

# Reduced-Order Nonlinear Observers via Contraction Analysis and Convex Optimization

Bowen Yi, Ruigang Wang, and Ian R. Manchester

**Abstract**—In this paper, we propose a new approach to design globally convergent reduced-order observers for nonlinear control systems via contraction analysis and convex optimization. Despite the fact that contraction is a concept naturally suitable for state estimation, the existing solutions are either local or relatively conservative when applying to physical systems. To address this, we show that this problem can be translated into an off-line search for a coordinate transformation after which the dynamics is (transversely) contracting. The obtained sufficient condition consists of some easily verifiable differential inequalities, which, on one hand, identify a very general class of “detectable” nonlinear systems, and on the other hand, can be expressed as computationally efficient convex optimization, making the design procedure more systematic. Connections with some well-established approaches and concepts are also clarified in the paper. Finally, we illustrate the proposed method with several numerical and physical examples, including polynomial, mechanical, electromechanical and biochemical systems.

**Index Terms**—state observer, nonlinear system, contraction analysis, convex optimization

## I. INTRODUCTION

Online state estimation of dynamical systems is of both practical and theoretical importance, since some system states are important for the purpose of control or monitoring but often unavailable due to technological or cost constraints. Apart from the control community, similar problems also arise in many other fields, *e.g.*, time-series prediction, machine learning and signal processing.

For linear systems, there are comprehensive results developed for observer design, see [25] for the asymptotic case, and recent papers [10, 33] achieving finite-time convergence. In contrast, constructive nonlinear observer is a more challenging task, which has been intensively studied for several decades in the control community, producing more than six research monographs on this topic [4, 5, 8, 13, 18, 32]. Despite many efforts, there are few tools available to design globally or semi-globally convergent observers for nonlinear *control* (a.k.a. time-varying or non-autonomous) systems, with notable exception of high-gain observers [18] and recently results on Kazantzis-Kravaris-Luenberger (KKL) observers [6]. To the best of our knowledge, these two methods are the most general and systematic frameworks for nonlinear observer design at present. In high-gain observers, a triangular form of the system dynamics with strong differential observability

is required to achieve semi-global convergence by high-gain injection, which, however, is widely recognized as harmful in engineering practice with the deleterious effect of amplification of unavoidable high-frequency noise [19]. KKL observers are motivated by Luenberger’s initial idea by adopting a coordinate change in order to get a linear stable error dynamics. Such an idea was extended to the nonlinear context in [17, 39], later elaborated in [2, 20], and finally applied to non-autonomous systems in [6], which provides an elegant theoretical framework to state estimation. From the numerical perspective, the last step in KKL observers involves *online* computation of an inverse mapping of smoothing operators, which can be formalized as non-convex online optimization, making it a daunting task to be solved in real time, particularly for high-dimensional systems. Besides, there are some recently developed observer design tools applicable to different particular classes of nonlinear systems, *e.g.*, immersion and invariance (I&I) observer [4], and parameter estimation-based observer [34, 47].

In this paper, we aim to provide a novel constructive approach to nonlinear observers via contraction analysis, a concept which was introduced to the control community in [22], and may historically date back to [21]. Contraction, also known as incremental exponential stability, studies the convergence between any pairs of solutions rather than a particular one, by studying the differential dynamics. Its extensions to constructive nonlinear feedback are studied in [28], see also [41] for its application in system identification. As clearly illustrated in [22], observer design is one of the initial motivations for contraction analysis. Indeed, using contraction analysis in observer design is not a new idea, and has similarities to the seminal work [24], where a coordinate change is involved in order to get a stable error dynamics. In [40], a technique similar to contraction analysis was used for constructive nonlinear stochastic observers. More recently, some necessary conditions on the existence of asymptotically convergent observers in the *original coordinate* for general nonlinear autonomous systems are proposed in [38], which is motivated by the study of contracting flows in Riemannian manifolds. It shows that the existence of a convergent observer for autonomous systems implies that the system vector field is strictly geodesically monotonic, tangentially to the output function level sets. However, from a constructive viewpoint the result in [38] can only guarantee local convergence, unless imposing a totally geodesic assumption, which is generally difficult to verify for a given system. There are also some results using contraction to design observers for specific classes of systems, *e.g.*, [1] for mechanical systems and [45] for port-

This work was supported by Australian Research Council.

The authors are with Australian Centre for Field Robotics and Sydney Institute for Robotics and Intelligent Systems, The University of Sydney, Sydney, NSW 2006, Australia. (email: {bowen.yi, ruigang.wang, ian.manchester}@sydney.edu.au)

Hamiltonian systems, whose nature is intrinsically local.

With these considerations, the following questions arise.

- 1) What can be the sufficient conditions to design a *globally* convergent observer for nonlinear control systems by studying contracting flows? Though an answer has already been partially reported in our previous work [26], we here consider reduced-order observer design and show that coordinate transformation provides sufficient degrees of freedom, in this way sharpening the existing results.
- 2) Is it possible to translate the obtained condition into offline convex optimization? When designing globally convergent observers, it is quite common to encounter partial differential equations (PDEs) or inequalities, thus requiring the user familiar with system structures or invoking some physical insights. To circumvent the difficulty, we give a convex representation in the paper, making it possible to take advantage of powerful numerical tools.
- 3) Is the set of systems identified in the paper large enough? To answer this question, we give links between the proposed method and some well-established concepts, or methods, including in the linear system context, strong differential observability, I&I observers and transverse contraction. As illustrations, we will show how to use the proposed framework to design reduced-order observers for several benchmark physical systems—the magnetic levitation (MagLev) model, the cart-pendulum system, and a biological reactor—and an academic example to verify the results with convex constraints.

The remainder of the paper is organized as follows. In Section II we formulate the problem of reduced-order observer design and give some preliminaries on contraction analysis. In Section III we start with the basic case that the observer shares the dimension of the unknown states, and present our general design framework. It is followed by the convex representation, as well as the generalized result removing the dimension constraint. In Section IV, we present further discussions on the proposed method and compare it with some existing methods. Some examples with simulation results are given in Section V. The paper is wrapped up with some concluding remarks.

*Caveat.* An abridged conference version of the paper can be found in [48].

## II. PROBLEM FORMULATION AND PRELIMINARIES

### A. Notations and symbols

*Notations.* All mappings are smooth enough, and when clear from the context, the arguments of mappings and the subindex of operators will be omitted. Given a matrix  $P(x)$  and a vector field  $f(x)$  with proper dimensions, we define the directional derivative as  $\partial_f P(x) = \sum_j \frac{\partial P(x)}{\partial x_j} f_j(x)$ . For a full-rank matrix  $g \in \mathbb{R}^{n \times m}$  ( $m < n$ ), we denote the generalized inverse as  $g^\dagger = [g^\top g]^{-1} g^\top$  and  $g^\perp$  a full-rank left-annihilator such that  $g^\perp g = 0$ . We use  $|\cdot|$  to denote Euclidean norms of vectors or matrices, and  $B_\epsilon(\mathcal{A}) = \{x \in \mathbb{R}^n \mid \inf_{y \in \mathcal{A}} |x - y| \leq \epsilon\}$  for a compact set  $\mathcal{A} \subset \mathbb{R}^n$  and some  $\epsilon > 0$ . We use  $\text{cl}(\cdot)$  to represent the closure of a set, and  $\text{Im}[\cdot]$  denotes the image of a given function. Given  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  we define the differential operator  $\nabla f := (\frac{\partial f}{\partial x})^\top$ , and use  $f(x)|_{x_a}^{x_b} = f(x_a) - f(x_b)$ . For square

matrices  $A$  and  $B$ , the notation  $A \succeq B$  indicates  $(A - B)$  positive semidefinite, and we use  $\text{sym}\{A\}$  to represent the symmetric part  $\frac{1}{2}(A + A^\top)$  of  $A$ . The set of all polynomials in  $x \in \mathbb{R}^n$  with real coefficients of degree  $m \in \mathbb{N}_+$  is written as  $\mathbb{R}_m[x]$ . A matrix-valued function  $A : \mathbb{R}^n \rightarrow \mathbb{R}_{>0}^{m \times m}$  is called uniformly bounded if  $a_1 I \preceq A(x) \preceq a_2 I$ ,  $\forall x$  with some  $a_2 \geq a_1 > 0$ .  $\lambda_i\{\cdot\}$ s denote the eigenvalues of square matrices. We use  $\phi^L$  to denote the left inverse of a mapping  $\phi$ , i.e.,  $\phi^L \circ \phi$  is an identity mapping.

*Nomenclature.* The variables  $x, y, u$  represent the unknown state, measurable state (output) and the input, respectively, which all live in Euclidean space, and we also write  $\chi := \text{col}(x, y)$ . For a variable  $(\cdot)$ , we assume its dimension being  $n_{(\cdot)}$ . For the dynamics of  $x$ , we use  $X(t; x_0)$  to denote its solution from the initial condition  $x_0$  at  $t = 0$ .

### B. Problem formulation

In this paper we consider state estimation of nonlinear control (time-varying) systems in the form of

$$\begin{aligned} \dot{x} &= f_x(x, y, u), \\ \dot{y} &= f_y(x, y, u), \end{aligned} \quad (1)$$

with unknown state  $x \in \mathbb{R}^{n_x}$ , measured state  $y \in \mathbb{R}^{n_y}$  and input  $u \in \mathbb{R}^{n_u}$ . We sometimes compactly write as

$$\dot{\chi} = f(\chi, u) \quad (2)$$

and  $n_\chi := n_x + n_y$ . Following the standard practice in observer design, we assume the considered input signal  $u : \mathbb{R}_+ \rightarrow \mathbb{R}^{n_u}$  belonging to a set  $\mathcal{U} \subset \mathcal{L}_\infty^{n_u}$ , such that the system (1) is forward complete. Namely, the solution exists uniquely for  $t \in [0, \infty)$ .

*Problem 1.* The objective is to design an observer

$$\begin{aligned} \dot{\xi} &= N(\xi, y, u), \\ \hat{x} &= H(\xi, y) \end{aligned} \quad (3)$$

with  $\xi \in \mathbb{R}^{n_\xi}$  the observer state, and  $\hat{x} \in \mathbb{R}^{n_x}$  the observer output, guaranteeing

$$\lim_{t \rightarrow \infty} |\hat{x}(t) - x(t)| = 0. \quad (4)$$

If there exists a function  $b : \mathbb{R}^{n_x} \times \mathbb{R}^{n_x} \rightarrow \mathbb{R}_{\geq 0}$  such that

$$|\hat{x}(t) - x(t)| \leq e^{-\lambda t} b(\hat{x}(0), x(0)), \quad b(x, x) = 0, \quad \lambda > 0,$$

then we call (3) an exponential observer with rate  $\lambda$ . When only estimating the unknown part  $x$  rather than the full system state  $(x, y)$ , we refer it as a reduced-order observer. In particular, the case  $n_x \leq n_\xi < n_\chi$  is clearly reduced-order.

*Remark 1:* We consider the class of systems (1) with the output partial states of the system. On one hand, this coordinate simplifies the following analysis in reduced-order observers design; on the other hand, it is quite general to cover several observable canonical forms and plenty of physical models, since the model  $\dot{x} = f(x, u)$ ,  $y = h(x)$  can be transformed into (1) by finding the complementary coordinates in many cases assuming that  $\nabla h(x)$  is full rank.

### C. Preliminaries on contraction analysis

Contraction analysis allows us to obtain convergence independent of knowledge of a particular solution [12, 22]. In contrast, a pre-defined solution is required in standard Lyapunov methods, and one then needs to construct a Lyapunov function of observation errors [36, 42]. This fact makes contraction analysis particularly tailored for observer design. Let us recall some definitions first.

*Definition 1: (Contraction)* Assume the nonlinear system

$$\dot{\mathbf{x}} = f(\mathbf{x}, t), \quad \mathbf{x} \in \mathbb{R}^n \quad (5)$$

is forward complete with the solution  $X(t; \mathbf{x}_0)$  from  $\mathbf{x}(0) = \mathbf{x}_0$ . The system is

- asymptotically contracting if  $\forall (x_a, x_b) \in \mathbb{R}^n \times \mathbb{R}^n$ ,

$$|X(t; x_a) - X(t; x_b)| \leq \kappa(|x_a - x_b|, t)$$

holds for any  $t \geq 0$  and some function  $\kappa$  of class  $\mathcal{KL}$ .

- contracting if the system is asymptotically contracting with  $\kappa(a, t) = ke^{-\lambda t}a$  for some  $k, \lambda > 0$ .

In contraction analysis we need to calculate the differential behaviour of an infinitesimal displacement  $\delta\mathbf{x}$  along the flow  $X(t; \mathbf{x}(0))$  generated by (5), then obtaining the differential dynamics

$$\frac{d}{dt}\delta\mathbf{x} = \frac{\partial f}{\partial \mathbf{x}}(\mathbf{x}, t)\delta\mathbf{x}, \quad (6)$$

which can be regarded as the first-order approximation along solutions. Then, the incremental stability of (5) is characterized by the stability of the linear time-varying (LTV) differential dynamics (6). A central result in [22] is that if there is a uniformly bounded metric  $M : \mathbb{R}^n \rightarrow \mathbb{R}_{>0}^{n \times n}$ , such that

$$\dot{M} + \frac{\partial f}{\partial \mathbf{x}}^\top M + M \frac{\partial f}{\partial \mathbf{x}} \preceq -2\lambda M \quad (\text{or } \prec 0), \quad (7)$$

then the system (5) is contracting (or asymptotically contracting). The interested reader may refer to [12, 22] for comprehensive introduction.

## III. MAIN RESULTS ON CONSTRUCTIVE SOLUTIONS

In this section, we study *constructive* solutions to Problem 1, as well as their convex representation.

### A. Contracting reduced-order observers

In this subsection, we will show that Problem 1 can be translated into contracting observers design, in which we aim to find mappings to achieve both *contraction* and *correctness*. Here, correctness refers to that the observer dynamics (3) admits a particular solution  $\xi_*(t)$  such that the associated observer output  $H(\xi_*, y)$  always equals to the unknown system state  $x(t)$ . To see this, assuming that Problem 1 is solvable, thus any different solutions  $\xi_a(t), \xi_b(t) \in \mathbb{R}^{n_\xi}$  guaranteeing

$$\begin{aligned} & \limsup_{t \rightarrow \infty} |H(\xi_a(t), y(t)) - H(\xi_b(t), y(t))| \\ & \leq \limsup_{t \rightarrow \infty} |H(\xi_a(t), y(t)) - x(t)| + |H(\xi_b(t), y(t)) - x(t)| \\ & = 0. \end{aligned}$$

Clearly, a sufficient condition to  $|H(\xi_a, y) - H(\xi_b, y)| \rightarrow 0$  is contraction of the observer dynamics in (3). Furthermore, in order to guarantee (4), there should exist a *particular solution* of (3) which establishes the connection between  $\xi$  and  $x$ . Intuitively, we construct, analyze, and implement an observer in a carefully selected  $\xi$ -coordinate, with  $H(\cdot)$  linking these two coordinates. It is summarized as follows.

