Continuous-time and sampled data stabilizers for nonlinear systems with input and measurement delays

Stefano Battilotti

Abstract—**In this paper, we propose continuous-time and sampled-data output feedback controllers for nonlinear multi-input multi-output systems with time-varying measurement and input delays, with no restriction on the bound or serious limitations on the growth of the nonlinearities. A state prediction is generated by chains of saturated high-gain observers with switching error-correction terms and the state prediction is used to stabilize the system with saturated controls. The observers reconstruct the unmeasurable states at different delayed time-instants, which partition the maximal variation interval of the time-varying delays. These delayed time instant depend both on the magnitude of the delays and the growth rate of the nonlinearities. We also discuss how to implement continuous-time stabilizers as zero-order discretizations to obtain sampled-data stabilizers.**

Index Terms—**Delay systems, time-varying measurement and input delays, dynamic state predictors, continuous-time and sampled-data output feedback controllers.**

I. Introduction

THE problem of reconstructing the unmeasurable state variables for stabilization by using the delayed output measurements is long-standing. In this case it is important to implement some kind of prediction based on the delayed measurements. For stable linear systems the problem has been solved in [36]. Nonlinear observer has been proposed in [23] for linearizable by additive output injection systems. A predictor based on a cascade of observers has been introduced with LMI techniques in [7]. For globally Lipschitz continuous invertible observability maps ([13] and [21]) the proposed observer consists of a chain of dynamic predictors that reconstruct the unmeasurable state vector at different delayed time-instants within the time-delay window introduced by the output measurements. Hence, the proposed nonlinear observer exhibits a chained structure that explicitly takes into account the magnitude of the output delay. Also globally Lipschitz conditions on the system are required in [15]. In all these papers linear predictors are used. A survey on observers with measurement delay is found in [34] while a predictor–based approach is extensively surveyed in [22].

Predictor–based results have been more recently obtained in [19] where a known compact absorbing set (plus some technical additional assumptions) is assumed for all the system trajectories. On the other hand, these dynamic predictors follow the structure of the ones introduced in [13] and [21].

Predictors, which are not implemented as dynamical filters, are designed in [17] under the assumption that either a) the expression of the state trajectories is explicitly known or b) the system is globally Lipschitz. In [18] the existence of predictor–based observers is shown under the hypothesis that the so-called predictor map is known exactly. The implementation of predictors containing integral terms (distributed predictors) may be computationally prohibitive for real-time applications, and the open-loop structure of the integral predictor makes it sensitive to uncertainties and modeling errors. In addition, it is not trivial to estimate $\mathbf{x}_t$ when the delay is not constant.

Actually, all the above cited results can be implemented only if the predictor map is available (this happens for linear systems, bilinear systems, chains of linear systems with input nonlinearities), except for [19] where a modified version of the chained predictors, introduced in [13] and [21], are used. Numerical and approximate predictors have been proposed in [16]. State predictors for nonlinear stable systems are studied in [5], removing globally Lipschitz and compact absorbing set assumptions by introducing techniques based on incremental homogeneity properties ([4]).

All the above papers consider constant measurement delays. The case of time-varying measurement delays has been considered in [37] although restricted to linear systems with piecewise constant measurement delays. A Lyapunov-Krasovskii approach is used in [14] and [39] and exponential error convergence is proved in [1]. However, applications are limited to small enough delays. A Razumichin approach has been used in [8], following [13], and in the more general context of multi-output systems in [9]. In the last two papers a chain of observer is used to achieve error convergence when this is not possible with only one observer (i.e. the measurement delay is too large). On the other hand, applications are limited to globally Lipschitz systems (as in [9]) or differentiable delays with additional rate restrictions (as in [14]). In [5] these limitations are overcome and global predictors are designed for nonlinear systems with bounded trajectories.

The reconstructed unmeasurable state variables are used for stabilization in the presence of delayed controls. A solution is to set to zero the input delay and then searching for upper bounds on the input delays that the closed-loop system can tolerate while still realizing the desired goal. This often involves Lyapunov-Krasovskii functionals (as discussed in [12] and [27], which often lead to satisfactory results when

This work was supported by the MIUR.

S. Battilotti is with Department of Computer, Control, and Management Engineering "Antonio Ruberti", Sapienza University of Rome, Via Ariosto 25, Italy.

the delay is small; see [26]. However, many applications have long input delays. In general, stabilization under long input delays needs control designs that use the value of the input delay, and in many cases, distributed delays are used, meaning the control uses all values of the state or input along some interval of past times; see [**?**], [29].

In [27] and a prediction based approach is used to construct globally asymptotically stabilizing control laws for time-varying systems using state-feedback. This approach differs from the classical reduction model approach or the prediction based approaches introduced by Krstic (as in [6], [22] and [17]) which also involve distributed terms. Several dynamic extensions are used, making it possible to obtain a prediction of the state variable without using distributed terms. Many contributions, including [3] and [1], use several dynamic extensions to carry out state prediction, but to the best of our knowledge, they do not apply to the problem we consider here and they use distributed terms. Our prediction stabilization technique applies to nonlinear Lipschitz systems, which is also the case for many prediction ones, e.g., [17].

Finally, the works [1], [10], and [41] were limited to linear time-invariant systems under additional eigenvalue conditions and controllability conditions or bounds on the delays, without robustness to uncertainty; [8] was also confined to linear time-invariant systems; [13] covered nonlinear systems under a globally drift-observability condition and [40] cover time-varying linear systems and give sufficient conditions for stabilizability under pseudo-predictor feedback using an integral delay system.

Most of the above papers are focused on the state-feedback problem with globally Lipschitz or linear dynamics. Only [11], [1] and [3] cover the output-feedback case with large delays but restricting to globally Lipschitz dynamics and only [20], [20] cover feedback linearizable systems but restricting to small delays. In this paper, we remove globally Lipschitz assumptions or linearity assumptions on the system by introducing techniques based on incremental homogeneity properties ([4]) and propose output-feedback stabilizers for multi-inputs multi-outputs nonlinear systems with time-varying measurement and input delays, with no restriction on the bounds. Following the idea of chains of linear observers ([9], [21]), we generate a state prediction by chains of nonlinear (high-gain) observers that reconstruct the unmeasurable state at different delayed time-instants, which partition the maximal variation interval of the time-varying delays. The number of observers is in general larger as the maximum delay is larger. Our remarkable improvement of this idea relies in the fact that the number of observers, in the presence of strong nonlinearities, should depend also on the growth rate of the nonlinearities. Stronger nonlinearities require a larger number of observers. The state prediction is used by a nonlinear controller to stabilize the system through the delayed control input. The novelty of our stabilizer is the use of a nonlinear (saturated) control law affected by a chain of nonlinear observers with saturated estimates and switching error-correction terms. Saturations (or alternatively rate limiters) take care of the strong nonlinearities of the system and avoid the peaking phenomenon (well-known for systems with no delays). Switching error-correction terms take care of the time-varying delays. Also incremental homogeneity properties introduce a novel and generalized technique for rescaling the controller's and observers' gains which is even new with regard to the controllers adopted in [4] for systems with no delays and has a key role in the stability properties of the closed-loop system. The nonlinear nature of the closed-loop system requires a quite technical approach and specific nonlinear analysis tools, which is another contribution of this paper. In particular, for the closed-loop stability analysis we introduce new classes of (logarithmic) Lyapunov-Razumichin functions which are particularly useful in the presence of nonlinear dynamics. Robustness and few other extensions are discussed in section IV-E. We also propose a zero-order hold discretization of the proposed continuous-time stabilizer to obtain a sampled-data output-feedback stabilizer (section V), taking advantage in the stability analysis that they both generate closed-loop trajectories which have the same values at the sampling times. Sampled-data predictors and controllers were studied in [2] and [17] under the above mentioned restrictions of continuous-time controllers while various other contributions achieve only *practical* asymptotic stability (see [35] for instance). By considering general nonlinearities and achieving asymptotic stability, our result is another remarkable contribution in the literature of sampled-data output-feedback for systems with delays. The novelty of our sample-data stabilizer is the use of a nonlinear (saturated) sampled-data control law affected by a chain of nonlinear sampled-data observers with saturated estimates and switching error-correction terms.

II. Notation

(**N1**) $\mathbb{R}^n$ (resp. $\mathbb{R}^{n\times s}$) is the set of n-dimensional real column vectors (resp. $n \times s$ matrices). $\mathbb{R}_{\geqslant}$ (resp. $\mathbb{R}^n_{\geqslant}$, $\mathbb{R}^{n\times s}_{\geqslant}$) denotes the set of non-negative real numbers (resp. vectors in $\mathbb{R}^n$, matrices in $\mathbb{R}^{n\times s}$, with non-negative real elements). $\mathbb{R}_{>}$ (resp. $\mathbb{R}^n_{>}$) denotes the set of positive real numbers (resp. vectors in $\mathbb{R}^n$ with real positive entries). $(\mathbb{R}^n)^*$ is the dual space of $\mathbb{R}^n$ (space of row vectors).

(**N2**) For any matrix $A \in \mathbb{R}^{p\times n}$ we denote by $A_{i,j}$ the (i,j)-th element of A and for any vector $v \in \mathbb{R}^n$ (or $v \in (\mathbb{R}^n)^*$) we denote by v_i the i-th element of v. Also, we may write vectors $v \in \mathbb{R}^n$ as $(v_1,\dots,v_n)^T$, vectors $w \in (\mathbb{R}^n)^*$ as $(w_1,\dots,w_n)$ and matrices $A \in \mathbb{R}^{s\times n}$ either as $A = [v_1,\dots,v_n]$ (i.e. by columns) or $A = [w_1^T,\dots,w_s^T]^T$ (i.e. by rows). I_n is the $n \times n$ identity matrix. Moreover,

$$\mathrm{diag}\{A_1,\dots,A_m\} := \begin{pmatrix} A_1 & 0 & \cdots & 0 \\ 0 & A_2 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & A_m \end{pmatrix}$$

where A_i is any matrix and the 0 blocks have suitable dimensions. We retain a similar notation for functions. Also, $|a|$ denotes the absolute value of $a \in \mathbb{R}$, $\|a\|$ denotes the euclidean norm of $a \in \mathbb{R}^n$ with $\|a\|_M := \sqrt{a^T M a}$ where $M \in \mathbb{R}^{n\times n}$, $\|A\|$ denotes the norm of $A \in \mathbb{R}^{n\times n}$ induced from $\|\cdot\|$. For any matrix A (resp. vector v) $\langle A\rangle$ denotes the matrix (resp. vector) with $\langle A\rangle_{i,j} := |A_{i,j}|$ (resp. with $\langle v\rangle_i := |v_i|$).

(**N3**) We denote by $\mathbf{C}^0(\mathscr{X},\mathscr{Y})$, $\mathscr{X}\subset\mathbb{R}^n$ and $\mathscr{Y}\subset\mathbb{R}^s$, the set of continuous functions $\alpha:\mathscr{X}\to\mathscr{Y}$. Moreover, $\mathcal{K}_0$ denotes the set of strictly increasing functions $\alpha\in\mathbf{C}^0(\mathbb{R}_\geqslant,\mathbb{R}_\geqslant)$, $\mathcal{K}$ denotes the set of functions $\alpha\in\mathcal{K}_0$ such that $\alpha(0)=0$, $\mathcal{K}_\infty$ denotes the set of functions $\alpha\in\mathcal{K}$ such that $\alpha(s)\to+\infty$ as $s\to+\infty$. Also, $\mathcal{L}$ denotes the set of strictly decreasing functions $\alpha\in\mathbf{C}^0(\mathbb{R}_\geqslant,\mathbb{R}_>)$ such that $\alpha(s)\to 0$ as $s\to+\infty$ and by $\mathcal{KL}$ denotes the set of functions $\alpha\in\mathbf{C}^0(\mathbb{R}_\geqslant\times\mathbb{R}_\geqslant,\mathbb{R}_\geqslant)$ such that $\alpha(s,\cdot)\in\mathcal{L}$ and $\alpha(\cdot,s)\in\mathcal{K}$ for each $s\in\mathbb{R}_\geqslant$. By $\alpha(s)=o(\beta(s))$ for $s\to p_0$ (where $p_0\in\mathbb{R}\cup\{\pm\infty\}$) we mean that $\lim_{s\to p_0}\frac{\alpha(s)}{\beta(s)}=0$.
(**N4**) For $\epsilon\in\mathbb{R}_>$, the group of dilations $\mathcal{G}=(\epsilon^{\mathfrak{r}},\diamond)$ is the set of elements $\epsilon^{\mathfrak{r}}:=(\epsilon^{\mathfrak{r}_1},\ldots,\epsilon^{\mathfrak{r}_n})^T\in\mathbb{R}^n$, $\mathfrak{r}\in\mathbb{R}^n$, with group operation $\epsilon^{\mathfrak{r}'}\diamond\epsilon^{\mathfrak{r}''}=\epsilon^{\mathfrak{r}'+\mathfrak{r}''}$ and identity element $\epsilon^{\mathbf{0}_n}:=\mathbf{1}_n:=(1,\ldots,1)^T$ where $\mathbf{0}_n:=(0,\ldots,0)^T$.

Also, we define the $\epsilon^{\mathfrak{r}}-$*dilation of* $v\in\mathbb{R}^n$ as the left group action $\diamond$ on $\mathbb{R}^n$ defined as $\epsilon^{\mathfrak{r}}\diamond v:=(\epsilon^{\mathfrak{r}_1}v_1,\cdots,\epsilon^{\mathfrak{r}_n}v_n)^T$. Similarly, we define the $\epsilon^{\mathfrak{r}}-$*dilation of* $w\in(\mathbb{R}^n)^*$ as the right group action $\diamond$ on $(\mathbb{R}^n)^*$ defined as $w\diamond\epsilon^{\mathfrak{r}}:=(\epsilon^{\mathfrak{r}_1}w_1,\cdots,\epsilon^{\mathfrak{r}_n}w_n)$.

By extension, we can define the *left* $\epsilon^{\mathfrak{r}}-$*dilation of* $A:=[w_1^T,\ldots,w_s^T]^T\in\mathbb{R}^{n\times s}$ as the left group action $\diamond$ on $\mathbb{R}^{n\times s}$ defined as $\epsilon^{\mathfrak{r}}\diamond A:=[\epsilon^{\mathfrak{r}_1}w_1^T,\ldots,\epsilon^{\mathfrak{r}_n}w_n^T]^T$ and the *right* $\epsilon^{\mathfrak{r}}-$*dilation of* $A:=[v_1,\ldots,v_n]\in\mathbb{R}^{s\times n}$ as the right group action $\diamond$ on $\mathbb{R}^{s\times n}$ defined as $A\diamond\epsilon^{\mathfrak{r}}:=[\epsilon^{\mathfrak{r}_1}v_1,\cdots,\epsilon^{\mathfrak{r}_n}v_1]$. The dilation's properties used in this paper are given in the appendix.
(**N5**) on $\mathbb{R}^n$ we introduce a partial ordering $\preceq$ as follows: for any pair of vectors $x,y\in\mathbb{R}^n$ we write $x\preceq y$ if and only if $x_i\leqslant y_i$ for all $i=1,\ldots,n$. We naturally extend this partial ordering on $\mathbb{R}^{n\times s}$: for any pair of matrices $A,B\in\mathbb{R}^{n\times s}$ $A\preceq B$ if and only if $A_{i,j}\leqslant B_{i,j}$ for all $i=1,\ldots,n,j=1,\ldots,s$. Also for any $A(x)\in\mathbb{R}^{n\times s}$, $x\in\mathbb{R}^m$, and compact $\mathcal{C}\subset\mathbb{R}^m$ we denote by $\sup_{x\in\mathcal{C}}A(x)$ any matrix A_M such that $A(x)\preceq A_M$ for all $x\in\mathbb{R}^m$.

On the other hand, for any pair of square matrices A,B we will write $A\leqslant B$ (resp. $A<B$) if and only if $A-B$ is negative semidefinite (resp. $A-B$ is negative definite).
(**N6**) A *saturation function* σ_l *with saturation levels* $l\in\mathbb{R}^n_>$ is a function $\sigma_l(x):=(\sigma_{l_1}(x_1),\ldots,\sigma_{l_n}(x_n))^T$, $x\in\mathbb{R}^n$, such that for each $i=1,\ldots,n$ and $x_i\in\mathbb{R}$:

$$\sigma_{l_i}(x_i)=\begin{cases}x_i & |x_i|\leqslant l_i\\ \mathrm{sign}(x_i)l_i & \text{otherwise.}\end{cases}\tag{1}$$

It is easy to prove the following inequalities:

$$\langle\sigma_l(x)-\sigma_l(y)\rangle\preceq 2\langle\sigma_l(x-y)\rangle\preceq 2l,\tag{2}$$
$$\langle\sigma_l(x)\rangle\preceq\langle x\rangle\tag{3}$$

for all $x,y\in\mathbb{R}^n$ and $l\in\mathbb{R}^n_>$.

