Another result which has been obtained is to consider the case in which the normal form is locally Lipschitz but not globally Lipschitz. Employing homogeneous in the bi-limit correction terms (see Andrieu et al., 2008) in combination with dynamics high-gain scaling, the following result have been obtained in Andrieu et al. (2009a). Note that in that case the observer provides a converging estimation only for bounded trajectories.

Theorem 6.5. *Assume that for all $x(0)$ in $\mathcal{X}_0$ the solution $t \mapsto x(t)$ is bounded for all positive time. Assume, moreover, that the system (1) admits a triangular normal form as in (58) with C constant and with the matrix function A and the mappings φ_i satisfying for all (z, y, t) in $\mathbb{R}^{n_z} \times \mathbb{R} \times \mathbb{R}$ and for all i in $\{1 \ldots, n_z\}$*

$$0 < \underline{a} < a_i(y, t) < \bar{a}, \tag{83}$$

$$\begin{aligned}\left|\varphi_i(y, \ldots, \underline{z}_{ia}, t) - \varphi_i(y, \ldots, z_{ib}, t)\right| \\ \le \bar{b}_0(y, z_a, t) \sum_{j=2}^{i} |z_{ja} - z_{jb}| + \bar{b}_\infty \sum_{j=2}^{i} |z_{ja} - z_{jb}|^{\alpha_{ij}} ,\end{aligned} \tag{84}$$

with

$$\bar{b}_0(y, z, t) \le \Gamma(y, t)\left(1 + \sum_{j=2}^{n_z} |z_j|^{v_j}\right) , \quad 0 \le v_j \le \frac{1}{j - 1} , \tag{85}$$

$$1 \le \alpha_{ij} < \frac{i}{j - 1} , \tag{86}$$

for some locally Lipschitz function Γ. Then, there exists a correction term k_{biHomo} and three positive real numbers (c_1, c_2, c_3) such that for all $\ell > \ell^*$, the system (60a)–(60b) is an asymptotic observer with

$$k(\hat{z}, y, t) = \mathcal{D}_\ell k_{\text{biHom}}(y - C\hat{z}_1) \tag{87}$$

with $\mathcal{D}_\ell$ defined as in (67), and

$$\dot{\ell} = \ell(c_1(c_2 - \ell) + c_3 \bar{b}_0(y, \hat{z}, t)), \qquad \ell(0) > c_2.$$

As shown in Andrieu et al. (2009a), this observer is robust.

6.4. Pure differentiators

In many works, researchers focused on the problem of estimating the derivatives (up to a certain order) of a time-varying output y. Such a problem can be recast as the problem of state-observation for dynamical systems (1) that can be put into the form (54), (55), in which the function φ_{n_z} is possibly t-dependent. Following the prescriptions of Section 6.2, one can simply design a high-gain observer of the form

$$\dot{\hat{z}} = A\hat{z} + \mathcal{D}_\ell K(y - C\hat{z}) \tag{88}$$

in which the copy of the term φ_{n_z} is ignored in the observer dynamics. The observer (88) is also often denoted as *dirty-derivative* observer, see, e.g., Teel and Praly (1994). In this case, practical convergence can be obtained if the state z evolves in a compact set (but output feedback stabilization at the origin can be achieved for controlled systems, see, e.g., Andrieu et al., 2008; Praly & Jiang, 1993).

Similarly, one can also implement an observer following the homogeneous approach in Section 6.3, that is of the form

$$\dot{\hat{z}} = A\hat{z} + \mathcal{D}_\ell k_{\text{Hom}}(y - C\hat{z}) \tag{89}$$

with k_{Hom} selected as in (76), and still neglecting φ_{n_z}. In this case, under the same boundedness assumption of the state z, exact finite-time convergence is obtained for ℓ large enough and for $\mathfrak{r} = -1$ (see, e.g., Chalanga, Kamal, Fridman, Bandyopadhyay, & Moreno, 2016).

6.5. Use of interconnection

An interesting development in the context of observer design for systems in triangular normal form (see Definition 6.2) is the use of interconnection and cascades of low-order differentiators. The main motivations can be summarized as follows:

- to improve the performances of high-gain observers (e.g., in terms of peaking phenomenon and sensitivity to measurement noise);
- to improve the limitations of homogeneous approaches when the homogeneity conditions (80) or (84) are not satisfied.

Concerning the performances of high-gain observers, recall that when the measured output y is affected by some measurement noise, namely $y = H(z_1, t) + \nu_y$ in place of (56a), the bound (70) is then modified into

$$|z(t) - \hat{z}(t)| \leq c\ell^{n_z-1} \exp(-\ell \lambda t)|z(0) - \hat{z}(0)| + \rho\ell^{n_z-1} \sup_{s \in [0,t]} |\nu_y(s)| \tag{90}$$

for all $t \geq 0$ and for some $\rho > 0$ independent of ℓ. The inequality (90) highlights, as previously discussed, the peaking phenomenon and the important effect of the measurement noise ν_y on the asymptotic estimation. Furthermore, it can be proved that for high-frequency measurement noise the asymptotic gain behaves actually as ℓ^{n_z}, see Astolfi, Marconi, Praly, and Teel (2016). Such considerations motivate for the use of more sophisticated designs, based on the use of an interconnection of lower-dimensional observers in order to increase the relative degree between the measurement noise ν_y and the estimates $\hat{z}$, and at the same time, to relax the homogeneity conditions (80) or (84) in the context of homogeneous theory. Different techniques that can be also employed in the context of performances improvements are also discussed in Section 8.3.

The main idea of a cascade structure for triangular normal forms consists in employing a cascade of $n_z - 1$ lower-dimensional differentiators of order 2, each of them feeding the successive differentiator. The observer takes the form

$$\begin{aligned}
\dot{\hat{z}}_{ii} &= F_i\left(y, t, \hat{z}_{12}, \hat{z}_{23}, \ldots, \hat{z}_{i(i+1)}\right) - \ell_i k_{i,1}\left(\hat{z}_{ii} - \hat{z}_{(i-1)i}\right) \\
\dot{\hat{z}}_{i(i+1)} &= F_{i+1}\left(y, t, \hat{z}_{12}, \hat{z}_{23}, \ldots, \hat{z}_{(i+1)(i+2)}\right) - \ell_i^2 k_{i,2}\left(\hat{z}_{ii} - \hat{z}_{(i-1)i}\right) \\
&\quad i = 1, \ldots, n_z - 1,
\end{aligned} \tag{91}$$

with the conventions $\hat{z}_{01} = y$ and $\hat{z}_{n_z(n_z+1)} = 0$, and where the ℓ_i's are positive real numbers that will be selected later on. The overall state dimension of the observer is $2n_z - 2$. The indexes of the observer variables are selected with the convention that $\hat{z}_{ii}$ is the first state-component of the ith block providing an estimate of the variable z_i, and $\hat{z}_{i(i+1)}$ is the second state-component of the ith block providing an estimate of the variable z_{i+1}. As a consequence, for each variable z_i, with $i = 2, \ldots, n_z - 1$, we have two different estimates.

Depending on the design employed for the functions $k_{i,1}, k_{i,2}$, we obtain two different classes of observers.

Linear gains: By selecting the gains $k_{i,j}$ in (91) linearly, namely

$$k_{ij}(s) = \mu_{i,j} s, \qquad \forall i = 1, \ldots, n_z - 1, \; j = 1, 2, \tag{92}$$

we obtain the so-called *low-power high-gain observer* which provides an extension of Theorem 6.1, see, e.g., Astolfi and Marconi (2015), Wang, Astolfi, Marconi, and Su (2017) with the following result.

Theorem 6.6. *Assume that the system (1) admits a triangular normal form (see Definition 6.2) with the mappings F and H satisfying (68) for all (z, y, t) in $\mathbb{R}^{n_z} \times \mathbb{R}^{n_y} \times \mathbb{R}$ for some positive real numbers $(\underline{a}, \bar{a}, \underline{c}, \bar{c}, \bar{b})$. Then, there exist positive real numbers $\mu_{i,j}$, $i = 1, \ldots, n_z - 1$, $j = 1, 2$, and $\ell^* \geq 1$ such that, for all $\ell > \ell^*$, the system (91), (92) is a robust tunable asymptotic observer with $\ell_i = \ell$ for all $i = 1, \ldots, n_z - 1$.*

In particular, the observer (91) is able to recover the same type of ISS bound as the one given in (90). However, since the relative degree between the output y and the estimate z_i is larger than one (in particular, when y in φ_i is substituted with $\hat{z}_{11}$), the sensitivity with respect to high-frequency measurement noise is improved, see, e.g., Astolfi, Marconi, Praly, and Teel (2018). Furthermore, by adding some saturation functions between the interconnections of each block, the peaking phenomenon can be removed when the state z of the plant evolves in a known compact set, see Astolfi et al. (2018).

Different versions of the low-power high-gain observer for systems in the prime form (54), (55) are the one proposed in Teel (2016) in which each block is implemented as a reduced-order high-gain observer (thus recovering an observer of total dimension $n_z - 1$), and the *cascade high-gain observer* proposed in Khalil (2017), in which all but the first block are reduced-order observers (thus obtaining an observer of total dimension n_z).

Nonlinear gains: As proposed in Andrieu et al. (2021) the gains $k_{i,j}$ in (91) can be designed following a mixed linear/homogeneous approach as

$$\begin{cases} k_{i,1}(s) = q(\kappa s), \\ k_{i,2}(s) = \text{sign}(s) + q(\kappa s), \end{cases} \quad q(s) = \text{sign}(\kappa s)|s|^{\frac{1}{2}} + s. \tag{93}$$

As explained at the beginning of Section 6.3, the observer dynamics must be understood as a differential inclusion. The following result can be stated.

Theorem 6.7. *Assume that the system (1) admits a continuous triangular normal form as in (58) with constant vectors A and C as in (55) and with the mappings φ which satisfy for all i in $\{1, \ldots, n_z\}$, for all (y, z_a, z_b) in $\mathbb{R} \times \mathbb{R}^{n_z} \times \mathbb{R}^{n_z}$ and $t > 0$,*

$$\left|\varphi_i(y, z_{2a}, \ldots, z_{ia}, t) - \varphi_i(y, z_{2b}, \ldots, z_{ib}, t)\right| \leq \bar{b}_0 + \bar{b}_1 \sum_{j=2}^{i} |z_{ja} - z_{jb}|. \tag{94}$$

Then, there exist positive real numbers $(\kappa, \ell_1, \ldots, \ell_{n_z-1})$ such that the system (91), (93), is a robust finite-time observer.

With respect to Theorem 6.4, we obtain much less restrictive conditions on the functions φ_i. In particular, the homogeneous inequality (80) is now replaced by (94) in which each function φ_i is bounded by a linear and a constant term. To the best of the authors' knowledge, a convergent observer of the form (60a) under the condition (94) cannot be obtained, thus motivating the interconnection structure (91) to be crucial for systems with such type of nonlinearities.

Note that an early version of the observer (91) in the pure differentiator context (see Section 6.4) can be found in Floquet and Barbot (2007), where the gains $k_{i,j}$ in (91) are selected following the homogeneous conditions with degree $\mathfrak{r} = -1$, namely as in (78). Therein, however, the interconnection term $\hat{z}_{(i+1)(i+1)}$ in the $\hat{z}_{i(i+1)}$ dynamics is not employed and each block is actually implemented as a pure sliding-mode differentiator of order 2. Another possibility is finally to select the gains $k_{i,j}$ in (91) following standard homogeneous theory, see Moreno (2021). In this case, the homogeneous version of the low-power high-gain observer (Astolfi & Marconi, 2015) is recovered and a result similar to Theorem 6.4 can be established.