*Problem 2. (Contracting observer)* Find smooth mappings  $N : \mathbb{R}^{n_\xi} \times \mathbb{R}^{n_y} \times \mathbb{R}^{n_u} \rightarrow \mathbb{R}^{n_\xi}$  and  $H : \mathbb{R}^{n_\xi} \times \mathbb{R}^{n_y} \rightarrow \mathbb{R}^{n_x}$  guaranteeing

**C1** (*contraction*) the observer dynamics (3) is contracting (or asymptotically contracting for the asymptotic case).

**C2** (*correctness*) the partial differential equation

$$\frac{\partial H}{\partial \xi}(\xi, y)N(\xi, y, u) + \frac{\partial H}{\partial y}(\xi, y)f_y(H(\xi, y), y, u) = f_x(H(\xi, y), y, u) \quad (8)$$

holds, and  $\text{Im}_{\xi \in \mathcal{A}}[H(\xi, y)] = \mathbb{R}^{n_x}$  uniformly in  $y$  for some open set  $\mathcal{A} \subset \mathbb{R}^{n_\xi}$ , in which the observer dynamics (3) is invariant.<sup>1</sup>  $\triangleleft$

*Proposition 1:* Assume the system states of (1) are bounded. If we can find mappings  $N(\xi, y, u)$  and  $H(\xi, y)$  satisfying **C1-C2**, then the observer (3) solves Problem 1.  $\triangleleft$

*Proof 1:* We only show the asymptotic case. For any  $(u, y)$ , invoking the invariance of (3) in  $\mathcal{A}$  and the assumption that the image space of  $H(\cdot, y)$  is the entire  $\mathbb{R}^{n_x}$ , then  $\forall x(0) \in \mathbb{R}^{n_x}$  we can always find a point  $\xi_*(0) \in \mathcal{A}$  such that

$$H(\xi_*(0), y(0)) = x(0). \quad (9)$$

Denote the solution of (3) as  $\xi_*(t)$  from the initial condition  $\xi_*(0)$ . The correctness (8) guarantees

$$H(\xi_*(t), y(t)) = x(t), \quad \forall t \geq 0,$$

if the initial conditions satisfy (9).

Since the system (3) is asymptotically contracting and  $\xi_*(t)$  is a particular solution, we have that for any solution  $\xi(t)$

$$\lim_{t \rightarrow \infty} |\xi(t) - \xi_*(t)| = 0. \quad (10)$$

Invoking the state boundedness assumption, we, without loss of generality, denote  $x \in \Omega_x \subset \mathbb{R}^{n_x}$ ,  $y \in \Omega_y \subset \mathbb{R}^{n_y}$  and  $\xi_* \in \Omega_\xi \subset \mathcal{A}$  with  $\text{Im}_{(\xi, y) \in \Omega_\xi \times \Omega_y} [H(\xi, y)] = \Omega_x$ , and all these sets are defined bounded. There always exists a moment  $T_1$  such that  $\xi(t) \in B_\epsilon(\text{cl}(\Omega_\xi))$  for  $t \geq T_1$  with a small  $\epsilon > 0$  due to the convergence (10) and  $\xi_*(t) \in \Omega_\xi$  for all  $t \geq 0$ . Then, for  $t \in [T_1, \infty)$  we have

$$\begin{aligned} \lim_{t \rightarrow \infty} |\hat{x}(t) - x(t)| &= \limsup_{t \rightarrow \infty} |H(\xi, y) - H(\xi_*, y)| \\ &\leq \lim_{t \rightarrow \infty} \max_{\xi \in B_\epsilon(\text{cl}(\Omega_\xi))} |\nabla_\xi H| |\xi(t) - \xi_*(t)| \\ &= 0, \end{aligned}$$

where we have used the state boundedness assumption, implying that  $|\nabla_\xi H|$  is bounded in the closure of  $\Omega_\xi$ .

In the sequel, we focus on the constructive solution and the convex representations to Problem 2, in this way solving Problem 1.

<sup>1</sup>That is  $\xi(0) \in \mathcal{A}$  implying  $\xi(t) \in \mathcal{A}, \forall t \geq 0$  for all  $y \in \mathbb{R}^{n_y}$ .

### B. General results for the case $n_\xi = n_x$

We start with the basic case selecting the observer dimension  $n_\xi = n_x$ . In Problem 2, it is shown that we may consider this problem expressing the dynamics (1) in a transformed coordinate, and then study its contraction property by regarding  $(u, y)$  as exogenous signals. Indeed, it is well known that the change of coordinate plays an important role in nonlinear observer design.

Let  $\phi : \mathbb{R}^{n_x} \times \mathbb{R}^{n_y} \rightarrow \mathbb{R}^{n_x}$  define the transformation

$$(x, y) \mapsto (z, y) := (\phi(x, y), y).$$

If we guarantee that  $\phi(\cdot, y)$  is bijective from  $\mathbb{R}^{n_x}$  to  $\mathbb{R}^{n_x}$ , then for each  $y$  there is a left inverse  $\phi^L : \mathbb{R}^{n_x} \times \mathbb{R}^{n_y} \rightarrow \mathbb{R}^{n_x}$  such that

$$\phi^L(\phi(x, y), y) = x. \quad (11)$$

Now the dynamics in the  $z$ -coordinate is written as

$$\dot{z} = \beta(z, y, u) \quad (12)$$

with

$$\begin{aligned} \beta(z, y, u) &:= f_z(\phi^L(z, y), y, u) \\ f_z(x, y, u) &= \Phi_x(x, y)f_x(x, y, u) + \Phi_y(x, y)f_y(x, y, u). \end{aligned} \quad (13)$$

and

$$\Phi_x(x, y) := \frac{\partial \phi}{\partial x}(x, y), \quad \Phi_y(x, y) := \frac{\partial \phi}{\partial y}(x, y). \quad (14)$$

For convenience, we also use the notation  $\Phi(\chi) := \nabla \phi^\top(\chi) = [\Phi_x(\chi), \Phi_y(\chi)]$ . Supposing we are able to find a transformation  $\phi(x, y)$  as described above such that the obtained system (12) is asymptotically contracting for any exogenous time-varying signals  $(u, y)$ , then it is trivial to solve Problem 2 by selecting  $N(\cdot) = \beta(\cdot)$  and  $H(\cdot) = \phi^L(\cdot)$ .

Based on the above intuitive idea, we are now in position to present the first main result of the paper on contracting reduced-order observers. Its convex representation is given in Section III-C, and the extension to the case  $n_\xi > n_x$  will be discussed in Section III-D.

*Theorem 1:* Consider the system (1), and assume that we can find mapping  $\phi : \mathbb{R}^{n_x} \rightarrow \mathbb{R}^{n_x}$  and uniformly bounded Riemannian metric  $M : \mathbb{R}^{n_x} \rightarrow \mathbb{R}_{>0}^{n_x \times n_x}$  satisfying

**A1** (diffeomorphism) the mapping  $\phi_y(\cdot) := \phi(\cdot, y)$  defines a diffeomorphism uniformly in  $y \in \mathbb{R}^{n_y}$ ;

**A2** (contraction)

$$\partial_{f_z} M \Big|_{\phi(\chi)} + M \frac{\partial f_z}{\partial x} \left[ \frac{\partial \phi}{\partial x} \right]^{-1} + \left[ \frac{\partial \phi}{\partial x} \right]^{-\top} \frac{\partial f_z}{\partial x}^\top M \prec 0 \quad (15)$$

(or  $\preceq -2\lambda M$ ,  $\lambda > 0$  for the exponential case) with  $f_z(\cdot)$  defined in (13), holds  $\forall (\chi, u) \in \mathbb{R}^{n_x} \times \mathcal{U}$ .

Then, there exists an  $n_x$ -dimensional reduced-order observer guaranteeing (4) globally.  $\triangleleft$

*Proof 2:* The assumption **A1** defines a global diffeomorphism, and we thus, according to the global inverse theorem [30], obtain that

$$\frac{\partial \phi}{\partial x}(x, y) = \left[ \frac{\partial \phi^L}{\partial z}(z, y) \right]_{z=\phi(\chi)}^{-1}. \quad (16)$$

The dynamics of  $z$  can be written as (12). Now design an observer (3) by selecting

$$N(\xi, y, u) = f_z(\phi^L(\xi, y), y, u), \quad H(\xi, y) = \phi^L(\xi, y). \quad (17)$$

The correctness condition **C2** can be verified trivially from the above construction. Therefore,  $z(t) = \phi(x(t), y(t))$  is a particular solution of the observer dynamics from  $z(0)$ .

Since  $\phi(\cdot, y)$  is a diffeomorphism for any  $y$  due to **A1**, in the sequel we only need to show the convergence of  $\xi$  to  $z$ . The dynamics of them are given by

$$\dot{\xi} = N(\xi, y, u), \quad \dot{z} = N(z, y, u),$$

where  $(u, y)$  are viewed as (available) exogenous signals. Towards this end, we first need to calculate the differential characterization along the solution of (3) with  $y(t)$  governed by the system (1). Note that the solution  $y(t) = Y(t; \chi_0, u)$  is an exogenous signal, which is independent of  $\xi(t)$ . Consider the initial conditions  $z(0)$  and  $\xi(0)$  are connected by a smooth curve  $\gamma : [0, 1] \rightarrow \mathbb{R}^{n_x}$ , such that  $\gamma(0) = z(0)$  and  $\gamma(1) = \xi(0)$ . The solution from  $\gamma(s)$ ,  $s \in [0, 1]$  under the exogenous signals  $(u, y)$  is denoted as  $\xi(t; \gamma(s), y, u)$ . Thus, we have the differential dynamics

$$\begin{aligned} \frac{d}{dt} \delta \xi &= \frac{\partial N}{\partial \xi}(\xi, y, u) \delta \xi \\ &= \frac{\partial f_z}{\partial \xi}(\phi^L(\xi, y), y, u) \delta \xi \\ &= \frac{\partial f_z}{\partial x}(\phi^L(\xi, y), y, u) \frac{\partial \phi^L}{\partial z}(\xi, y) \delta \xi \\ &\stackrel{(16)}{=} \frac{\partial f_z}{\partial x}(\hat{x}, y, u) \left[ \frac{\partial \phi}{\partial x}(\hat{x}, y) \right]^{-1} \delta \xi, \end{aligned}$$

with the infinitesimal displacement  $\delta \xi = \frac{\partial \xi(t; \gamma(s), y, u)}{\partial s}$ .

Define a differential Lyapunov function as  $V(\xi, \delta \xi) = \delta \xi^\top M(\xi) \delta \xi$ , yielding

$$\begin{aligned} \dot{V}(\xi, \delta \xi) &= \delta \xi^\top \left( \partial_{f_z} M + 2 \text{sym} \left( M \frac{\partial f_z}{\partial x} \left( \frac{\partial \phi}{\partial x} \right)^{-1} \right) \right) \delta \xi \\ &\stackrel{(15)}{\leq} 0 \end{aligned}$$

by replacing  $x$  with  $\hat{x} = \phi^L(\xi, y)$ , and thus we have  $\lim_{t \rightarrow \infty} V(\xi(t), \delta \xi(t)) = 0$ . Therefore, we obtain

$$\begin{aligned} &\limsup_{t \rightarrow \infty} |\xi(t) - z(t)| \\ &\leq \lim_{t \rightarrow \infty} c \int_0^1 V(\xi(t; \gamma(s), y, u), \delta \xi(t; \gamma(s), y, u)) ds = 0, \end{aligned}$$

for some  $c > 0$ . It completes the proof.

*Remark 2:* We adopt in the assumption **A2** a Riemannian metric  $\delta \xi^\top M(\xi) \delta \xi$  to design an observer, which is very general for systems living in Euclidean space under some smoothness assumption. It can be generalized using a Finsler metric  $F(\xi, \delta \xi)$  [12], then guaranteeing the induced distance monotonically decreasing. It is particularly useful if system states live in Riemannian manifolds, e.g., matrix Lie groups.

*Remark 3:* It is underlined here that the sufficient condition given in Theorem 1 only relies on the existence of the mapping  $\phi(x, y)$ , though the left inverse  $\phi^L$  is needed when realizing the

design. Assuming that we have already found  $\phi(x, y)$ , its left inverse is given by the solution of the optimization  $\phi^L(\xi, y) = \operatorname{argmin}_{\hat{x} \in \mathbb{R}^{n_x}} |\phi(\hat{x}, y) - \xi|^2$ , which may be implemented by off-line searching for analytic solutions, or on-line searching for numerical results, see [20] for an explicit formula.

In the above analysis, since we have used the inverse function theorem, the correctness condition **C2** is satisfied automatically. Some further discussions about Theorem 1 will be given in Section IV.

### C. Convex sets of nonlinear contracting observers

The sufficient condition—Assumptions **A1** and **A2**—in Theorem 1 is clearly non-convex in general, since it depends in a nonlinear way on the left inverse of the unknown function  $\phi(x, y)$ . In this subsection, we will present some *convex representations*, which are motivated by, and extend to the reduced-order case, our previous results in [26]. Indeed, searching for  $\phi(\cdot)$  and  $M(\cdot)$  can be viewed as an optimization problem on an infinite-dimensional function space. If we could obtain convex sufficient conditions, one may verify these differential equations and inequalities using the sum-of-square relaxation, thus making the observer design numerically tractable.

The condition **A1** needs to check the non-convex function  $\det\{\nabla_x \phi\}$ . We have its convex approximation below.

*Lemma 1:* The convex condition **H1** is sufficient to **A1**.

**H1** (the monotone condition) For any  $y \in \mathbb{R}^{n_y}$ ,

$$\frac{\partial \phi}{\partial x}(x, y) + \left[ \frac{\partial \phi}{\partial x}(x, y) \right]^\top \succeq kI_{n_x}, \quad k > 0. \quad (18)$$

*Proof 3:* Since the partial derivative operator is linear, the inequality (18) is convex in  $\phi(x, y)$ . Following [41, Theorem 1] or the Hadamard's global inverse theorem [30, Theorem 2], we can verify the diffeomorphism of  $\phi_y(\cdot) = \phi(\cdot, y)$ .

Apart from the convexity of **H1**, another benefit of the monotonicity we imposed relies on the existing efficient algorithms to calculate the inverse mapping  $\phi^L$  when implementing the observer, *e.g.*, the Newton method and the ellipsoid method [31, 41].