III. The class of systems and problem statement

We consider continuous-time nonlinear systems with delayed measurements $\mathbf{y}$ and inputs $\mathbf{u}$:

$$\dot{\mathbf{x}}_t = A\mathbf{x}_t+B\mathbf{u}_{t-c}+\phi(\mathbf{x}_t),\ t\geqslant -c-2d_\infty,\tag{4}$$
$$\mathbf{y}_t = C\mathbf{x}_{t-\mathbf{d}_t},\ t\geqslant 0\tag{5}$$

with state $\mathbf{x}_t\in\mathbb{R}^n$, measurements $\mathbf{y}_t\in\mathbb{R}^p$, continuous measurement delay $\mathbf{d}_t\in\mathbb{R}_\geqslant$, known up to time t and bounded by a known constant d_∞ and known constant input delay $c\in\mathbb{R}_\geqslant$ (see section IV-E2 for time-varying input delay $\mathbf{c}_t$). The assumption that the delays are known is realistic in many applications. The input $\mathbf{u}_t$ is set to zero for $t\leqslant c$. We assume that ϕ is locally Lipschitz (see section IV-E1 for locally Lipschitz output nonlinearities ψ: $\mathbf{y}_t=C\mathbf{x}_{t-\mathbf{d}_t}+\psi(\mathbf{x}_{t-\mathbf{d}_t})$). The matrices A,B,C have the form

$$A=\mathrm{diag}\{A_1,\ldots,A_m\},\ B=\mathrm{diag}\{B_1,\ldots,B_m\},\\ C=\mathrm{diag}\{C_1,\ldots,C_p\},\tag{6}$$

where (A_i,B_i) are in Brunowski form and $C_i=(1,0,\ldots,0)$. The problem we want to solve in this paper is to design continuous-time stabilizers of (4) using the output information $\mathbf{y}_t$ and sampled-data stabilizers using the sampled output information $\mathbf{y}_{t_h}$, $t_h:=hT$ ($h\in\mathbb{N}$ and $T\in\mathbb{R}_>$ the sampling period).

IV. Continuous-time stabilizers

The continuous-time stabilizer we propose consists of a controller together with a certain number of chained observers. These observers are chained in the sense that each observer in the chain computes the estimate of the state of the controlled process, delayed by a sufficiently small relative amount, and hands over a certain amount of information (like its own estimate) to the next one in the chain. The approach of using chained sub-predictors for coping with large delays is not new ([13], [9]). The novelty here is to consider the measurement and control delays $\mathbf{d}_t$ and c forming together a large delay $\mathbf{d}_t+c$ (from the last received measurement to the first applied control action) and the partition of the delay interval $[-c,d_\infty]$ into an increasing sequence of points $\{p^{(j)}\}_{j=1,\ldots,\nu}$, which determines the number ν of sub-predictors. Another important novelty is that ν depends not only on how large is the delay but also on the growth rate of the nonlinearities of the controlled process (tunable chain length ν). According to this partition, each observer of the chain computes an estimate of the delayed state $\mathbf{x}_t^{(j)}:=\mathbf{x}_{t-p^{(j-1)}}$, $j=2,\ldots,\nu+1$, denoted by $\hat{\mathbf{x}}_t^{(j)}$. The first element of the observer chain is an observer which computes the estimate $\hat{\mathbf{x}}_t^{(\nu+1)}$ of the (maximally) delayed state $\mathbf{x}_t^{(\nu+1)}:=\mathbf{x}_{t-d_\infty}$ and the last element of the observer chain is an observer which computes the estimate $\hat{\mathbf{x}}_t^{(2)}$ of the state $\mathbf{x}_t^{(2)}:=\mathbf{x}_{t+c}$ (i.e. a c-step prediction). The control action is defined by processing this last estimate so that, when delayed by c at the input $\mathbf{u}_{t-c}$ of the system, it corresponds to the estimate of $\mathbf{x}_t$. The partition of the interval $[-c,d_\infty]$ into a sequence of points $\{p^{(j)}\}_{j=1,\ldots,\nu}$ is made precise as follows.

Definition *4.1: A real sequence* $\{p^{(j)}\}_{j=1,\ldots,\nu}$ *is a* δ*-fine partition of an interval* $[a,b]\subset\mathbb{R}$, $\delta\in\mathbb{R}_>$, *if* $\nu=\lceil\frac{b-a}{\delta}\rceil+1$, $p^{(j)}:=a+(j-1)\delta$ *for* $j=1,\ldots,\nu-1$ *and* $p^{(\nu)}:=b$.
Notice that the number N depends on the refinement δ of the partition and $p^{(\nu)}-p^{(\nu-1)}\leqslant\delta$ with $p^{(\nu)}-p^{(\nu-1)}=\delta$ if and only if $\frac{b-a}{\delta}$ is integer. In what follows, we consider δ-fine partitions $\{p^{(j)}\}_{j=1,\ldots,\nu}$ of the interval $[-c,d_\infty]$ including the point 0 and an auxiliary extra point $p^{(\nu+1)}>d_\infty$ such that $p^{(\nu+1)}-p^{(\nu)}\leqslant\delta$ and we assume that $p^{(\nu_0)}=0$ for some $\nu_0\in\{1,\ldots,\nu\}$ (we will say that *the partition is extended and centered at* 0).

A. The observer chain with tunable length

Each observer of the chain, say the j-th observer of the chain, manipulates a certain amount of information, according to the relative values of the delay $\mathbf{d}_t$ with respect to the partition of $[-c, d_\infty]$: typically, when $\mathbf{d}_t$ is large the observer will process the estimate $\hat{\mathbf{x}}_t^{(j+1)}$ handed over by the preceding observer in the chain, while for small values of $\mathbf{d}_t$ the observer will use the available outputs $\mathbf{y}_t$ and, if necessary, past outputs $\mathbf{y}_s,\ s \leqslant t$. Different data processing of the above type determine different innovations for each observer to guarantee convergence of the estimate to the delayed state. As already stated, we assume that $\mathbf{d}_t$ is bounded by d_∞ and continuous. A useful property of continuous delays is the following. Let $\mathbf{y}_t^*$ be the undelayed outputs, i.e. $\mathbf{y}^*_{t-\mathbf{d}_t} = \mathbf{y}_t$.

Lemma *4.1: If $\mathbf{d}_t$ is continuous, when for each $t > d_\infty$ and $\Delta : \mathbf{d}_t \leqslant \Delta \leqslant t\ \exists \bar{t} \leqslant t : \mathbf{y}_{\bar{t}} = \mathbf{y}^*_{t-\Delta}$.*
In other words, when the delay $\mathbf{d}_t$ is continuous, past measurements are available for processing continuously in time up to t.

Let's get into the technical structure of each observer in the chain. Let $\mathfrak{f}^{(o)} \in \mathbb{R}^n$, $\mathfrak{r} \in \mathbb{R}^n_>, \epsilon, l^{(o)} \in \mathbb{R}_>$ and diagonal positive definite $\Gamma^{(o)} \in \mathbb{R}^{n\times n}$ be design parameters. Moreover, in accordance with the notation $\mathbf{x}_t^{(j+1)} := \mathbf{x}_{t-p^{(j)}}$ set $\mathbf{u}_t^{(j+1)} := \mathbf{u}_{t-p^{(j)}}$. The observer chain is described by

$$\begin{aligned}\dot{\hat{\mathbf{x}}}_t^{(j)} &= A\hat{\mathbf{x}}_t^{(j)} + B\mathbf{u}_{t-c}^{(j)} \\ &+ \phi\left(\sigma_{\lambda^{(o)}(\epsilon)}\left(\hat{\mathbf{x}}_t^{(j)}\right)\right) + P^{(o)^{-1}}C^T R^{(o)}\mathbf{z}_t^{(j)}, \\ &\quad j = 2,\ldots,\nu+1,\ t \geqslant 0,\end{aligned} \tag{7}$$

with saturation function $\sigma_{\lambda^{(o)}(\epsilon)}$ and saturation levels $\lambda^{(o)}(\epsilon) := l^{(o)}\epsilon^{\mathfrak{r}}$, matrices

$$\begin{aligned}&P^{(o)} = (I_n - G^{(o)}A^T)^T \diamond \epsilon^{-2\mathfrak{r}} \diamond (I_n - G^{(o)}A^T), \\ &R^{(o)} = C(\epsilon^{-\mathfrak{r}} \diamond G^{(o)} \diamond \epsilon^{-\mathfrak{r}})C^T,\ G^{(o)} = \epsilon^{\mathfrak{f}^{(o)}} \diamond \Gamma^{(o)} \diamond \epsilon^{\mathfrak{f}^{(o)}}, \end{aligned}\tag{8}$$

and innovations $\mathbf{z}_t^{(j)}$ defined as follows:

▫ for $j = \nu_0 + 1, \ldots, \nu + 1$

$$\mathbf{z}_t^{(j)} := \begin{cases} \mathbf{y}_{t^{(j)}} - C\hat{\mathbf{x}}^{(j)}_{t-\mathbf{s}_t^{(j)}} & \text{if } \mathbf{d}_t \in [0, p^{(j-1)}), \\ \mathbf{y}_t - C\hat{\mathbf{x}}^{(j)}_{t-\mathbf{s}_t^{(j)}} & \text{if } \mathbf{d}_t \in [p^{(j-1)}, p^{(j)}], \\ C(\hat{\mathbf{x}}_t^{(j+1)} - \hat{\mathbf{x}}^{(j)}_{t-\mathbf{s}_t^{(j)}}) & \text{if } \mathbf{d}_t \in (p^{(j)}, p^{(\nu+1)}] \end{cases} \tag{9}$$

(where $\mathbf{y}_{t^{(j)}}$ is the past output at $t^{(j)} \in [0, t]$ such that $t^{(j)} - \mathbf{d}_{t^{(j)}} = t - p^{(j-1)}$: $t^{(j)}$ does exist by lemma 4.1) with delay

$$\mathbf{s}_t^{(j)} := \begin{cases} 0 & \text{if } \mathbf{d}_t \in [0, p^{(j-1)}), \\ \mathbf{d}_t - p^{(j-1)} & \text{if } \mathbf{d}_t \in [p^{(j-1)}, p^{(j)}], \\ p^{(j)} - p^{(j-1)} & \text{if } \mathbf{d}_t \in (p^{(j)}, p^{(\nu+1)}], \end{cases} \tag{10}$$

▫ for $j = 2, \ldots, \nu_0$

$$\mathbf{z}_t^{(j)} := C(\hat{\mathbf{x}}_t^{(j+1)} - \hat{\mathbf{x}}^{(j)}_{t-\mathbf{s}_t^{(j)}}) \tag{11}$$

$$\mathbf{s}_t^{(j)} := p^{(j)} - p^{(j-1)}. \tag{12}$$

Each observer is initialized as follows:

$$\hat{\mathbf{x}}_\theta^{(j)} := 0, \forall \theta \in [-c - 2d_\infty, 0] \tag{13}$$

(this particular initialization is motivated by sake of simplicity, otherwise we would have a slightly more involved design in the proof of the main theorem: see step (III) after (94)).The length ν of the chain depends not only on the magnitude of the delays but also on the nonlinearities of the system and it is a critical parameter in our design.

Remark *4.1:* Notice that when $\mathbf{d}_t \in [0, p^{(j-1)})$, $j = \nu_0 + 1, \ldots, \nu + 1$, the past outputs $\mathbf{y}_{t^{(j)}}$ ($t^{(j)} < t$, where $t^{(j)} = t + \mathbf{d}_{t^{(j)}} - p^{(j-1)}$) is processed for the innovation $\mathbf{z}_t^{(j)}$. The estimate $\mathbf{x}_t^{(j)}$ is not delayed ($\mathbf{s}_t^{(j)} = 0$). Notice that for the implementation of this step we need the past outputs $\mathbf{y}_{t^{(j)}}$ ($t^{(j)} < t$) and this requires the continuity of $\mathbf{d}_t$. This is the only point for which the continuity of $\mathbf{d}_t$ is needed. If $\mathbf{d}_t$ is not continuous the output $\mathbf{y}_{t^{(j)}}$ may be not available for processing. In this case we may think to reconstruct the value $\mathbf{y}_{t^{(j)}}$ from the past outputs (exactly or approximately using for instance sinc-functions). As it appears from (9), (11), the chained structure is given by the estimate $\hat{\mathbf{x}}_t^{(j+1)}$ of $\mathbf{x}_t^{(j+1)}$ computed by the $(j+1)$-th observer in the chain and handed over to the j-th observer only either when $\mathbf{d}_t \in (p^{(j)}, p^{(\nu+1)}]$ ((9)) and for the observers which compute state interpolations (i.e. past values of the state: (11)) or for the observers which compute state predictions (i.e. future values of the state: (11)).

Remark *4.2:* Notice that each observer (7) is a copy of the system (4), delayed by the amount $p^{(j-1)}$, with saturated estimates $\sigma_{\lambda^{(o)}(\epsilon)}(\hat{\mathbf{x}}_t^{(j)})$ and updated by the innovation process $\mathbf{z}_t^{(j)}$, weighted by the gain matrix $P^{(o)^{-1}}C^T R^{(o)}$. The gain matrix is defined as a suitable dilated transformation with parameter ϵ, which follows very naturally from the incremental homogeneity assumptions on the process nonlinearities f which will be introduced in the section IV-C. The importance of saturating the estimates when trying to reconstruct the state of a nonlinear system with delay-free measurements has been pointed out in various works since the late 90's. Here, we prove the important fact that also in the presence of measurement delays we need to process saturated estimates.

B. The controller

Let $\mathfrak{f}^{(s)} \in \mathbb{R}^n$, $l^{(s)} \in \mathbb{R}_>$ and diagonal positive definite $\Gamma^{(s)} \in \mathbb{R}^{n\times n}$ be design parameters. The controller is defined as

$$\begin{aligned}\mathbf{u}_t &:= -R^{(s)}B^T P^{(s)}(I_n - A^T G^{(s)}) \times \\ &\quad \times \sigma_{\lambda^{(s)}(\epsilon)}\Big((I_n - A^T G^{(s)})^{-1}\hat{\mathbf{x}}_t^{(2)}\Big)\end{aligned} \tag{14}$$

with saturation function $\sigma_{\lambda^{(s)}(\epsilon)}$ and saturation levels $\lambda^{(s)}(\epsilon) := l^{(s)}\epsilon^{\mathfrak{r}}$ (in general $\neq \lambda^{(o)}(\epsilon)$) and

$$\begin{aligned}&P^{(s)} = (I_n - A^T G^{(s)})^{-T} \diamond \epsilon^{-2\mathfrak{r}} \diamond (I_n - A^T G^{(s)})^{-1}, \\ &R^{(s)} = B^T(\epsilon^{\mathfrak{r}} \diamond G^{(s)} \diamond \epsilon^{\mathfrak{r}})B,\ G^{(s)} = \epsilon^{\mathfrak{f}^{(s)}} \diamond \Gamma^{(s)} \diamond \epsilon^{\mathfrak{f}^{(s)}}. \end{aligned}\tag{15}$$

Notice how in (14) $\hat{\mathbf{x}}_t^{(2)}$ provides an estimate of $\mathbf{x}_t^{(2)} = \mathbf{x}_{t+c}$ and the control, as well as the estimates, are saturated with different levels $\lambda^{(s)}(\epsilon) \neq \lambda^{(o)}(\epsilon)$.

Remark *4.3:* Notice that the controller (14) comes out from the composition of a linear controller with the saturation $\sigma_{\lambda^{(s)}(\epsilon)}(\cdot)$. The linear controller is characterized by a gain matrix $R^{(s)}B^T P^{(s)}$ defined as a suitable dilated transformation with parameter ϵ, which follows very naturally from the incremental homogeneity assumptions on the process nonlinearities f which will be introduced in the section IV-C. The importance of saturating the control when trying to asymptotically stabilize a delay-free nonlinear system by output feedback has been pointed out since the late 90's. Here, we prove the important fact that also in the presence of delays it is important to saturate the (delayed) control action.

C. Main assumptions and results

Our assumptions on the system (4) are the following (see a review of incremental homogeneity in appendix A).

(H0) (*forward completeness*): *the trajectories* $\mathbf{x}_t$ *of (4) satisfy the following inequality: there exist* $\mu \in \mathbb{R}_>$ *and continuously differentiable and proper* $U : \mathbb{R}^n \to \mathbb{R}_\geqslant$ *and* $\kappa \in \mathcal{K}_\infty$ *such that for all* $t \geqslant -c - 2d_\infty$

$$\dot{U}(\mathbf{x}_t)|_{(4)} \leqslant \mu U(\mathbf{x}_t) + \kappa(\|\mathbf{u}_{t-c}\|), \tag{16}$$

(H1) (*state feedback design*): *for some degrees* $\mathfrak{f}^{(s)} \in \mathbb{R}^n$ *and weights* $\mathfrak{r} \in \mathbb{R}^n_>$ *such that*

$$\begin{aligned}&\mathfrak{f}^{(s)}_{j-1} \leqslant \hat{\mathfrak{f}}^{(s)}_j \leqslant \mathfrak{f}^{(s)}_j,\ j = 2, \ldots, n,\\ &\hat{\mathfrak{f}}^{(s)}_1 := \mathfrak{f}^{(s)}_1,\ \hat{\mathfrak{f}}^{(s)}_j := \mathfrak{r}_j - \mathfrak{r}_{j-1} - \mathfrak{f}^{(s)}_{j-1},\ j = 2, \ldots, n,\end{aligned} \tag{17}$$

f is homogeneous in the upper bound with quadruples $(\mathfrak{r}, \mathfrak{r} + \mathfrak{f}^{(s)}, \hat{\mathfrak{f}}^{(s)}, \Phi^{(s)}(x))$ *and lower triangular* $\Phi^{(s)}(0)$,

(H2) (*observer design*): *for some degrees* $\mathfrak{f}^{(o)} \in \mathbb{R}^n$ *and weights* $\mathfrak{r} \in \mathbb{R}^n_>$ *such that*

$$\begin{aligned}&2\mathfrak{f}^{(o)}_j - \mathfrak{f}^{(o)}_{j-1} \leqslant \hat{\mathfrak{f}}^{(o)}_{j-1} \leqslant \mathfrak{f}^{(o)}_{j-1},\ j = 2, \ldots, n,\\ &\hat{\mathfrak{f}}^{(o)}_j := \mathfrak{r}_{j+1} - \mathfrak{r}_j - \mathfrak{f}^{(o)}_{j+1},\ j = 1, \ldots, n-1,\ \hat{\mathfrak{f}}^{(o)}_n := \mathfrak{f}^{(o)}_n,\end{aligned} \tag{18}$$

ϕ is incrementally homogeneous in the upper bound with quadruples $(\mathfrak{r}, \mathfrak{r} + \hat{\mathfrak{f}}^{(o)}, \mathfrak{f}^{(o)}, \Phi^{(o)}(x', x''))$ *and lower triangular* $\Phi^{(o)}(0,0)$.