Finally, note that when the system (1) has (uniformly) bounded trajectories and admits a continuous triangular normal form as in (58) with constant vectors A and C as in (55) and without any further assumption of regularity on φ, it is possible to follow the approach proposed in Bernard, Praly, and Andrieu (2017). This consists in building a finite-time observer by interconnecting n homogeneous observers of degree $\mathfrak{r} = -1$ (see Section 6.3.1), with the ith block of dimension i estimating $(x_1, \ldots, x_i)$. Indeed, thanks to triangularity, the estimate of $(x_1, \ldots, x_i)$ provided by the ith block, which converges in finite-time, can be used to feed $(\varphi_1, \ldots, \varphi_i)$ in the $(i+1)th$ block and the homogeneous correction terms with $\mathfrak{r} = -1$ allows to handle the bounded nonlinearity φ_{i+1} on the last line. This observer is also shown to be uniformly asymptotically stable and robust.

7. Observers from backward distinguishability

When D. Luenberger published his first results concerning the design of observers for linear systems in Luenberger (1964), his idea was to look for an invertible change of coordinates transforming the plant dynamics

$$\dot{x} = Ax, \qquad y = Cx,$$

into a form

$$\dot{z} = \Lambda z + \Gamma y$$

with Λ Hurwitz, for which a trivial observer is simply made of a copy of the dynamics

$$\dot{\hat{z}} = \Lambda \hat{z} + \Gamma y .$$

Indeed, the estimation error in the z-coordinates then evolves simply along the contracting dynamics

$$\frac{d}{dt}(\hat{z} - z) = \Lambda(\hat{z} - z)$$

and $\hat{x}$ is obtained from $\hat{z}$ by inverting the transformation. He proved that when the pair (A, C) is observable with $n_y = 1$, this is always possible via a linear stationary transformation $z = Tx$ with $n_z = n_x$, for any Hurwitz matrix Λ in $\mathbb{R}^{n_x \times n_x}$ with no common eigenvalues with A, and Γ any vector in $\mathbb{R}^{n_x}$ such that the pair (Λ, Γ) is controllable. This is based on the fact that the Sylvester equation

$$TA = \Lambda T + \Gamma C \tag{95}$$

admits in this case a solution that is unique and invertible.

Some researchers have therefore tried to reproduce Luenberger's original methodology on nonlinear systems, i.e., find a (uniformly injective) transformation into a Hurwitz form

$$\dot{z} = \Lambda z + \Gamma(y, t) \quad , \quad y = H(z, t) \tag{96}$$

with Λ Hurwitz, and associated observer

$$\dot{\hat{z}} = \Lambda \hat{z} + \Gamma(y, t) . \tag{97}$$

Unlike in Section 5, this procedure is not a linearization of the plant dynamics, since the output function H in (96) can be any nonlinear function (see Kazantzis & Kravaris, 1998, Remark 4). This crucial difference leads to far less restrictive conditions on the system, and because the corresponding observer (97) is a simple copy of the dynamics, it is not even necessary to have an explicit expression for H.

7.1. Autonomous systems

The case of autonomous nonlinear systems was first proposed and analyzed in a general context by Shoshitaishvili (1992). It was rediscovered later by Kazantzis and Kravaris (1998) who gave a local analysis close to an equilibrium point under conditions relaxed later on in Krener and Xiao (2003). The localness as well as most of the restrictive assumptions were then by-passed in Andrieu and Praly (2006) leading to the so-called *KKL-observers* or *nonlinear Luenberger observers*, which is also strongly related to the observer proposed in Kreisselmeier and Engel (2003). More precisely, in Andrieu and Praly (2006), the authors investigate the possibility of transforming an autonomous system (2) into a Hurwitz autonomous form (96) with $\Gamma(y, t) = \Gamma(y)$. This raises the question of finding, for some integer n_z, a C^1 function $T : \mathbb{R}^{n_x} \to \mathbb{R}^{n_z}$ verifying

$$\frac{dT}{dx}(x)f(x) = \Lambda T(x) + \Gamma(h(x)) \qquad \forall x \in \mathcal{X} \tag{98}$$

with Λ some Hurwitz matrix of dimension n_z and $\Gamma : \mathbb{R}^{n_y} \to \mathbb{R}^{n_z}$ some continuous function. The existence of such a transformation is shown for any Hurwitz matrix Λ with sufficiently negative eigenvalues and for some well-chosen functions Γ under the only assumption that the system is backward-complete[7] in $\mathcal{X}$ (Andrieu & Praly, 2006, Theorem 2). Of course, this is not enough since, as we saw in Theorem 3.1, it is required that T be uniformly injective on $\mathcal{X}$ to deduce from the estimate $\hat{z}$ of $T(x)$ an estimate of x. It is shown in Andrieu and Praly (2006, Theorem 3) that injectivity of T is achieved for almost any diagonal complex Hurwitz matrix Λ of dimension[8] $(n_x + 1)n_y$ and for any Γ verifying some growth condition, under a *backward-distinguishability* assumption.

Definition 7.1. Given an open set $\mathcal{O} \supset \mathcal{X}$, the system (2) is *backward $\mathcal{O}$-distinguishable* on $\mathcal{X}$ if for any $(x_{a,0}, x_{b,0}) \in \mathcal{X} \times \mathcal{X}$ such that $x_{a,0} \neq x_{b,0}$, and for any solution (x_a, x_b) to (2) initialized at $(x_{a,0}, x_{b,0})$, there exists $t < 0$ such that both x_a and x_b are defined in $\mathcal{O}$ on $[t, 0]$ and

$$h(x_a(t)) \neq h(x_b(t)).$$

In other words, two different states in $\mathcal{X}$ can be distinguished in $\mathcal{O}$ from the past values of the output. In the case where $\mathcal{X}$ is bounded, the result of Andrieu and Praly (2006) can be simplified as follows.

Theorem 7.1 (*Andrieu & Praly, 2006*). *Assume that $\mathcal{X}$ is bounded and there exists an open bounded subset $\mathcal{O}$ of $\mathbb{R}^{n_x}$ containing $cl(\mathcal{X})$ such that the system (2) is backward $\mathcal{O}$-distinguishable on $\mathcal{X}$. Then, there exists a strictly positive number ℓ and a set $\mathcal{R}$ of zero Lebesgue measure in $\mathbb{C}^{n_x+1}$ such that denoting*

$$\Omega = \left\{ \lambda \in \mathbb{C} \ : \ \mathfrak{R}(\lambda) < -\ell \right\},$$

for any $(\lambda_1, \ldots, \lambda_{n_x+1})$ in $\Omega^{n_x+1} \setminus \mathcal{R}$, there exists a C^1 map $T : \mathcal{O} \to \mathbb{R}^{(n_x+1) \times n_y}$ uniformly injective on $\mathcal{X}$ and verifying (98) with

$$\Lambda = \tilde{\Lambda} \otimes I_{n_y}, \qquad \Gamma(y) = (\tilde{\Gamma} \otimes I_{n_y}) y,$$

---

[7] Any solution exiting $\mathcal{X}$ in finite backward time must cross the boundary of $\mathcal{X}$. See Andrieu and Praly (2006, Definition 1).

[8] Separating the real/imaginary parts, the observer is thus of dimension $2(n_x + 1)n_y$ on $\mathbb{R}$.

and

$$\tilde{\Lambda} = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_{n_x+1} \end{pmatrix}, \qquad \tilde{\Gamma} = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}. \tag{99}$$

It is therefore possible to design an observer for an autonomous non-linear system (2) under the weak assumption of backward-distinguishability. Note that with a stronger assumption of strong differential observability of order m (see Definition 6.1), and still in a bounded set, it is also proved in Andrieu and Praly (2006, Theorem 4) that the injectivity of T is ensured for any m eigenvalues $\lambda_i = \ell \tilde{\lambda}_i$ with ℓ sufficiently large and $\tilde{\lambda}_i < 0$.

Assuming differential observability, it is shown in Andrieu (2014) that the observer is exponential in the sense of (6). Moreover, it can be shown that the observer obtained is robust since it satisfies an asymptotic gain property as in Eq. (10).

Note that Theorem 7.1 recommends to filter each output with $n_x + 1$ generic complex eigenvalues in Ω. Separating real and imaginary parts, this gives a real implementation of dimension $2(n_x + 1)$ for each output. Actually, ongoing works show that the result extends to any controllable real pair $(\tilde{\Lambda}, \tilde{\Gamma})$ with $\tilde{\Lambda} \in \mathbb{R}^{2n_x+1}$ diagonalizable and a generic choice of complex conjugate or real eigenvalues in Ω. Therefore, this allows a real implementation of dimension $2n_x + 1$ (instead of $2(n_x + 1)$) for each output and a more general choice of filters.

7.2. Time-varying/controlled systems

After first steps in Rotella and Zambettakis (2013), Trumpf (2007) for linear time-varying systems, an extension of the Luenberger design to nonlinear controlled systems was considered in Engel (2007), following the ideas of Kreisselmeier and Engel (2003). In Engel (2007), injectivity of a time-varying transformation T is proved only under a so-called "finite-complexity" assumption, originally introduced in Kreisselmeier and Engel (2003) for autonomous systems. Unfortunately, this property is very restrictive and hard to check. Besides, no indication about the needed dimension n_z is given. Those problems were then overcome more recently in Bernard and Andrieu (2019) with results of existence and injectivity of a time-varying transformation under more standard and constructive observability assumptions.

More precisely, consider $n_z \in \mathbb{N}$, a Hurwitz matrix $\Lambda \in \mathbb{R}^{n_z \times n_z}$ and a vector $\Gamma \in \mathbb{R}^{n_z \times n_y}$. In order to transform a general time-varying system (1) into the Hurwitz form (96) with[9] $\Gamma(y, t) = \Gamma y$, we need to find a transformation $T : \mathbb{R}^{n_x} \times [0, +\infty) \to \mathbb{R}^{n_z}$ such that for any x in $\mathcal{X}$ and any time $t \in [0, +\infty)$,

$$\frac{\partial T}{\partial x}(x, t) f(x, t) + \frac{\partial T}{\partial t}(x, t) = \Lambda T(x, t) + \Gamma h(x, t) . \tag{100}$$

According to Theorem 3.1, (97) then gives an observer for (1) if besides T becomes injective uniformly in time and in space at least after a certain time.

It is shown in Bernard and Andrieu (2019, Lemma 1) that a solution to (100) always exists if solutions to (1) initialized in $\mathcal{X}$ do not explode in finite backward time. Note that this latter assumption can always be made to hold if $\mathcal{X}$ is bounded by saturating f outside of $\mathcal{X}$. Then, still under a backward-distinguishability condition, but this time in finite time, "generic" injectivity is proved in Bernard and Andrieu (2019, Theorem 3).

Theorem 7.2 (*Bernard & Andrieu, 2019*). *Assume that $\mathcal{X}$ is bounded and the system (1) is backward-distinguishable in time t_d on $\mathcal{X}$, i.e. for any $t_0 \geq t_d$ and any pair of solutions (x_a, x_b) to (1) verifying $(x_a(t_0), x_b(t_0)) \in \mathcal{X} \times \mathcal{X}$ with $x_a(t_0) \neq x_b(t_0)$, there exists $s \in [t_0 - t_d, t_0]$ such that*

$$h(x_a(s), s) \neq h(x_b(s), s) .$$

Then there exists a set $\mathcal{R}$ of zero-Lebesgue measure in $\mathbb{C}^{n_x+1}$ such that for any $(\lambda_1, \ldots, \lambda_{n_x+1})$ in $\Omega^{n_x+1} \setminus \mathcal{R}$ with $\Omega = \{\lambda \in \mathbb{C}, \Re(\lambda) < 0\}$, there exists a C^1 map $T : \mathcal{X} \to \mathbb{R}^{(n_x+1) \times n_y}$ verifying (100) with

$$\Lambda = \tilde{\Lambda} \otimes I_{n_y}, \qquad \Gamma = \tilde{\Gamma} \otimes I_{n_y},$$

and $\tilde{\Lambda}, \tilde{\Gamma}$ defined in (99), such that $T(\cdot, t)$ is injective on $\mathcal{X}$ for $t > t_d$.