Now we move forward to the condition **A2** and study its convex representation. Note that the contraction analysis was done in the transformed coordinate  $z = \phi(\chi)$ , in which the dynamics is given in (13). Its second equation  $f_z(\chi, u) = \Phi(\chi)f(\chi, u)$  shows that  $f_z(\cdot)$  is in the “image space” of a given vector field  $f(\chi, u)$ . An important point is that the relationship between  $\phi(\cdot)$  and  $f_z(\cdot)$  is linear. If we regard  $f_z(\cdot)$  as a *new decision variable*, and jointly search for the transformation  $\phi(\chi)$  and the mapping  $f_z(\chi, u)$ , we may be able to obtain a convex set of reduced-order contracting observers.

In terms of the new decision variable  $f_z(\cdot)$  involved, we additionally—compared with **A2**—require the following correctness condition, which is a variant of **C2** in our design framework. It is a linear equality constraint with respect to  $\phi(\chi)$  and  $f_z(\chi, u)$ , thus convex.

**H2** (the modified correctness)

$$\Phi_x(\chi)f_x(\chi, u) + \Phi_y(\chi)f_y(\chi, u) = f_z(\chi, u) \quad (19)$$

for all  $(\chi, u) \in \mathbb{R}^{n_x} \times \mathbb{R}^{n_u}$ .

Then, regarding the convex representation of the contraction condition **A2**, we have the following.

*Lemma 2:* Under **H1**, the following convex conditions **H3** or **H4**, together with **H2**, imply the condition **A2**, *i.e.*,

$$\left[ \mathbf{H2} \wedge (\mathbf{H3} \vee \mathbf{H4}) \right] \implies \mathbf{A2}.$$

**H3** By restricting the transformation

$$\phi(x, y) = Px + \varphi(y) \quad (20)$$

with a smooth mapping  $\varphi: \mathbb{R}^{n_y} \rightarrow \mathbb{R}^{n_x}$ , then search for  $P \succ 0$ ,  $\varphi(\cdot)$ ,  $f_z(\cdot)$  and  $Q$  verifying

$$F(x, y, u) + F^\top(x, y, u) + Q \preceq 0, \quad (21)$$

with  $F(x, y, u) := \frac{\partial f_z}{\partial x}(x, y, u)$ , and  $Q \succ 0$  (or  $Q = 2\lambda P$  for the exponential case with  $\lambda > 0$ ).

**H4** Find  $\phi(x, y)$ ,  $Q$ ,  $f_z(\cdot)$ , metric  $P \succ 0$  and  $r > 0$  verifying

$$\begin{bmatrix} 2\operatorname{sym}(\Phi_x - \frac{r}{2}F) - P - rQ & \Phi_x + \frac{r}{2}F \\ (\Phi_x + \frac{r}{2}F)^\top & P \end{bmatrix} \succeq 0, \quad (22)$$

with  $Q \succ 0$  (or  $Q = 2\lambda\Phi_x^\top P^{-1}\Phi_x$  for contraction).  $\triangleleft$

*Proof 4:* The conditions (21) and (22) are state-dependent linear matrix inequalities (LMIs), which are jointly convex in  $\phi$  (or  $\varphi$ ),  $f_z$ ,  $P$  and quasi-convex in  $\lambda$  for the exponential case. In **A2**, the function  $f_z$  is defined as (13). In the convex representation, we regard  $f_z$  as a free mapping to be searched for, which, thus, should satisfy the correctness condition **H2**.

The remainder of the proof is to verify the inequality in **A2**. First, we consider the assumption **H3** with the parameterization  $\phi(x, y) = Px + \varphi(y)$ , thus

$$\Phi_x = \frac{\partial \phi}{\partial x}(x, y) = P.$$

Then, it is straightforward to get<sup>2</sup>

$$\begin{aligned} (21) \iff & F + F^\top \preceq -2\lambda\Phi_x^\top P^{-1}\Phi_x \\ \iff & P^{-1}F\Phi_x^{-1} + \Phi_x^{-\top}F^\top P^{-1} \preceq -2\lambda M \\ \iff & \partial_{f_z} P^{-1} + P^{-1}F\Phi_x^{-1} + \Phi_x^{-\top}F^\top P^{-1} \preceq -2\lambda M \\ & \stackrel{M:=P^{-1}}{\iff} \mathbf{A2}, \end{aligned}$$

verifying the first convex sufficient condition in **H3**.

For the case in **H4**, the inequality (22) is equivalent to

$$\left( \Phi_x - \frac{r}{2}F \right) + \left( \Phi_x - \frac{r}{2}F \right)^\top - \left| \Phi_x + \frac{r}{2}F \right|_{P^{-1}}^2 - P - rQ \succeq 0. \quad (23)$$

As show in [41, Section III], the inequality (23) is a sufficient condition to

$$\Phi_x^\top P^{-1}F + F^\top P^{-1}\Phi_x \preceq -Q. \quad (24)$$

Namely, using the inequalities

$$\Phi_x^\top MF + F^\top M\Phi_x = \frac{1}{r} \left[ |\Phi_x + \frac{r}{2}F|_M^2 - |\Phi_x - \frac{r}{2}F|_M^2 \right]$$

<sup>2</sup>We give the details of the exponential case, and the asymptotic case follows in a similar way.

and  $-c^\top P^{-1}c \leq b^\top P b - 2b^\top c$  with  $M = P^{-1}$  successively, we get (24) from (23). It yields

$$(24) \quad \begin{aligned} &\iff \Phi_x^\top P^{-1}F + F^\top P^{-1}\Phi_x \preceq -2\lambda\Phi_x^\top P^{-1}\Phi_x \\ &\iff P^{-1}F\Phi_x^{-1} + \Phi_x^{-\top}F^\top P^{-1} \preceq -2\lambda P^{-1} \\ &\iff \partial_{f_z}P^{-1} + P^{-1}F\Phi_x^{-1} + \Phi_x^{-\top}F^\top P^{-1} \preceq -2\lambda M \\ &\stackrel{M:=P^{-1}}{\iff} \mathbf{A2}, \end{aligned}$$

where in the second equivalence we have used the assumption **H1** to guarantee the existence of the inverse matrix of  $\Phi_x$ . It completes the proof.

*Remark 4:* We will show, via examples, that the convex condition **H3** is quite general to cover several physical models. It enjoys another merit of simple left inverse mapping  $\phi^L(\xi, y) = P^{-1}[\xi - \varphi(y)]$ . Besides, the condition **H3** can be generalized into the form  $\phi(x, y) = P(y)x + \varphi(y)$ , with  $P(\cdot)$  dependent on  $y$ . For this case, (21) becomes the convex constraint  $F(x, y, u) + F^\top(x, y, u) + \dot{P}(y) + Q \preceq 0$ .

To accomplish the reduced-order observer design, we would like to search  $\phi(\chi)$  and  $f_z(\chi, u)$  simultaneously over finite dimensional space. Thus, we parameterize them as

$$\phi(\chi) = \sum_{i=1}^{q_1} \theta_i \phi_i(\chi), \quad f_z(\chi, u) = \sum_{i=q_1+1}^{q_2} \theta_i f_{z,i}(\chi, u) \quad (25)$$

with basis mappings  $\phi_i$  and  $f_{z,i}$  ( $i = 1, \dots, q_2$ ) smooth enough, and  $q_1, q_2 \in \mathbb{N}_+$  verifying  $q_1 < q_2$ . The role of parameterization is to reformulate the obtained infinite dimensional optimization over function space into a finite dimensional convex problem, which is efficiently solvable.

We are in position to present the second main result of the paper about a convex set of contracting reduced-order observers, which is a *constructive* solution to Problem 2.

*Theorem 2:* Given the nonlinear system (1), for the case  $n_\xi = n_x$  a convex set of contracting reduced-order observers is given by (25) together with the observer dynamics

$$\dot{\hat{\xi}} = f_z(\hat{x}, y, u), \quad \hat{x} = \phi^L(\hat{\xi}, y) \quad (26)$$

with  $\phi^L(\cdot, y)$  the left inverse of  $\phi(\cdot, y)$  solving Problem 2, where  $\theta \in \Theta$  is defined by the convex constraints **H1**, **H2**, and either **H3** or **H4**.

*Proof 5:* As already shown above, the condition **H1** guarantees the existence of inverse mapping  $\phi^L(\cdot)$  globally. Following Theorem 1 and Lemmata 1-2, we get the claim.

#### D. Extension to the immersion case $n_\xi > n_x$

In the previous subsections, we considered the case that the observer dynamics shares the dimension of the systems state. It, however, is openly recognized that by considering injective immersion may allow to take into account a larger class of nonlinear control systems, being widely used in observer design with various purposes.

In this subsection, we extend our results to the case  $n_\xi > n_x$ . The motivation is threefold.

- Sometimes there exists observable singularity for the system to be estimated. By increasing observer dimensions, we may be able to guarantee the injectivity of changes of coordinates, see [2, 37] and [8, Chapter 4].

- New mappings will appear in (19), the new degree of freedom from which relaxes the correctness condition.
- For a given nonlinear system, it may provide a strictly larger set, compared to the case of  $n_\xi = n_x$ , of convergent observers, among which we may select an optimal design in some sense, in this way enhancing performance.<sup>3</sup>

Following the idea in Section III-A, we would like to find  $\phi : \mathbb{R}^{n_x} \times \mathbb{R}^{n_y} \rightarrow \mathbb{R}^{n_\xi}$  with  $n_\xi > n_x$ . However, two difficulties arise, namely, (i)  $\nabla_x \phi(x, y)$  is not square, thus the inverse function theorem not applicable; and (ii) the image space  $\mathcal{I}_\xi$  of  $\phi(\cdot, y)$  may be only an *open subset* of  $\mathbb{R}^{n_\xi}$ , and thus the left inverse  $\phi^L(\cdot, y)$  is not defined globally. A possible approach to deal with this issue is to construct an extended left mapping  $\phi_e^L : \mathbb{R}^{n_\xi} \times \mathbb{R}^{n_y}$  satisfying  $\phi_e^L(z, y) = \phi^L(z, y)$ ,  $\forall (z, y) \in \mathcal{I}_\xi \times \mathbb{R}^{n_y}$ . It is usually applied to the case in which the transformation  $\phi(\cdot)$  has been *already known*. However, it is not suitable in our situation, since our target is to find the mapping  $\phi(\cdot)$ .

In order to be able to keep consistent with the proposed constructive method, as well as to circumvent the difficulties above, we introduce an *augmentation* of the  $x$ -coordinate into the extended system state  $(x, w)$  given by

$$\dot{w} = f_w(x, w, y, u) \quad (27)$$

with  $w \in \mathbb{R}^{n_w}$ ,  $n_w = n_\xi - n_x$  and  $x_e := \text{col}(x, w)$ .<sup>4</sup> The following lemma indicates that the augmented state  $w$  should not change the “detectability” of the system (1).

*Lemma 3:* Assume that there exists an exponential observer for the system (1).

- 1) If the augmented system (27) is contracting, then there exists an observer for the extended state  $(x, w)$ .
- 2) Conversely, for the case that the given system is autonomous (*i.e.*  $u = 0$ ) with all states bounded, assuming that the  $(x, w)$ -system has an exponential observer (32) with a smooth Lyapunov function  $V(x_e, y, \xi)$  such that

$$a_1|\xi - \phi(x_e, y)|^2 \leq V(x_e, y, \xi) \leq a_2|\xi - \phi(x_e, y)|^2 \\ \overbrace{\dot{V}(x_e, y, \xi)} \leq -\lambda V(x_e, y, \xi)$$

with  $a_1, a_2, \lambda > 0$ , then the system (27) is contracting.

*Proof 6:* For the first claim, it can be proved by cascading the existing observer with

$$\dot{\hat{w}} = f_w(\hat{x}, \hat{w}, y, u)$$

with  $\hat{x}$  the estimate of  $x$  from the existing observer, and invoking the solution continuity of differential equations, and the state boundedness.

Regarding the second claim, we consider the autonomous case  $u = 0$ . According to [3, Proposition 4], for any  $(x_e, y, \delta x_e, \delta y) \in \mathbb{R}^{n_\xi + n_y} \times T\mathbb{R}^{n_\xi + n_y}$ , the norm of infinitesimal displacement  $|\text{col}(\delta x_e, \delta y)|$  is monotonically decaying along trajectories, tangentially to some output function level

<sup>3</sup>A similar problem—parameterizing all convergent observers for linear time-invariant systems—was comprehensively studied in [14].

<sup>4</sup>In [7], complementary operation of full column-rank Jacobian is used to express an observer in the preferred coordinate assuming that an observer exists *already*.

sets. That is, in our case if we fix the differential output  $\delta y(t) \equiv 0$ , then the obtained LTV system

$$\begin{pmatrix} \delta \dot{x} \\ \delta \dot{w} \end{pmatrix} = \begin{pmatrix} \frac{\partial f_x}{\partial x}(x, y, 0) & 0 \\ \frac{\partial f_w}{\partial x}(x, w, y, 0) & \frac{\partial f_w}{\partial w}(x, w, y, 0) \end{pmatrix} \begin{pmatrix} \delta x \\ \delta w \end{pmatrix} \quad (28)$$

satisfies

$$\left\| \begin{bmatrix} \delta x(t) \\ \delta w(t) \end{bmatrix} \right\|^2 \leq a_0 \exp(-\lambda t) \left\| \begin{bmatrix} \delta x(0) \\ \delta w(0) \end{bmatrix} \right\|^2,$$

for some  $a_0 > 0$ . The inequality holds for any initial conditions, and thus we consider a particular one  $\delta x(0) = 0$ , for which we have  $\delta x(t) = 0, \forall t \geq 0$  along (28). Then, it yields

$$\delta \dot{w} = \frac{\partial f_w}{\partial w}(x, w, y, 0) \delta w$$

admitting

$$|\delta w(t)|^2 \leq a_0 \exp(-\lambda t) \left| \frac{\partial \phi_e(x_e(0), y(0))}{\partial w} \delta w(0) \right|^2$$

uniformly along any trajectories. It implies that the augmented dynamics (27) should be contracting.

The second claim shows the necessity to impose contraction properties with respect to  $w$  when introducing the augmentation (27), a sufficient condition for which is to find a uniformly bounded metric  $M_w : \mathbb{R}^{n_w} \rightarrow \mathbb{R}^{n_w \times n_w}$  satisfying

$$\partial_{f_w} M_w + M_w \frac{\partial f_w}{\partial w} + \frac{\partial f_w}{\partial w}^\top M_w \preceq -2\lambda_w M_w, \quad \lambda_w > 0. \quad (29)$$

Indeed, we can always find a vector field  $f_w$  satisfying (29), the simplest one among which may refer to

$$f_w = Aw + f_a(x, y, u) \quad (30)$$

with Hurwitz matrix  $A$  and smooth function  $f_a$ .