(H3) *(state feedback performances recovery):* $\mathfrak{f}^{(o)}_n > \mathfrak{f}^{(s)}_n$.

Remark *4.4:* Assumptions (H1) and (H2) are enough general for coping with large classes of nonlinear systems: the nonlinearities must satisfy some incremental homogeneity conditions, one for state-feedback design (H1) and one for observer design (H2). The additional condition (H3) is a fast recovery condition (through state reconstruction) of the closed-loop performances achieved by state-feedback and couples the state-feedback design with the observer design. Output feedback controllers are obtained from the state-feedback controllers by processing the state estimates instead of the true (unknown) values of the state. Notice that $\Phi^{(s)}(0)$ (resp. $\Phi^{(o)}(0,0)$) is required to be lower triangular, which implies that f, when at least once differentiable, has a lower triangular linearization at 0. This implies that the linearization of (4) at 0 is controllable. Assumptions based on incremental homogeneity similar to (H1)-(H3) have been considered in [4] for designing controllers for systems with no delays. In this paper, we consider more general control and observer structures than the ones introduced in [4] with *ad hoc* techniques for the choice of the gain matrices and saturation levels as well as for the closed-loop stability analysis. It is not difficult to check for assumptions (H1) and (H2). In general, this kind of assumptions amount to solve a set of algebraic inequalities in the unknowns $\mathfrak{r} \in \mathbb{R}^n_>$ and $\mathfrak{f}^{(\cdot)} \in \mathbb{R}^n$. For example the system

$$\begin{aligned}\dot{\mathbf{x}}_{1,t} &= \mathbf{x}_{2,t} + \mathbf{x}_{1,t}\\ \dot{\mathbf{x}}_{2,t} &= -\mathbf{x}_{1,t} + (1 - \mathbf{x}^2_{1,t})\mathbf{x}_{2,t} + \mathbf{u}_{t-c}\end{aligned} \tag{19}$$

satisfies all the assumptions (H1)-(H3) with $\phi(\mathbf{x}) = (\mathbf{x}_1, -\mathbf{x}_1 + (1 - \mathbf{x}^2_1)\mathbf{x}_2)^T$, $\mathfrak{r} = (1,3)^T$, $\mathfrak{f}^{(s)} = (1,1)^T$, $\mathfrak{f}^{(o)} = (4,2)^T$ and suitable $\Phi^{(s)}(x)$ and $\Phi^{(o)}(x', x'')$ (which we leave to the reader) with lower triangular $\Phi^{(s)}(0)$ and $\Phi^{(o)}(0,0)$.

Assumption (H0) is a standard assumption for forward completeness and it can be relaxed by requiring that the trajectories of (4) satisfy (16) only up to time $t = c$ (i.e. forward completeness for the open-loop system). This kind of assumption is needed to ensure that in the absence of control input (up to time $t = \gamma$) the state trajectories do not explode to infinity. For instance, assumption (H0) holds for (19) with $U(x) = \|x\|^2$, $\mu = 3$ and $\kappa = s^2$.

Remark *4.5:* A consequence of (17) and (18) is that the numbers $\mathfrak{f}^{(s)}_j$, $j = 1, \ldots, n$, form a non-decreasing sequence while the numbers $\mathfrak{f}^{(o)}_j$, $j = 1, \ldots, n$, form a non-increasing sequence and in the overall by (H3) we have

$$\mathfrak{f}^{(s)}_1 \leqslant \ldots \leqslant \mathfrak{f}^{(s)}_n < \mathfrak{f}^{(o)}_n \leqslant \ldots \leqslant \mathfrak{f}^{(o)}_1. \tag{20}$$

When $m > 1$ in (6) the monotone condition (20) may be somehow restrictive. Under this regard (H1) and (H2) can be weakened in such a way that, with m_i being the dimension of the matrix A_i in (6),

$$\mathfrak{f}^{(s)}_{\sum_{i=1}^{r} m_i} \leqslant \ldots \leqslant \mathfrak{f}^{(s)}_{\sum_{i=1}^{r+1} m_i} < \mathfrak{f}^{(o)}_{\sum_{i=1}^{s+1} m_i} \leqslant \ldots \leqslant \mathfrak{f}^{(o)}_{\sum_{i=1}^{s} m_i}. \tag{21}$$

for all $r, s = 0, \ldots, m-1$ (with $\sum_{j=1}^{0} = 1$), i.e the monotone condition holds only for chains with length m_i.

Since $\preceq$ is a partial ordering on $\mathbb{R}^n$, the monotone condition (20) induces also a partial ordering on the group of dilations. As a matter of fact,

$$\mathfrak{v} \preceq \mathfrak{w} \Rightarrow \epsilon^{\mathfrak{v}} \preceq \epsilon^{\mathfrak{w}} \tag{22}$$

if $\epsilon > 1$ (i.e. expanding dilations) and $\mathfrak{v} \preceq \mathfrak{w} \Rightarrow \epsilon^{\mathfrak{w}} \preceq \epsilon^{\mathfrak{v}}$ if $\epsilon < 1$ (i.e. contracting dilations). In this way, for expanding dilations we have on account of (20)

$$\epsilon^{\mathfrak{f}^{(s)}_1}\mathbf{1}_n = \epsilon^{\mathfrak{f}^{(s)}_1\mathbf{1}_n} \preceq \epsilon^{\mathfrak{f}^{(s)}} \preceq \epsilon^{\mathfrak{f}^{(s)}_n\mathbf{1}_n} = \epsilon^{\mathfrak{f}^{(s)}_n}\mathbf{1}_n, \tag{23}$$

$$\epsilon^{\mathfrak{f}^{(o)}_n}\mathbf{1}_n = \epsilon^{\mathfrak{f}^{(o)}_n\mathbf{1}_n} \preceq \epsilon^{\mathfrak{f}^{(o)}} \preceq \epsilon^{\mathfrak{f}^{(o)}_1\mathbf{1}_n} = \epsilon^{\mathfrak{f}^{(o)}_1}\mathbf{1}_n. \tag{24}$$

The first important stabilization result of this paper is the following.

Theorem *4.1: Let* $\mathcal{C} \subset \mathbb{R}^n$ *be a given compact set. Under assumptions* (H0)-(H3) *there exist diagonal positive definite* $\Gamma^{(l)} \in \mathbb{R}^{n\times n}$, $l^{(l)} \in \mathbb{R}_>$, $l \in \{s, o\}$, $\epsilon \in \mathbb{R}_>$, $\delta \in \mathbb{R}_>$ *and a* δ*-fine partition* $\{p^{(j)}\}_{j=1,\ldots,\nu}$ *of* $[-c, d_\infty]$, *extended and centered at* 0, *such that the solutions* $(\mathbf{x}_t, \hat{\mathbf{x}}^{(j)}_t)$, $j = 2, \ldots, \nu+1$, *of (4), (5), (7), (14), with* $\mathbf{x}_{-c-2d_\infty} \in \mathcal{C}$, *are bounded for all* $t \geqslant -c - 2d_\infty$ *and* $\lim_{t\to+\infty} \mathbf{x}_t = 0$.

The continuous-time controller (7), (14) guarantees asymptotic stability of (4) for all initial conditions $\mathbf{x}_{-c-2d_\infty} \in \mathcal{C}$, where $\mathcal{C}$ is an *a priori* given compact set. In this sense our controller (7), (14) *semi-globally* asymptotically stabilizes (4). Boundedness and convergence are uniform (in the sense of $\mathcal{KL}$ functions) as pointed out at the end of the proof of the theorem.

D. Proof of theorem 4.1

Boundedness analysis. To prove boundedness of trajectories, we will construct a Lyapunov-Razumichin function for the closed-loop system. First, some preliminaries. For any $\nu \in \mathbb{N}$ let $b^{(j)}$, $j = 2, \ldots, \nu+1$, be real numbers such that

$$
\begin{aligned}
&2 > b^{(\nu+1)} > 1, \\
&2 - \frac{1}{b^{(j+1)}} > b^{(j)} > 1, \; j = 2, \ldots, \nu,
\end{aligned} \tag{25}
$$

and $a \in (0,1)$ and $k \in (1, +\infty)$ be such that

$$
2 - 3a - \frac{1}{b^{(j+1)}} > k b^{(j)}(1+a), \; j = 2, \ldots, \nu \tag{26}
$$

(such a, k exist by continuity and (25)). Let $l_\infty^{(s)}, \Gamma^{(s)}$ and $l_\infty^{(o)}, \Gamma^{(o)}$ be as in lemmas A.4 and A.3 (with a as in (26)) and define

$$
l^{(s)} := \frac{l^{(o)}}{\|I_n + A^T \Gamma^{(s)}\|} \tag{27}
$$

with any $l^{(o)} \in (0, l_\infty^{(o)}]$ such that $l^{(s)} \leqslant l_\infty^{(s)}$. Let $\kappa \in \mathcal{K}_\infty$ be as in (H0) and $\omega \in \mathcal{K}_\infty$ be such that (recall that $\mathfrak{f}_n^{(o)} > \mathfrak{f}_n^{(s)}$ by (H3))

$$
\begin{aligned}
\omega(\epsilon) &= o(\epsilon^{\mathfrak{f}_n^{(o)} - \mathfrak{f}_n^{(s)}}) \text{ as } \epsilon \to +\infty \qquad (28)\\
\ln \epsilon &= o(\omega(\epsilon)) \text{ as } \epsilon \to +\infty. \qquad (29)
\end{aligned}
$$

(for instance, $\omega(\epsilon) = \epsilon^{(\mathfrak{f}_n^{(o)} - \mathfrak{f}_n^{(s)})/2}$). Introduce $\delta \in \mathcal{L}$ such that

$$
\begin{aligned}
\delta(\epsilon) &= o(e^{-\epsilon - l^{(s)2}\omega(\epsilon)}) \text{ as } \epsilon \to +\infty, \qquad (30)\\
\delta(\epsilon) &= o\left(\frac{1}{\kappa(e^\epsilon)}\right) \text{ as } \epsilon \to +\infty, \qquad (31)
\end{aligned}
$$

(for instance, $\delta(\epsilon) = \frac{e^{-l^{(s)2}\omega(\epsilon)}}{\kappa(e^\epsilon)}$ if κ is superlinear). Choose a $\delta(\epsilon)$-fine partition $\{p^{(j)}(\epsilon)\}_{j=1,\ldots,\nu(\epsilon)}$ of the interval $[-c, d_\infty]$, extended and centered at 0. The partition $\{p^{(j)}(\epsilon)\}_{j=1,\ldots,\nu(\epsilon)}$ depends on ϵ in that the number of points $\nu(\epsilon)$ depends on ϵ and it is finer and finer as larger ϵ is with $\lim_{\epsilon\to+\infty} \delta(\epsilon) = 0$ and $\lim_{\epsilon\to+\infty} \nu(\epsilon) = +\infty$. By the definition of the delay $\mathbf{s}_t^{(j)}$ in (10), (12) and since $\{p^{(j)}\}_{j=1,\ldots,\nu(\epsilon)}$ is $\delta(\epsilon)$-fine,

$$
\mathbf{s}_t^{(j)} \leqslant \delta(\epsilon), \; \forall t \geqslant 0. \tag{32}
$$

Throughout the proof, for simplicity and if not explicitly needed, we will omit the argument (ϵ) of the functions $p^{(j)}$, ν, δ, ω and $\lambda^{(l)}$, $l \in \{s, o\}$. In general, time functions will be denoted in boldface, any other function with greek letters, matrices with capital letters and numbers with small letters. The parameter ϵ, which determines the gain matrices (15) and (8), varies in between $(1, +\infty)$ and will be left free until properly chosen at the end of the proof to satisfy all the intermediate conditions. In particular, its minimum guaranteed value (denoted ϵ_∞ in the proof) will be increased at certain key points of the proof to satisfy all the required conditions. Consider the delayed state equation

$$
\dot{\mathbf{x}}_t^{(2)} \;=\; A\mathbf{x}_t^{(2)} + B\mathbf{u}_{t-c}^{(2)} + \phi(\mathbf{x}_t^{(2)}) \tag{33}
$$

for $t \geqslant 0$ and notice that, from the definition (14) of the control $\mathbf{u}_t$,

$$
\mathbf{u}_t^{(2)} = -R^{(s)} B^T P^{(s)} \left(\mathbf{x}_{t+c}^{(2)} + \mathbf{w}_{t+c}^{(2)}\right) \tag{34}
$$

with exogenous input

$$
\begin{aligned}
\mathbf{w}_t^{(2)} = &(I_n - A^T G^{(s)}) \Big\{ \sigma_{\lambda^{(s)}(\epsilon)} \left((I_n - A^T G^{(s)})^{-1} \widehat{\mathbf{x}}_t^{(2)} \right) \\
&- (I_n - A^T G^{(s)})^{-1} \mathbf{x}_t^{(2)} \Big\}.
\end{aligned} \tag{35}
$$

The following lemma establishes a bound for the dynamics of the state $\mathbf{x}_t^{(2)}$ evaluated through a suitable Lyapunov function $V^{(s)}$.

Lemma *4.2: Consider the equations (33), (34) and (35):*

$$
\begin{aligned}
\dot{\mathbf{x}}_t^{(2)} &= (A - BR^{(s)}B^T P^{(s)})\mathbf{x}_t^{(2)} \\
&\quad - BR^{(s)} B^T P^{(s)} \mathbf{w}_t^{(2)} + \phi(\mathbf{x}_t^{(2)})
\end{aligned} \tag{36}
$$

for $t \geqslant 0$ with $P^{(s)}$ and $R^{(s)}$ defined in (15) and $\Gamma^{(s)}$ given in lemma A.4. If $V^{(s)}(\mathbf{x}^{(2)}) := \|\epsilon^{-\mathfrak{r}} \diamond {S^{(s)}}^{-1} \mathbf{x}^{(2)}\|^2$ with $S^{(s)} := (I_n - A^T G^{(s)})^{-1}$ then for $t \geqslant 0$

$$
\begin{aligned}
\dot{V}^{(s)}|_{(36)} \leqslant &-(1-a) \min_{j=1,\ldots,N} \{\Gamma_{j,j}^{(s)}\} \epsilon^{2\mathfrak{f}_1^{(s)}} V^{(s)}(\mathbf{x}_t^{(2)}) \\
&+ \|\epsilon^{-\mathfrak{r}+\mathfrak{f}^{(s)}} \diamond S^{(s)} \mathbf{w}_t^{(2)}\|_{\Gamma^{(s)}}^2 \\
&+ \gamma_1 \left(\mathbf{x}_t^{(2)}, {S^{(s)}}^{-1} \sigma_{\lambda^{(s)}}(S^{(s)} \mathbf{x}_t^{(2)}) \right)
\end{aligned} \tag{37}
$$

for suitable function γ_1 such that $\gamma_1(\mathbf{v}, \mathbf{v}) = 0$ for all $\mathbf{v} \in \mathbb{R}^n$ and $a \in (0,1)$ introduced in (26).