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 Lie derivatives up to a certain order is injective with respect to x (weak differential observability). Actually, under a strong differential observability assumption on (1), it is shown in Bernard and Andrieu (2019, Theorem 2) that, with sufficiently fast eigenvalues, n_z can be chosen according to the order of differential observability of each output.

7.3. Computation of the transformation

All in all, the main difficulty of the KKL observers lies in the computation of the map T, let alone its left-inverse $\mathcal{T}$. From a theoretical point of view, an injective solution $T : \mathcal{O} \to \mathbb{R}^{(n_x+1) \times n_y}$ to (98) shown to exist in Theorem 7.1, is written explicitly as

$$T(x) = \int_{-\infty}^{0} e^{-\Lambda \tau} \Gamma(h(\check{X}(x, \tau))) \, d\tau , \tag{101}$$

where $\check{X}(x, \tau)$ is the solution at time τ and initialized at x to modified dynamics $\dot{x} = \check{f}(x)$ with $f = \check{f}$ on $\mathcal{X}$. A similar expression exists for the time-varying context (see Bernard & Andrieu, 2019, Lemma 1) Unfortunately, the use of those explicit expression (101) is not easy since it necessitates to integrate backwards a differential equation at each time step. Several examples in Bernard (2019, Section 6.3) show how the map T can sometimes be computed without relying on the integral formulas. Otherwise, an approximation strategy was proposed in Marconi and Praly (2008) in the context of output regulation, and a learning-based strategy was presented in da Costa Ramos, Meglio, Morgenthaler, da Silva, and Bernard (2020) in order to learn offline the map T via neural networks.

Note that when the time dependence of (1) comes from an input u, namely we have (3), then the map $T(\cdot, t)$ depends implicitly on the whole past trajectory of u and is thus not computable offline. An alternative would be to keep the stationary transformation obtained for some constant value of u (for instance the drift system at $u \equiv 0$) and prove that the additional terms due to the presence of u do not prevent convergence. In the context of input-affine systems (57), this can be done under very strong assumptions of uniform observability and strong differential observability of order n_x of f, namely in the same context as Gauthier and Bornard (1981) allowing to transform the dynamics in the triangular form (56b) (see Bernard & Andrieu, 2019, Theorem 4).

8. About the implementation of an observer

8.1. The left-inversion problem

The previous sections have shown that it is possible, under certain conditions, to build an observer by transforming the plant dynamics into a favorable form for which an observer is known. It follows that the dynamics of the plant and of the observer are not expressed in the same coordinates and often evolve in spaces of different dimensions. In order to obtain an estimate for the system state, i.e., compute $\mathcal{T}$ (or even sometimes to write the observer dynamics), it is necessary to invert the injective transformation T. But even if T is stationary, this inversion is difficult in practice, when an analytical expression for a

---

[9] We could have considered the more general Hurwitz form (96), but taking Γ linear and stationary is sufficient to obtain satisfactory results.

global inverse is not available. Indeed, the inversion usually relies on the resolution of a minimization problem of the type

$$\hat{x} = \min_{x \in \mathcal{X}} |T(x) - \hat{z}|$$

with a heavy computational cost and the risk of local minima.

In the case in which T is a diffeomorphism on an open set $\mathcal{O}$ containing $\mathcal{X}$, one may hope to avoid this minimization by implementing the observer (4) directly in the x-coordinates with

$$\dot{\hat{x}} = \left(\frac{dT}{dx}(\hat{x}) \right)^{-1} \mathcal{F}(T(\hat{x}), y, t) . \tag{102}$$

This is done for instance in Astolfi and Praly (2013, 2017), Bernard, Praly, and Andrieu (2018), Ciccarella, Mora, and Germani (1993), Gauthier et al. (1992), Maggiore and Passino (2003), Mora, Germani, and Manes (2000), Zeitz (1987). But even in this apparently simple case, the observer (102) must be treated carefully. Indeed, although x remains in $\mathcal{O}$ where the Jacobian of T is invertible, there is no guarantee that $\hat{x}$ will remain in $\mathcal{O}$ as well, in particular during transient behaviors where the solutions $t \mapsto \hat{z}(t)$ of (4) may leave the image set $T(\mathcal{O})$. For instance it is shown in Bernard et al. (2018) that the observer proposed in Gauthier et al. (1992) admits solutions that explode in finite-time. This can be solved either by modifying the observer dynamics $\mathcal{F}$ to constrain the state in $T(\mathcal{O})$ as in Astolfi and Praly (2013, 2017), Maggiore and Passino (2003), but some convexity properties are needed to guarantee that the convergence is preserved; or by keeping the observer dynamics $\mathcal{F}$ unchanged but modifying T outside of $\mathcal{X}$ so that $T(\mathcal{O}) = \mathbb{R}^m$, as proposed in Bernard et al. (2018). But there is no systematic method to perform this image extension in practice.

Now, when $n_z > n_x$, i.e. the Jacobian is rectangular and T is at best an injective immersion, (102) cannot be implemented. A possible solution was then proposed in Bernard et al. (2018). The main idea consists in extending T into a surjective diffeomorphism T_e, by adding $n_z - n_x$ fictitious states w to x, such that

$$T_e(x, 0) = T(x) \qquad \forall x \in \mathcal{X}$$

and implementing

$$\overbrace{\begin{bmatrix} \dot{\hat{x}} \\ \dot{\hat{w}} \end{bmatrix}} = \left(\frac{dT_e}{d(x,w)}(\hat{x}, \hat{w}) \right)^{-1} \mathcal{F}(T_e(\hat{x}, \hat{w}), y, t) .$$

Another path is to implement Newton-like or gradient-like algorithms in parallel to (4), namely add dynamics of the type

$$\dot{\hat{x}} = \mu K(\hat{x})(\hat{z} - T(\hat{x}))$$

where $K(x) = \frac{dT}{dx}(x)^\top$ or $K(x) = \left(\frac{dT}{dx}(x)^\top \frac{dT}{dx}(x) \right)^{-1} \frac{dT}{dx}(x)^\top$ as proposed in Astolfi and Possieri (2019), Menini, Possieri, and Tornambe (2017) for high gain observers. But the convergence of the obtained observer is only local. Similarly, continuation algorithms which "follow" the minimum of $x \mapsto |T(x) - \hat{z}|^2$ can be used under a convexity assumption like in Hammouri, Ahmed, and Othman (2018), but again the convergence is only local.

Finally, one may mention an alternative route followed in Bernard and Marconi (2020), which consists in implementing, in parallel to (102), an independent practical observer which is used to reset $\hat{x}$ whenever it gets too close to the boundary of $\mathcal{O}$. Here, only a finite-number of approximate left-inversions of T need to be carried out and the resulting convergence is global.

8.2. Taking into account state constraints

In many applications, the state of the plant is known to evolve in a given set or manifold $\mathcal{X}$, describing the set of interest of the variables for the considered set of initial conditions $\mathcal{X}_0$. Examples include positive systems (Farina & Rinaldi, 2011), quaternion or other Lie groups (Bonnabel, 2007; Hamel & Mahony, 2006; Martin & Salaün, 2010; Salcudean, 1991), general manifolds admitting symmetries (Aghannan

& Rouchon, 2003). When the observer is implemented in other co-ordinates $z = T(x)$, also z is thus known to remain in the manifold $\mathcal{Z} = T(\mathcal{X})$. But unless the observer is designed in a way that incorporates these constraints, $\hat{x}$ (resp. $\hat{z}$) could in principle leave the set $\mathcal{X}$ (resp. $\mathcal{Z}$) for some time, because of transient dynamics or disturbances. A typical example is the peaking phenomenon occurring in high-gain observers (Khalil & Praly, 2013). During these excursions outside of $\mathcal{X}$ (resp. $\mathcal{Z}$), the estimate is known to be inaccurate and cannot be exploited. Worse, the observer maps $\mathcal{F}$ and $\mathcal{T}$ may not be defined or this may lead to instability in the context of output-feedback. Besides, it is advantageous from a numerical implementation point of view to specify beforehand the range of each variable and to have this range as small as possible to increase precision. This is not possible if the estimates are not guaranteed to remain in (or close to) $\mathcal{X}$ (resp. $\mathcal{Z}$).

A simple idea to address some of these issues is to saturate or to project $\hat{x}$ into $\mathcal{X}$, while preserving the observer's inner dynamics $\mathcal{F}$, namely change only $\mathcal{T}$. However, the inner observer state $\hat{z}$ can still venture far from its "nominal" set $\mathcal{Z}$. A natural way to overcome these issues is to constrain $\hat{z}$ in a given set $\hat{\mathcal{Z}} \supset \mathcal{Z}$. This has nevertheless to be done with care to preserve the observer convergence and performance. In control problems, combining convergence with invariance constraints is typically handled through control barrier functions (Ames, Xu, Grizzle, & Tabuada, 2016), leading to the so-called *safety control*. However, those methods do not transpose easily to the estimation field because the state, and thus the value of the Lyapunov function, are unknown. Instead, two different strategies may be followed :

- either the observer dynamics $\mathcal{F}$ are designed from the start to ensure simultaneously convergence of the estimation error and invariance of $\hat{\mathcal{Z}}$;
- or the observer dynamics $\mathcal{F}$ are first designed along general unconstrained methods, and then modified in a second step to ensure the estimate remains in a particular set $\hat{\mathcal{Z}}$.

Along the first path, we may mention invariant observers (Aghannan & Rouchon, 2003) in which the observer dynamics are designed so that to respect some state space symmetries. Similar ideas have been developed in the filtering context in Bonnabel (2007), Martin and Salaün (2010) or in Hamel and Mahony (2006) for state estimation for rotation matrices and more generally for Lie groups (Barrau & Bonnabel, 2016; Lageman, Trumpf, & Mahony, 2009; Mahony, Hamel, & Pflimlin, 2008; Mahony, Hamel, & Trumpf, 2020; Phogat & Chang, 2020). This is also the path which is followed in Back and Astolfi (2008), Bonnabel, Astolfi, and Sepulchre (2011) for positive systems in which the positive orthant is made invariant along the observer dynamics.

On the other hand, along the second path, projection-based solutions have been proposed for high-gain observers in Maggiore and Passino (2003) for Lur'e type systems in Tanwani, Brogliato, and Prieur (2014), for Kalman filters (Berkane, Tayebi, & Teel, 2018; Simon, 2010), within the framework of *projected dynamical systems* in Heemels, Camlibel, and Heertjes (2020) for exponentially convergent observers with linear correction terms, and in Astolfi, Bernard, Postoyan, and Marconi (2021) for any type of observer with quadratic Lyapunov functions under general nonlinear convex constraints. Actually, instead of projecting $\hat{z}$ or $\mathcal{F}$, Astolfi et al. (2021) also proposes to consider the constraint information as an additional pseudo-measurement and add a correction term in $\mathcal{F}$ in a way that does not increase the derivative of the Lyapunov function (to preserve convergence) but that ensure $\hat{z}$ remains in $\hat{\mathcal{Z}}$.