Now, it is straightforward to extend Theorems 1-2 to this case. We only give here the convex variant of the latter. Similarly, we may parameterize mappings  $f_{ze} : \mathbb{R}^{n_\xi + n_y + n_u} \rightarrow \mathbb{R}^{n_\xi}$ ,  $\phi : \mathbb{R}^{n_\xi + n_y + n_u} \rightarrow \mathbb{R}^{n_\xi}$  and  $f_w : \mathbb{R}^{n_\xi + n_y + n_u} \rightarrow \mathbb{R}^{n_w}$ , and matrices  $P_e \in \mathbb{R}_{>0}^{n_\xi \times n_\xi}$  and  $M_w \in \mathbb{R}_{>0}^{n_w \times n_w}$  as

$$\begin{aligned} f_{ze} &= \sum_{i=q_1+1}^{q_2} \theta_i f_{ze,i}(x_e, y, u), & \phi &= \sum_{i=1}^{q_1} \theta_i \phi_i(x_e, y) \\ f_w &= \sum_{i=q_2+1}^{q_3} \theta_i f_{w,i}(x_e, y, u) \end{aligned} \quad (31)$$

for some  $q_i \in \mathbb{N}_+$  and  $q_i > q_j$  ( $i > j$ ). For convenience, we also define the gradients

$$F_e(x_e, y, u) = \frac{\partial f_{ze}}{\partial x_e}(x_e, y, u), \quad \Phi_{xe}(x_e, y) = \frac{\partial \phi}{\partial x_e}(x_e, y).$$

**Proposition 2:** Consider the system (1) with a fixed augmentation vector field  $f_w$  satisfying (29). For the case  $n_\xi > n_x$  a convex set of contracting reduced-order observers is given by (31), together with the observer dynamics

$$\dot{\xi} = f_{ze}(\hat{x}, \hat{w}, y, u), \quad \begin{bmatrix} \hat{x} \\ \hat{w} \end{bmatrix} = \phi^L(\xi, y), \quad (32)$$

with  $\phi^L(\cdot, y)$  the left inverse of  $\phi(x_e, y)$  and  $\theta \in \Theta$  defined by the convex constraints **H1'**, **H2'** and either **H3'** or **H4'**.

**H1'** (monotonicity)  $\forall y \in \mathbb{R}^{n_y}$

$$\frac{\partial \phi}{\partial x_e}(x, w, y) + \left[ \frac{\partial \phi}{\partial x_e}(x, w, y) \right]^\top \succeq kI_{n_\xi}, \quad k > 0.$$

**H2'** (correctness)

$$\begin{aligned} \Phi_x(x_e, y) f_x(\chi, u) + \Phi_y(x_e, y) f_y(\chi, u) \\ + \Phi_w(x_e, y) f_w(x_e, y, u) = f_{ze}(x_e, y, u) \end{aligned} \quad (33)$$

with  $\Phi_w = \nabla_w \phi^\top(x, w, y, u)$ .

**H3'** By restricting  $\phi(x, w, y) = P_e \text{col}(x, w) + \varphi(y)$  with  $\varphi : \mathbb{R}^{n_y} \rightarrow \mathbb{R}^{n_\xi}$ , and requiring

$$F_e(x_e, y, u) + F_e^\top(x_e, y, u) + Q_e \preceq 0, \quad (34)$$

with  $P_e \succ 0$ , and  $Q_e \succ 0$  (or  $Q_e = 2\lambda P_e$  for the exponential case with  $\lambda > 0$ ).

**H4'** The stronger contraction condition

$$2\text{sym}(\Phi_{xe} - \frac{r}{2} F_e) - |\Phi_{xe} + \frac{r}{2} F_e|_{P_e^{-1}}^2 - P_e - rQ_e \succeq 0,$$

with  $r > 0$ , and  $Q_e \succ 0$  (or  $Q_e = 2\lambda \Phi_{xe}^\top P_e^{-1} \Phi_{xe}$  for contraction with rate  $\lambda > 0$ ).  $\triangleleft$

*Proof 7:* The proof follows *mutatis mutandis* the one of Theorem 2 and the analysis above.

In the above proposition, we fix the vector field  $f_w$  and then search for mappings  $\phi$  and  $f_{ze}$ , since *bi-linear* operations appear in both the correctness condition (33) and the contraction condition (29) if we regard both  $f_w$  and  $M_w$  as decision variables, making the optimization problem non-convex. Then, we get a state-dependent bilinear matrix inequality (BMI) problem. An alternative to maintain convexity is to fix  $\Phi_w$ , then providing the degree of freedom to search for the vector field  $f_w$ . It is clear that (33) relaxes the correctness condition (19) by involving additional terms. On the other hand, the excessive coordinates may provide more observer candidates, thus making it possible to achieve performance enhancement by carefully selecting among them.

**Remark 5:** If the augmented  $w$ -dynamics is in the form of (30) with  $f_a$  only dependent on  $(u, y)$ , the proposed coordinate change is similar to the so-called *filtered transformation* introduced in [29]. Such a technique has been proved very useful in *adaptive observer* design, *i.e.*, estimating unknown states under parameter uncertainty.

## IV. FURTHER RESULTS AND DISCUSSIONS

In this section, we give further discussions about the main results in Section III, and show the connections with some existing results in the literature. Some remarks are in order.

**Remark 6:** For the case  $n_\xi = n_x$ , it is sometimes unnecessary to require the transformation  $\phi(\cdot, y)$  a bijection from  $\mathbb{R}^{n_x} \rightarrow \mathbb{R}^{n_x}$ . A possible case is that the image space is an open subset in  $\mathbb{R}^{n_x}$ , *i.e.*,  $\phi(\mathbb{R}^{n_x}, y) \subset \mathbb{R}^{n_x}$ . For such a case, image extension or projection operator may be useful to continue design [7, 37]. However, we adopt **H1**, on one hand, to circumvent the possible difficulty from non-subjectivity,

and on the other hand, to obtain a *convex* condition with computational consideration.

*Remark 7:* The obtained reduced-order observer is given in the  $z$ -coordinate, but the correctness can also be verified in the original coordinate. We may write the observer (26) in the  $x$ -coordinate, though *unimplementable*, which is given by

$$\begin{aligned}\dot{\hat{x}} &= \frac{\partial \phi^L}{\partial z}(\xi, y) f_z(\phi^L(\xi, y), y, u) + \frac{\partial \phi^L}{\partial y}(\xi, y) \dot{y} \\ &= \left[ \frac{\partial \phi}{\partial x}(\hat{x}, y) \right]^{-1} f_z(\hat{x}, y, u) + \frac{\partial \phi^L}{\partial y}(\xi, y) \dot{y} \\ &= \left[ \frac{\partial \phi}{\partial x}(\hat{x}, y) \right]^{-1} \left( \frac{\partial \phi}{\partial x}(\hat{x}, y) f_x(\hat{x}, y, u) \right. \\ &\quad \left. + \frac{\partial \phi}{\partial y}(\hat{x}, y) f_y(\hat{x}, y, u) \right) + \frac{\partial \phi^L}{\partial y}(\xi, y) \dot{y} \\ &= f_x(\hat{x}, y, u) + [\Phi_x^{-1} \Phi_y] (f_y(\hat{x}, y, u) - f_y(x, y, u)),\end{aligned}\quad (35)$$

where in the last equation we have used the fact that

$$\begin{bmatrix} x \\ y \end{bmatrix} \mapsto T(x, y) := \begin{bmatrix} \phi(x, y) \\ y \end{bmatrix}\quad (36)$$

is another diffeomorphism, and thus using the inverse function theorem again yields

$$\frac{\partial \phi^L}{\partial y}(z, y) \Big|_{z=\phi(x, y)} = - \left[ \frac{\partial \phi}{\partial x}(x, y) \right]^{-1} \frac{\partial \phi}{\partial y}(x, y).$$

The last line of (35) verifies the invariance of the proposed observer, *i.e.*,  $\hat{x}(t) \equiv x(t)$ ,  $\forall t \geq 0$ , if  $x(0) = \hat{x}(0)$ , since the last term  $(f_y(\hat{x}, y, u) - f_y(x, y, u))$  plays the role of “updated (or innovation) term”. Note that, however, the observer dynamics (35) cannot be realized due to the unavailability of  $\dot{y}$ , equivalently  $f_y(x, y, u)$ , and we should implement it in the  $z$ -coordinate instead.

We next show the connections among the proposed designs and some existing results.

#### A. Linear time-invariant systems

It is easy to verify that the conditions either in Theorem 1 or Theorem 2 (H3) are equivalent to the detectability of linear time-invariant (LTI) systems, summarized below.

*Proposition 3:* Consider the LTI system

$$\begin{bmatrix} \dot{x} \\ \dot{y} \end{bmatrix} = \underbrace{\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}}_{:=A} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} B_1 \\ B_2 \end{bmatrix} u\quad (37)$$

with unknown state  $x \in \mathbb{R}^{n_x}$ , input  $u \in \mathbb{R}^{n_u}$  and output  $y \in \mathbb{R}^{n_y}$ . The conditions **A1-A2** are the necessary and sufficient condition to the detectability of the system (37).  $\triangleleft$

*Proof 8:* For linear systems, we consider linear transformation in the form of

$$\phi(x, y) = T_1 x + T_2 y$$

for some matrices  $T_1$  and  $T_2$ , with  $T_1$  full rank. The sufficient part is trivial, since **A1-A2** imply the existence of a linear Luenberger observer<sup>5</sup>

$$\dot{z} = Fz + Dy + Eu, \quad \hat{x} = T_1^{-1}(z - T_2 y)$$

for some matrices  $D, E$ , and  $F$  is Hurwitz. It is equivalent to detectability in the LTI context, see for example [25, Section III]. Regarding the necessary part, we have

Detectability of (37)

$$\begin{aligned}\iff & \exists L \in \mathbb{R}^{(n_x+n_y) \times n_y} \quad \text{s.t. } \lambda_i\{A + L[0 \ I_{n_y}]\} \in \mathbb{C}_- \\ \iff & \exists \tilde{L} \in \mathbb{R}^{n_x \times n_y} \quad \text{s.t. } \lambda_i\{A_{11} + \tilde{L}A_{12}\} \in \mathbb{C}_-\end{aligned}$$

where in the second equivalence we have used [9, Lemma 1], and  $\lambda_i\{\cdot\}$ s denote eigenvalues of square matrices. Now we select the mapping

$$\phi(x, y) = x + \tilde{L}y$$

with  $\tilde{L}$  guaranteeing  $A_{11} + \tilde{L}A_{12}$  a Hurwitz matrix. It is straightforward to verify that  $\phi(\cdot)$  satisfies **A1-A2**.

Indeed, the obtained reduced-order observer for LTI systems is precisely the observer proposed in the pioneering paper [24, Theorem 4, pp. 77].

#### B. Strongly differentially observable systems

In this subsection, we show the set of systems identified by Theorem 1 is larger than the set of strongly differentially observable systems. It is well known that high-gain observers are applicable to this class of nonlinear systems. For simplicity, we consider the single-output ( $n_y = 1$ ) autonomous system

$$\dot{x} = f_x(x, y), \quad \dot{y} = f_y(x, y).\quad (38)$$

*Definition 2: (Strong differential observability [13])* The single-output system (38) is strongly differentially observable, if there exists  $n_\ell \in \mathbb{N}$  such that  $\mathcal{H}_\ell^e(\chi) = (y, \mathcal{H}_\ell^T(\chi))^T$  is an injective immersion to  $\chi = \text{col}(x, y)$ , where  $\mathcal{H}_\ell(\chi) = [f_y(\chi), L_f f_y(\chi), \dots, L_f^{n_\ell-2} f_y(\chi)]^T$ .

The above definition can be equivalently expressed as  $\mathcal{H}_\ell(\chi)$  is an injective immersion to  $x$  uniformly in  $y$ . We have the following.

*Proposition 4:* If the system (38) is strongly differentially observable with  $n_\ell = n_x + n_y$ , then the function

$$\phi(x, y) = \mathcal{H}_\ell(x, y) - \ell \Lambda y$$

satisfies **A2** semi-globally, where  $\ell > 0$ ,  $\Lambda = [\lambda_1, \dots, \lambda_{n_x}]^T$  with all the roots of  $s^{n_x} + \lambda_1 s^{n_x-1} + \dots + \lambda_{n_x}$  in  $\mathbb{C}_-$ .

*Proof 9:* We may verify that the above construction yields  $\dot{z} = f_z(x, y)$  with

$$f_z = \ell Q (\mathcal{H}_\ell(x, y) - \Lambda y) + \ell Q \Lambda y + b(x, y)$$

and

$$Q = \begin{bmatrix} -\lambda_1 & 1 & & \\ \vdots & & \ddots & \\ \vdots & & & 1 \\ -\lambda_{n_x} & 0 & \dots & 0 \end{bmatrix}, \quad b(x, y) = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ L_f^{n_x-1} f_y(x, y) \end{bmatrix}.$$

<sup>5</sup>Here, we refer to Luenberger’s initial idea by involving a change of coordinate in [24]. We adopt the qualifier “initial” since what is nowadays called “Luenberger observer” is different from [24], see [5] for more discussions.

We note that

$$\begin{aligned} \frac{\partial f_z}{\partial x} \left[ \frac{\partial \phi}{\partial x} \right]^{-1} &= \frac{\partial \beta}{\partial z} \Big|_{z=\phi(x,y)} \\ &= \ell Q + \frac{\partial \phi^L}{\partial z} \Big|_{z=\phi(x,y)} \\ &= \ell Q + \frac{\partial b}{\partial x} \left( \frac{\partial \mathcal{H}_\ell}{\partial x} \right)^{-1}. \end{aligned}$$

Since  $Q$  is a Hurwitz matrix, there exists a matrix  $M \succ 0$  such that  $MQ + Q^\top M = -I$ . By selecting a sufficiently large  $\ell > 0$ , and noting  $\mathcal{H}_\ell$  independent of the parameter  $\ell$ , we guarantee **A2** semi-globally.

*Remark 8:* The above result can be extended to more general cases—multi-input and nonautonomous. Furthermore,  $n_\ell = n_\chi$  is the most widely studied case in high-gain observers. If a given system is strongly differentially observable with  $n_\ell > n_\chi$ , we need to add an augmentation dynamics as done in Section III-D.