Proof: Set $\widetilde{\mathbf{x}} := S^{(s)}\mathbf{x}$ with $\Sigma^{(s)} := (I - A^T\Gamma^{(s)})^{-1}$ and $\Sigma_{\text{inv}}^{(s)} := I_n + A^T \Gamma^{(s)}$ (see also definitions in (126)) and for simplicity we will omit the time subscript $_t$ and the superscript $^{(2)}$. Using (118) and the group properties of the dilations together with their inverse (125), commutative (122) and associative (123) properties given in the appendix, it is easy to obtain

$$
\begin{aligned}
&R^{(s)} B^T P^{(s)} = \\
&= B^T (\epsilon^{\mathfrak{r}} \diamond \epsilon^{\mathfrak{f}^{(s)}} \diamond \Gamma^{(s)} \diamond \epsilon^{\mathfrak{r}} \diamond \epsilon^{\mathfrak{f}^{(s)}}) BB^T \diamond \epsilon^{-2\mathfrak{r}} \diamond S^{(s)} \\
&= (B^T \diamond \epsilon^{\mathfrak{r}+\mathfrak{f}^{(s)}}) \Gamma^{(s)} (BB^T \diamond \epsilon^{-\mathfrak{r}+\mathfrak{f}^{(s)}}) S^{(s)}.
\end{aligned} \tag{38}
$$

(BB^T by (114) and $\Gamma^{(s)}$ by definition are diagonal). We have (with operator Δ defined in (111))

$$
\begin{aligned}
\dot{V}^{(s)}|_{(36)} = &\overbrace{2\left(\epsilon^{-\mathfrak{r}} \diamond \widetilde{\mathbf{x}}\right)^T \epsilon^{-\mathfrak{r}} \diamond \Big\{ S^{(s)} A {S^{(s)}}^{-1} \widetilde{\mathbf{x}}}^{(\mathbf{I})} \\
&\overbrace{- B(B^T \diamond \epsilon^{\mathfrak{r}+\mathfrak{f}^{(s)}}) \Gamma^{(s)} (BB^T \diamond \epsilon^{-\mathfrak{r}+\mathfrak{f}^{(s)}}) \widetilde{\mathbf{x}} \Big\}}^{(\mathbf{I})} \\
&\overbrace{+ 2\left(\epsilon^{-\mathfrak{r}} \diamond \widetilde{\mathbf{x}}\right)^T \epsilon^{-\mathfrak{r}} \diamond S^{(s)} \phi({S^{(s)}}^{-1} \sigma_{\lambda^{(s)}}(\widetilde{\mathbf{x}}))}^{(\mathbf{II})} \\
&\overbrace{+ 2\left(\epsilon^{-\mathfrak{r}} \diamond \widetilde{\mathbf{x}}\right)^T \epsilon^{-\mathfrak{r}} \diamond S^{(s)} \Delta\phi({S^{(s)}}^{-1} \widetilde{\mathbf{x}}, {S^{(s)}}^{-1} \sigma_{\lambda^{(s)}}(\widetilde{\mathbf{x}}))}^{(\mathbf{III})} \\
&\overbrace{- 2\left(\epsilon^{-\mathfrak{r}} \diamond \widetilde{\mathbf{x}}\right)^T \epsilon^{-\mathfrak{r}} \diamond B \times}^{(\mathbf{IV})} \\
&\overbrace{\times \left(B^T \diamond \epsilon^{\mathfrak{r}+\mathfrak{f}^{(s)}}) \Gamma^{(s)} (BB^T \diamond \epsilon^{-\mathfrak{r}+\mathfrak{f}^{(s)}}) S^{(s)} \mathbf{w} \right)}^{(\mathbf{IV})}.
\end{aligned} \tag{39}
$$

We find an upper bound for the terms under graphs in (39). Let's begin with the bracketed term ($\mathbf{I}$). On account of (119) with remark A.1 and the group properties of the dilations together with commutative (122) and associative (123) properties

$$\begin{aligned}(\mathbf{I}) &= 2\left(\epsilon^{-\mathfrak{r}+\mathfrak{f}^{(s)}}\diamond\widetilde{\mathbf{x}}\right)^T\Big\{\epsilon^{-\mathfrak{r}-\mathfrak{f}^{(s)}}\diamond\Big(A[S^{(s)^{-1}}-I_n]\\ &+A+[S^{(s)}-I_n]AS^{(s)^{-1}}\Big)\diamond\epsilon^{\mathfrak{r}-\mathfrak{f}^{(s)}}\\ &-BB^T\Gamma^{(s)}BB^T\Big\}\left(\epsilon^{-\mathfrak{r}+\mathfrak{f}^{(s)}}\diamond\widetilde{\mathbf{x}}\right)\\ &=2\left(\epsilon^{-\mathfrak{r}+\mathfrak{f}^{(s)}}\diamond\widetilde{\mathbf{x}}\right)^T\Big\{\epsilon^{-\mathfrak{r}-\mathfrak{f}^{(s)}}\diamond\Big(-AA^TG^{(s)}AA^T\\ &+A+[S^{(s)}-I_n]AS^{(s)^{-1}}\Big)\diamond\epsilon^{\mathfrak{r}-\mathfrak{f}^{(s)}}\\ &-BB^T\Gamma^{(s)}BB^T\Big\}\left(\epsilon^{-\mathfrak{r}+\mathfrak{f}^{(s)}}\diamond\widetilde{\mathbf{x}}\right). \qquad (40)\end{aligned}$$

Using (116), (121) with remark A.1 and the incremental homogeneity properties given in (i) and (iii) of lemma A.2, we obtain

$$(\mathbf{I})\leqslant 2\Big\langle\epsilon^{-\mathfrak{r}+\mathfrak{f}^{(s)}}\diamond\widetilde{\mathbf{x}}\Big\rangle^T\Big\{-\Gamma^{(s)}+A+\Big(\sum_{j=1}^{n-1}(A^T\Gamma^{(s)})^j\Big)A\Sigma_{\text{inv}}^{(s)}\Big\}\times$$
$$\times\Big\langle\epsilon^{-\mathfrak{r}+\mathfrak{f}^{(s)}}\diamond\widetilde{\mathbf{x}}\Big\rangle:=\|\epsilon^{-\mathfrak{r}+\mathfrak{f}^{(s)}}\diamond\widetilde{\mathbf{x}}\|^2_{-2\Gamma^{(s)}+N^{(s)}+N^{(s)T}}. \qquad (41)$$

Let's consider the bracketed term ($\mathbf{II}$) in (39). First of all, notice that on account of the incremental homogeneity property (iv) given in lemma A.2

$$\Big\langle\epsilon^{-\mathfrak{r}}\diamond S^{(s)^{-1}}\sigma_{\lambda^{(s)}}(\widetilde{\mathbf{x}})\Big\rangle\preceq\Sigma_{\text{inv}}^{(s)}\Big\langle\epsilon^{-\mathfrak{r}}\diamond\sigma_{\lambda^{(s)}}(\widetilde{\mathbf{x}})\Big\rangle \qquad (42)$$

and by the properties (2) of saturation functions (recall that $\lambda^{(s)}(\epsilon):=l^{(s)}\epsilon^{\mathfrak{r}}$)

$$\|\Sigma_{\text{inv}}^{(s)}\Big\langle\epsilon^{-\mathfrak{r}}\diamond\sigma_{\lambda^{(s)}}(\widetilde{\mathbf{x}})\Big\rangle\|\leqslant\|\Sigma_{\text{inv}}^{(s)}\mathbf{1}_n\|l^{(s)}. \qquad (43)$$

It follows straightforwardly that

$$\Phi^{(s)}\Big|_{x'=\epsilon^{-\mathfrak{r}}\diamond S^{(s)^{-1}}\sigma_{\lambda^{(s)}}(\widetilde{\mathbf{x}})}\preceq\sup_{\substack{x\in\mathbb{R}^n:\\ \|x\|\leqslant\|\Sigma_{\text{inv}}^{(s)}\mathbf{1}_n\|l^{(s)}}}\Phi^{(s)}(x)$$

(the matrix $\Phi^{(s)}(x')$ is introduced in (H1) and $\max\Phi^{(s)}(x')$ in the notation section). Using the group properties of the dilations, the incremental homogeneity properties (42) and (ii) and (iv) given in lemma A.2, the partial ordering (22) with $\hat{\mathfrak{f}}^{(s)}\preceq\mathfrak{f}^{(s)}$ (on account of (17)) and the property (3) of the saturation functions

$$\begin{aligned}(\mathbf{II})&\leqslant 2\big\langle\epsilon^{-\mathfrak{r}}\diamond\widetilde{\mathbf{x}}\big\rangle^T\Big\langle\epsilon^{-\mathfrak{r}}\diamond S^{(s)}\phi(S^{(s)^{-1}}\sigma_{\lambda^{(s)}(\epsilon)}(\widetilde{\mathbf{x}}))\Big\rangle\\ &\leqslant 2\Big\langle\epsilon^{-\mathfrak{r}+\mathfrak{f}^{(s)}}\diamond\widetilde{\mathbf{x}}\Big\rangle^T\Sigma^{(s)}\Phi^{(s)}\Big|_{x'=\epsilon^{-\mathfrak{r}}\diamond S^{(s)^{-1}}\sigma_{\lambda^{(s)}}(\widetilde{\mathbf{x}})}\times\\ &\times\Big\langle\epsilon^{-\mathfrak{r}+\hat{\mathfrak{f}}^{(s)}}\diamond S^{(s)^{-1}}\sigma_{\lambda^{(s)}}(\widetilde{\mathbf{x}})\Big\rangle\\ &\leqslant 2\Big\langle\epsilon^{-\mathfrak{r}+\mathfrak{f}^{(s)}}\diamond\widetilde{\mathbf{x}}\Big\rangle^T\Sigma^{(s)}\sup_{\substack{x'\in\mathbb{R}^n:\\ \|x'\|\leqslant\|\Sigma_{\text{inv}}^{(s)}\mathbf{1}_n\|l^{(s)}}}\Phi^{(s)}(x')\times\\ &\times\Sigma_{\text{inv}}^{(s)}\Big\langle\epsilon^{-\mathfrak{r}+\mathfrak{f}^{(s)}}\diamond\sigma_{\lambda^{(s)}}(\widetilde{\mathbf{x}})\Big\rangle\\ &\leqslant 2\Big\langle\epsilon^{-\mathfrak{r}+\mathfrak{f}^{(s)}}\diamond\widetilde{\mathbf{x}}\Big\rangle^T\Sigma^{(s)}\sup_{\substack{x'\in\mathbb{R}^n:\\ \|x'\|\leqslant\|\Sigma_{\text{inv}}^{(s)}\mathbf{1}_n\|l^{(s)}}}\Phi^{(s)}(x')\times\\ &\times\Sigma_{\text{inv}}^{(s)}\Big\langle\epsilon^{-\mathfrak{r}+\mathfrak{f}^{(s)}}\diamond\widetilde{\mathbf{x}}\Big\rangle:=\|\epsilon^{-\mathfrak{r}+\mathfrak{f}^{(s)}}\diamond\widetilde{\mathbf{x}}\|^2_{M^{(s)}+M^{(s)T}}. \qquad (44)\end{aligned}$$

Let's consider the bracketed term ($\mathbf{III}$) in (39). Using the group properties of the dilations and Young's inequality

$$\begin{aligned}(\mathbf{III})&\leqslant 2\big\langle\epsilon^{-\mathfrak{r}}\diamond\widetilde{\mathbf{x}}\big\rangle^T\times\\ &\times\Big\langle\epsilon^{-\mathfrak{r}}\diamond S^{(s)}\Delta\phi(S^{(s)^{-1}}\widetilde{\mathbf{x}},S^{(s)^{-1}}\sigma_{\lambda^{(s)}}(\widetilde{\mathbf{x}}))\Big\rangle\\ &\leqslant a\|\epsilon^{-\mathfrak{r}+\mathfrak{f}^{(s)}}\diamond\widetilde{\mathbf{x}}\|^2\\ &+\frac{1}{a}\|\epsilon^{-\mathfrak{r}-\mathfrak{f}^{(s)}}\diamond S^{(s)}\Delta\phi(S^{(s)^{-1}}\widetilde{\mathbf{x}},S^{(s)^{-1}}\sigma_{\lambda^{(s)}}(\widetilde{\mathbf{x}}))\|^2\\ &:=a\|\epsilon^{-\mathfrak{r}+\mathfrak{f}^{(s)}}\diamond\widetilde{\mathbf{x}}\|^2+\gamma_1\left(\mathbf{x},S^{(s)^{-1}}\sigma_{\lambda^{(s)}}(S^{(s)}\mathbf{x})\right) \qquad (45)\end{aligned}$$

where the function γ_1 is such that $\gamma_1(\mathbf{v},\mathbf{v})=0$ for all $\mathbf{v}\in\mathbb{R}^n$. Eventually, we majorize the bracketed term ($\mathbf{IV}$) in (39). On account of the group properties of the dilations, commutative (122) (recall that BB^T by (114) and $\Gamma^{(s)}$ by definition are diagonal) and associative (123) properties and Young's inequality

$$\begin{aligned}(\mathbf{IV})\ &\leqslant\ \|\epsilon^{-\mathfrak{r}+\mathfrak{f}^{(s)}}\diamond\widetilde{\mathbf{x}}\|^2_{\Gamma^{(s)}}+\|\epsilon^{-\mathfrak{r}+\mathfrak{f}^{(s)}}\diamond S^{(s)}\mathbf{w}\|^2_{BB^T\Gamma^{(s)}BB^T}\\ &\leqslant\ \|\epsilon^{-\mathfrak{r}+\mathfrak{f}^{(s)}}\diamond\widetilde{\mathbf{x}}\|^2_{\Gamma^{(s)}}+\|\epsilon^{-\mathfrak{r}+\mathfrak{f}^{(s)}}\diamond S^{(s)}\mathbf{w}\|^2_{\Gamma^{(s)}}. \qquad (46)\end{aligned}$$

After collecting (40)-(46), since $\Gamma^{(s)}$ is given according to lemma A.4 so that (131) holds true and, on account of the partial ordering (23),

$$\begin{aligned}\|\epsilon^{-\mathfrak{r}+\mathfrak{f}^{(s)}}\diamond\widetilde{\mathbf{x}}\|^2_{\Gamma^{(s)}}&=\|\epsilon^{\mathfrak{f}^{(s)}}\diamond(\epsilon^{-\mathfrak{r}}\diamond S^{(s)}\mathbf{x})\|^2_{\Gamma^{(s)}}\\ &\geqslant\min_{j=1,\dots,N}\{\Gamma^{(s)}_{j,j}\}\epsilon^{2\mathfrak{f}_1^{(s)}}V^{(s)}(\mathbf{x}), \qquad (47)\end{aligned}$$

we obtain (37). ∎

Next, we evaluate an upper bound for the second right-hand term in (37). To this aim, let $S^{(s)}$ and $\Sigma_{\text{inv}}^{(s)}$ be as in lemma 4.2 and set (see also definitions in (126))

$$S^{(o)}\ :=\ I_n-G^{(o)}A^T,\ \Sigma_{\text{inv}}^{(o)}:=(I-\Gamma^{(o)}A^T)^{-1}. \qquad (48)$$

Lemma *4.3: For $t\geqslant 0$:*

$$\begin{aligned}&\|\epsilon^{-\mathfrak{r}+\mathfrak{f}^{(s)}}\diamond S^{(s)}\mathbf{w}_t^{(2)}\|^2_{\Gamma^{(s)}}\\ &\leqslant 4n\max_{j=1,\dots,n}\{\Gamma^{(s)}_{j,j}\}\Big[\epsilon^{2\mathfrak{f}_n^{(s)}}\min\Big\{{l^{(s)}}^2,\epsilon^{2(\mathfrak{f}_1^{(o)}-\mathfrak{f}_n^{(o)})}\times\\ &\times\|\Sigma^{(s)}\Sigma_{inv}^{(o)}\|^2\left\|\epsilon^{-\mathfrak{r}}\diamond S^{(o)}(\mathbf{x}_t^{(2)}-\hat{\mathbf{x}}_t^{(2)})\right\|^2\Big\}\\ &+\gamma_2\left(\mathbf{x}_t,S^{(s)^{-1}}\sigma_{\lambda^{(s)}}(S^{(s)}\mathbf{x}_t)\right) \qquad (49)\end{aligned}$$

for some function γ_2 such that $\gamma_2(\mathbf{v},\mathbf{v})=0$ for all $\mathbf{v}\in\mathbb{R}^n$.

Proof: From (35)

$$\mathbf{w}_t^{(2)}=S^{(s)^{-1}}\Big\{\sigma_{\lambda^{(s)}}\left(S^{(s)}\hat{\mathbf{x}}_t^{(2)}\right)-S^{(s)}\mathbf{x}_t^{(2)}\Big\}. \qquad (50)$$

From now on, for simplicity, we will omit the superscript $^{(2)}$, the subscript $_t$ and set $\mathbf{e}:=\mathbf{x}-\hat{\mathbf{x}}$. From (50) it follows

$$\Big\langle S^{(s)}\mathbf{w}\Big\rangle\preceq\overbrace{\Big\langle\sigma_{\lambda^{(s)}}\left(S^{(s)}(\mathbf{x}-\mathbf{e})\right)-\sigma_{\lambda^{(s)}}(S^{(s)}\mathbf{x})\Big\rangle}^{(\mathbf{I})}$$
$$+\overbrace{\Big\langle\sigma_{\lambda^{(s)}}(S^{(s)}\mathbf{x})-S^{(s)}\mathbf{x}\Big\rangle}^{(\mathbf{II})} \qquad (51)$$

and using the property (2) of saturation functions

$$(\mathbf{I})\preceq 2\Big\langle\sigma_{\lambda^{(s)}}\left(S^{(s)}S^{(o)^{-1}}\left(S^{(o)}\mathbf{e}\right)\right)\Big\rangle. \qquad (52)$$

By the incremental homogeneity property (iv) given in lemma A.1 and (iv) given in lemma A.2, following (52) we have

$$
\begin{aligned}
(\mathbf{I}) &\preceq \Big\langle \sigma_{\lambda^{(s)}}\Big(S^{(s)}\epsilon^{\mathfrak{r}-\mathfrak{f}^{(o)}}\diamond \Sigma^{(o)}_{inv}\left(\epsilon^{-\mathfrak{r}+\mathfrak{f}^{(o)}}\diamond S^{(o)}\mathbf{e}\right)\Big)\Big\rangle \\
&\preceq \Big\langle \sigma_{\lambda^{(s)}}\left(\epsilon^{\mathfrak{r}}\diamond \Sigma^{(s)}\epsilon^{-\mathfrak{f}^{(o)}}\diamond \Sigma^{(o)}_{inv}\left(\epsilon^{-\mathfrak{r}+\mathfrak{f}^{(o)}}\diamond S^{(o)}\mathbf{e}\right)\right)\Big\rangle \\
&= \epsilon^{\mathfrak{r}}\diamond\Big\langle \sigma_{l^{(s)}\mathbf{1}_n}\left(\Sigma^{(s)}\epsilon^{-\mathfrak{f}^{(o)}}\diamond \Sigma^{(o)}_{inv}\left(\epsilon^{-\mathfrak{r}+\mathfrak{f}^{(o)}}\diamond S^{(o)}\mathbf{e}\right)\right)\Big\rangle \quad (53)
\end{aligned}
$$