8.3. Tuning and characterization of performances

8.3.1. Analysis of the behavior of observers

From the practical point of view, it is in general important not only to guarantee that an observer is asymptotic, but also that its

performances are satisfactory in terms of transient response (convergent speed) and steady-state behavior in presence of measurement noise and possibly model uncertainties (Astolfi, Postoyan, & Nešić, 2020). For linear systems, a certain number of different linear tools are available for both synthesis and design, such as Laplace and Fourier transform, H_∞ or $\mathcal{L}_2$ gain designs (based on the minimization on an input–output gain and the bounded real lemma) or Kalman filtering stochastic approach, see, e.g. Kalman and Bucy (1961), just to cite a few. However, it is also very well known that certain limitations cannot be overcome with standard linear tools, see, e.g. Seron, Braslavsky, and Goodwin (2012).

For nonlinear observers, one may gain in flexibility in using nonlinear/hybrid tools, but it has to be stressed that it is also much harder, in general, to analyze the desired aforementioned performances characteristics. As a matter of fact, the tool which is typically employed is the one of Lyapunov functions, allowing both synthesis and design, with the performances that are characterized in terms of classical ISS bounds, see, e.g., Shim and Liberzon (2016). For the majority of the observer design techniques presented in these survey, such a Lyapunov analysis can be easily developed and it is not hard to verify that most of observers are indeed *robust* according to the definition given in the introduction, see inequality (10). However, it is also important to remind that such bounds are in general conservative, because Lyapunov analysis reflects the (worst-case) asymptotic gain of the estimation error (10), failing to capture any frequency information. In this sense, the analysis of the behavior in presence of high-frequency measurement noise developed in the context of high-gain observers (Astolfi et al., 2016, 2018) is a remarkable attempt in such a direction. On the other hand, Lyapunov analysis may also be used as a design tool to improve certain properties, as shown in Andrieu, Prieur, Tarbouriech, and Zaccarian (2016), where a hybrid scheme is employed to reduce the peaking phenomenon of high-gain observers.

8.3.2. Use of variable gains and multi-observers

A typical limitation in observer design is the requirement of having fast convergence and a small influence of the measurement noise in steady-state. Intuitively, a "large-gain" (in terms of magnitude) is needed to achieve fast convergence, but a "small-gain" is desired to minimize the effect of the measurement noise on the observer-dynamics. Variable gains can therefore be employed to improve performance either in terms of sensitivity to measurement noise, or in increasing the speed of convergence of the observer. This can be achieved with continuous-time gain adaptation, for instance, by means of Riccati or Lyapunov based designs such as in Kalman filters (Kalman & Bucy, 1961), Kalman-like observers, see, e.g., Boizot et al. (2010), Chitour (2002), and in high-gain observers (Sanfelice & Praly, 2011b), or employing switching strategies between different gains in a hybrid/switching context, see, e.g., Ahrens and Khalil (2009), Clement (2001), Efimov, Polyakov, Levant, and Perruquetti (2018).

In the same spirit of employing different gains, another methodology consists in using a bank of observers and to estimate online the "best-one", i.e. the one providing the more accurate estimate. Such an approach has been investigated, for instance, in Astolfi, Postoyan, and Nešić (2020), Chong, Nešić, Postoyan, and Kuhlmann (2015), Esfandiari and Shakarami (2019), Krener and Duarte (1996), Postoyan, Hamid, and Daafouz (2015).

Finally, we recall also Engel and Kreisselmeier (2002), Li and Sanfelice (2015), Menold, Findeisen, and Allgower (2003), Ríos and Teel (2016) where different observers with different gains are combined to obtain finite-time estimation.

8.3.3. Use of filters

A different route that it is often used in observer design to improve the performance in presence of measurement noise is based on the use of filters in order to reduce its effect. Although filtering the sole measured output may introduce phase-lag in the estimates, the use of filters for the output-error estimation error may improve the sensitivity properties of the observer while preserving its asymptotic convergence in nominal "noise-free" conditions (see Astolfi, Zaccarian, & Jungers, 2021). A well-known example of such an approach is the so-called Proportional-Integral Observer (PI-observer) in which both the output estimation error and its integral are typically employed in the observer design, enlightening remarkable improvements, see, e.g., Beale and Shafai (1989), Busawon and Kabore (2001), Niemann, Stoustrup, Shafai, and Beale (1995). The so-called "low-power high-gain observer" approach investigated in Astolfi and Marconi (2015), Astolfi et al. (2018) founds on the same principle of augmenting the relative-degree between the measurement noise and the state estimate, in order to increase the filtering properties at high-frequencies. In this case, the filters are directly inherited from the cascade structure, see Section 6.5. The use of more general low-pass filters have also been investigated in the context of high-gain observers (see, e.g., Astolfi, Zaccarian, & Jungers, 2021; Tréangle, Farza, & M'Saad, 2019) and homogeneous observers (see, e.g., Jbara, Levant, & Hanan, 2021) presented in Sections 6.2 and 6.3.

Based on similar ideas, we recall also the recent idea in which the observer is designed by using integrals of the output combined with a moving-horizon strategy, see Menini, Possieri, and Tornambè (2019), Menini, Possieri, and Tornambe (2021).

Finally, recent developments showed how the use of saturation (also denoted as Stubborn observer) or dead-zone with adapting thresholds may be employed to improve the sensitivity with respect to sporadic impulsive noise and biased noise, see, e.g., Alessandri and Zaccarian (2018), Astolfi, Jungers, and Zaccarian (2021), Cocetti, Tarbouriech, and Zaccarian (2018).

8.4. Output sampling and continuous–discrete observers

In many engineering applications, the output y is sampled and only available intermittently at specific time instances. A first natural idea is to approximately discretize the continuous-time observer designed for the continuous-time plant model, or to design a discrete-time observer for an approximate discretized model of the plant. This typically leads to practical observers, and attention must be payed to the discretization method in order not to destroy the convergence properties of the observers, see, e.g., Arcak and Nešić (2004).

When the sampling time is not sufficiently small for such discretizations, or when asymptotic convergence is sought, specific *continuous–discrete observers* need to be designed, considering the time elapsed between successive sampling events. Generally, they consist either of a continuous-time predictor made of the copy of the plant dynamics with discrete updates of the estimate at each sampling time, or of a continuous-time observer with a sample-and-hold innovation term.

In this second case, convergence of such observers was first shown under the assumption that the sampling period is sufficiently small (this strategy is also known in literature as *emulation* approach), for linear observers in Sur and Paden (1997), high-gain observers in Ahmed-Ali and Lamnabhi-Lagarrigue (2011), Deza et al. (1992), Farza et al. (2014), and contractive observers in Manchester (2018). Using the hybrid formalism proposed in Goebel, Sanfelice, and Teel (2012), general conditions in the context of network communications have been proposed in Postoyan and Nesic (2011), Postoyan and Nešić (2012) for robust nonlinear observers and reduced-order observers. More sophisticated designs have been developed for arbitrary constant sampling period (Dinh, Andrieu, Nadri, & Serres, 2015; Raff & Allgöwer, 2007), and more generally for sporadic measurements, using sample-and-hold strategies (Ahmed-Ali, Burlion, Lamnabhi-Lagarrigue, & Hann, 2014; Raff, Kogel, & Allgower, 2008) or discrete updates of the estimate at the sampling times (Ahmed-Ali, Postoyan, & Lamnabhi-Lagarrigue, 2009; Bernard & Sanfelice, 2022; Etienne, Hetel, Efimov, & Petreczky, 2017; Ferrante, Gouaisbaut, Sanfelice, & Tarbouriech, 2016; Li, Phillips, &

Sanfelice, 2017; Mazenc, Andrieu, & Malisoff, 2015; Sferlazza, Tarbouriech, & Zaccarian, 2019), among others.

Actually, when we know that the time elapsed between two successive sampling events always belongs to a certain set, the observer can be modeled by an autonomous hybrid system (Goebel et al., 2012). Via this formalism, it is shown in Bernard and Sanfelice (2022) that the discrete updates of the estimate at the sampling time, should be designed as for a discrete observer of an equivalent discrete-time system, made of the plant dynamics sampled at the sampling times. For linear systems, this allows to use the gains of a discrete Kalman filter (Bernard & Sanfelice, 2022), or recover the matrix inequalities obtained for instance in Ferrante et al. (2016), Raff and Allgöwer (2007). For Lipschitz nonlinear systems, such tools may still be used when the dynamics are in the convex hull of linear dynamics (Dinh et al., 2015).

Last but not least, one should also mention the literature of event-triggered observers when the sampling times of the output can be triggered based on a policy defined by the observer or by a *smart-sensor* (i.e. a sensor equipped with computational capacities) that decides when to transmit the information. We briefly recall (Etienne, Di Gennaro, & Barbot, 2017; Shi, Chen, & Darouach, 2016; Shi, Chen, & Shi, 2014) concerning the use of event-triggering strategies and Andrieu, Nadri, Serres, and Vivalda (2015), Rabehi, Meslem, and Ramdani (2021) for self-triggering strategies. See also Petri, Postoyan, Astolfi, Nešić, and Heemels (2021) and references therein.

8.5. Adaptive observers

It may happen that some parameters of the plant dynamics (1) are unknown or not precisely known. That is why a lot of researchers devoted their attention to the design of observers for systems of the form

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

where $\theta \in \mathbb{R}^{n_\theta}$ is a set of parameters characterizing the plant's dynamics and which is unknown: we talk about *adaptive observers*. Within this context, two main approaches may be pursued.

Joint estimation: The first one consists in considering the constant parameters θ as an additional state with a pure integrator dynamics, i.e., $\dot{\theta} = 0$ (state augmentation). In such a case, one can recast the observation problem for the extended state (x, θ) and re-apply, for instance, one of the technique reviewed in this survey. In doing so, joint-observation requires the appropriate observability assumption to hold for the full state (x, θ). See, for instance, Afri, Andrieu, Bako, and Dufour (2016) for the case of KKL observers.

Disjoint estimation: A second approach is to consider the estimation problem disjoint from the "identification" problem of the parameters and follows typical adaptive control techniques Åström and Wittenmark (2013), Ioannou and Sun (1995), Krstic, Kokotovic, and Kanellakopoulos (1995), Ortega, Aranovskiy, Pyrkin, Astolfi, and Bobtsov (2020). In doing so, one is typically interested in the convergence of the state estimation error without any requirement/guarantee of the correct estimation of the parameter θ. Such an approach has been extensively studied over the last decades with "ad-hoc" solutions based on the particular structure/properties of the plant's dynamics. See, for instance, in the following (non-exhaustive) list of references: Bastin and Gevers (1988), Carroll and Lindorff (1973), Cho and Rajamani (1997), Farza, M'Saad, Maatoug, and Kamoun (2009), Luders and Narendra (1973), Marino (1990), Tyukin, Steur, Nijmeijer, and Van Leeuwen (2013), Zhang (2002).

We remark that as discussed in Besancon and Țiclea (2017), under some persistence of excitation assumptions, the two approaches may be interchangeable, namely when joint estimation can be employed, then a disjoint strategy can be pursued, and *vice-versa*.

8.6. Disturbance observers and extended state observers

In some control applications, such as disturbance estimation and output feedback stabilization, it is commonly assumed that the system (1) is transformed into a system of dimension n_z of the form

$$\dot{z} = Az + \varphi(t, z, v_z), \qquad y = Cz$$

with (A, φ, C) being a triplet in prime form (see (55)), and v_z being an unknown input to be estimated. By considering v_z as a fictitious state (as often done in adaptive control, see Section 8.5), the dynamics of which is modeled as $\dot{v}_z = 0$, one can then design a differentiator (following the prescriptions of Section 6) of order $n_z + 1$ in order to have an approximate estimation of v_z and at the same time increase the performances of the observers in terms of asymptotic gain with respect to v_z. Such an approach is often denoted as *Extended State Observer* (shortly ESO), see, e.g. Guo and Zhao (2011) and references therein, or *extended high-gain observer*, see, e.g., Nazrulla and Khalil (2010). Design philosophy and convergence proofs follow the same principles presented in Sections 6.2 and 6.3. Concerning other classes of *disturbance observers*, we refer to the survey Radke and Gao (2006) and references therein.