### C. Connections with the full-order observer in [26]

Our previous paper [26] uses a similar underlying mechanism to design full-order contracting observer for the system

$$\dot{\mathbf{x}} = f(\mathbf{x}, u), \quad y = h(\mathbf{x}) \quad \mathbf{x} \in \mathbb{R}^n, \quad y \in \mathbb{R}^{n_y}. \quad (39)$$

To this end, we aim to find a coordinate change  $z = \phi(\mathbf{x})$  with  $\dot{z} = f_z(\mathbf{x}, y, u)$  with some function  $f_z$ , which may be non-unique, but satisfying

- C1'** (correctness)  $f_z(\mathbf{x}, h(\mathbf{x}), u) = \Phi(\mathbf{x})f(\mathbf{x}, u)$ , with  $\Phi := \frac{\partial \phi}{\partial \mathbf{x}}$  full rank;
- C2'** (contraction)  $\Phi^\top P^{-1}F + F^\top P^{-1}\Phi \prec 0$ , with constant  $P \succ 0$  and  $F := \frac{\partial}{\partial \mathbf{x}} f(\mathbf{x}, y, u)$  and viewing  $(u, y)$  as exogenous signals.

Then, a full-order observer in the *original coordinate* is given by

$$\dot{\hat{\mathbf{x}}} = \Phi(\hat{\mathbf{x}})^{-1} f(\hat{\mathbf{x}}, y, u). \quad (40)$$

Its extension to discrete time and sampled-data versions were also considered in [26].

In Theorem 1 and Proposition 2 we extend the results in [26] to the reduced-order observer design in transformed coordinates with arbitrary dimension  $n_\xi \geq n_x$ , thus enlarging the domain of applicability to more nonlinear systems. Here are some key differences between them.

- For the full-order case in [26], the system dynamics in the transformed  $z$ -coordinate is partially or virtually contracting [44], since the output  $y$  is explicitly a function of the state  $z$ . But it is not the case for the reduced-order observer design in this paper.
- The convex condition **H3**, unlike the case in [26], introduces a nonlinear output injection term  $\varphi(y)$ , making it quite general. Indeed, coordinate changes in the form of  $\phi(x) = Px + \varphi(y)$  are widely adopted in the field of nonlinear observers [4].
- An advantage of the full-order design [26] is its absence of computing the left inverse mapping, and there is no need to guarantee the “integrability” of  $\Phi(\mathbf{x})$ .

### D. Connections with I&I observers

To the best of our knowledge, the most comprehensive design approach of nonlinear reduced-order observers until now may refer to the I&I observers, proposed in [16] and later elaborated in [4, Chapter 5], by means of rendering attractive an appropriately selected invariant manifold in extended state space. The key extra degree of freedom given in I&I observers relies on a coordinate transformation in order to be able to get a (generally nonlinear) asymptotically stable error dynamics. The procedure is similar to the one done in Section III-A. In this subsection, we will clarify the connections and differences between the proposed designs and I&I observers. Let us recall the main result in I&I observers.<sup>6</sup>

*Proposition 5:* [16] Consider the system (1) with forward complete solutions. Suppose that there exist mappings  $\rho : \mathbb{R}^{n_\xi} \times \mathbb{R}^{n_y} \rightarrow \mathbb{R}^{n_\xi}$ ,  $\varphi : \mathbb{R}^{n_x} \times \mathbb{R}^{n_y} \rightarrow \mathbb{R}^{n_\xi}$  with a left inverse  $\varphi^L(\cdot, y)$  such that the following holds.

- I1** For all  $(y, \xi) \in \mathbb{R}^{n_y} \times \mathbb{R}^{n_\xi}$ ,  $\det(\nabla_\xi \rho) \neq 0$ .
- I2** The system

$$\begin{aligned} \dot{e} &= \left[ \frac{\partial \varphi}{\partial y}(x, y) f_y(x, y, u) + \frac{\partial \varphi}{\partial x}(x, y) f_x(x, y, u) \right] \Big|_{x=\hat{x}}^{x=x} \\ &\quad + \frac{\partial \rho}{\partial y} [f_y(x, y, u) - f_y(\hat{x}, y, u)] \end{aligned} \quad (41)$$

is asymptotically stable at the origin uniformly in  $x, y$  and  $u$  with  $\hat{x} := \varphi^L(\varphi(x, y) + e, y)$ . Then, selecting

$$\begin{aligned} N(\cdot) &= \left( \frac{\partial \rho}{\partial \xi} \right)^{-1} \left[ \frac{\partial \varphi}{\partial x}(\hat{x}, y) f_x(\hat{x}, y, u) \right. \\ &\quad \left. + \frac{\partial \varphi}{\partial y}(\hat{x}, y) f_y(\hat{x}, y, u) - \frac{\partial \rho}{\partial y}(\xi, y) f_y(\hat{x}, y, u) \right] \\ H(\cdot) &= \varphi^L(\rho(\xi, y), y). \end{aligned}$$

makes the system (3) a convergent observer.  $\triangleleft$

We are now ready to show the connections between I&I reduced-order observers and the proposed contracting observer design in Section III. In order to simplify the analysis, instead of **I1** we additionally require a slightly stronger condition

$$|\det(\nabla_\xi \rho)| > \rho_0, \quad \rho_0 > 0. \quad (42)$$

Then we have a global inverse  $\rho^L : \mathbb{R}^{n_\xi} \times \mathbb{R}^{n_y} \rightarrow \mathbb{R}^{n_\xi}$  such that  $\rho^L(\rho(\eta, y), y) = \eta$  for any  $\eta \in \mathbb{R}^{n_\xi}$ .

We have the following.

*Proposition 6:* Consider the system (1) admitting an I&I observer identified in Proposition 5, assuming that the mappings satisfy (42) and **I2** with  $n_\xi = n_x$ . Then, a variant of **A2**, namely

- A2'** The LTV system  $\frac{d}{dt} \delta \xi = \frac{\partial f_z}{\partial x}(x, y, u) \left[ \frac{\partial \phi}{\partial x}(x, y) \right]^{-1} \delta \xi$  is asymptotically stable uniformly along any possible trajectory of  $(x, y, u)$ .

holds, and the given system has a contracting reduced-order observer in the sense of Theorem 1. Furthermore, the obtained observer *exactly* coincides with the I&I observer by selecting  $\phi(x, y) = \rho^L(\varphi(x, y), y)$ .

<sup>6</sup>In [16] all the transformations, *i.e.*  $\rho$  and  $\varphi$ , are allowed to be time-varying functions. We here adopt the time-invariant case to simplify the presentation.

*Proof 10:* For convenience, we define a new intermediate variable as  $\eta := \varphi(x, y)$ . Invoking (42) and the left invertibility of  $\varphi(x, y)$ , we define a mapping  $\phi^L : \mathbb{R}^{n_x} \times \mathbb{R}^{n_y} \rightarrow \mathbb{R}^{n_x}$  as

$$\phi^L(z, y) := \varphi^L(\rho(z, y), y),$$

and with some substitutions we have

$$\phi^L(z, y)|_{z=\phi(x, y)} = \varphi^L(\rho(\phi(x, y), y), y) = x,$$

thus  $\phi^L$  being the left inverse of the composite function  $\phi(\cdot)$ .

In **I2**, the variable  $e$  is, indeed, the estimation error in the  $\eta$ -coordinate, *i.e.*,  $e = \rho(\xi, y) - \varphi(x, y)$ . Since the mapping  $\rho(\cdot)$  identifies a diffeomorphism, one may equivalently define the error in the  $\xi$ -coordinate as

$$e_z := \xi - \rho^L(\varphi(x, y), y),$$

implying that **I2** is equivalent to the uniform asymptotic stability of the dynamics of  $e_z$  with respect to the origin.

The dynamics of  $e_z$  is given by

$$\begin{aligned} \dot{e}_z &= N(\xi, y, u) - \left[ \frac{\partial \rho^L}{\partial \eta} \right] \left[ \frac{\partial \varphi}{\partial x} f_x(x, y, u) + \frac{\partial \varphi}{\partial y} f_y(x, y, u) \right] \\ &\quad - \frac{\partial \rho^L}{\partial y} f_y(x, y, u) \\ &= N(\xi, y, u) - \left[ \frac{\partial \rho}{\partial z} \right]^{-1} \left[ \frac{\partial \varphi}{\partial x} f_x + \frac{\partial \varphi}{\partial y} f_y \right] + \left[ \frac{\partial \rho}{\partial z} \right]^{-1} \frac{\partial \rho}{\partial y} f_y \\ &= \left( \left( \frac{\partial \rho}{\partial z}(\phi(x, y), y) \right)^{-1} \times \right. \\ &\quad \left. \left( \frac{\partial \varphi}{\partial x} f_x(x, y, u) + \frac{\partial \varphi}{\partial y} f_y(x, y, u) - \frac{\partial \rho}{\partial y} f_y(x, y, u) \right) \right) \Big|_{\hat{x}} \\ &= \left( \frac{\partial \phi}{\partial x}(x, y) f_x(x, y, u) + \frac{\partial \phi}{\partial y}(x, y) f_y(x, y, u) \right) \Big|_{\hat{x}} \\ &= f_z(\hat{x}, y, u) - f_z(x, y, u) \\ &= f_z(\phi^L(\phi(x, y) + e_z, y), y, u) - f_z(x, y, u) \\ &= \int_0^1 \frac{\partial f_z}{\partial x} \left[ \frac{\partial \phi}{\partial x} \right]^{-1} (\phi(x, y) + \mu e_z, y, u) d\mu \cdot e_z \end{aligned} \quad (43)$$

where in the third equation we substitute the mapping  $N(\cdot)$  in **I2**, with  $\hat{x} = \phi^L(\phi(x, y) + e_z, y)$ .

Now we consider a line segment  $\gamma(s)$  between two initial conditions  $\xi(0)$  and  $z(0) = \phi(x(0), y(0))$  defined by the parameterization

$$\gamma(s) = s\xi(0) + (1-s)z(0), \quad s \in [0, 1].$$

We denote  $Z(t; \gamma(s), y, u)$  as the solution of the dynamical system

$$\dot{z} = f_z(\phi^L(z, y), y, u)$$

under the exogenous signals  $(u, y)$  from the initial condition  $\gamma(s)$  with  $s \in [0, 1]$ . We define the infinitesimal displacement  $\delta\xi$  as  $\delta\xi = \frac{\partial Z(t; \gamma(s), y, u)}{\partial s}$ , the dynamics of which is governed by the system in **A2'**, with  $x(t) = \phi^L(Z(t; \gamma(s), y, u), y)$ . From the definition of  $\delta\xi$ , we have

$$\xi(0) = z(0) + \delta\xi(0).$$

Invoking (43), it then yields  $\xi(t) \equiv z(t) + \delta\xi(t)$ , and noting the uniform asymptotic stability of the  $e_z$ -system, thus

$$|\delta\xi(t)| \leq \kappa(|\delta\xi(0)|, t), \quad t \geq 0$$

for some function  $\kappa$  of class  $\mathcal{KL}$ .

Some further discussions are in order.

- 1) The proposed designs—including Theorems 1–2—provide “more constructive” solutions to the problem of reduced-order observer design, compared with the I&I methodology. This is because the conditions in the proposed methods only constitute of some differential inequalities, but enjoy sufficient degrees of freedom. In particular, the results in Theorem 2 and Proposition 2 involve some computationally efficient convex conditions.
- 2) In Theorem 1, we need to search for two mappings—the coordinate transformation  $\phi(x, y)$  and the metric  $M$ , and in Theorem 2 we need one more mapping  $f_z$  to achieve convex relaxation. In contrast, when designing I&I observers we need to search simultaneously for four mappings—the coordinate change  $\varphi(x, y)$ , its left inverse  $\varphi^L(z, y)$ , another mapping  $\rho$  and a Lyapunov function (to show convergence).
- 3) If the mapping  $\rho(\cdot)$  satisfies (42) instead of **II**, as clearly shown in the proof of Proposition 6, it is unnecessary to introduce the mapping  $\rho$  in reduced-order observer design. We are able to “gather” all the degrees of freedom from  $\rho$  and  $\varphi$  in the transformation  $\phi$ .

### E. Connections with transverse contraction

In this subsection, we study the link between the proposed designs and transverse contraction, a notion introduced in [27] to analyze attractive limit cycles. It will be shown that designing a reduced-order observer is equivalent to find a coordinate transformation in which the system is transversely contracting.

We slightly extend [27, Definition 1] as follows.

*Definition 3: (transverse contraction)* The forward complete system (2) is said to be transversely contracting with rate  $\lambda > 0$  (or transversely asymptotically contracting) w.r.t. the function  $\psi$ , where  $\psi : \mathbb{R}^{n_x} \rightarrow \mathbb{R}^r$  ( $1 \leq r < n_x$ ), if for any initial condition pair  $(\chi_a, \chi_b) \in \mathbb{R}^{2n_x}$  and input  $u$  we have

$$|\psi(\chi(t; \chi_a, u)) - \psi(\chi(t; \chi_b, u))| \leq e^{-\lambda t} b(\chi_a, \chi_b) \quad (44)$$

for some function  $b(\cdot) \geq 0$  with  $b(\chi, \chi) = 0$  or

$$|\psi(\chi(t; \chi_a, u)) - \psi(\chi(t; \chi_b, u))| \leq \kappa(|\chi_a - \chi_b|, t) \quad (45)$$

for transverse asymptotic contraction, with  $\kappa$  of class  $\mathcal{KL}$ .  $\triangleleft$

It is clear that the definition in [27], tailored for orbital stability, is a particular case of Definition 3. Considering the simple case that a given autonomous system has a unique attractive limit cycle represented by the function  $q(\chi) = 0$  with  $q : \mathbb{R}^{n_x} \rightarrow \mathbb{R}^r$ , we recover the above definition by choosing  $\psi$  as  $q(\chi)$ . We provide a sufficient condition to achieve transverse asymptotic contraction below. The condition is motivated by the concept of *asymptotic phase* in dynamical systems and also used in [28, Corollary 2] to verify the existence of asymptotic invariant manifold, also similar to what we obtain as follows.