(recall that $\lambda^{(s)}(\epsilon) = l^{(s)}\epsilon^{\mathfrak{r}}$). Using the partial ordering (23) it follows from (51) and (53) that

$$
\begin{aligned}
&\langle \epsilon^{-\mathfrak{r}+\mathfrak{f}^{(s)}}\diamond S^{(s)}\mathbf{w}\rangle \\
&\quad\preceq 2\epsilon^{\mathfrak{f}^{(s)}_n}\Big\langle \sigma_{\mathbf{1}_n l^{(s)}}\left(\epsilon^{\mathfrak{f}^{(o)}_1-\mathfrak{f}^{(o)}_n}\Sigma^{(s)}\Sigma^{(o)}_{inv}\left(\epsilon^{-\mathfrak{r}}\diamond S^{(o)}\mathbf{e}\right)\right)\Big\rangle \\
&\quad+\widetilde{\gamma}_2\left(\mathbf{x}, {S^{(s)}}^{-1}\sigma_{\lambda^{(s)}}(S^{(s)}\mathbf{x})\right) \quad (54)
\end{aligned}
$$

where the function $\widetilde{\gamma}_2$ majorizes $(\mathbf{II})$ in (51) and it is such that $\widetilde{\gamma}_2(\mathbf{v},\mathbf{v}) = 0$ for all $\mathbf{v}\in\mathbb{R}^n$. From the inequality $\|\langle\sigma_{c\mathbf{1}_n}(\mathbf{v})\rangle\|^2 \leqslant n\min\{c^2,\|\mathbf{v}\|^2)$ for all $\mathbf{v}\in\mathbb{R}^n$ eventually we get (49). ■

Set $\mathbf{e}^{(j)} := \mathbf{x}^{(j)} - \hat{\mathbf{x}}^{(j)}$, $j = 2,\dots,\nu+1$. Consider the observer chain

$$
\begin{aligned}
\dot{\hat{\mathbf{x}}}^{(j)}_t &= A\hat{\mathbf{x}}^{(j)}_t + B\mathbf{u}^{(j)}_{t-c} + \phi\left(\sigma_{\lambda^{(o)}}(\hat{\mathbf{x}}^{(j)}_t)\right) \quad (55)\\
&\quad + {P^{(o)}}^{-1}C^T R^{(o)}\mathbf{z}^{(j)}_t, j = 2,\dots,\nu+1,
\end{aligned}
$$

for $t\geqslant 0$ and notice that, from the definition (9), (11) of the innovation $\mathbf{z}^{(j)}_t$,

$$\mathbf{z}^{(j)}_t = C(\mathbf{e}^{(j)}_t + \mathbf{q}^{(j)}_t) \quad (56)$$

with exogenous input

$$
\mathbf{q}^{(j)}_t = \begin{cases} -\mathbf{e}^{(j)}_t + \mathbf{e}^{(j)}_{t-\mathbf{s}^{(j)}_t} & \text{if } \mathbf{d}_t\in[0,p^{(j)}],\\ -\mathbf{e}^{(j)}_t + \mathbf{e}^{(j)}_{t-\mathbf{s}^{(j)}_t} - \mathbf{e}^{(j+1)}_t & \text{if } \mathbf{d}_t\in(p^{(j)},p^{(\nu+1)}]\end{cases} \quad (57)
$$

and delay

$$
\mathbf{s}^{(j)}_t = \begin{cases} 0 & \text{if } \mathbf{d}_t\in[0,p^{(j-1)}),\\ \mathbf{d}_t - p^{(j-1)} & \text{if } \mathbf{d}_t\in[p^{(j-1)},p^{(j)}],\\ p^{(j)}-p^{(j-1)} & \text{if } \mathbf{d}_t\in(p^{(j)},p^{(\nu+1)}]\end{cases} \quad (58)
$$

if $j = \nu_0+1,\dots,\nu+1$ and

$$\mathbf{q}^{(j)}_t = -\mathbf{e}^{(j)}_t + \mathbf{e}^{(j)}_{t-\mathbf{s}^{(j)}_t} - \mathbf{e}^{(j+1)}_t, \quad (59)$$

$$\mathbf{s}^{(j)}_t = p^{(j)} - p^{(j-1)} \quad (60)$$

if $j = 2,\dots,\nu_0$.

Next, consider the equations for $\mathbf{x}^{(j)}_t$, $j = 2,\dots,\nu+1$:

$$\dot{\mathbf{x}}^{(j)}_t = A\mathbf{x}^{(j)}_t + B\mathbf{u}^{(j)}_{t-c} + \phi(\mathbf{x}^{(j)}_t) \quad (61)$$

for $t\geqslant 0$. The following lemma establishes a bound for the dynamics of the estimation error $\mathbf{e}^{(j)}_t$ evaluated through a suitable Lyapunov function $V^{(o)}$. To this aim, let $S^{(o)}$ and $\Sigma^{(o)}_{\text{inv}}$ be as in (48) and set $\Sigma^{(o)} := I_n + \Gamma^{(o)}A^T$.

Lemma *4.4: Consider the equations (61), (55):*

$$
\begin{aligned}
\dot{\mathbf{x}}^{(j)}_t &= A\mathbf{x}^{(j)}_t + B\mathbf{u}^{(j)}_{t-c} + \phi(\mathbf{x}^{(j)}_t)\\
\dot{\hat{\mathbf{x}}}^{(j)}_t &= A\hat{\mathbf{x}}^{(j)}_t + B\mathbf{u}^{(j)}_{t-c} + \phi\left(\sigma_{\lambda^{(o)}(\epsilon)}(\hat{\mathbf{x}}^{(j)}_t)\right) \quad (62)\\
&\quad + {P^{(o)}}^{-1}C^T R^{(o)} C\left(\mathbf{e}^{(j)}_t + \mathbf{q}^{(j)}_t\right)
\end{aligned}
$$

for $t\geqslant 0$ with $P^{(o)}$ and $R^{(o)}$ defined in (8) and $\Gamma^{(o)}$ given in lemma A.3. If $V^{(o)}(\mathbf{e}^{(j)}) := \|\epsilon^{-\mathfrak{r}}\diamond S^{(o)}\mathbf{e}^{(j)}\|^2$ then for all $t\geqslant 0$

$$
\begin{aligned}
\dot{V}^{(o)}|_{(62)} &\leqslant -a\min_{j=1,\dots,n}\{\Gamma^{(o)}_{j,j}\}\,\epsilon^{2\mathfrak{f}^{(o)}_n}V^{(o)}(\mathbf{e}^{(j)}_t)\\
&\quad-\left[2-3a-\frac{1}{b^{(j)}}\right]\|\epsilon^{-\mathfrak{r}+\mathfrak{f}^{(o)}}\diamond S^{(o)}\mathbf{e}^{(j)}_t\|^2_{\Gamma^{(o)}} \quad (63)\\
&\quad+b^{(j)}\|\epsilon^{-\mathfrak{r}+\mathfrak{f}^{(o)}}\diamond \mathbf{q}^{(j)}_t\|^2_{\Gamma^{(o)}} + \gamma_3\left(\mathbf{x}^{(j)}_t, \sigma_{\lambda^{(o)}(\epsilon)}(\mathbf{x}^{(j)}_t)\right).
\end{aligned}
$$

for suitable function γ_3 such that $\gamma_3(\mathbf{v},\mathbf{v}) = 0$ for all $\mathbf{v}\in\mathbb{R}^n$ and $a\in(0,1)$ and $b^{(j)}\in(1,2)$ introduced in (25), (26).

Proof: For simplicity, we will omit the subscript $_t$, the superscript $^{(j)}$ and set $\widetilde{\mathbf{e}} = S^{(o)}\mathbf{e}$. Notice that, using (117) and the group properties of the dilations together with their commutative (122) and associative (123) properties given in the appendix,

$$
\begin{aligned}
&{P^{(o)}}^{-1}C^T R^{(o)} =\\
&= {S^{(o)}}^{-1}\diamond\epsilon^{2\mathfrak{r}}\diamond C^T C(\epsilon^{-\mathfrak{r}}\diamond\epsilon^{\mathfrak{f}^{(o)}}\diamond\Gamma^{(o)}\diamond\epsilon^{-\mathfrak{r}}\diamond\epsilon^{\mathfrak{f}^{(o)}})C^T\\
&= {S^{(o)}}^{-1}(\epsilon^{\mathfrak{r}+\mathfrak{f}^{(o)}}\diamond C^T C)\Gamma^{(o)}(\epsilon^{-\mathfrak{r}+\mathfrak{f}^{(o)}}\diamond C^T) \quad (64)
\end{aligned}
$$

($C^T C$ by (114) and $\Gamma^{(o)}$ by definition are diagonal). We have (with operator Δ defined in (111))

$$
\begin{aligned}
\dot{V}^{(o)}|_{(62)} &= \overbrace{2\left(\epsilon^{-\mathfrak{r}}\diamond\widetilde{\mathbf{e}}\right)^T\epsilon^{-\mathfrak{r}}\diamond\Big\{S^{(o)}AS^{(o)^{-1}}\widetilde{\mathbf{e}}}^{(\mathbf{I})}\\
&\qquad\overbrace{-(\epsilon^{\mathfrak{r}+\mathfrak{f}^{(o)}}\diamond C^T C)\Gamma^{(o)}(\epsilon^{-\mathfrak{r}+\mathfrak{f}^{(o)}}\diamond C^T)CS^{(o)^{-1}}\widetilde{\mathbf{e}}\Big\}}^{(\mathbf{I})}\\
&+\overbrace{2(\epsilon^{-\mathfrak{r}}\diamond\widetilde{\mathbf{e}})^T\epsilon^{-\mathfrak{r}}\diamond S^{(o)}\Delta\phi\Big(\sigma_{\lambda^{(o)}(\epsilon)}(\mathbf{x}),\sigma_{\lambda^{(o)}(\epsilon)}(\mathbf{x}-S^{(o)^{-1}}\widetilde{\mathbf{e}})\Big)}^{(\mathbf{II})}\\
&+\overbrace{2(\epsilon^{-\mathfrak{r}}\diamond\widetilde{\mathbf{e}})^T\epsilon^{-\mathfrak{r}}\diamond S^{(o)}\Delta\phi\left(\mathbf{x},\sigma_{\lambda^{(o)}(\epsilon)}(\mathbf{x})\right)}^{(\mathbf{III})}\\
&-\overbrace{2(\epsilon^{-\mathfrak{r}}\diamond\widetilde{\mathbf{e}})^T\epsilon^{-\mathfrak{r}}\diamond(\epsilon^{\mathfrak{r}+\mathfrak{f}^{(o)}}\diamond C^T C)\Gamma^{(o)}(\epsilon^{-\mathfrak{r}+\mathfrak{f}^{(o)}}\diamond C^T)C\mathbf{q}}^{(\mathbf{IV})}. \quad (65)
\end{aligned}
$$

We find an upper bound for the terms under graphs in (65). Let's begin with the bracketed term $(\mathbf{I})$. On account of (119) and the group properties of the dilations together with commutative (122) and associative (123) properties

$$
\begin{aligned}
(\mathbf{I}) &= 2(\epsilon^{-\mathfrak{r}+\mathfrak{f}^{(o)}}\diamond\widetilde{\mathbf{e}})^T\Big\{\epsilon^{-\mathfrak{r}-\mathfrak{f}^{(o)}}\diamond\Big([S^{(o)}-I_n]A + A \quad (66)\\
&\quad+S^{(o)}A[S^{(o)^{-1}}-I_n]\Big)\diamond\epsilon^{\mathfrak{r}-\mathfrak{f}^{(o)}} - C^T C\Gamma^{(o)}C^T C\Big\}(\epsilon^{-\mathfrak{r}+\mathfrak{f}^{(o)}}\diamond\widetilde{\mathbf{e}})\\
&= 2(\epsilon^{-\mathfrak{r}+\mathfrak{f}^{(o)}}\diamond\widetilde{\mathbf{e}})^T\Big\{\epsilon^{-\mathfrak{r}-\mathfrak{f}^{(o)}}\diamond\Big(-A^T AG^{(o)}A^T A + A +\\
&\quad+S^{(o)}A[S^{(o)^{-1}}-I_n]\Big)\diamond\epsilon^{\mathfrak{r}-\mathfrak{f}^{(o)}} - C^T C\Gamma^{(o)}C^T C\Big\}(\epsilon^{-\mathfrak{r}+\mathfrak{f}^{(o)}}\diamond\widetilde{\mathbf{e}})
\end{aligned}
$$

and using (116), (120) and the incremental homogeneity properties given in (i) and (iii) of lemma A.1.

$$
\begin{aligned}
&\leqslant 2\langle\epsilon^{-\mathfrak{r}+\mathfrak{f}^{(o)}}\diamond\widetilde{\mathbf{e}}\rangle^T\Big\{-\Gamma^{(o)} + A + \Sigma^{(o)}A\sum_{j=1}^{n-1}(\Gamma^{(o)}A^T)^j\Big)\Big\}\times\\
&\times\langle\epsilon^{-\mathfrak{r}+\mathfrak{f}^{(o)}}\diamond\widetilde{\mathbf{e}}\rangle := \|\epsilon^{-\mathfrak{r}+\mathfrak{f}^{(o)}}\diamond\widetilde{\mathbf{e}}\|^2_{-2\Gamma^{(o)}+N^{(o)}+N^{(o)T}}. \quad (67)
\end{aligned}
$$

Let's consider the bracketed term $(\mathbf{II})$ in (65). By the properties (2) of saturation functions (recall that $\lambda^{(o)}(\epsilon) = l^{(o)}\epsilon^{\mathfrak{r}}$)

$$\|\epsilon^{-\mathfrak{r}} \diamond \sigma_{\lambda^{(o)}}(\mathbf{v})\| \leqslant \|\mathbf{1}_n\| l^{(o)} = n l^{(o)},\ \forall \mathbf{v} \in \mathbb{R}^n.$$

It follows straightforwardly that for all $\mathbf{x}$ and $\widetilde{\mathbf{e}}$

$$\Phi^{(o)}(x', x'')\Big|_{\substack{x' = \epsilon^{-\mathfrak{r}} \diamond \sigma_{\lambda^{(o)}}(\mathbf{x}) \\ x'' = \epsilon^{-\mathfrak{r}} \diamond \sigma_{\lambda^{(o)}}(\mathbf{x} - S^{(o)-1}\widetilde{\mathbf{e}})}} \preceq \sup_{\substack{x', x'' \in \mathbb{R}^n: \\ \|x'\|, \|x''\| \leqslant n l^{(o)}}} \Phi^{(o)}(x', x'')$$

(the matrix $\Phi^{(o)}(x', x'')$ is introduced in (H2) and $\max \Phi^{(o)}(x', x'')$ in the notation section). Using the group properties of the dilations, the property (2) of saturation functions, the incremental homogeneity properties (ii) and (iv) given in lemma A.1 and since the linear operator $S^{(o)}$ commutes with the operator Δ

$$\begin{aligned}
(\mathbf{II}) &\leqslant 2\left\langle \epsilon^{-\mathfrak{r}} \diamond \widetilde{\mathbf{e}} \right\rangle^T \times \\
&\quad \times \left\langle \epsilon^{-\mathfrak{r}} \diamond \Delta(S^{(o)}\phi)\Big(\sigma_{\lambda^{(o)}}(\mathbf{x}), \sigma_{\lambda^{(o)}}(\mathbf{x} - S^{(o)-1}\widetilde{\mathbf{e}})\Big) \right\rangle \\
&\leqslant 2\left\langle \epsilon^{-\mathfrak{r}+\mathfrak{f}^{(o)}} \diamond \widetilde{\mathbf{e}} \right\rangle^T \Sigma^{(o)} \Phi^{(o)}\Big|_{\substack{x' = \epsilon^{-\mathfrak{r}} \diamond \sigma_{\lambda^{(o)}}(\mathbf{x}) \\ x'' = \epsilon^{-\mathfrak{r}} \diamond \sigma_{\lambda^{(o)}}(\mathbf{x} - S^{(o)-1}\widetilde{\mathbf{e}})}} \times \\
&\quad \times \left\langle \epsilon^{-\mathfrak{r}+\mathfrak{f}^{(o)}} \diamond \Delta\sigma_{\lambda^{(o)}}\left(\mathbf{x}, \mathbf{x} - S^{(o)-1}\widetilde{\mathbf{e}}\right) \right\rangle \\
&\leqslant 4\left\langle \epsilon^{-\mathfrak{r}+\mathfrak{f}^{(o)}} \diamond \widetilde{\mathbf{e}} \right\rangle^T \Sigma^{(o)} \sup_{\substack{x', x'' \in \mathbb{R}^n: \\ \|x'\|, \|x''\| \leqslant n l^{(o)}}} \Phi^{(o)}(x', x'') \times \\
&\quad \times \Sigma^{(o)}_{\text{inv}} \left\langle \epsilon^{-\mathfrak{r}+\mathfrak{f}^{(o)}} \diamond \widetilde{\mathbf{e}} \right\rangle := \|\epsilon^{-\mathfrak{r}+\mathfrak{f}^{(o)}} \diamond \widetilde{\mathbf{e}}\|^2_{M^{(o)}+M^{(o)T}}. \quad (68)
\end{aligned}$$