8.7. Use of observers in feedback control

One of the main motivations for the use of an observer in practical problems is given by feedback control when only a part of the state can be measured. In practice, a common approach consists in designing first a state-feedback design, and then replace the state by its estimate provided by an observer. For linear systems, it is not difficult to show that the poles of the closed-loop dynamics coincide with the union of the poles given by the state-feedback design and those of the observer. Such a property is also known as *separation-principle* (see, e.g., Chen, 1984).

For nonlinear systems the use of observers in feedback control is more involved due to possible nonlinear phenomena such as finite-escape time (see Esfandiari & Khalil, 1987). Furthermore, the control itself may destroy the convergence properties of the observers in lack of uniform observability properties, e.g., Sacchelli, Brivadis, Andrieu, Serres, and Gauthier (2020). In such a context, high-gain observers played a fundamental role in establishing the first "nonlinear separation principle" in output feedback control, see, e.g., Andrieu and Praly (2009), Atassi and Khalil (1999), Shim and Heon (2000), Teel and Praly (1994). For triangular normal form, the use of homogeneous techniques is also well investigated, see, e.g., Andrieu et al. (2008), Levant (2003, 2005), Qian (2005a, 2005b), Qian and Lin (2006), Yang and Lin (2004).

In practice, most of the time, the combination of an asymptotic observer and a stabilizing state-feedback design is not sufficient to guarantee the overall stability of an equilibrium of the closed-loop system and stronger properties such as uniform observability and ISS (or iISS) of the feedback/observer are usually needed, see, e.g., Andrieu and Praly (2009). Overall, at the time of the publication of this article, the problem of output-feedback control is far from being completely understood and is an open field of research.

9. Conclusions

In this survey we have reviewed the main methods to design an asymptotic observer for finite dimensional nonlinear dynamical systems, that have been developed in the last six decades. They are summarized in Table 1. These methods have been classified based on the observability property available on the given dynamical system. Such an approach allows to obtain a general and unifying point of view on this topic. Although in its sixties, this research theme is still young and active. This is illustrated by the recent publications to which we have referred in the course of this paper. However, some difficulties remain unsolved and we do not yet have a complete solution to the

observation problem in the nonlinear context. For example, it remains difficult to characterize the existence of observers for systems with inputs.

Also, even if the theory presented all along the paper may guarantee the existence of an observer, its construction and practical implementation for (possibly large-dimensional) systems can be difficult. For instance, KKL observers provide a very general answer to the observation problem, but the explicit construction of the transformation T remains problematic apart from particular applications (Bernard & Praly, 2020; Malaizé et al., 2012). Also, the inversion of the map T in most of the designs still remains a numerical obstacle and leads practitioners to prefer observers written in the plant coordinates, such as the EKF, even at the price of local convergence. Moreover, the optimal selection of the observer's parameters (in order to guarantee some desired performances) remains generically an open problem.

This explains why state observers for nonlinear systems are not systematically implemented in real-life applications. However, many examples show their relevance (see for instance (Ali, Hoang, Hussain, & Dochain, 2015), see also the *soft sensors* industry) and we hope that this survey will participate in their dissemination. It is also important to highlight that in many practical applications, discrete-time observers, such as Extended Kalman Filters (and its modifications, such as Unscented EKF, Extended EFK or Ensemble EKF), moving horizon observers or particle filters, are directly designed and implemented. An exhaustive overview of discrete-time approaches is out of the scope of this article but of great interest for applications.

Some theoretical problems were not addressed during this study and may constitute future research themes. For instance, designing observers for more complex dynamical systems (such as hybrid, of large/infinite dimension, or with delayed measurements) is not covered by this survey. The use of observers for feedback control in the presence of observability singularities is also of great interest and a rich open research problem.

Finally, an approach that seems promising is the joint use of these methods in combination with techniques resulting from data analysis and supervised learning methods. Along these lines, some preliminary results can be found in da Costa Ramos et al. (2020), Manchester (2018).

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

Afri, C., Andrieu, V., Bako, L., & Dufour, P. (2016). State and parameter estimation: A nonlinear luenberger observer approach. *IEEE Transactions on Automatic Control*, *62*(2), 973–980.

Aghannan, N., & Rouchon, P. (2003). An intrinsic observer for a class of Lagrangian systems. *IEEE Transactions on Automatic Control*, *48*(6), 936–945.

Ahmed-Ali, T., Burlion, L., Lamnabhi-Lagarrigue, F., & Hann, C. (2014). A sampled-data observer with time-varying gain for a class of nonlinear systems with sampled-measurements. In *IEEE conference on decision and control* (pp. 316–321).

Ahmed-Ali, T., & Lamnabhi-Lagarrigue, F. (2011). High gain observer design for some networked control systems. *IEEE Transactions on Automatic Control*, *57*(4), 995–1000.

Ahmed-Ali, T., Postoyan, R., & Lamnabhi-Lagarrigue, F. (2009). Continuous–discrete adaptive observers for state affine systems. *Automatica*, *45*(12), 2986–2990.

Ahrens, J. H., & Khalil, H. K. (2009). High-gain observers in the presence of measurement noise: A switched-gain approach. *Automatica*, *45*(4), 936–943.

Alamir, M. (2007). Nonlinear observers and applications. In *Lecture notes in control and information sciences: vol. 363*, (pp. 139–179). Springer.

Alessandri, A., & Rossi, A. (2015). Increasing-gain observers for nonlinear systems: Stability and design. *Automatica*, *57*, 180–188.

Alessandri, A., & Zaccarian, L. (2018). Stubborn state observers for linear time-invariant systems. *Automatica*, *88*, 1–9.

Ali, J. M., Hoang, N. H., Hussain, M. A., & Dochain, D. (2015). Review and classification of recent observers applied in chemical process systems. *Computers & Chemical Engineering*, *76*, 27–41.

Ames, A. D., Xu, X., Grizzle, J. W., & Tabuada, P. (2016). Control barrier function based quadratic programs for safety critical systems. *IEEE Transactions on Automatic Control*, *62*(8), 3861–3876.

Andrieu, V. (2014). Convergence speed of nonlinear luenberger observers. *SIAM Journal on Control and Optimization*, *52*(5), 2831–2856.

Andrieu, V., Astolfi, D., & Bernard, P. (2021). Observer design via interconnections of second-order mixed sliding-mode/linear differentiators. *International Journal of Robust and Nonlinear Control*, *31*(9), 3631–3657.

Andrieu, V., Besançon, G., & Serres, U. (2013). Observability necessary conditions for the existence of observers. In *IEEE conference on decision and control* (pp. 4442–4447).

Andrieu, V., Jayawardhana, B., & Praly, L. (2016). Transverse exponential stability and applications. *IEEE Transactions on Automatic Control*, *61*(11), 3396–3411.

Andrieu, V., Nadri, M., Serres, U., & Vivalda, J.-C. (2015). Self-triggered continuous–discrete observer with updated sampling period. *Automatica*, *62*, 106–113.

Andrieu, V., & Praly, L. (2006). On the existence of a kazantzis–kravaris / luenberger observer. *SIAM Journal on Control and Optimization*, *45*(2), 432–456.

Andrieu, V., & Praly, L. (2009). A unifying point of view on output feedback designs for global asymptotic stabilization. *Automatica*, *45*(8), 1789–1798.

Andrieu, V., Praly, L., & Astolfi, A. (2008). Homogeneous approximation, recursive observer design, and output feedback. *SIAM Journal on Control and Optimization*, *47*(4), 1814–1850.

Andrieu, V., Praly, L., & Astolfi, A. (2009a). High gain observers with updated gain and homogeneous correction terms. *Automatica*, *45*(2), 422–428.

Andrieu, V., Praly, L., & Astolfi, A. (2009b). Homogeneity in the bi-limit as a tool for observer and feedback design. In *IEEE conference on decision and control*.

Andrieu, V., Praly, L., & Astolfi, A. (2009c). Homogeneity in the bi-limit as a tool for observer and feedback design. In *Proceedings of the 48th IEEE conference on decision and control (CDC) held jointly with 2009 28th Chinese control conference* (pp. 1050–1055). IEEE.

Andrieu, V., Prieur, C., Tarbouriech, S., & Zaccarian, L. (2016). A hybrid scheme for reducing peaking in high-gain observers for a class of nonlinear systems. *Automatica*, *72*, 138–146.

Arcak, M., & Kokotović, P. (2001). Nonlinear observers: a circle criterion design and robustness analysis. *Automatica*, *37*(12), 1923–1930.

Arcak, M., & Nešić, D. (2004). A framework for nonlinear sampled-data observer design via approximate discrete-time models and emulation. *Automatica*, *40*(11), 1931–1938.

Astolfi, D., Alessandri, A., & Zaccarian, L. (2021). Stubborn and dead-zone redesign for nonlinear observers and filters. *IEEE Transactions on Automatic Control*, *66*(2), 667–682.

Astolfi, D., Bernard, P., Postoyan, R., & Marconi, L. (2021). Constrained state estimation for nonlinear systems: a redesign approach based on convexity. *IEEE Transactions on Automatic Control*, 1.

Astolfi, D., & Marconi, L. (2015). A high-gain nonlinear observer with limited gain power. *IEEE Transactions on Automatic Control*, *60*(11), 3059–3064.

Astolfi, D., Marconi, L., Praly, L., & Teel, A. (2016). Sensitivity to high-frequency measurement noise of nonlinear high-gain observers. *IFAC-PapersOnLine*, *49*(18), 862–866.

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

Astolfi, A., & Ortega, R. (2003). Immersion and invariance : a new tool for stabilization and adaptive control of nonlinear systems. *IEEE Transactions on Automatic Control*, *48*(4), 590–606.

Astolfi, D., & Possieri, C. (2019). Design of local observers for autonomous nonlinear systems not in observability canonical form. *Automatica*, *103*, 443–449.

Astolfi, D., Postoyan, R., & Nešić, D. (2020). Uniting observers. *IEEE Transactions on Automatic Control*, *65*(7), 2867–2882.

Astolfi, D., & Praly, L. (2013). Output feedback stabilization for SISO nonlinear systems with an observer in the original coordinate. In *IEEE conference on decision and control* (pp. 5927–5932).

Astolfi, D., & Praly, L. (2017). Integral action in output feedback for multi-input multi-output nonlinear systems. *IEEE Transactions on Automatic Control*, *62*(4), 1559–1574.

Astolfi, D., Zaccarian, L., & Jungers, M. (2021). On the use of low-pass filters in high-gain observers. *Systems & Control Letters*, *148*, Article 104856.

Åström, K. J., & Wittenmark, B. (2013). *Adaptive control*. Courier Corporation.

Atassi, A. N., & Khalil, H. K. (1999). A separation principle for the stabilization of a class of nonlinear systems. *IEEE Transactions on Automatic Control*, *44*(9), 1672–1687.

Back, J., & Astolfi, A. (2008). Design of positive linear observers for positive linear systems via coordinate transformations and positive realizations. *SIAM Journal on Control and Optimization*, *47*(1), 345–373.

Back, J., & Seo, J. H. (2004). Immersion of non-linear systems into linear systems up to output injection : Characteristic equation approach. *Internation Journal of Control*, *77*(8), 723–734.