*Proposition 7:* Consider the system (2) forward invariant in a compact set  $\mathcal{E} := \mathcal{E}_x \times \mathcal{E}_y$ . Suppose there exists a mapping

$\psi : \mathbb{R}^{n_x} \rightarrow \mathbb{R}^r$  ( $1 \leq r < n_x$ ) with  $\nabla \psi$  full rank, and a uniformly bounded metric  $P(\chi) \in \mathbb{R}_{>0}^{n_x \times n_x}$  such that

$$\frac{\partial \psi}{\partial \chi} \left( \partial_f P + P(\chi) \frac{\partial f}{\partial \chi} + \frac{\partial f^\top}{\partial \chi} P(\chi) \right) \frac{\partial \psi^\top}{\partial \chi} < 0 \quad (46)$$

holds. Then,

- 1) the system (2) is transversely asymptotically contracting with respect to  $\psi$ .
- 2) it admits a trivial functional observer  $\dot{\hat{\chi}} = f(\hat{\chi}, u)$  with respect to the function  $\psi$ .
- 3) if there exists a class  $\mathcal{K}_\infty$  function  $\rho_y$  (parameterised by  $y$ ) satisfying the injectivity  $|x_a - x_b| \leq \rho_y(|\psi(x_a, y) - \psi(x_b, y)|)$ ,  $\forall (x_a, x_b) \in B_\epsilon^2(\mathcal{E}_x)$  uniformly in  $y$  for some  $\epsilon > 0$ , the estimate

$$\hat{x} = \psi^L(\psi(\hat{\chi}), y) := \underset{\eta \in B_\epsilon(\mathcal{E})}{\operatorname{argmin}} |\psi(\eta, y) - \psi(\hat{\chi})| \quad (47)$$

guarantees (4).  $\triangleleft$

Its proof is given in Appendix A. We here make some comparisons between the results in Section III and transverse contraction.

- 1) The mapping (36) with  $\phi$  defined in Theorem 1, is a diffeomorphism, and thus we may rewrite the full system dynamics in the  $(z, y)$ -coordinate as

$$\Sigma_z : \dot{z} = f_z(z, y, u), \quad \dot{y} = f_y(z, y, u).$$

If the system (2) has a convergent reduced-order observer in the sense of Theorem 1, the condition A2 makes the system  $\Sigma_z$  transversely asymptotically contracting w.r.t. the linear mapping  $\psi$  and the metric defined as

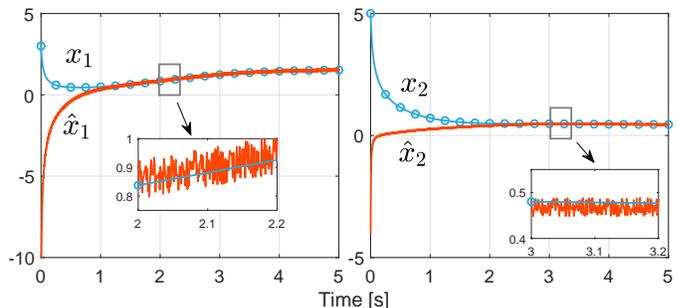
$$\psi = \begin{bmatrix} I_{n_x} & 0_{n_y \times n_y} \end{bmatrix} \begin{bmatrix} z \\ y \end{bmatrix}, \quad P = \begin{bmatrix} P_z(z) & P_{yz}(z, y) \\ P_{yz}^\top(z, y) & P_y(z, y) \end{bmatrix}$$

where  $P_z(z)$  defined in Theorem 1 (with a slight abuse of notations), and  $P_{yz}$  and  $P_{yy}$  are any mappings which guarantee  $P$  a uniformly bounded metric.

- 2) The system (2) in its original coordinate does not verify transverse contraction. The main difference between Theorem 1 and Proposition 7 is the way how we use  $y$  and its estimate  $\hat{y}$ . In the latter, we *do not* try to make full use of the information contained in the measured output when designing the functional observer  $\dot{\hat{\chi}} = f(\hat{\chi}, u)$ , making the sufficient condition relatively conservative. Meanwhile, transverse contraction implies the existence of a functional observer.
- 3) Some definitions similar to transverse contraction have been studied in the literature. In [12], the authors introduce horizontal contraction, a notion defined on *tangent space*, to study the incremental stability along specific distributions. Here, the transverse contraction is defined on state space. In [15] the concept of semi-contraction is recently proposed, which can be regarded as a particular case. In Appendix B, we give an alternative sufficient condition to transverse contraction.

## V. EXAMPLES

In this section, we use the proposed methods to design reduced-order observers for four examples, including a numerical one and three physical benchmarks.



**Fig. 1:** Simulation results of the polynomial system (48)

### A. A polynomial example

The first example is a polynomial system studied in [11, 26], with full-order observers presented therein. We use this example to show how to apply Theorem 2 to design an observer via convex optimization. The dynamics is given by

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} x_1 - \frac{1}{3}x_1^3 - x_1x_2^2 \\ x_1 - x_2 - \frac{1}{3}x_2^3 - x_2x_1^2 \end{bmatrix}, \quad \dot{y} = x_1. \quad (48)$$

We aim to design a reduced-order observer using H3 in Theorem 2, with mappings  $P \in \mathbb{R}_{>0}^{2 \times 2}$ ,  $f_z$  and  $\varphi : \mathbb{R} \rightarrow \mathbb{R}^2$  to be determined. We search polynomials using a positivstellensatz [35] as

$$f_{z1}(x, y), f_{z2}(x, y) \in \mathbb{R}_3[x, y], \quad \varphi_1(y), \varphi_2(y) \in \mathbb{R}_2[y]$$

and select  $\lambda = 1$ . A reduced-order observer is found, via Yalmip [23] and Mosek, in 0.61 seconds on a desktop with 3.00 GHz Intel Core i7-9700 CPU and 3 GB RAM. The obtained mappings are  $P = \operatorname{diag}(0.6370, 0.6369)$  and  $\varphi(y) = \operatorname{col}(-2.1872y, -0.6368y)$ , with the observer given by

$$\dot{\xi} = \begin{bmatrix} 0.6370(\hat{x}_1 - \frac{1}{3}\hat{x}_1^3 - \hat{x}_1\hat{x}_2^2) \\ 0.6369(\hat{x}_1 - \hat{x}_2 - \frac{1}{3}\hat{x}_2^3 - \hat{x}_2\hat{x}_1^2) \end{bmatrix}, \quad \hat{x} = P^{-1}(\xi - \varphi(y)).$$

We show in Fig. 1 the simulations with initial conditions  $x(0) = [3 \ 5]^\top$ ,  $y(0) = -4$  and  $\xi(0) = [0 \ 0]^\top$ , where high-frequency “measurement noise” has been added in the output channel. It is generated by the block “Uniform Random Number” in Matlab/Simulink, with sample time 0.001 and the interval  $[-0.02, 0.02]$ . As we can see, the estimates converge to their true values after a short transient stage, and the observer is not sensitive to measurement noise.

### B. Magnetic levitation system

In the second example, we study an electromechanical system—magnetic levitation (MagLev) model, which is given by

$$\begin{bmatrix} \dot{\lambda} \\ \dot{q} \\ \dot{p} \end{bmatrix} = \begin{bmatrix} \frac{R}{k}(q - c)\lambda + u \\ \frac{1}{m}p \\ \frac{1}{2k}\lambda^2 - mg \end{bmatrix} \quad (49)$$

with  $\lambda$  the flux linkage,  $q \in (-\infty, c)$  the position, considering physical constraints,  $p$  the momenta and  $u$  the input voltage, where  $R, m, c, k > 0$  are some physical parameters [46]. Here, we assume only the position  $y = q$  measurable, with unknown states  $x = \operatorname{col}(\lambda, p)$ .<sup>7</sup>

<sup>7</sup>See [46] for a more challenging task of sensorless estimation, assuming only current  $i := \frac{1}{k}(c - q)\lambda$  measured.

The MagLev model is indeed a polynomial system, and thus the convex formulations in Theorem 2 are applicable. However, we provide an alternative of *recursive* searching design. In the first step, we consider a two-dimensional system with  $\lambda$  as the unknown state, and the measured state  $y$ . Then using **H3** to search for parameter  $P_1 > 0$  and mapping  $\varphi_1 : \mathbb{R} \rightarrow \mathbb{R}$  under the constraint  $q < c$ . It is trivial to get an admissible solution  $P_1 = 1$  and  $\varphi_1 = 0$ , and we are able to design an observer to get exponentially convergent estimate of  $\lambda$ . In the second step, we regard  $p$  as the unmeasured state,  $y$  the output and  $\lambda$  a “known” input, and then solve the convex optimization in Theorem 2 again. We summarize the result below.

*Proposition 8:* Consider the MagLev model (49) with  $q < c$ . The observer

$$\dot{\hat{\xi}} = \begin{bmatrix} \frac{R}{k}(y-c)\xi_1 + u \\ -\frac{\ell}{m}\xi_2 - \frac{\ell^2}{m}y - mg + \frac{1}{2k}\xi_1^2 \end{bmatrix}, \quad \hat{x} = \xi + \begin{bmatrix} 0 \\ \ell y \end{bmatrix} \quad (50)$$

with any gain  $\ell > 0$ , guarantees (4) globally.  $\triangleleft$

In [45] the authors present an observer, via contraction analysis, with only  $y = q$  measured, as done here. However, that design only guarantees local convergence. We give the simulation results of both the proposed observer (50) and the full-order observer in [45] with all physical parameters borrowed from [46], and  $\ell = 0.5$  for (50) and  $r_0 = 5$ ,  $k_0 = 600$  for the design in [45]. We consider the system initial condition  $x(0) = [0.003, 0, 0]^\top$ , and two sets of observer initial conditions. In the first, we assume having a good initial guess, *i.e.*,  $\hat{\xi}(0) = [0, 0]^\top$  for the proposed observer and  $\hat{x}(0) = [0.0035, 0, 0]^\top$  for the other. As we can see, both of them get satisfactory performance. In the second, we consider large initial condition differences, *i.e.*,  $\xi(0) = [0.010, 0.010]^\top$  and  $\hat{x}(0) = [0.006, 0.002, 0.0078]^\top$ . For this case, our proposed observer states still converge to the true ones, but the observer in [45] diverges, showing its locality.

### C. Inverted pendulum on cart system

In this subsection, we consider a mechanical model of cart-pendulum system, the model of which is given by

$$\dot{q} = \Psi(q)p, \quad \dot{p} = \Psi^\top(q)[\nabla V(q) - G(q)u] \quad (51)$$

with

$$\Psi(q) = \begin{bmatrix} \frac{\sqrt{m}}{\sqrt{m-b^2\cos^2(q_1)}} & 0 \\ -b\cos(q_1) & \frac{1}{\sqrt{m}} \end{bmatrix}, \quad G = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

and  $V(q) = a\cos(q_1)$ , where  $q \in \mathbb{S} \times \mathbb{R}$  is the generalized position—the pendulum angle  $q_1$  and the horizontal position of the car,  $p \in \mathbb{R}^2$  is a momenta-like variable with  $\Psi(q)p$  the generalized velocities, with parameters  $a, b, m > 0$ . We assume that only  $y = q$  is measured, and the task is to estimate the internal state  $x = p$ , equivalently the speed. It is well known that high-gain observer is applicable to this task with a *semi-global* domain of attraction by increasing the adaptation gain. We here use **H3** to obtain a globally convergent observer *without* high-gain injection.

In the design, we fix  $P = I_2$  and get the transformation  $\phi(x, y) = x + \varphi(y)$ , with  $\varphi(y)$  to be found. Note that  $\Psi(q)$  is

full rank, and the dynamics of  $y$  is affine in  $x$ . The condition **H3** motivates us to select  $\varphi(y)$  satisfying the PDE

$$\frac{\partial \varphi}{\partial y}(y)\Psi(y) = -\lambda I_2 \quad (52)$$

with  $\lambda > 0$ . Using the symbolic solver, we get a feasible solution

$$\varphi(y) = -\lambda \begin{bmatrix} \int_0^{y_1} \sqrt{1 - \frac{b^2}{m} \cos^2(s)} ds \\ \frac{b}{\sqrt{m}} \sin(y_1) + \sqrt{m}y_2 \end{bmatrix}. \quad (53)$$

We are ready to present the observer design as follows.

*Proposition 9:* Consider the system (51) with  $q$  measured. The reduced-order observer

$$\begin{aligned} \dot{\hat{\xi}} &= -\lambda \hat{\xi} + \lambda \varphi(y) + \Psi^\top(y)[\nabla V(y) - Gu] \\ \hat{x} &= \xi - \varphi(y). \end{aligned} \quad (54)$$

guarantees  $\lim_{t \rightarrow \infty} |\hat{x}(t) - p(t)| = 0$  exponentially.

*Proof 11:* Since the function (53) is the solution of (52), we thus have  $F^\top + F + 2\lambda I_2 = 0$  satisfying the condition **H3**. Invoking

$$\begin{aligned} \frac{d}{dt}(\phi(x, y)) &= \frac{\partial \varphi}{\partial y}(y)\Psi(y)x + \Psi^\top(y)[\nabla V(y) + Gu] \\ &= -\lambda \phi(x, y) + \lambda \varphi(y) + \Psi^\top(y)[\nabla V(y) + Gu], \end{aligned}$$

thus we get the observer above, such that  $\hat{x}$  is an exponentially convergent estimate of  $p = \Psi^{-1}\dot{q}$ .

We also give the simulation results in Fig. 3 with parameters  $m = 1, a = 1, b = 0.1$ , initial conditions  $\xi(0) = [0, 0]^\top$ ,  $p(0) = [0.4, 0.3]^\top$  and  $q(0) = [\frac{1}{2}\pi - 0.1, -0.1]$ , and  $u(t) = 0.2\cos(t)$ , using different adaptation gains. Clearly, a large parameter  $\lambda$  implies a fast convergence speed.