Let's consider the bracketed term $(\mathbf{III})$ in (65). Using Young's inequality and the group properties of the dilations,

$$\begin{aligned}
(\mathbf{III}) &\leqslant 2\left\langle \epsilon^{-\mathfrak{r}} \diamond \widetilde{\mathbf{e}} \right\rangle^T \left\langle \epsilon^{-\mathfrak{r}} \diamond \Delta(S^{(o)}\phi)(\mathbf{x}, \sigma_{\lambda^{(o)}}(\mathbf{x})) \right\rangle \\
&\leqslant a\|\epsilon^{-\mathfrak{r}+\mathfrak{f}^{(o)}} \diamond \widetilde{\mathbf{e}}\|^2 + \frac{1}{a}\|\epsilon^{-\mathfrak{r}-\mathfrak{f}^{(o)}} \diamond \Delta(S^{(o)}\phi)(\mathbf{x}, \sigma_{\lambda^{(o)}}(\mathbf{x}))\|^2 \\
&:= a\|\epsilon^{-\mathfrak{r}+\mathfrak{f}^{(o)}} \diamond \widetilde{\mathbf{e}}\|^2 + \gamma_3(\mathbf{x}, \sigma_{\lambda^{(o)}}(\mathbf{x})), \quad (69)
\end{aligned}$$

where the function γ_3 is such that $\gamma_3(\mathbf{v}, \mathbf{v}) = 0$ for all $\mathbf{v} \in \mathbb{R}^n$. Eventually, we majorize the bracketed term $(\mathbf{IV})$ in (65). On account of the group properties of the dilations, using the commutative (122) (recall that C^TC and $\Gamma^{(o)}$ are diagonal) and associative (123) properties with (113) and Young's inequality

$$\begin{aligned}
&\leqslant \frac{1}{b^{(j)}}\|\epsilon^{-\mathfrak{r}+\mathfrak{f}^{(o)}} \diamond \widetilde{\mathbf{e}}\|^2_{\Gamma^{(o)}} + b^{(j)}\|\epsilon^{-\mathfrak{r}+\mathfrak{f}^{(o)}} \diamond \mathbf{q}\|^2_{C^TC\Gamma^{(o)}C^TC} \\
&\leqslant \frac{1}{b^{(j)}}\|\epsilon^{-\mathfrak{r}+\mathfrak{f}^{(o)}} \diamond \widetilde{\mathbf{e}}\|^2_{\Gamma^{(o)}} + b^{(j)}\|\epsilon^{-\mathfrak{r}+\mathfrak{f}^{(o)}} \diamond \mathbf{q}\|^2_{\Gamma^{(o)}}. \quad (70)
\end{aligned}$$

Collecting (66)-(70), since $\Gamma^{(o)}$ is given according to lemma A.3 so that (128) holds true, on account of the partial ordering (23),

$$\begin{aligned}
\|\epsilon^{-\mathfrak{r}+\mathfrak{f}^{(o)}} \diamond \widetilde{\mathbf{e}}\|^2_{\Gamma^{(o)}} &= \|\epsilon^{\mathfrak{f}^{(o)}} \diamond (\epsilon^{-\mathfrak{r}} \diamond S^{(o)}\mathbf{e})\|^2_{\Gamma^{(o)}} \\
&\geqslant \min_{j=1,\ldots,N}\{\Gamma^{(o)}_{j,j}\}\epsilon^{2\mathfrak{f}^{(o)}_n} V^{(o)}(\mathbf{e}), \quad (71)
\end{aligned}$$

we obtain (63). ■

Mimicking the proof of lemma 4.4 and using the partial ordering (24) we obtain also the following useful inequality.

Lemma *4.5: For $t \geqslant 0$*

$$\begin{aligned}
\left\langle \epsilon^{-\mathfrak{r}+\mathfrak{f}^{(o)}} \diamond S^{(o)}\dot{\mathbf{e}}^{(\nu+1)}_t \right\rangle &\preceq \epsilon^{3\mathfrak{f}^{(o)}_1}\Pi\left\langle \epsilon^{-\mathfrak{r}} \diamond S^{(o)}\mathbf{e}^{(\nu+1)}_t \right\rangle \\
&\quad + \gamma_4\left(\mathbf{x}^{(\nu+1)}_t, \sigma_{\lambda^{(o)}}(\mathbf{x}^{(\nu+1)}_t)\right)
\end{aligned}$$

and

$$\begin{aligned}
\left\langle \epsilon^{-\mathfrak{r}+\mathfrak{f}^{(o)}} \diamond S^{(o)}\dot{\mathbf{e}}^{(j)}_t \right\rangle &\preceq \epsilon^{3\mathfrak{f}^{(o)}_1}\Pi\Big[\sum_{h=j}^{j+1}\left\langle \epsilon^{-\mathfrak{r}} \diamond S^{(o)}\mathbf{e}^{(h)}_t \right\rangle \\
&+ \left\langle \epsilon^{-\mathfrak{r}} \diamond S^{(o)}\mathbf{e}^{(j)}_{t-\tau^{(j)}} \right\rangle\Big] + \gamma_4\left(\mathbf{x}^{(j)}_t, \sigma_{\lambda^{(o)}}(\mathbf{x}^{(j)}_t)\right)
\end{aligned}$$

if $j = 2, \ldots, \nu$, where

$$\Pi := \Sigma^{(o)}\left[A + \sup_{\substack{x', x'' \in \mathbb{R}^n: \\ \|x'\|, \|x''\| \leqslant n l^{(o)}}} \Phi^{(o)}(x', x'')\right]\Sigma^{(o)}_{inv} + \Gamma^{(o)}$$

and for suitable function γ_4 such that $\gamma_4(\mathbf{v}, \mathbf{v}) = 0$ for all $\mathbf{v} \in \mathbb{R}^n$.

Next, using lemma 4.5 we compute an upper bound for the second right-hand term in (63).

Lemma *4.6: For $t \geqslant \delta$:*

$$\|\epsilon^{-\mathfrak{r}+\mathfrak{f}^{(o)}} \diamond \mathbf{q}^{(\nu+1)}_t\|^2_{\Gamma^{(o)}} = 0 \quad (72)$$

and

$$\begin{aligned}
&\|\epsilon^{-\mathfrak{r}+\mathfrak{f}^{(o)}} \diamond \mathbf{q}^{(j)}_t\|^2_{\Gamma^{(o)}} \\
&\quad \leqslant (1+a)\|\epsilon^{-\mathfrak{r}+\mathfrak{f}^{(o)}} \diamond S^{(o)}\mathbf{e}^{(j+1)}_t\|^2_{\Gamma^{(o)}} \\
&\quad +r\epsilon^{6\mathfrak{f}^{(o)}_1}\delta\int_{t-\delta}^{t}\Big(\sum_{h=j}^{j+1}V^{(o)}(\mathbf{e}^{(h)}_\theta) + V^{(o)}(\mathbf{e}^{(j)}_{\theta-\mathbf{s}^{(j)}_\theta})\Big)d\theta \\
&\quad +\gamma_5\left(\mathbf{x}^{(j)}_t, \sigma_{\lambda^{(o)}}(\mathbf{x}^{(j)}_t)\right) \ \text{ if } j = 2, \ldots, \nu, \quad (73)
\end{aligned}$$

with

$$r := 8\left(1 + \frac{1}{a}\right)\max_{j=1,\ldots,n}\{\Gamma^{(o)}_{j,j}\}\|\Pi\|^2$$

and suitable function γ_5 such that $\gamma_5(\mathbf{v}, \mathbf{v}) = 0$ for all $\mathbf{v} \in \mathbb{R}^n$.

Proof: We begin with $\mathbf{q}^{(\nu+1)}_t$ for which, by its definition in (57) and since by construction $p^{(\nu)} = d_\infty$ with $\mathbf{d}_t \in [0, d_\infty]$ for all $t \geqslant 0$, we have $\mathbf{q}^{(\nu+1)}_t \equiv 0$, i.e. (72).

Now, let $j = 2, \ldots, N$. In this case $\mathbf{d}_t \notin [0, p^{(j)}]$ for all $t \geqslant 0$ so that $\mathbf{q}^{(j)}_t \not\equiv 0$. On account of the definition (57) and (64) with the bound (32) on the delay $\mathbf{s}^{(j)}_t$

$$\begin{aligned}
\left\langle \epsilon^{-\mathfrak{r}+\mathfrak{f}^{(o)}} \diamond \mathbf{q}^{(j)}_t \right\rangle &\preceq \langle \epsilon^{-\mathfrak{r}+\mathfrak{f}^{(o)}} \diamond S^{(o)}\mathbf{e}^{(j+1)}_t \rangle \\
&+ \int_{t-\mathbf{s}^{(j)}_t}^{t}\left\langle \epsilon^{-\mathfrak{r}+\mathfrak{f}^{(o)}} \diamond S^{(o)}\dot{\mathbf{e}}^{(j)}_s \right\rangle ds \quad (74)
\end{aligned}$$

for all $t \geqslant \delta$. Using lemma 4.5, the Jensen's inequality

$$\left\|\int_{t-\mathbf{s}_t}^{t}\mathbf{v}_\theta d\theta\right\|^2 \leqslant \tau\int_{t-s_\infty}^{t}\|\mathbf{v}_\theta d\theta\|^2,\ \forall t \geqslant s_\infty \quad (75)$$

for any bounded delay function $\mathbf{s}_t$ such that $\sup_{t\geqslant 0}\mathbf{s}_t := s_\infty$, the bound (32) on $\mathbf{s}^{(j)}_t$ and Young's inequality with (113), we obtain (73). ■

Next, we define a candidate Lyapunov-Razumichin function for the overall system (33), (55) as follows:

$$W^{(s,o)}(\mathbf{x}^{(2)},\mathbf{e}^{\otimes}) := V^{(s)}(\mathbf{x}^{(2)}) + \frac{1}{\omega}\ln\left(1+W^{(o)}(\mathbf{e}^{\otimes})\right) \quad (76)$$

where $\mathbf{e}^{\otimes} := (\mathbf{e}^{(2)},\dots,\mathbf{e}^{(\nu+1)})$, $\omega \in \mathcal{K}_\infty$ has been introduced in (28), (29) and

$$W^{(o)}(\mathbf{e}^{\otimes}) := \epsilon^{2(\mathfrak{f}_1^{(o)}-\mathfrak{f}_n^{(o)})}\sum_{j=2}^{\nu+1}\frac{1}{k^{j-2}}V^{(o)}(\mathbf{e}^{(j)}) \quad (77)$$

with $k \in (1,+\infty)$ introduced in (26). The function $W^{(s,o)}$ is continuously differentiable, positive definite and proper. On account of the condition (26) on the numbers $a, b^{(j)}$, $j=2,\dots,\nu+1$, and lemmas 4.2, 4.4 and 4.6, denoting for simplicity $V^{(s)}(\mathbf{x}_t^{(2)})$ with $\mathbf{V}_t^{(s)}$, $W^{(s,o)}(\mathbf{x}_t^{(2)},\mathbf{e}_t^{\otimes})$ with $\mathbf{W}_t^{(s,o)}$ and $W^{(o)}(\mathbf{e}_t^{\otimes})$ with $\mathbf{W}_t^{(o)}$, after straightforward but lengthy calculations we obtain for $t \geqslant \delta$

$$\begin{aligned}
&\dot{\mathbf{W}}^{(s,o)}|_{(33),(55)}\\
&\leqslant -(1-a)\epsilon^{2\mathfrak{f}_1^{(s)}} - a\frac{\epsilon^{2\mathfrak{f}_n^{(o)}}}{\omega}\frac{\mathbf{W}_t^{(o)}}{1+\mathbf{W}_t^{(o)}}\\
&+4n\max_{j=1,\dots,n}\{\Gamma_{j,j}^{(s)}\}\,\epsilon^{2\mathfrak{f}_n^{(s)}}\min\left\{{l^{(s)}}^2, \|\Sigma^{(s)}\Sigma_{inv}^{(o)}\|^2\mathbf{W}_t^{(o)}\right\}\\
&+\frac{2r}{(1+\mathbf{W}_t^{(o)})}\frac{\epsilon^{6\mathfrak{f}_1^{(o)}}\delta}{\omega}\int_{t-\delta}^{t}((1+k)\mathbf{W}_\theta^{(o)}+(\nu-1)\mathbf{W}_{\theta-\mathbf{s}_\theta^{(j)}}^{(o)})d\theta\\
&+\gamma_8\left(\mathbf{x}_t, {S^{(s)}}^{-1}\sigma_{\lambda^{(s)}}(S^{(s)}\mathbf{x}_t)\right)\\
&+\sum_{j=2}^{\nu+1}\gamma_9^{(j)}\left(\mathbf{x}_{t-c-p^{(j-1)}}^{(2)}, \sigma_{\lambda^{(o)}}(\mathbf{x}_{t-c-p^{(j-1)}}^{(2)})\right), \quad (78)
\end{aligned}$$

where γ_8 and $\gamma_9^{(j)}$, $j=2,\dots,\nu+1$, are suitable functions such that $\gamma_8(\mathbf{v},\mathbf{v}) = \gamma_9^{(j)}(\mathbf{v},\mathbf{v}) = 0$ for all $\mathbf{v}\in\mathbb{R}^n$ (we used the fact that $\mathbf{x}_t^{(j)} = \mathbf{x}_{t-c-p^{(j-1)}}^{(2)}$). If we show that for all $t \geqslant \delta$ such that $\mathbf{W}_t^{(s,o)} \leqslant {l^{(s)}}^2$ we have

$$\begin{aligned}\mathbf{W}_{t+\theta}^{(s,o)} \leqslant \mathbf{W}_t^{(s,o)}, \forall\theta\in[-c-d_\infty,0]\\
\Rightarrow \dot{\mathbf{W}}_t^{(s,o)}|_{(33),(55)} \leqslant 0, \quad (79)\end{aligned}$$

by a Razumichin-type argument we prove the following boundedness result on the trajectories of (33), (55):

$$\begin{aligned}\mathbf{W}_\theta^{(s,o)} \leqslant {l^{(s)}}^2, \forall\theta\in[-c-d_\infty,\delta]\\
\Rightarrow \mathbf{W}_t^{(s,o)} \leqslant {l^{(s)}}^2\ \forall t\geqslant\delta \quad (80)\end{aligned}$$

i.e. the trajectories of (33), (55) are contained for all $t \geqslant \delta$ in some compact set $\mathcal{D}$ as long as they are contained in $\mathcal{D}$ for all $t \in [-c-d_\infty,\delta]$ or which is the same $\mathcal{D}$ is invariant for (33), (55) after $t=\delta$. To this aim, first of all, notice that if

$$\mathbf{x}\in\mathbb{R}^n : V^{(s)}(\mathbf{x}) \leqslant {l^{(s)}}^2 \quad (81)$$

then by definition of $V^{(s)}$

$$-l^{(s)}\epsilon^{\mathfrak{r}} \preceq S^{(s)}\mathbf{x} \preceq l^{(s)}\epsilon^{\mathfrak{r}} \quad (82)$$

and, also, by the incremental property (iv) given in lemma A.1 and (27)

$$\begin{aligned}
\|\epsilon^{-\mathfrak{r}}\diamond\mathbf{x}\|^2 &= \left\|\epsilon^{-\mathfrak{r}}\diamond\left\langle {S^{(s)}}^{-1}(S^{(s)}\mathbf{x})\right\rangle\right\|^2 \quad (83)\\
&\leqslant \left\|\Sigma_{inv}^{(s)}\left\langle\epsilon^{-\mathfrak{r}}\diamond S^{(s)}\mathbf{x}\right\rangle\right\|^2\\
&\leqslant \|\Sigma_{inv}^{(s)}\|^2 V^{(s)}(\mathbf{x}) \leqslant \|\Sigma_{inv}^{(s)}\|^2 {l^{(s)}}^2 = {l^{(o)}}^2,
\end{aligned}$$

which implies

$$-l^{(o)}\epsilon^{\mathfrak{r}} \preceq \mathbf{x} \preceq l^{(o)}\epsilon^{\mathfrak{r}}. \quad (84)$$

We also notice that

$$\begin{aligned}
\mathbf{W}_{t+\theta}^{(s,o)} \leqslant \mathbf{W}_t^{(s,o)} \leqslant {l^{(s)}}^2, \forall\theta\in[-c-d_\infty,0] \quad (85)\\
\Rightarrow \mathbf{W}_{t+\theta}^{(o)} \leqslant \omega e^{{l^{(s)}}^2\omega}\mathbf{V}_t^{(s)} + e^{{l^{(s)}}^2\omega}\mathbf{W}_t^{(o)},\\
\forall\theta\in[-c-d_\infty,0].
\end{aligned}$$