Barrau, A., & Bonnabel, S. (2016). The invariant extended Kalman filter as a stable observer. *IEEE Transactions on Automatic Control*, *62*(4), 1797–1812.

Bastin, G., & Gevers, M. R. (1988). Stable adaptive observers for nonlinear time-varying systems. *IEEE Transactions on Automatic Control*, *33*(7), 650–658.

Beale, S., & Shafai, B. (1989). Robust control system design with a proportional integral observer. *International Journal of Control*, *50*(1), 97–111.

Berkane, S., Tayebi, A., & Teel, A. R. (2018). Hybrid constrained estimation for linear time-varying systems. In *57th IEEE conference on decision and control*.

Bernard, P. (2019). *Lecture notes on control and information sciences, Observer design for nonlinear systems. Vol. 4*. Springer.

Bernard, P., & Andrieu, V. (2019). Luenberger observers for non autonomous nonlinear systems. *IEEE Transactions on Automatic Control*, *64*(1), 270–281.

Bernard, P., & Marconi, L. (2020). Hybrid implementation of observers in plant's coordinates with a finite number of approximate inversions and global convergence. *Automatica*, *111*, Article 108654.

Bernard, P., Mimmo, N., & Marconi, L. (2021). On the semi-global stability of an EK-like filter. *IEEE Control Systems Letters*, *5*(5), 1771–1776.

Bernard, P., & Praly, L. (2020). Estimation of position and resistance of a sensorless PMSM : a nonlinear luenberger approach for a non-observable system. *IEEE Transactions on Automatic Control*.

Bernard, P., Praly, L., & Andrieu, V. (2017). Observers for a non-Lipschitz triangular form. *Automatica*, *82*, 301–313.

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

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

Bernard, P., & Sanfelice, R. G. (2022). Observer design for hybrid dynamical systems with approximately known jump times. *Automatica*, (to appear) URL HAL https://hal.archives-ouvertes.fr/hal-02187411.

Besançon, G. (1999). Further results on high gain observers for nonlinear systems. In *IEEE conference on decision and control*. Vol. 3 (pp. 2904–2909).

Besançon, G., & Bornard, G. (1997). On characterizing a class of observer forms for nonlinear systems. In *European control conference*.

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

Besancon, G., & Ticlea, A. (2007). An immersion-based observer design for rank-observable nonlinear systems. *IEEE Transactions on Automatic Control*, *52*(1), 83–88.

Besancon, G., & Țiclea, A. (2017). On adaptive observers for systems with state and parameter nonlinearities. *IFAC-PapersOnLine*, *50*(1), 15416–15421.

Blondel, P., Postoyan, R., Raël, S., Benjamin, S., & Desprez, P. (2018). Nonlinear circle-criterion observer design for an electrochemical battery model. *IEEE Transactions on Control Systems Technology*, *27*(2), 889–897.

Boizot, N., Busvelle, E., & Gauthier, J.-P. (2010). An adaptive high-gain observer for nonlinear systems. *Automatica*, *46*(9), 1483–1488.

Bonnabel, S. (2007). Left-invariant extended Kalman filter and attitude estimation. In *2007 46th IEEE Conference on decision and control* (pp. 1027–1032). IEEE.

Bonnabel, S., Astolfi, A., & Sepulchre, R. (2011). Contraction and observer design on cones. In *2011 50th IEEE conference on decision and control and European control conference* (pp. 7147–7151). IEEE.

Bonnabel, S., & Slotine, J.-J. (2015). A contraction theory-based analysis of the stability of the deterministic extended Kalman filter. *IEEE Transactions on Automatic Control*, *60*(2), 565–569.

Bornard, G., Celle-Couenne, F., & Gilles, G. (1995). *Nonlinear systems, modeling and estimation 1* (pp. 173–216). US: Springer.

Bornard, G., Couenne, N., & Celle, F. (1989). Regularly persistent observers for bilinear systems. In J. Descusse, M. Fliess, A. Isidori, & D. Leborgne (Eds.), *New trends in nonlinear control theory* (pp. 130–140). Berlin, Heidelberg: Springer Berlin Heidelberg.

Bossane, D., Rakotopara, D., & Gauthier, J.-P. (1989). Local and global immersion into linear systems up to output injection. In *IEEE conference on decision and control* (pp. 2000–2004).

Boutayeb, M., Rafaralahy, H., & Darouach, M. (1997). Convergence analysis of the extended Kalman filter used as an observer for nonlinear deterministic discrete-time systems. *IEEE Transactions on Automatic Control*, *42*(4), 581–586.

Brivadis, L., Andrieu, V., Chabanon, É., Gagniere, E., Lebaz, N., & Serres, U. (2020). New dynamical observer for a batch crystallization process based on solute concentration. *Journal of Process Control*, *87*, 17–26.

Bullinger, E., & Allgower, F. (1997). An adaptive high-gain observer for nonlinear systems. In *Proceedings of the 36th IEEE conference on decision and control. Vol. 5* (pp. 4348–4353). IEEE.

Busawon, K. K., & Kabore, P. (2001). Disturbance attenuation using proportional integral observers. *International Journal of Control*, *74*(6), 618–627.

Carroll, R., & Lindorff, D. (1973). An adaptive observer for single-input single-output linear systems. *IEEE Transactions on Automatic Control*, *18*(5), 428–435.

Chalanga, A., Kamal, S., Fridman, L. M., Bandyopadhyay, B., & Moreno, J. A. (2016). Implementation of super-twisting control: Super-twisting and higher order sliding-mode observer-based approaches. *IEEE Transactions on Industrial Electronics*, *63*(6), 3677–3685.

Chen, C.-T. (1984). *Linear system theory and design*. CBS College Publishing.

Chitour, Y. (2002). Time-varying high-gain observers for numerical differentiation. *IEEE Transactions on Automatic Control*, *47*(9), 1565–1569.

Cho, Y. M., & Rajamani, R. (1997). A systematic approach to adaptive observer synthesis for nonlinear systems. *IEEE Transactions on Automatic Control*, *42*(4), 534–537.

Chong, M. S., Nešić, D., Postoyan, R., & Kuhlmann, L. (2015). Parameter and state estimation of nonlinear systems using a multi-observer under the supervisory framework. *IEEE Transactions on Automatic Control*, *60*(9), 2336–2349.

Ciccarella, G., Mora, M. Dalla, & Germani, A. (1993). A luenberger-like observer for nonlinear systems. *International Journal of Control*, *57*(3), 537–556.

Clement, G. (2001). An interpolation method for gain-scheduling. In *Proceedings of the 40th IEEE conference on decision and control (Cat. No. 01CH37228). Vol. 2* (pp. 1310–1315). IEEE.

Cocetti, M., Tarbouriech, S., & Zaccarian, L. (2018). On dead-zone observers for linear plants. In *2018 Annual American control conference* (pp. 5138–5143). IEEE.

Cruz-Zavala, E., & Moreno, J. A. (2019). Levant's arbitrary-order exact differentiator: A Lyapunov approach. *IEEE Transactions on Automatic Control*, *64*(7), 3034–3039.

da Costa Ramos, L., Meglio, F. Di, Morgenthaler, V., da Silva, L. F. Figueira, & Bernard, P. (2020). Numerical design of Luenberger observers for nonlinear systems. In *IEEE conference on decision and control* (pp. 5435–5442).

Deza, F., Busvelle, E., Gauthier, J.-P., & Rakotopara, D. (1992). High gain estimation for nonlinear systems. *Systems & Control Letters*, *18*(4), 295–299.

Dinh, T. N., Andrieu, V., Nadri, M., & Serres, U. (2015). Continuous-discrete time observer design for Lipschitz systems with sampled measurements. *IEEE Transactions on Automatic Control*, *60*(3), 787–792.

Efimov, D., Polyakov, A., Levant, A., & Perruquetti, W. (2018). Convergence acceleration for observers by gain commutation. *International Journal of Control*, *91*(9), 2009–2018.

Efimov, D., Raïssi, T., Chebotarev, S., & Zolghadri, A. (2013). Interval state observer for nonlinear time varying systems. *Automatica*, *49*(1), 200–205.

Emelyanov, S. V., Korovin, S. K., Nikitin, S. V., & Nikitina, M. G. (1989). Observers and output differentiators for nonlinear systems. *Doklady Akademii Nauk*, *306*, 556–560.

Engel, R. (2007). Nonlinear observers for Lipschitz continuous systems with inputs. *International Journal of Control*, *80*(4), 495–508.

Engel, R., & Kreisselmeier, G. (2002). A continuous-time observer which converges in finite time. *IEEE Transactions on Automatic Control*, *47*(7), 1202–1204.

Esfandiari, F., & Khalil, H. K. (1987). Observer-based design of uncertain systems: recovering state feedback robustness under matching conditions. In *Proceedings of allerton conference*, Monticello, IL pp. 97—106.

Esfandiari, K., & Shakarami, M. (2019). Bank of high-gain observers in output feedback control: Robustness analysis against measurement noise. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*.

Etienne, L., Di Gennaro, S., & Barbot, J.-P. (2017). Periodic event-triggered observation and control for nonlinear Lipschitz systems using impulsive observers. *International Journal of Robust and Nonlinear Control*, *27*(18), 4363–4380.

Etienne, L., Hetel, L., Efimov, D., & Petreczky, M. (2017). Observer synthesis under time-varying sampling for Lipschitz nonlinear systems. *Automatica*, *85*, 433–440.

Fan, X., & Arcak, M. (2003). Observer design for systems with multivariable monotone nonlinearities. *Systems & Control Letters*, *50*(4), 319–330.

Farina, L., & Rinaldi, S. (2011). *Positive linear systems: theory and applications. Vol. 50*. John Wiley & Sons.

Farza, M., M'Saad, M., Fall, M. L., Pigeon, E., Gehan, O., & Busawon, K. (2014). Continuous-discrete time observers for a class of MIMO nonlinear systems. *IEEE Transactions on Automatic Control*, *59*(4), 1060–1065.

Farza, M., M'Saad, M., Maatoug, T., & Kamoun, M. (2009). Adaptive observers for nonlinearly parameterized class of nonlinear systems. *Automatica*, *45*(10), 2292–2299.

Farza, M., Oueder, M., Abdennour, R. B., & M'Saad, M. (2011). High gain observer with updated gain for a class of MIMO nonlinear systems. *International Journal of Control*, *84*(2), 270–280.

Ferrante, F., Gouaisbaut, F., Sanfelice, R. G., & Tarbouriech, S. (2016). State estimation of linear systems in the presence of sporadic measurements. *Automatica*, *73*, 101–109.

Fliess, M. (1990). Generalized controller canonical form for linear and nonlinear dynamics. *IEEE Transactions on Automatic Control*, *35*(9), 994–1001.

Fliess, M., & Kupka, I. (1983). A finiteness criterion for nonlinear input-output differential systems. *SIAM Journal on Control and Optimization*, *21*(5), 721–728.

Floquet, T., & Barbot, J. P. (2007). Super twisting algorithm-based step-by-step sliding mode observers for nonlinear systems with unknown inputs. *International Journal of Systems Science*, *38*(10), 803–815.

Franquiz, F. J., Muñoz, J. D., Udrea, B., & Balas, M. J. (2018). Optimal range observability maneuvers of a spacecraft formation using angles-only navigation. *Acta Astronautica*, *153*, 337–348.

Gauthier, J., & Bornard, G. (1981). Observability for any *u*(*t*) of a class of nonlinear systems. *IEEE Transactions on Automatic Control*, *26*(4), 922–926.

Gauthier, J.-P., Hammouri, H., & Othman, S. (1992). A simple observer for nonlinear systems applications to bioreactors. *IEEE Transactions on Automatic Control*, *37*(6), 875–880.