### D. A biological reactor

The last example is used to illustrate the case  $n_\xi > n_x$  in Subsection III-D. We now consider a model of reaction in [37]

$$\dot{x} = -\frac{1}{k}\mu(x)y, \quad \dot{y} = \mu(x)y, \quad (55)$$

with the states  $(x, y) \in \mathbb{R}_{\geq 0}^2$ , a parameter  $k > 0$  and  $\mu(\cdot)$  a non-negative smooth function such that  $\mu(0) = 0$ , and we refer the reader to [37] for the physical meanings of states and parameters. We here adopt the non-monotonic growth function

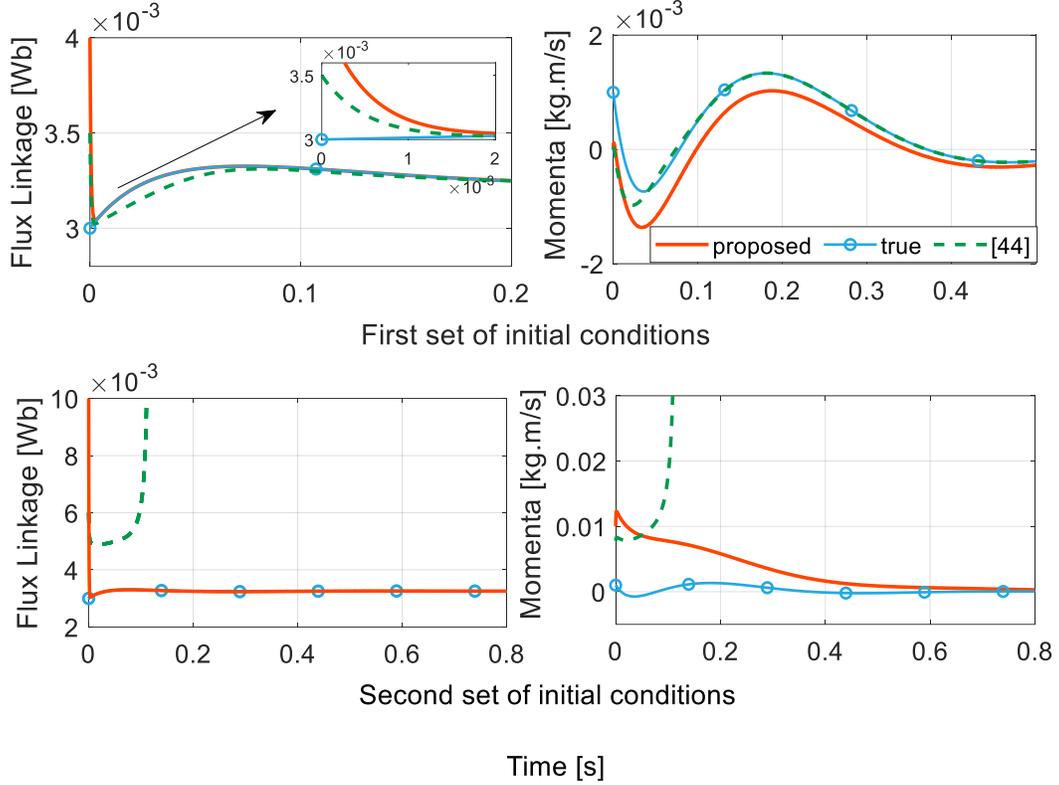
$$\mu(x) = rx \left(1 - \frac{x}{c}\right), \quad x \in [0, c]$$

with  $r, c > 0$  to characterize some inhibition effect of the reactor. To simplify the presentation, we assume all the positive constants equal to one. It was shown in [37] that  $[y, f_y]$  is not injective as soon as  $\mu(x)$  is non-monotonic, thus with the need to increase the observer dimension. Here, we follow Proposition 2 to provide a new design as follows.

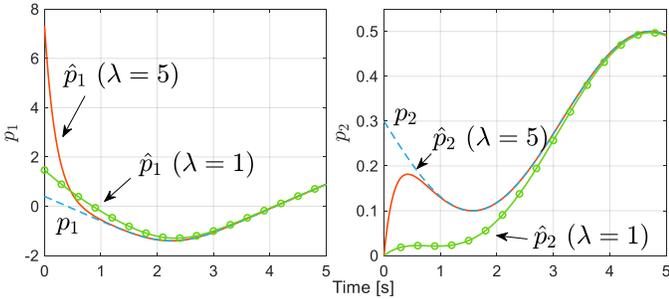
*Proposition 10:* Consider the system (55). The observer

$$\dot{\hat{\xi}} = -\hat{\xi} + \begin{bmatrix} \ln y \\ 2y + 1 \\ y + y^2 \\ \frac{1}{2}\xi_2 - \frac{1}{2}(\xi_2^2 - 4(\xi_3 - \xi_1 + \ln y))^{\frac{1}{2}} \end{bmatrix}, \quad \hat{x} = \xi_4 - y$$

guarantees  $\lim_{t \rightarrow \infty} |\hat{x}(t) - x(t)| = 0$ .



**Fig. 2:** The unmeasured states and their estimates from the proposed observer (50) and the structure preserving observer in [45], which is a local design, for the MagLev model (49)



**Fig. 3:** Estimation performance of the cart-pendulum system

*Proof 12:* We design an augmentation  $\dot{w} = Aw + f_a(y)$  with  $A = -I_{3 \times 3}$ , the initial condition constraints  $w_2(0) = w_3(0) = 0$ ,  $w_1(0) = (x(0) + y(0))^2 + \ln(y(0))$ . It should be underscored here that such initial conditions are *not* used in the observer implementation. It is straightforward to verify for all  $t \geq 0$

$$2z_4(t) = w_2(t) - [(w_2^2(t) - 4(w_3(t) - w_1(t) + \ln(y(t))))^{\frac{1}{2}},$$

with  $z_4 := x + y$ . By selecting the transformation as  $\phi(x, w, y) = \text{col}(w, x + y) \in \mathbb{R}^4$ , it is straightforward to verify **H1'**, **H2'** and **H3'**.

Our design is compared with the high-gain observer (HGO)

proposed in [37], which is given by

$$\dot{\xi} = \begin{bmatrix} \xi_2 \\ -\xi_3 \min(e^{\xi_1}, \bar{\xi}_1) \\ \text{sat}_{m, M} \left( e^{\xi_1} (2\xi_2^2 - \xi_3 \tilde{\phi}(\xi)) \right) \end{bmatrix} + \begin{bmatrix} -3\ell \\ -3\ell^2 \\ 3\ell \xi_3 \mathbb{1}_{\xi \leq \ln \bar{\xi}_1} + \ell^3 \max(e^{-\xi_1}, \frac{1}{\bar{\xi}_1}) \end{bmatrix} e_y$$

$$\hat{x} = \begin{bmatrix} e^{\xi_1} \\ \frac{1}{2}(1 - \tilde{\phi}(\chi)) \end{bmatrix}, \quad e_y := \xi_1 - \ln y$$

with some parameters  $\bar{\xi}_1, m, M, \xi_2^*, P(\xi_2) := \xi_2((\xi_2/\xi_2^*)^2 - 3\xi_2/\xi_2^* + 1)$  and a function

$$\tilde{\phi}(\xi) = \begin{cases} -1 & \text{if } \xi_2 > \xi_2^* \text{ and } \xi_3 < -\xi_2 \\ (\xi_2/\xi_2^*)^2 - 3\xi_2/\xi_2^* + 1 & \text{if } \xi_2 \in (0, \xi_2^*) \text{ and } \xi_3 < P(\xi_2) \\ \xi_3/\xi_2 & \text{otherwise.} \end{cases}$$

Clearly, the HGO is much more complicated, stemmed from Lipschitsian extensions. These two observers are compared via simulations, with the initial conditions  $[0, 0, 0, 0.5]^T$  in the proposed contracting observer, and  $[1, 0.1, 0]^T$ , together with the gain  $\ell = 3$  for the HGO. Some simulations were done for the cases with and without measurement noise, shown in Fig. 4. It is well known that the convergence speed of HGOs mainly depends on the adaptation gain  $\ell > 0$ . In order to make a fair comparison to evaluate their effects to noise, we tuned them, via changing  $\ell$ , such that their convergence times are similar. As expected, our new design overperforms the HGO, in particular in the presence of measurement noise, since the underlying mechanism of HGOs makes them highly sensitive to measurement noise. We underline that the effect of noise in the HGO becomes increasingly serious, since the studied trajectory verifies  $x_2(t) \rightarrow 0$  and  $\xi_2$  is its estimate, with

$\xi_2$  appearing in the denominator of  $\bar{\phi}(\xi)$ . Besides, peaking phenomenon is clearly observed in the HGO, which usually requires some extra modification, *e.g.*, adding saturation function or projector [18], to avoid deleterious effects.

## VI. CONCLUDING REMARKS

In this paper we address the problem of reduced-order observer design for nonlinear control systems by means of contraction analysis and convex optimization. As we have shown, searching for a change of coordinate—after which the system is (transversely) contracting—plays a key role in the constructive procedure, as well as providing sufficient degrees of freedom to characterize “detectability” for nonlinear systems. Then we formulate the obtained sufficient condition into a convex optimization problem by introducing the image as a new decision variable. We apply the main results to linear systems and strongly differentially observable nonlinear systems, showing generality of the proposed framework. Links to I&I observers and transverse contraction are also clarified in the paper. At the end, we apply the method to several academic and practical models as illustrations.

Along this research line, the following problems are under-way.

- If the correctness is infeasible in some sense, the proposed method is still able to provide estimation with ultimate bounded errors. It is also interesting to study interval observers using the introduced framework.
- After getting a convex sets of contracting observers, it is of practical interest to select the best or mix them in the presence of unavoidable measurement noise.
- For nonlinear nonautonomous systems, the necessary condition on the existence of an asymptotic observer is still open. If it exists, it is natural to study the relationship between the necessary condition and the condition **A2**.
- Some practical models are not uniformly observable, but may be asymptotically estimated when the system trajectories satisfy additional excitation conditions. It is interesting to generalize the proposed framework to this case.
- In the recent work [6], the exogenous input  $u(t)$  is involved in a coordinate transformation. If we allow this degree of freedom in the proposed method, it will definitely extend the realm of applicability. At the moment, such a time-varying transformation in our framework only makes sense conceptually, but with difficulty to implement it.

## APPENDIX

### A. THE PROOF OF PROPOSITION 7

Invoking Wazewski theorem [5] and the fact that Euclidean space is contractible, we obtain that the Jacobian completion of  $\nabla\psi(\chi)$  is solvable, *i.e.*, there exists a mapping  $\tilde{g} : \mathbb{R}^{n_\chi} \rightarrow \mathbb{R}^{n_\chi \times (n_\chi - r)}$  such that

$$\det \left[ \frac{\partial \psi}{\partial \chi}^\top(\chi) \quad \tilde{g}(\chi) \right] \neq 0.$$

We can use Gram-Schmidt process to orthonormalize the column vectors of  $\left[ \nabla\psi(\chi) \quad \tilde{g}(\chi) \right]$  with the standard inner

product in Euclidean space, and the obtained last  $(n_\chi - r)$  vectors are denoted as  $g(\chi)$ . It is clear that  $g(\chi)$  is uniformly full rank, and each of its column  $g_i(\chi)$  verifies

$$g_i(\chi) \in \ker \left( \frac{\partial \psi}{\partial \chi}(\chi) \right), \quad \forall i \in \{1, \dots, n_\chi - r\}. \quad (56)$$

Now we select an arbitrary piecewise continuous  $\mathcal{C}^1$  path  $\gamma : [0, 1] \rightarrow \mathbb{R}^{n_\chi}$  between any two point  $\chi_a = \gamma(0)$  and  $\chi_b = \gamma(1)$ . Consider a family of auxiliary systems

$$\frac{d}{dt}\Gamma(s, t) = f(\Gamma(s, t), u(t)) + g(\Gamma(s, t))v(s, t), \quad (57)$$

with  $\Gamma(s, 0) = \gamma(s)$ ,  $\Gamma(\cdot) \in \mathbb{R}^{n_\chi}$  and the *control input*  $v(s, t)$ , where we regard  $u(t)$  as an exogenous time-varying signal. By calculating its differential dynamics, we have

$$\begin{aligned} \frac{d}{dt} \frac{d}{ds} \Gamma(s, t) &= g(\Gamma(s, t)) \frac{d}{ds} v(s, t) \\ &+ \left( \frac{\partial f}{\partial \chi}(\Gamma(s, t), u(t)) + \sum_{i=1}^{n_\chi - r} \frac{\partial g_i}{\partial \chi} v_i(s, t) \right) \frac{d}{ds} \Gamma(s, t) \end{aligned} \quad (58)$$

According to [28, Theorem 1] and the full-rank property of  $g(\cdot)$ , there exists a differential controller  $\frac{d}{ds} v(s, t) = \delta v(\Gamma(s, t), \frac{d}{ds} \Gamma(s, t), t)$  such that the origin of the differential dynamics (58) is asymptotically stable uniformly in  $s \in [0, 1]$  with a quadratic Lyapunov function. Then, the controller  $v(s, t) = - \int_s^1 \delta v(\Gamma(\sigma, t), \frac{d}{ds} \Gamma(\sigma, t), t) d\sigma$  makes the close loop incrementally asymptotically stable, thus

$$\lim_{t \rightarrow \infty} |\Gamma(1, t) - \Gamma(0, t)| = 0. \quad (59)$$

On the other side, since  $\Gamma(1, 0) = \hat{\chi}(0)$  and  $v(1, t) = 0$ , we have

$$\Gamma(1, t) = \hat{\chi}(t). \quad (60)$$

We also have

$$\begin{aligned} &\frac{d}{dt} \left( \psi(\chi(t)) - \psi(\Gamma(0, t)) \right) \\ &= \frac{\partial \psi}{\partial \chi}(\chi(t)) f(\chi(t), u) \\ &\quad - \frac{\partial \psi}{\partial \chi}(\Gamma(0, t)) \left( f(\Gamma(0, t), u) + g(\Gamma(0, t)) v(0, t) \right) \\ &\stackrel{(56)}{=} \frac{\partial \psi}{\partial \chi}(\chi(t)) f(\chi(t), u) - \frac{\partial \psi}{\partial \chi}(\Gamma(0, t)) f(\Gamma(0, t), u), \end{aligned} \quad (61)$$

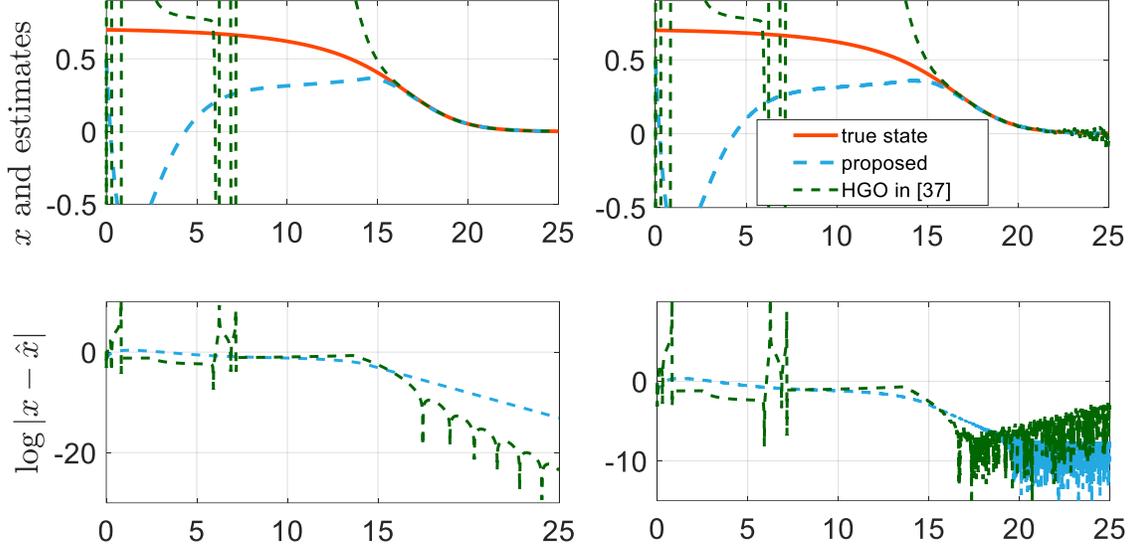
together with  $\Gamma(0, 0) = \chi(0)$  it implying

$$\psi(\chi(t)) = \psi(\Gamma(0, t)), \quad \forall t \geq 0. \quad (62)$$

Combining (59), (60) and (62), we prove the first claim. The claim 2) is trivial to verify by fixing  $\chi = \chi(t; \chi_a, u)$  and  $\hat{\chi} = \chi(t; \chi_b, u)$ , verifying the second claim.