With all this in mind, let $t \geqslant \delta$ be such that $\mathbf{W}_{t+\theta}^{(s,o)} \leqslant \mathbf{W}_t^{(s,o)} \leqslant {l^{(s)}}^2$ for $\theta\in[-c-d_\infty,0]$. From the definition of $W^{(s,o)}$ in (76) we have $\mathbf{V}_{t+\theta}^{(s)} \leqslant {l^{(s)}}^2$ for $\theta\in[-c-d_\infty,0]$ and since by definition $0 \leqslant c+p^{(j-1)} \leqslant c+d_\infty$ for all $j=2,\dots,\nu+1$, it follows that

$$\|\epsilon^{-\mathfrak{r}}\diamond S^{(s)}\mathbf{x}_{t-c-p^{(j-1)}}^{(2)}\|^2 = \mathbf{V}_{t-c-p^{(j-1)}}^{(s)} \leqslant {l^{(s)}}^2$$

for each $j=2,\dots,\nu+1$. On account of the conclusions in (82) and (84)

$$\begin{aligned}
\gamma_8\left(\mathbf{x}_t, {S^{(s)}}^{-1}\sigma_{\lambda^{(s)}}(S^{(s)}\mathbf{x}_t)\right) = 0,\\
\gamma_9^{(j)}\left(\mathbf{x}_{t-c-p^{(j-1)}}^{(2)}, \sigma_{\lambda^{(o)}}(\mathbf{x}_{t-c-p^{(j-1)}}^{(2)})\right) = 0 \quad (86)
\end{aligned}$$

for each $j=2,\dots,\nu+1$. From (78) using (32) and (85), we get (by re-introducing the argument (ϵ) in ω,δ and ν)

$$\begin{aligned}
&\mathbf{W}_{t+\theta}^{(s,o)} \leqslant \mathbf{W}_t^{(s,o)}, \forall\theta\in[-c-d_\infty,0]\\
&\Rightarrow \dot{\mathbf{W}}_t^{(s,o)}|_{(33),(55)} \leqslant -\epsilon^{2\mathfrak{f}_1^{(s)}}\Big[\rho(\epsilon)\mathbf{V}_t^{(s)}+\pi(\epsilon)\frac{\mathbf{W}_t^{(o)}}{1+\mathbf{W}_t^{(o)}}\\
&\qquad -\chi(\epsilon)\min\{{l^{(s)}}^2, d\mathbf{W}_t^{(o)}\}\Big]
\end{aligned}$$

with $d := \|\Sigma^{(s)}\Sigma_{inv}^{(o)}\|^2$ and

$$\begin{aligned}
\rho(\epsilon) &:= 1-a-2r\epsilon^{6\mathfrak{f}_1^{(o)}-2\mathfrak{f}_1^{(s)}}\delta^2(\epsilon)(k+\nu(\epsilon))e^{{l^{(s)}}^2\omega(\epsilon)},\\
\pi(\epsilon) &:= \frac{\epsilon^{2(\mathfrak{f}_n^{(o)}-\mathfrak{f}_1^{(s)})}}{\omega(\epsilon)}\left[a-2r\epsilon^{6\mathfrak{f}_1^{(o)}-2\mathfrak{f}_n^{(o)}}\delta^2(\epsilon)(k+\nu(\epsilon))e^{{l^{(s)}}^2\omega(\epsilon)}\right],\\
\chi(\epsilon) &:= 4n\max_{j=1,\dots,n}\{\Gamma_{j,j}^{(s)}\}\,\epsilon^{2(\mathfrak{f}_n^{(s)}-\mathfrak{f}_1^{(s)})}.
\end{aligned}$$

Consider the function $\xi : \mathbb{R}_\geqslant\times\mathbb{R}_\geqslant\to\mathbb{R}$

$$\xi(\mathbf{v},\mathbf{w}) := \rho(\epsilon)\mathbf{v} + \pi(\epsilon)\frac{\mathbf{w}}{1+\mathbf{w}} - \chi(\epsilon)\min\{{l^{(s)}}^2, d\mathbf{w}\}. \quad (87)$$

As a consequence of definition 4.1

$$\begin{aligned}
&\lim_{\epsilon\to+\infty}\delta(\epsilon)(k+\nu(\epsilon))\\
&= \lim_{\epsilon\to+\infty}\delta(\epsilon)\left(k+\lceil\frac{c+d_\infty}{\delta(\epsilon)}\rceil+1\right) < c+d_\infty+1
\end{aligned}$$

and on account of (H3) and the asymptotic conditions (30), (28) and (29),

$$\lim_{\epsilon\to+\infty}\rho(\epsilon)=1-a>0,\ \lim_{\epsilon\to+\infty}\frac{\varphi(\epsilon)}{\chi(\epsilon)}=+\infty,$$

and, moreover, $\chi(\epsilon)>0$ for all ϵ. so that all the conditions of lemma A.5 are met. By this lemma there exists $\epsilon_\infty>1$ such that for all $\epsilon\geqslant\epsilon_\infty$: $\xi(\mathbf{v},\mathbf{w})\geqslant 0$ for all $\mathbf{v},\mathbf{w}\in\mathbb{R}_\geqslant$. Therefore, for any $\epsilon\geqslant\epsilon_\infty$ we recover (79) and, therefore, (80). Under this regard we are left to guarantee that for all $\epsilon\geqslant\epsilon_\infty$ (with possibly larger ϵ_∞)

$$\mathbf{W}_\theta^{(s,o)}\leqslant {l^{(s)}}^2,\forall\theta\in[-c-d_\infty,\delta] \tag{88}$$

i.e. the trajectories of (33), (55) are contained for all $t\in[-c-d_\infty,\delta]$ in some compact set $\mathcal{D}$. We do this in four steps.

(I) From (16) of (H0) and since $\mathbf{u}_{t-c}^{(\nu+1)}\equiv 0$ for $t\leqslant c+d_\infty$ and $\mathbf{x}_{-c-2d_\infty}\in\mathcal{C}$ (as stated in theorem 4.1), we have with $L(\mathbf{v}):=\ln(1+U(\mathbf{v}))$ and for all $\theta\in[-c-d_\infty,\delta]$

$$L(\mathbf{x}_\theta^{(\nu+1)})\ \leqslant\ L(\mathbf{x}_{-c-2d_\infty})+\delta^*\mu\leqslant\max_{\mathbf{x}\in\mathcal{C}}L(\mathbf{x})+\delta^*\mu, \tag{89}$$

where $\delta^*\in\mathcal{L}$ is defined as $\delta^*:=\delta+2d_\infty+c$. Analogously, for each $j=2,\dots,\nu$ and for all $\theta\in[-c-d_\infty,\delta]$

$$L(\mathbf{x}_\theta^{(j)})\ \leqslant\ L(\mathbf{x}_{-c-p^{(j-1)}}^{(\nu+1)})+\delta^*\mu+\delta\max_{t\in[0,\delta]}\kappa(\|\mathbf{u}_t\|^2) \tag{90}$$

where we used the fact that $\mathbf{x}_{-c-p^{(j-1)}}^{(\nu+1)}=\mathbf{x}_{-c-d_\infty}^{(j)}$. Moreover, on account of (38) and the definition of $\mathbf{u}_t$ in (14)

$$\sup_{t\in[0,+\infty)}\|\mathbf{u}_t\|^2\leqslant l^{(s)}\|B^T\Gamma^{(s)}BB^T\epsilon^{\mathfrak{r}+2\mathfrak{f}}\|. \tag{91}$$

Using (31), (89)-(91) with the fact that $-c-d_\infty\leqslant -c-p^{(j-1)}\leqslant 0$ for all $j=2,\dots,\nu+1$, we can increase (if necessary) ϵ_∞ such that for all $\epsilon\geqslant\epsilon_\infty$ and for all $\theta\in[-c-d_\infty,\delta]$

$$\begin{aligned}&L(\mathbf{x}_\theta^{(\nu+1)})\leqslant\max_{\mathbf{x}\in\mathcal{C}}L(\mathbf{x})+(2d_\infty+c+1)\mu:=a_\infty,\\&L(\mathbf{x}_\theta^{(j)})\leqslant 2a_\infty,\ j=2,\dots,\nu.\end{aligned} \tag{92}$$

(II) Since on account of (iv) of lemma A.2

$$V^{(s)}(\mathbf{x})=\|\epsilon^{-\mathfrak{r}}\diamond S^{(s)}\mathbf{x}\|^2\leqslant\|\Sigma_{inv}^{(s)}\|^2\|\epsilon^{-\mathfrak{r}}\diamond\mathbf{x}\|^2,\forall\mathbf{x},$$

using (92) together with the properness of L, we can always (if necessary) increase ϵ_∞ such that for all $\epsilon\geqslant\epsilon_\infty$ and for all $\theta\in[-c-d_\infty,\delta]$

$$V^{(s)}(\mathbf{x}_\theta^{(j)})\leqslant\frac{{l^{(s)}}^2}{2},\ j=2,\dots,\nu+1 \tag{93}$$

and therefore, by the conclusions in (84), for all $\theta\in[-c-d_\infty,\delta]$

$$-l^{(o)}\epsilon^{\mathfrak{r}}\preceq\mathbf{x}_\theta^{(j)}\preceq l^{(o)}\epsilon^{\mathfrak{r}},\ j=2,\dots,\nu+1. \tag{94}$$

(III) On account of (iv) of lemma A.1 and the partial ordering (24)

$$V^{(o)}(\mathbf{x})=\|\epsilon^{-\mathfrak{r}}\diamond S^{(o)}\mathbf{x}\|^2\leqslant\epsilon^{2(\mathfrak{f}_1^{(o)}-\mathfrak{f}_n^{(o)})}\|\Sigma^{(o)}\|^2\|\epsilon^{-\mathfrak{r}}\diamond\mathbf{x}\|^2,\forall\mathbf{x},$$

and since $\hat{\mathbf{x}}_\theta^{(j)}\equiv 0$ for $\theta\in[-c-2d_\infty,0]$ (by the initialization (13)), we have

$$V^{(o)}(\mathbf{e}_\theta^{(j)})=V^{(o)}(\mathbf{x}_\theta^{(j)}),\ j=2,\dots,\nu+1, \tag{95}$$

for all $\theta\in[-c-d_\infty,0]$. Recalling the definition (77) of $W^{(o)}$ with $\sum_{j=2}^{+\infty}\frac{1}{k^j}\leqslant\frac{k}{k-1}$ (since $k>1$) and making use of the asymptotic properties (28) and (29) with (92), we can increase (if necessary) ϵ_∞ such that for all $\epsilon\geqslant\epsilon_\infty$ and for all $\theta\in[-c-d_\infty,0]$:

$$\begin{aligned}&\frac{1}{\omega}\ln\left(1+W^{(o)}(\mathbf{e}_\theta^{\otimes})\right)\\&\leqslant\frac{1}{\omega}\ln\left(1+\frac{k}{k-1}\epsilon^{2(\mathfrak{f}_1^{(o)}-\mathfrak{f}_n^{(o)})}\max_{\mathbf{x}:L(\mathbf{x})\leqslant 2a_\infty}V^{(o)}(\mathbf{x})\right)\leqslant\frac{{l^{(s)}}^2}{2}.\end{aligned} \tag{96}$$

(IV) Eventually, it is possible to increase (if necessary) ϵ_∞ such that (96) holds for all $\epsilon\geqslant\epsilon_\infty$ and also for all $\theta\in[0,\delta]$. Indeed, by integration of (63) over $[0,\delta]$ for $j=2,\dots,\nu+1$ with the definition of $\mathbf{d}_t$ in (57) and (59), since the function γ_3 has null contribution by virtue of (94) and moreover $\mathbf{q}_t^{(\nu+1)}\equiv 0$, we get for all $\theta\in[0,\delta]$

$$V^{(o)}(\mathbf{e}_\theta^{(\nu+1)})\leqslant V^{(o)}(\mathbf{e}_0^{(\nu+1)}), \tag{97}$$

and, using additionally commutative (122) and associative (123) properties with the partial ordering (24), we get for all $\theta\in[0,\delta]$ and $j=2,\dots,\nu$

$$\begin{aligned}V^{(o)}(\mathbf{e}_\theta^{(j)})&\leqslant V^{(o)}(\mathbf{e}_0^{(j)})+16\epsilon^{2\mathfrak{f}_1^{(o)}}\max_{j=1,\dots,n}\{\Gamma_{j,j}^{(o)}\}\times\\&\times\int_0^\delta(V^{(o)}(\mathbf{e}_t^{(j+1)})+V^{(o)}(\mathbf{e}_t^{(j)})+V^{(o)}(\mathbf{e}_{t-\mathbf{s}_t^{(j)}}^{(j)}))dt.\end{aligned} \tag{98}$$

Upon noticing that for $\theta\in[-c-d_\infty,0]$: $V^{(o)}(\mathbf{e}_\theta^{(j)})=V^{(o)}(\mathbf{x}_\theta^{(j)})\leqslant\frac{{l^{(s)}}^2}{2}$ by (93) and the initialization (13), the inequalities (97) and (98) give place to

$$\tau_\infty^{(\nu+1)}\ \leqslant\ \frac{{l^{(s)}}^2}{2}, \tag{99}$$

$$\begin{aligned}\tau_\infty^{(j)}\ &\leqslant\ \frac{{l^{(s)}}^2}{2}+16\epsilon^{2\mathfrak{f}_1^{(o)}}\max_{j=1,\dots,n}\{\Gamma_{j,j}^{(o)}\}\times\\&\times\delta(\tau_\infty^{(j+1)}+2\tau_\infty^{(j)}+{l^{(s)}}^2),\ j=2,\dots,\nu,\end{aligned} \tag{100}$$

where $\tau_\infty^{(j)}:=\max_{\theta\in[0,\delta]}V^{(o)}(\mathbf{e}_\theta^{(j)})$. Using the asymptotic properties (29), (30) of δ, we can assume ϵ_∞ (increased if necessary) such that

$$32\epsilon^{2\mathfrak{f}_1^{(o)}}\delta(\epsilon)\max_{j=1,\dots,n}\{\Gamma_{j,j}^{(o)}\}\leqslant q$$

for all $\epsilon\geqslant\epsilon_\infty$ with $q\in(0,1)$ such that

$$p:=\frac{1}{2}\frac{1+q}{1-q}<1. \tag{101}$$

Eventually from (100) we obtain $\tau_\infty^{(j)}\leqslant p({l^{(s)}}^2+\tau_\infty^{(j+1)})$ for $j=2,\dots,\nu$. These recursive inequalities with (99) can be solved backwards starting from (99) to give with (101)

$$\tau_\infty^{(j)}:=\max_{\theta\in[0,\delta]}V^{(o)}(\mathbf{e}_\theta^{(j)})\leqslant{l^{(s)}}^2\left(2+\frac{1}{1-p}\right) \tag{102}$$

for all $j=2,\dots,\nu+1$ and $\epsilon\geqslant\epsilon_\infty$. These inequalities are used like in step **(III)** to meet (96) for all $\epsilon\geqslant\epsilon_\infty$ (with increased ϵ_∞ if necessary) and for all $\theta\in[0,\delta]$. The steps **(I)**-**(IV)** prove (88) and, as a consequence of (80),

$$\mathbf{W}_t^{(s,o)}\leqslant{l^{(s)}}^2,\ \forall t\geqslant 0. \tag{103}$$

In particular, by definition of $W^{(s,o)}$ we have

$$\mathbf{V}_t^{(s)} \leqslant {l^{(s)}}^2,\ \forall t \geqslant 0. \tag{104}$$

State and estimates asymptotic convergence analysis. Once the boundedness condition (103) has been obtained, it is easy to prove asymptotic convergence to zero of $\mathbf{x}_t$ (and all the errors $\mathbf{e}_t^{(j)}$). For later use, notice that (103) implies that $\mathbf{x}_t^{(2)}$ and $\mathbf{e}_t^{(j)}$, $j = 2, \ldots, \nu+1$ are bounded for all $t \geqslant 0$. We begin with (63) for $j = \nu + 1$. Recalling that the function γ_3 in (63) has null contribution on account of (104) and the conclusions in (84) form (81) and moreover, $\mathbf{q}_t^{(\nu+1)} \equiv 0$, we obtain $\lim_{t\to+\infty} V^{(o)}(\mathbf{e}_t^{(\nu+1)}) = 0$. By induction assume $\lim_{t\to+\infty} V^{(o)}(\mathbf{e}_t^{(j+1)}) = 0$ for some $j = 2, \ldots, \nu$. Using (63) with (73) and recalling that the functions γ_3 in (63) and, respectively, γ_5 in (73) have null contribution on account of (104) and the conclusions in (82), on application of the Razumichin-type theorem 1 of [38] (in particular formula (32)) with exogenous input $V^{(o)}(\mathbf{e}_t^{(j+1)})$, we obtain the existence of $\beta^{(o)} \in \mathcal{KL}$ and $\rho^{(o)} \in \mathcal{KL}$ such that $V^{(o)}(\mathbf{e}_t^{(j)}) \leqslant \beta^{(o)}(V^{(o)}(\mathbf{e}_0^{(j)}), t) + \rho^{(o)}(\sup_{\theta\in[0,t]} V^{(o)}(\mathbf{e}_\theta^{(j+1)}))$ for all $t \geqslant 0$. This, upon the induction hypothesis on $V^{(o)}(\mathbf{e}_t^{(j+1)})$, implies $\lim_{t\to+\infty} V^{(o)}(\mathbf{e}_t^{(j)}) = 0$. It follows by induction that $\lim_{t\to+\infty} V^{(o)}(\mathbf{e}_t^{(j)}) = 0$ for all $j = 2, \ldots, \nu+1$. Finally, consider (37) with (49). Recalling that the functions γ_1 in (37) and γ_2 in (49) have null contribution on account of (104) and the conclusions in (82) form (81), we obtain the existence of $\beta^{(s)} \in \mathcal{KL}$ and $\rho^{(s)} \in \mathcal{KL}$ such that $V^{(s)}(\mathbf{x}_t^{(2)}) \leqslant \beta^{(s)}(V^{(s)}(\mathbf{x}_0^{(2)}), t) + \rho^{(s)}(\sup_{\theta\in[0,t]} V^{(o)}(\mathbf{e}_\theta^{(2)}))$ for all $t \geqslant 0$. This, with $\lim_{t\to+\infty} V^{(o)}(\mathbf{e}_t^{(2)}) = 0$, implies $\lim_{t\to+\infty} V^{(s)}(\mathbf{x}_t^{(2)}) = 0$ and, therefore, $\lim_{t\to+\infty} \mathbf{x}_t^{(2)} = \lim_{t\to+\infty} \mathbf{x}_t = 0$. Convergence and boundedness are uniform by the same theorem 1 of [38].