Gauthier, J.-P., & Kupka, I. (2001). *Deterministic observation theory and applications*. Cambridge University Press.

Goebel, R., Sanfelice, R. G., & Teel, A. R. (2012). *Hybrid dynamical systems: modeling stability, and robustness*. Princeton, NJ: Princeton University Press.

Gouzé, J. L., Rapaport, A., & Hady-Sadok, M. Z. (2000). Interval observers for uncertain biological systems. *Ecological Modelling*, *133*, 45–56.

Guay, M. (2002). Observer linearization by output-dependent time-scale transformations. *IEEE Transactions on Automatic Control*, *47*(10), 1730–1735.

Guo, B.-Z., & Zhao, Z.-L. (2011). On the convergence of an extended state observer for nonlinear systems with uncertainty. *Systems & Control Letters*, *60*(6), 420–430.

Hamel, T., & Mahony, R. (2006). Attitude estimation on SO(3) based on direct inertial measurements. In *Proceedings 2006 IEEE international conference on robotics and automation, 2006. ICRA 2006* (pp. 2170–2175). IEEE.

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

Hammouri, H., & De Leon Morales, J. (1990). Observer synthesis for state-affine systems. In *IEEE conference on decision and control* pp. 784–785.

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

Hammouri, H., Vu, H. G., & Yahoui, H. (2013). Local observer for infinitesimally observable nonlinear systems. *International Journal of Control*, *86*(4), 579–590.

Heemels, W. P. M. H., Camlibel, M. K., & Heertjes, M. F. (2020). Oblique projected dynamical systems and incremental stability under state constraints. *IEEE Control Systems Letters*, *4*(4), 1060–1065.

Hinson, B. T., Binder, M. K., & Morgansen, K. A. (2013). Path planning to optimize observability in a planar uniform flow field. In *2013 American control conference* (pp. 1392–1399).

Hinson, B. T., & Morgansen, K. A. (2013). Observability optimization for the nonholonomic integrator. In *2013 American control conference* (pp. 4257–4262).

Hu, G.-D. (2006). Observers for one-sided Lipschitz non-linear systems. *IMA Journal of Mathematical Control and Information*, *23*(4), 395–401.

Ibrir, S. (2009). LPV approach to continuous and discrete nonlinear observer design. In *Proceedings of the 48h IEEE Conference on decision and control (CDC) held jointly with 2009 28th chinese control conference* (pp. 8206–8211). IEEE.

Ibrir, S., & Diop, S. (1999). On continuous time differentiation observers. In *1999 European control conference* (pp. 2402–2406). IEEE.

Ioannou, P. A., & Sun, J. (1995). *Robust adaptive control*. Prentice Hall, Inc.

Isidori, A. (2017). *Lectures in feedback design for multivariable systems*. Springer.

Jazwinski, A. H. (1970). *Stochastic processes and filtering theory*. Academic Press.

Jbara, A., Levant, A., & Hanan, A. (2021). Filtering homogeneous observers in control of integrator chains. *International Journal of Robust and Nonlinear Control*, *31*(9), 3658–3685.

Jennings, L. S., Fernando, T. L., & Trinh, H. M. (2011). Existence conditions for functional observability from an eigenspace perspective. *IEEE Transactions on Automatic Control*, *56*(12), 2957–2961.

Jouan, P. (2003). Immersion of nonlinear systems into linear systems modulo output injection. *SIAM Journal on Control and Optimization*, *41*(6), 1756–1778.

Kalman, R. E., & Bucy, R. S. (1961). New results in linear filtering and prediction theory. *Journal of Basic Engineering*, *108*, 83–95.

Kang, W., Krener, A. J., Xiao, M., & Xu, L. (2013). A survey of observers for nonlinear dynamical systems. In *Data assimilation for atmospheric, oceanic and hydrologic applications (Vol. II)* (pp. 1–25). Springer.

Karagiannis, D., & Astolfi, A. (2005). Nonlinear observer design using invariant manifolds and applications. In *IEEE conference on decision and control and european control conference*.

Kazantzis, N., & Kravaris, C. (1998). Nonlinear observer design using Lyapunov's auxiliary theorem. *Systems & Control Letters*, *34*, 241–247.

Keller, H. (1987). Nonlinear observer by transformation into a generalized observer canonical form. *Internation Journal of Control*, *46*(6), 1915–1930.

Khalil, H. K. (2017). Cascade high-gain observers in output feedback control. *Automatica*, *80*, 110–118.

Khalil, H. K., & Praly, L. (2013). High-gain observers in nonlinear feedback control. *International Journal of Robust and Nonlinear Control*, *24*.

Khan, A., Xie, W., Bo, Z., & Liu, L.-W. (2021). A survey of interval observers design methods and implementation for uncertain systems. *Journal of the Franklin Institute*.

Kou, S. R., Elliott, D. L., & Tarn, T. J. (1975). Exponential observers for nonlinear dynamic systems. *Information and Control*, *29*(3), 204–216.

Kravaris, C. (2016). Functional observers for nonlinear systems. *IFAC Symposium on Nonlinear Control Systems NOLCOS*, *49*(18), 505–510.

Kreisselmeier, G., & Engel, R. (2003). Nonlinear observers for autonomous lipshitz continuous systems. *IEEE Transactions on Automatic Control*, *48*(3), 451–464.

Krener, A. J. (2003). The convergence of the extended Kalman filter. In *Directions in mathematical systems theory and optimization* (pp. 173–182). Springer.

Krener, A. J., & Duarte, A. (1996). A hybrid computational approach to nonlinear estimation. In *Proceedings of 35th IEEE conference on decision and control. Vol. 2* (pp. 1815–1819). IEEE.

Krener, A. J., & Isidori, A. (1983). Linearization by output injection and nonlinear observers. *Systems & Control Letters*, *3*, 47–52.

Krener, A. J., & Respondek, W. (1985). Nonlinear observers with linearizable dynamics. *SIAM Journal on Control and Optimization*, *23*(2), 197–216.

Krener, A. J., & Xiao, M. (2003). Nonlinear observer design in the Siegel domain. *SIAM Journal on Control and Optimization*, *41*(3), 932–953.

Krishnamurthy, R., Khorrami, F., & Chandra, R. S. (2003). Global high-gain-based observer and backstepping controller for generalized output-feedback canonical form. *IEEE Transactions on Automatic Control*, *48*(12), 2277–2283.

Krstic, M., Kokotovic, P. V., & Kanellakopoulos, I. (1995). *Nonlinear and adaptive control design*. John Wiley & Sons, Inc.

Lageman, C., Trumpf, J., & Mahony, R. (2009). Gradient-like observers for invariant dynamics on a Lie group. *IEEE Transactions on Automatic Control*, *55*(2), 367–377.

Lei, H., Wei, J., & Lin, W. (2005). A global observer for observable autonomous systems with bounded solution trajectories. In *Proceedings of the 44th IEEE conference on decision and control* (pp. 1911–1916). IEEE.

Levant, A. (2001). Higher-order sliding modes and arbitrary-order exact robust differentiation. In *Proceedings of the European control conference* (pp. 996–1001).

Levant, A. (2003). Higher-order sliding modes, differentiation and output-feedback control. *International Journal of Control*, *76*(9–10), 924–941.

Levant, A. (2005). Homogeneity approach to high-order sliding mode design. *Automatica*, *41*(5), 823–830.

Li, Y., Phillips, S., & Sanfelice, R. G. (2017). Robust distributed estimation for linear systems under intermittent information. *IEEE Transactions on Automatic Control*, *63*(4), 973–988.

Li, Y., & Sanfelice, R. G. (2015). A finite-time convergent observer with robustness to piecewise-constant measurement noise. *Automatica*, *57*, 222–230.

Lin, H., Zhai, G., & Antsaklis, P. J. (2003). Set-valued observer design for a class of uncertain linear systems with persistent disturbance. In *American control conference*.

Lohmiller, Winfried, & Slotine, Jean-Jacques E. (1998). On contraction analysis for non-linear systems. *Automatica*, *34*(6), 683–696.

Lopez-Ramirez, F., Polyakov, A., Efimov, D., & Perruquetti, W. (2018). Finite-time and fixed-time observer design: Implicit Lyapunov function approach. *Automatica*, *87*, 52–60.

Luders, G., & Narendra, K. S. (1973). An adaptive observer and identifier for a linear system. *IEEE Transactions on Automatic Control*, *18*(5), 496–499.

Luenberger, D. (1964). Observing the state of a linear system. *IEEE Transactions on Military Electronics*, *8*, 74–80.

Maggiore, M., & Passino, K. M. (2003). A separation principle for a class of non-UCO systems. *IEEE Transactions on Automatic Control*, *48*(7), 1122–1133.

Mahony, R., Hamel, T., & Pflimlin, J.-M. (2008). Nonlinear complementary filters on the special orthogonal group. *IEEE Transactions on Automatic Control*, *53*(5), 1203–1218.

Mahony, R., Hamel, T., & Trumpf, J. (2020). Equivariant systems theory and observer design. arXiv preprint arXiv:2006.08276.

Malaizé, J., Praly, L., & Henwood, N. (2012). Globally convergent nonlinear observer for the sensorless control of surface-mount permanent magnet synchronous machines. In *IEEE conference on decision and control*.

Manchester, I. R. (2018). Contracting nonlinear observers: Convex optimization and learning from data. In *2018 Annual American control conference* (pp. 1873–1880). IEEE.

Marconi, L., & Praly, L. (2008). Uniform practical nonlinear output regulation. *IEEE Transactions on Automatic Control*, *53*, 1184–1202.

Marino, R. (1990). Adaptive observers for single output nonlinear systems. *IEEE Transactions on Automatic Control*, *35*(9), 1054–1058.

Martin, Philippe, & Salaün, Erwan (2010). Generalized multiplicative extended Kalman filter for aided attitude and heading reference system. In *AIAA guidance, navigation, and control conference* (pp. 8300).

Mazenc, F., Andrieu, V., & Malisoff, M. (2015). Design of continuous-discrete observers for time-varying nonlinear systems. *Automatica*, *57*, 135–144.

Mazenc, F., & Bernard, O. (2011). Interval observers for linear time-invariant systems with disturbances. *Automatica*, *47*(1), 140–147.

McShane, E. J. (1934). Extension of range of functions. *American Mathematical Society. Bulletin*, *40*(12), 837–842.

Menini, L., Possieri, C., & Tornambe, A. (2017). A "practical" observer for nonlinear systems. In *IEEE conference on decision and control* (pp. 3015–3020).

Menini, L., Possieri, C., & Tornambè, A. (2019). Observers for linear systems by the time integrals and moving average of the output. *IEEE Transactions on Automatic Control*, *64*(12), 4859–4874.

Menini, L., Possieri, C., & Tornambe, A. (2021). On the use of the time-integrals of the output in observer design for non-linear autonomous system. *IEEE Transactions on Automatic Control*.

Menold, P. H., Findeisen, R., & Allgower, F. (2003). Finite time convergent observers for nonlinear systems. In *42nd IEEE international conference on decision and control (IEEE Cat. No. 03CH37475). Vol. 6* (pp. 5673–5678). IEEE.

Mora, M. Dalla, Germani, A., & Manes, C. (2000). Design of state observers from a drift-observability property. *IEEE Transactions on Automatic Control*, *45*(8), 1536–1540.

Moreno, J. (2021). Levant's differentiator by interconnection of low order blocks. In *Third IFAC conference on modelling, identification and control of nonlinear systems*.

Nazrulla, S., & Khalil, H. K. (2010). Robust stabilization of non-minimum phase nonlinear systems using extended high-gain observers. *IEEE Transactions on Automatic Control*, *56*(4), 802–813.