Regarding the last item, invoking the forward invariance assumption, we have  $\psi(\chi) \in \psi(\mathcal{E})$  for all  $t \geq 0$ . Recalling the convergence (45), for any  $\epsilon > 0$  we can always find a moment  $t_\epsilon$  such that  $\psi(\hat{\chi}(t)) \in B_\epsilon(\mathcal{E})$ ,  $\forall t \geq t_\epsilon$ . In terms of injectivity, the mapping defined in (47) guarantees

$$\psi(\hat{\chi}) \in B_\epsilon(\mathcal{E})$$



**Fig. 4:** Simulation results of the reactor model with comparison to the high-gain observer in [37]

implying

$$|\psi^L(\psi(\hat{\chi}), y) - \psi^L(\psi(\chi), y)| \leq \rho_y(|\psi(\hat{\chi}) - \psi(\chi)|).$$

It, then, yields

$$|\hat{x}(t) - x(t)| \leq \rho_y(|\psi(\hat{\chi}(t)) - \psi(\chi(t))|), \quad t \geq t_e.$$

Combining (45), we complete the proof.  $\square$

#### B. AN ALTERNATIVE CONDITION TO TRANSVERSE CONTRACTION

In this section, we give an alternative sufficient condition to transverse contraction, by using a *semi-definite* Riemannian metric  $W(\chi)$ . A similar problem arises in [43].

*Proposition 11:* Consider the system (2) forward invariant in  $\mathcal{E}$  under input  $u(t)$ . Suppose there exists a positive semi-definite  $W : \mathbb{R}^{n_x} \rightarrow \mathbb{R}_{\geq 0}^{n_x \times n_x}$  with rank  $r > 0$ , parameterized as  $W(\chi) := \Psi(\chi)P(\chi)\Psi^\top(\chi)$  with  $\Psi(\chi) \in \mathbb{R}^{n_x \times r}$  and  $P(\chi) \in \mathbb{R}^{r \times r}$  both full rank.<sup>8</sup> If the inequality

$$\partial_f W(\chi) + \frac{\partial f^\top}{\partial \chi} W(\chi) + W(\chi) \frac{\partial f}{\partial \chi} < 0 \quad (63)$$

holds, and  $\nabla \Psi_i(\chi) = (\nabla \Psi_i(\chi))^\top$  for  $i = 1, \dots, r$ , then the given system is transversely asymptotically contracting w.r.t. the function  $\psi$  defined by  $\psi(\chi) = \int_0^1 (\Psi(s\chi))^\top \chi ds$ .

*Proof 13:* We choose a differential Lyapunov function

$$V(\chi, \delta\chi) = \delta\chi^\top W(\chi) \delta\chi,$$

and define

$$\delta\chi_h := \Psi(\chi)\Psi(\chi)^\dagger \delta\chi,$$

and

$$\delta\chi_v := [I - \Psi(\chi)\Psi(\chi)^\dagger] \delta\chi.$$

<sup>8</sup>A positive semi-definite matrix can always be represented in such a parameterization.

For each  $(\chi, \delta\chi) \in \mathbb{R}^{n_x} \times T\mathbb{R}^{n_x}$ , we have

$$\delta\chi = \delta\chi_h + \delta\chi_v,$$

and then  $V(\chi, \delta\chi) = \delta\chi_h^\top W(\chi) \delta\chi_h$ , and thus  $V(\chi, \delta\chi) = V(\chi, \delta\chi_h)$ . The associated differential dynamics is  $\frac{d}{dt} \delta\chi = \frac{\partial f(\chi, u)}{\partial \chi} \delta\chi$ , and for any nonvanishing  $\delta\chi$ ,

$$\dot{V}(\chi, \delta\chi) = \delta\chi^\top (\partial_f W + \frac{\partial f}{\partial \chi} W + W \frac{\partial f^\top}{\partial \chi}) \delta\chi < 0,$$

thus

$$\lim_{t \rightarrow \infty} V(\chi(t), \delta\chi(t)) = 0. \quad (64)$$

It is obvious that  $\nabla \psi(\chi) = \Psi^\top(\chi)$ . Consider a differential curve  $\gamma : [0, 1] \rightarrow \mathbb{R}^{n_x}$  with  $\gamma(s) \in \mathcal{E}$ , we parameterize two different initial conditions as  $\chi_a(0) = \gamma(0)$  and  $\chi_b(0) = \gamma(1)$ . Then, it yields from (64) the following implications.

$$\begin{aligned} \Rightarrow \lim_{t \rightarrow \infty} \int_0^1 \left| \frac{d}{ds} \chi(t; \gamma(s), u)^\top \Psi(\chi(t; \gamma(s), u)) \right|_{P(\cdot)} ds &= 0 \\ \Rightarrow \lim_{t \rightarrow \infty} \frac{d}{ds} \psi(\chi(t; \gamma(s), u)) &= 0 \\ \Rightarrow \lim_{t \rightarrow \infty} |\psi(\chi(t; \chi_a, u)) - \psi(\chi(t; \chi_b, u))| &= 0 \end{aligned}$$

where in the third implication we have used  $\det(P(x)) \neq 0$  and the boundedness of  $P(\cdot)$ . It completes the proof.

#### ACKNOWLEDGEMENT

The authors would like to express their gratitude to three reviewers for their careful reading of our manuscript, and many thoughtful comments that helped improve its clarity.

#### REFERENCES

- [1] N. Aghannan and P. Rouchon, An intrinsic observer for a class of Lagrangian systems, *IEEE Trans. Automatic Control*, vol. 48, pp. 936–945, 2003.

- [2] V. Andrieu and L. Praly, On the existence of a Kazantzis-Kravaris/Luenberger observer, *SIAM J. Control and Optimization*, vol. 45, pp. 432–456, 2006.
- [3] V. Andrieu, G. Besançon and U. Serres, Observability necessary conditions for the existence of observers, *IEEE Conf. Decision and Control*, Florence, Italy, pp. 4442–4447, 2013.
- [4] A. Astolfi, D. Karagiannis and R. Ortega, *Nonlinear and Adaptive Control with Applications*, Springer, London, 2007.
- [5] P. Bernard, *Observer Design for Nonlinear Systems*, Springer, Switzerland, 2019.
- [6] P. Bernard and V. Andrieu, Luenberger observers for nonautonomous nonlinear systems, *IEEE Trans. Automatic Control*, vol. 64, pp. 270–281, 2018.
- [7] P. Bernard, V. Andrieu and L. Praly, Expressing an observer in preferred coordinates by transforming an injective immersion into a surjective diffeomorphism, *SIAM J. Control and Optimization*, vol. 56, pp. 2327–2352, 2018.
- [8] G. Besançon (Ed.), *Nonlinear Observers and Applications*, Springer, Berlin, 2007.
- [9] G. Besançon and H. Hammouri, On uniform observation of nonuniformly observable systems, *Systems & Control Letters*, vol. 29, pp. 9–19, 1996.
- [10] R. Engel and G. Kreisselmeier, A continuous-time observer which converges in finite time, *IEEE Trans. Automatic Control*, vol. 47, no. 7, pp. 1202–1204, 2002.
- [11] X. Fan and M. Arcak, Observer design for systems with multivariable monotone nonlinearities, *Systems & Control Letters*, vol. 50, pp. 319–330, 2003.
- [12] F. Forni and R. Sepulchre, A differential Lyapunov framework for contraction analysis, *IEEE Trans. Automatic Control*, vol. 59, pp. 614–628, 2014.
- [13] J.P. Gauthier and I. Kupka, *Deterministic Observation Theory and Applications*, Cambridge University Press, New York, 2001.
- [14] G.C. Goodwin and R.H. Middleton, The class of all stable unbiased state estimators, *Systems & Control Letters*, vol. 13, pp. 161–163, 1989.
- [15] S. Jafarpour, P. Cisneros-Velarde and F. Bullo, Weak and semi-contraction theory with application to network systems, *ArXiv Preprint*, 2020. (arXiv:2005.09774)
- [16] D. Karagiannis, D. Carnevale and A. Astolfi, Invariant manifold based reduced-order observer design for nonlinear systems, *IEEE Trans. Automatic Control*, vol. 53, pp. 2602–2614, 2008.
- [17] N. Kazantzis and C. Kravaris, Nonlinear observer design using Lyapunov’s auxiliary theorem, *Systems & Control Letters*, vol. 34, no. 5, pp. 241–247, 1998.
- [18] H.K. Khalil, *High-gain Observers in Nonlinear Feedback Control*, SIAM, Philadelphia, 2017.
- [19] H.K. Khalil and L. Praly, High-gain observer in nonlinear feedback control, *Int. J. Robust and Nonlinear Control*, vol. 24, no. 6, pp. 993–1015, 2014.
- [20] G. Kreisselmeier and R. Engel, Nonlinear observers for autonomous Lipschitz continuous systems, *IEEE Trans. Automatic Control*, vol. 48, pp. 451–464, 2003.
- [21] D. Lewis, Metric properties of differential equations, *American Journal of Mathematics*, pp. 294–312, 1949.
- [22] W. Lohmiller and J.-J.E. Slotine, On contraction analysis for non-linear systems, *Automatica*, vol. 34, pp. 638–696, 1998.
- [23] J. Lofberg, Yalmip: A toolbox for modeling and optimization in MATLAB, *Proc. of the CACSD Conf.*, Taipei, Taiwan, 2004.
- [24] D.G. Luenberger, Observing the state of a linear system, *IEEE Trans. Mil. Electron.*, vol. 8, pp. 74–80, 1964.
- [25] D.G. Luenberger, An introduction to observers, *IEEE Trans. Automatic Control*, vol. 16, no. 6, pp. 596–602, 1971.
- [26] I.R. Manchester, Contracting nonlinear observers: Convex optimization and learning from data, *American Control Conf.*, Milwaukee, USA, pp. 1873–1880, June 27–29, 2018.
- [27] I.R. Manchester and J.-J.E. Slotine, Transverse contraction criteria for existence, stability, and robustness of a limit cycle, *Systems & Control Letters*, vol. 63, pp. 32–38, 2014.
- [28] I.R. Manchester and J.-J.E. Slotine, Control contraction metrics: Convex and intrinsic criteria for nonlinear feedback design, *IEEE Trans. Automatic Control*, vol. 62, pp. 3046–3053, 2017.
- [29] R. Marino and P. Tomei, *Nonlinear Control Design: Geometric, Adaptive and Robust*, Prentice-Hall, London, 1995.
- [30] J.D. Miller, Some global inverse function theorems, *J. of Math. Analysis and Appl.*, vol. 100, pp. 375–384, 1984.
- [31] A. Nemirovski, Prox-method with rate of convergence of  $O(1/t)$  for variational inequalities with Lipschitz continuous monotone operators and smooth convex-concave saddle point problems, *SIAM J. Optim.*, vol. 15, pp. 229–252, 2004.
- [32] H. Nijmeijer and T.I. Fossen, (Eds.), *New Directions in Nonlinear Observer Design*, Springer-Verlag, London, 1998.
- [33] R. Ortega, A. Bobtsov, N. Nikolae, J. Schiffer and D. Dochain, Generalized parameter estimation-based observers: Application to power systems and chemical-biological reactors, *Automatica*, vol. 129, 109635, 2021.
- [34] R. Ortega, A. Bobtsov, A. Pyrkin and S. Aranovskiy, A parameter estimation approach to state observation of nonlinear systems, *Systems & Control Letters*, vol. 85, pp. 84–94, 2015.
- [35] P.A. Parrilo, Semidefinite programming relaxations for semialgebraic problems, *Math. Program.*, vol. 96, pp. 293–320, 2003.
- [36] L. Praly, On observers with state independent error Lyapunov function, *Proc. IFAC Symp. Nonlinear Control Syst. (NOLCOS)*, St. Petersburg, Russia, pp. 1349–1354, 2001.
- [37] A. Rapaport and A. Maloum, Design of exponential observers for nonlinear systems by embedding, *Int. J. Robust and Nonlinear Control*, vol. 14, pp. 273–288, 2004.
- [38] R.G. Sanfelice and L. Praly, Convergence of nonlinear observers on  $\mathbb{R}^n$  with a Riemannian metric (Part I), *IEEE Trans. Automatic Control*, vol. 57, pp. 1709–1722, 2011.
- [39] A. Shoshitaishvili, On control branching systems with degenerate linearization, *IFAC Symp. Nonlinear Control Syst.*, pp. 495–500, 1992.
- [40] T.-J. Tarn and Y. Rasis, Observers for nonlinear stochastic systems, *IEEE Trans. Automatic Control*, vol. 21, pp. 441–448, 1976.
- [41] M.M. Tobenkin, I.R. Manchester and A. Megretski, Convex parameterization and fidelity bounds for nonlinear identification and reduced-order modelling, *IEEE Trans. Automatic Control*, vol. 62, no. 7, pp. 3679–3686, 2017.
- [42] J. Tsiniias, Further results on the observers design problem, *Systems & Control Letters*, vol. 14, pp. 411–418, 1990.
- [43] L. Wang, F. Forni, R. Ortega, Z. Liu and H. Su, Immersion and invariance stabilization of nonlinear systems via virtual and horizontal contraction, *IEEE Trans. Automatic Control*, vol. 62, pp. 4017–4022, 2017.
- [44] W. Wang and J.-J.E. Slotine, On partial contraction analysis for coupled nonlinear oscillators, *Biological Cybernetics*, vol. 92, pp. 38–53, 2005.
- [45] A. Yaghmaei and M.J. Yazdanpanah, Structure preserving observer design for port-Hamiltonian systems, *IEEE Trans. Automatic Control*, vol. 64, pp. 1214–1220, 2019.
- [46] B. Yi, R. Ortega, H. Siguerdidjane, J.E. Machado and W. Zhang, On generation of virtual outputs via signal injection: Application to observer design for electromechanical systems, *European J Control*, vol. 54, pp. 129–139, 2020.
- [47] B. Yi, R. Ortega and W. Zhang, On state observers for nonlinear systems: A new design and a unifying framework, *IEEE Trans. Automatic Control*, vol. 64, pp. 1193–1200, 2019.
- [48] B. Yi, R. Wang and I.R. Manchester, Reduced-order observers for nonlinear systems via contraction analysis and convex optimisation, *American Control Conf.*, New Orleans, USA, May 26–28, 2021.