Niemann, H. H., Stoustrup, J., Shafai, B., & Beale, S. (1995). LTR design of proportional-integral observers. *International Journal of Robust and Nonlinear Control*, *5*(7), 671–693.

Ortega, R., Aranovskiy, S., Pyrkin, A. A., Astolfi, A., & Bobtsov, A. A. (2020). New results on parameter estimation via dynamic regressor extension and mixing: Continuous and discrete-time cases. *IEEE Transactions on Automatic Control, 66*(5), 2265–2272.

Perruquetti, W., Floquet, T., & Moulay, E. (2008). Finite-time observers: application to secure communication. *IEEE Transactions on Automatic Control, 53*(1), 356–360.

Petri, E., Postoyan, R., Astolfi, D., Nešić, D., & Heemels, W. P. M. H. (2021). Event-triggered observer design for linear systems. In *60th IEEE conference on decision and control.*

Phogat, K. S., & Chang, D. E. (2020). Invariant extended Kalman filter on matrix Lie groups. *Automatica, 114,* Article 108812.

Plestan, F., & Glumineau, A. (1997). Linearization by generalized input-output injection. *Systems & Control Letters, 31,* 115–128.

Postoyan, R., Hamid, M. H. A., & Daafouz, J. (2015). A multi-observer approach for the state estimation of nonlinear systems. In *2015 54th IEEE conference on decision and control* (pp. 1793–1798). IEEE.

Postoyan, R., & Nesic, D. (2011). A framework for the observer design for networked control systems. *IEEE Transactions on Automatic Control, 57*(5), 1309–1314.

Postoyan, R., & Nešić, D. (2012). On emulated nonlinear reduced-order observers for networked control systems. *Automatica, 48*(4), 645–652.

Praly, L. (2001). On observers with state independent error Lyapunov function. *IFAC Proceedings Volumes, 34*(6), 1349–1354.

Praly, L. (2003). Asymptotic stabilization via output feedback for lower triangular systems with output dependent incremental rate. *IEEE Transactions on Automatic Control, 48*(6), 1103–1108.

Praly, L. (2015). Observers for nonlinear systems. In J. Baillieul, & T. Samad (Eds.), *Encyclopedia of systems and control.* Springer.

Praly, L., & Jiang, Z.-P. (1993). Stabilization by output feedback for systems with ISS inverse dynamics. *Systems & Control Letters, 21*(1), 19–33.

Proychev, T. P., & Mishkov, R. L. (1993). Transformation of nonlinear systems in observer canonical form with reduced dependency on derivatives of the input. *Automatica, 29*(2), 495–498.

Qian, C. (2005). A homogeneous domination approach for global output feedback stabilization of a class of nonlinear systems. In *Proceedings of the American control conference.*

Qian, C. (2005). A homogeneous domination approach for global output feedback stabilization of a class of nonlinear systems. In *IEEE American control conference* (pp. 4708–4715).

Qian, C., & Lin, W. (2006). Recursive observer design, homogeneous approximation, and nonsmooth output feedback stabilization of nonlinear systems. *IEEE Transactions on Automatic Control, 51*(9).

Rabehi, D., Meslem, N., & Ramdani, N. (2021). Finite-gain $\mathcal{L}_1$ event-triggered interval observers design for continuous-time linear systems. *International Journal of Robust and Nonlinear Control, 31*(9), 4131–4153.

Radke, A., & Gao, Z. (2006). A survey of state and disturbance observers for practitioners. In *American control conference* (pp. 6–11).

Raff, T., & Allgöwer, F. (2007). Observers with impulsive dynamical behavior for linear and nonlinear continuous-time systems. In *IEEE conference on decision and control* (pp. 4287–4292).

Raff, T., Kogel, M., & Allgower, F. (2008). Observer with sample-and-hold updating for Lipschitz nonlinear systems with nonuniformly sampled measurements. In *2008 American control conference* (pp. 5254–5257).

Raghavan, S., & Hedrick, J. K. (1994). Observer design for a class of nonlinear systems. *International Journal of Control, 59*(2), 515–528.

Rajamani, R. (1998). Observers for Lipschitz nonlinear systems. *IEEE Transactions on Automatic Control, 43*(3), 397–401.

Rapp, K., & Nyman, P.-O. (2004). Stability properties of the discrete-time extended Kalman filter. *IFAC Symposium on Nonlinear Control Systems, 37*(13), 1377–1382.

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

Respondek, W., Pogromski, A., & Nijmeijer, H. (2004). Time scaling for observer design with linearizable error dynamics. *Automatica, 40,* 277–285.

Ríos, H., & Teel, A. R. (2016). A hybrid observer for fixed-time state estimation of linear systems. In *2016 IEEE 55th conference on decision and control* (pp. 5408–5413). IEEE.

Rotella, F., & Zambettakis, I. (2013). On functional observers for linear time-varying systems. *IEEE Transactions on Automatic Control, 58*(5), 1354–1360.

Rubio Scola, I., Besançon, G., & Georges, D. (2018). Optimizing Kalman optimal observer for state affine systems by input selection. *Automatica, 93,* 224–230.

Rudolph, J., & Zeitz, M. (1994). A block triangular nonlinear observer normal form. *Systems & Control Letters, 23*(1), 1–8.

Sacchelli, L., Brivadis, L., Andrieu, V., Serres, U., & Gauthier, J.-P. (2020). Dynamic output feedback stabilization of non-uniformly observable dissipative systems. *IFAC-PapersOnLine, 53*(2), 4923–4928.

Salaris, P., Cognetti, M., Spica, R., & Giordano, P. R. (2019). Online optimal perception-aware trajectory generation. *IEEE Transactions on Robotics, 35*(6), 1307–1322.

Salcudean, S. (1991). A globally convergent angular velocity observer for rigid body motion. *IEEE Transactions on Automatic Control, 36*(12), 1493–1497.

Sanfelice, R. G., & Praly, L. (2011a). Convergence of nonlinear observers on $\mathbb{R}^n$ with a Riemannian metric (Part I). *IEEE Transactions on Automatic Control, 57*(7), 1709–1722.

Sanfelice, R. G., & Praly, L. (2011b). On the performance of high-gain observers with gain adaptation under measurement noise. *Automatica, 47,* 2165–2176.

Sanfelice, R. G., & Praly, L. (2015). Convergence of nonlinear observers on $\mathbb{R}^n$ with a Riemannian metric (part II). *IEEE Transactions on Automatic Control, 61*(10), 2848–2860.

Seron, M. M., Braslavsky, J. H., & Goodwin, G. C. (2012). *Fundamental limitations in filtering and control.* Springer Science & Business Media.

Sferlazza, A., Tarbouriech, S., & Zaccarian, L. (2019). Time-varying sampled-data observer with asynchronous measurements. *IEEE Transactions on Automatic Control, 64*(2), 869–876.

Shi, D., Chen, T., & Darouach, M. (2016). Event-based state estimation of linear dynamic systems with unknown exogenous inputs. *Automatica, 69,* 275–288.

Shi, D., Chen, T., & Shi, L. (2014). Event-triggered maximum likelihood state estimation. *Automatica, 50*(1), 247–254.

Shim, H., & Heon, J. H. Jin (2000). Non-linear output feedback stabilization on a bounded region of attraction. *International Journal of Control, 73*(5), 416–426.

Shim, H., & Liberzon, D. (2016). Nonlinear observers robust to measurement disturbances in an ISS sense. *IEEE Transactions on Automatic Control, 61*(1), 48–61.

Shim, H., & Praly, L. (2003). Remarks on equivalence between full order and reduced order nonlinear observers. In *42nd IEEE international conference on decision and control. Vol. 6* (pp. 5837–5840). IEEE.

Shoshitaishvili, A. (1992). On control branching systems with degenerate linearization. *IFAC Symposium on Nonlinear Control Systems,* 495–500.

Simon, D. (2010). Kalman filtering with state constraints: a survey of linear and nonlinear algorithms. *IET Control Theory & Applications, 4*(8), 1303–1318.

Song, Y., & Grizzle, J. W. (1992). The extended Kalman filter as a local asymptotic observer for nonlinear discrete-time systems. *American Control Conference, 29,* 3365–3369.

Sontag, E. D. (2008). Input to state stability: Basic concepts and results. In *Nonlinear and optimal control theory* (pp. 163–220). Springer.

Sontag, E. D., & Wang, Y. (1997). Output-to-state stability and detectability of nonlinear systems. *Systems & Control Letters, 29*(5), 279–290.

Sur, J., & Paden, B. (1997). Observers for linear systems with quantized output. In *Annual American control conference* (pp. 3012–3016).

Tanwani, A., Brogliato, B., & Prieur, C. (2014). Stability and observer design for lur'e systems with multivalued, nonmonotone, time-varying nonlinearities and state jumps. *SIAM Journal on Control and Optimization, 52*(6), 3639–3672.

Teel, A. R. (2016). Further variants of the astolfi/marconi high-gain observer. In *2016 American control conference* (pp. 993–998). IEEE.

Teel, A., & Praly, L. (1994). Global stabilizability and observability imply semi-global stabilizability by output feedback. *Systems & Control Letters, 22*(5), 313–325.

Thau, F. E. (1973). Observing the state of non-linear dynamic systems. *International Journal of Control, 17*(3), 471–479.

Tornambe, A. (1989). Use of asymptotic observers having high gains in the state and parameter estimation. In *IEEE Conference on decision and control. Vol. 2.* (pp. 1791–1794).

Tréangle, C., Farza, M., & M'Saad, M. (2019). Filtered high gain observer for a class of uncertain nonlinear systems with sampled outputs. *Automatica, 101,* 197–206.

Trumpf, J. (2007). Observers for linear time-varying systems. *Linear Algebra and its Applications, 425,* 303–312.

Tsinias, J. (1989). Observer design for nonlinear systems. *Systems & Control Letters, 13*(2), 135–142.

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

van den Berg, F. W. J., Hoefsloot, H. C. J., Boelens, H. F. M., & Smilde, A. K. (2000). Selection of optimal sensor position in a tubular reactor using robust degree of observability criteria. *Chemical Engineering Science, 55*(4), 827–837.

Vijayaraghavan, K. (2014). Generalised observer design for dissipative Lipschitz nonlinear systems in the presence of measurement noise. *International Journal of Control, 87*(11), 2273–2285.

Wang, L., Astolfi, D., Marconi, L., & Su, H. (2017). High-gain observers with limited gain power for systems with observability canonical form. *Automatica, 75,* 16–23.

Yang, B., & Lin, W. (2004). Homogeneous observers, iterative design, and global stabilization of high-order nonlinear systems by smooth output feedback. *IEEE Transactions on Automatic Control, 49*(7), 1069–1080.

Zeitz, M. (1984). Observability canonical (phase-variable) form for nonlinear time-variable systems. *International Journal of Systems Science, 15*(9), 949–958.

Zeitz, M. (1987). The extended luenberger observer for nonlinear systems. *Systems & Control Letters, 9*(2), 149–156.

Zemouche, A., Boutayeb, M., & Bara, G. I. (2008). Observers for a class of Lipschitz systems with extension to H^∞ performance analysis. *Systems & Control Letters, 57*(1), 18–27.

Zemouche, A., Zhang, F., Mazenc, F., & Rajamani, R. (2018). High-gain nonlinear observer with lower tuning parameter. *IEEE Transactions on Automatic Control, 64*(8), 3194–3209.

Zhang, Q. (2002). Adaptive observer for multiple-input-multiple-output (MIMO) linear time-varying systems. *IEEE Transactions on Automatic Control, 47*(3), 525–529.

Zimmer, G. (1994). State observation by on-line minimization. *International Journal of Control, 60*(4), 595–606.