E. Extensions: output nonlinearities, time-varying input delays, multiple delays and robustness

1) Output nonlinearities: Theorem 4.1 can be extended by including output nonlinearities in our model (4) as $\mathbf{y}_t := C\mathbf{x}_{t-\mathbf{d}_t} + \psi(\mathbf{x}_{t-\mathbf{d}_t})$, smooth ψ, and at the same time adding incremental homogeneity assumptions on ψ in (H2) as follows: (H2b) $C^T\psi$ *is incrementally homogeneous in the upper bound with quadruples* $(\mathfrak{r}, \mathfrak{r} - \mathfrak{f}^{(o)}, \mathfrak{f}^{(o)}, C^T\Psi^{(o)}(x', x''))$, *with* ${\Psi^{(o)}}^T(0,0)\Psi^{(o)}(0,0) < C^TC$.

The condition ${\Psi^{(o)}}^T(0,0)\Psi^{(o)}(0,0) < C^TC$ is a sector-condition on the linear approximation of ψ. For instance, a saturated output $\mathbf{y}_t := \sigma_l(C\mathbf{x}_{t-\mathbf{d}_t})$ meets (H2b). The definition of the innovation $\mathbf{z}_t^{(j)}$ in each observer (7) must be changed by replacing $C\widehat{\mathbf{x}}_t^{(j)}$ with $C\widehat{\mathbf{x}}_t^{(j)} + \psi\Big(\sigma_{\lambda^{(o)}(\epsilon)}\Big(\widehat{\mathbf{x}}_t^{(j)}\Big)\Big)$.

2) Time-varying input delays: Theorem 4.1 can be also extended to the case of time-varying delays. In this case, we assume that $\mathbf{c}_t$, the input delay, is continuous and bounded by some known c_∞ and the functions $\{\mathbf{d}_t, \mathbf{c}_t\}$ known up to time t. The partition$\{p^{(j)}\}_{j=1,\ldots,\nu}$ for defining the observer chain is applied on the interval $[-c_\infty, d_\infty]$. Theorem 4.1 remains true (but the proof is a slightly more lengthy and tedious) by replacing c with c_∞ and changing the control law (14) as follows:

$$\begin{aligned}\mathbf{u}_t \ &:= \ -R^{(s)}B^TP^{(s)}(I_n - A^TG^{(s)}) \times \\ &\quad\ \times \sigma_{\lambda^{(s)}(\epsilon)}\Big((I_n - A^TG^{(s)})^{-1}\widehat{\mathbf{x}}_t^{(j+1)}\Big) \\ &\quad\ \text{if } -\mathbf{c}_t \in [p^{(j)}, p^{(j+1)}).\end{aligned}$$

Notice that the control $\mathbf{u}_t$ changes according to the relative position of $-\mathbf{c}_t$ with respect to the partition of $[-c_\infty, 0]$.

3) Multiple input and measurement delays: More realistically, for our MIMO system (4), (5) we may consider multiple input and measurement delays as follows. The input vector $\mathbf{u}_{t-c}$ is replaced by the vector $(\mathbf{u}_{1,t-c_1}, \cdots, \mathbf{u}_{m,t-c_m})^T$, for multiple delays $c_1, \cdots, c_m$, and the measurement vector $\mathbf{y}_t$ is replaced by the vector $(C_1\mathbf{x}_{t-\mathbf{d}_{1,t}}, \cdots, C_p\mathbf{x}_{t-\mathbf{d}_{p,t}})^T$, for multiple delays $\mathbf{d}_{1,t}, \cdots, \mathbf{d}_{p,t}$. Theorem 4.1 can be extended by including multiple delays by simply re-defining each controller component $\mathbf{u}_{j,t}$, $j = 1, \ldots, m$, as in (14) by using the delay c_j and, similarly, each innovation component $\mathbf{z}_{j,t}$, $j = 1, \ldots, p$, in the observer chain as in (9) by using the delay $\mathbf{d}_{j,t}$.

4) Robustness w.r.t. uncertainties and disturbances: The controller (7), (14), in particular the chain of observers (7), relies on the perfect knowledge of the nonlinear function $\phi(\cdot)$, This may cause a lack of robustness. If we adopt a more general model (inclusive of uncertainties and disturbances)

$$\begin{aligned}\dot{\mathbf{x}}_t \ &= \ A\mathbf{x}_t + B\mathbf{u}_{t-c} + \phi(\mathbf{x}_t, \omega_t),\ t \geqslant -c - 2d_\infty, \\ \mathbf{y}_t \ &= \ C\mathbf{x}_{t-\mathbf{d}_t} + D\omega_{t-\mathbf{d}_t},\ t \geqslant 0\end{aligned} \tag{105}$$

where ω_t is a time-varying disturbance/uncertanty (norm-bounded by ω_∞), it is possible to robustify the controller (7), (14) as follows: while (14) remains the same, (7) is modified by replacing $\phi(\sigma_{\lambda^{(o)}(\epsilon)}(\widehat{\mathbf{x}}_t^{(j)}))$ with $\phi(\sigma_{\lambda^{(o)}(\epsilon)}(\widehat{\mathbf{x}}_t^{(j)}), 0)$. By slightly strengthening the incremental homogeneity assumptions (H1)-(H2) in such a way to include the effect of the variable ω on $\phi(\mathbf{x}, \omega)$ it is possible to prove a disturbance-to-state stability result for the closed-loop system resulting from (105).

F. Example and simulations

For testing our stabilizer we consider the system

$$\begin{aligned}&\dot{\mathbf{x}}_{1,t} = \mathbf{x}_{2,t} \\ &\dot{\mathbf{x}}_{2,t} = -\mathbf{x}_{1,t} + (1 - \mathbf{x}_{1,t}^2)\mathbf{x}_{2,t} + \mathbf{u}_{t-1},\ \mathbf{y}_t = \mathbf{x}_{1,t-\mathbf{d}_t}\end{aligned} \tag{106}$$

The measurements are taken over intervals of the form $[1.1h, 1.1h+1]$ for $h = 0, 1, \ldots$ and are supplied at a high rate during the subsequent time interval $[1.1h+1, 1.1(h+1)]$. Correspondingly, the measurement delay profile is $\mathbf{d}_t$ as follows: $\mathbf{d}_t = t - 1.1h$ if $t \in [1.1h, 1.1h+1]$ and $\mathbf{d}_t = 1 - 10(t - 1.1h - 1)$ if $t \in [1.1h+1, 1.1(h+1)]$, $h = 0, 1, \ldots$, and it is bounded by $d_\infty = 1$. Moreover, the input delay is $c = 1$. System (106) satisfies assumptions (H0)-(H3) of theorem 4.1 with $\mathfrak{r}_1 = 1/8$, $\mathfrak{r}_2 = 3/8$, $\mathfrak{f}_1^{(s)} = \mathfrak{f}_2^{(s)} = 1/8$, $\mathfrak{f}_1^{(o)} = 1/2$ and $\mathfrak{f}_2^{(o)} = 1/4$. A stabilizer has been designed according to our procedure and a simulation has been worked out with initial conditions $\mathbf{x}_{-c-2d_\infty} = (-5, -4)^T$. With such state initial conditions (a square initialization region $\mathcal{C}$

with side 10 has been guaranteed) an observer chain with $\nu = 11$ is sufficient for our aims. The interval $[-1, 1]$ has been partitioned into 10 subintervals with equal length 0.2 and points $p_j = -1+0.2(j-1)$, $j = 1, \ldots, 12$ (with the extra point $p_{12} := 1.2$). The saturation levels of the estimates are set with $l^{(s)} = 0.05$ and $l^{(o)} = 0.1$, the diagonal elements of $\Gamma^{(s)}$ are respectively 1 and 10 (see lemma A.4), the diagonal elements of $\Gamma^{(o)}$ are respectively 10 and 1 (see lemma A.3). The closed-loop state trajectories $\mathbf{x}_t$ together with the prediction errors $\mathbf{e}_t^{(2)}$ are shown versus time in Fig. 1.

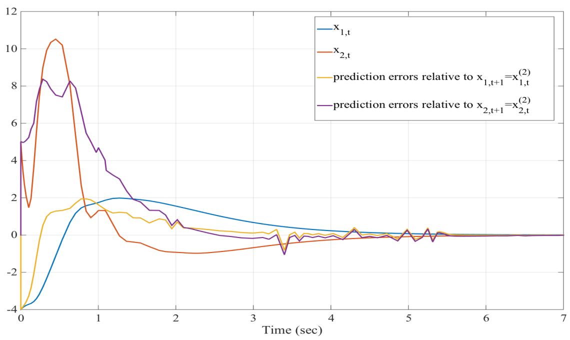

Fig. 1. Closed-loop state trajectories $\mathbf{x}_t = (\mathbf{x}_{1,t}, \mathbf{x}_{2,t})^T$ and prediction errors relative to $\mathbf{x}_{t+1} = (\mathbf{x}_{1,t+1}, \mathbf{x}_{2,t+2})^T$.

V. Sampled-data stabilizers

The design of continuous-time stabilizer for (4), (5) given in the previous section suggests naturally the way of designing a sampled-data stabilizer for (4), (5). This will consists of a sampled-data controller and a chain of sampled-data observers with sampling period T. Sampled-data stabilizers can be naturally obtained from particular classes of continuous-time stabilizers as follows. Let

$$\begin{aligned} \mathbf{u}_t &= \alpha(\hat{\mathbf{x}}_{t_h}), \\ \dot{\hat{\mathbf{x}}}_t &= A\hat{\mathbf{x}}_t + \beta(\hat{\mathbf{x}}_{t_h}^{(t_0,\ldots,t_k)}, \mathbf{y}_{t_h}^{(t_0,\ldots,t_k)}),\ t \in [t_h, t_{h+1}), \end{aligned} \tag{107}$$

$h, k \in \mathbb{N}$, $k \leqslant h$, be a continuous-time stabilizer for (4), (5) with $t_h := hT$, locally Lipschitz continuous functions α, β and $\mathbf{v}_t^{(t_0,\ldots,t_k)} := (\mathbf{v}_t, \mathbf{v}_{t-t_0}, \cdots, \mathbf{v}_{t-t_k})$. A sampled-data stabilizer for (4), (5) is obtained as a zero-order hold discretization of (107) with $\mathbf{u}_t = \alpha(\hat{\mathbf{x}}_{t_h})$ for $t \in [t_h, t_{h+1})$ and

$$\hat{\mathbf{x}}_{t_{h+1}} = A_T\hat{\mathbf{x}}_{t_h} + B_T\beta(\hat{\mathbf{x}}_{t_h}^{(t_0,\ldots,t_k)}, \mathbf{y}_{t_h}^{(t_0,\ldots,t_k)}),\ h \in \mathbb{N}, \tag{108}$$

where $A_T = e^{AT}$ and $B_T = \int_0^T e^{As} ds$. The stability analysis (boundedness and asymptotic convergence) of (4), (5), (108) is carried out through the stability analysis of (4), (5), (107) (therefore, following the proof of theorem 4.1) since the trajectories of (108) and (107) coincide at the sampling times.

With this in mind, as a first step we design a continuous-time stabilizer for (4), (5) having the form (107). From this we obtain a sampled-data stabilizer for (4), (5) according to the zero-order hold discretization procedure pointed out in (108). The δ-fine partition $\{p^{(j)}\}_{j=\ldots,\nu}$ of the interval $[-c, d_\infty]$ is chosen so that each point $p^{(j)}$ (and therefore δ) is a multiple of the sampling time T. For this reason, exactly as δ in the proof of theorem 4.1, the sampling period T will depend on the parameter ϵ and, therefore, both on the magnitude of the delays and on the growth rate of the nonlinearities. Our continuous-time stabilizer consists of a (zero-hold in the period T) controller and a chain of ν continuous-time observers (switching after each period T). Let $P^{(j)}, R^{(j)}, G^{(j)}$, $j \in \{s, o\}$, be as in (8) and (15). The observer chain is described by:

$$\begin{aligned} \dot{\hat{\mathbf{x}}}_t^{(j)} &= A\hat{\mathbf{x}}_t^{(j)} + B\mathbf{u}_{t-c}^{(j)} \\ &+ \phi\left(\sigma_{\lambda^{(o)}(\epsilon)}\left(\hat{\mathbf{x}}_{t_h}^{(j)}\right)\right) + {P^{(o)}}^{-1}C^T R^{(o)}\mathbf{z}_{t_h}^{(j)}, \\ & j = 2, \ldots, \nu+1,\ t \in [t_h, t_{h+1}), \end{aligned} \tag{109}$$

with innovations $\mathbf{z}_t^{(j)}$ and delays $\mathbf{s}_t^{(j)}$ defined as in (9), (10) for $j = \nu_0+1, \ldots, \nu+1$, where now $\mathbf{y}_{t^{(j)}}$ is the past output at $t^{(j)} := \max\{t_k \in [0, t]\colon t_k - \mathbf{d}_{t_k} \leqslant t - p^{(j-1)}\}$, and in (11), (12) for $j = 2, \ldots, \nu_0$. Each observer is initialized as in (13). The controller is defined as

$$\begin{aligned} \mathbf{u}_t &:= -R^{(s)}B^T P^{(s)}(I_n - A^T G^{(s)}) \times \\ & \times\sigma_{\lambda^{(s)}(\epsilon)}\left((I_n - A^T G^{(s)})^{-1}\hat{\mathbf{x}}_{t_h}^{(2)}\right),\ t \in [t_h, t_{h+1}). \end{aligned} \tag{110}$$

It is easy to check that (109), (110) has the form (107). The main result of this section is the following and it is proved along the lines of the proof of theorem 4.1.

Theorem *5.1: Let $\mathcal{C} \subset \mathbb{R}^n$ be a given compact set. Under assumptions* (H0)-(H3) *there exist diagonal positive definite $\Gamma^{(j)} \in \mathbb{R}^{n\times n}$, $l^{(j)} \in \mathbb{R}_>$, $j \in \{s, o\}$, $\epsilon, \delta, T \in \mathbb{R}_>$ and a δ-fine partition $\{p^{(j)}\}_{j=-1,\ldots,\nu}$ of $[-c, d_\infty]$, extended and centered at* 0*, such that the solutions $(\mathbf{x}_t, \hat{\mathbf{x}}_t^{(j)})$, $j = 2\ldots, \nu+1$, of (4), (5), (109), (110), with $\mathbf{x}_{-c-2d_\infty} \in \mathcal{C}$, are bounded for all $t \geqslant -c-2d_\infty$ and $\lim_{t\to+\infty}\mathbf{x}_t = 0$.*

The continuous-time controller (109), (110) has the form (107) and *semi-globally* asymptotically stabilizes (4). Also in this case boundedness and convergence results are uniform (in the sense of $\mathcal{KL}$ functions). The sampled-data stabilizer, obtained from a zero-order hold discretization of the continuous-time (109), (110) as pointed out in (108), semi-globally asymptotically stabilizes (4), since the trajectories of (109), (110) and its sampled-data counterpart coincide at the sampling times.

The problem can be studied in the framework of non-uniform sampling and the sampling period T may be variable. The only additional hypothesis to be taken into account is a positive lower bound for T (no Zeno phenomena) while the upper bound for T is determined as inas in theorem 5.1.

Appendix

The notion of (incremental) homogeneity in a generalized sense has been introduced in [4] in the context of (semi-)global stabilization and observer design problems. Here we recall this notion in a slightly more general form. Let

$$(\Delta\phi)(x', x'') := \phi(x') - \phi(x'') \tag{111}$$

and if ϕ is the identity function we simply write $\Delta(x', x'') := x' - x''$.

Definition *A.1: A parametric function $\phi(\epsilon) \in \mathbf{C}^0(\mathbb{R}^n, \mathbb{R}^l)$, $\epsilon \in \mathbb{R}_>$, is said to be incrementally homogeneous (in the* *generalized sense: g.i.h.) with quadruple* $(\mathfrak{r}, \mathfrak{d}, \mathfrak{h}, \Phi(x', x''))$ *if there exist* $\mathfrak{d} \in \mathbb{R}^l, \mathfrak{h} \in \mathbb{R}^n, \mathfrak{r} \in \mathbb{R}^n_>$ *and* $\Phi \in \mathbf{C}^0(\mathbb{R}^n \times \mathbb{R}^n, \mathbb{R}^{l\times n})$ *such that for all* $\epsilon \in \mathbb{R}_>$ *and* $x', x'' \in \mathbb{R}^n$

$$(\Delta\phi(\epsilon))(\epsilon^{\mathfrak{r}} \diamond x', \epsilon^{\mathfrak{r}} \diamond x'') = \epsilon^{\mathfrak{d}} \diamond \left(\Phi(x', x'')\Delta(\epsilon^{\mathfrak{h}} \diamond x', \epsilon^{\mathfrak{h}} \diamond x'')\right)$$

When the variation Δ of $\phi(\epsilon)$ is computed in between the dilated points $x' := x \in \mathbb{R}^n$ and $x'' := 0$, with $\phi(\epsilon)(0) = 0$, we say $\phi(\epsilon)$ *is homogeneous (in the generalized sense: g.h.) with quadruple* $(\mathfrak{r}, \mathfrak{d}, \mathfrak{h}, \Phi'(x))$ with $\Phi'(x) := \Phi(x, 0)$.

Example *A.1:* The parametric function $\phi(\epsilon) : x \in \mathbb{R}^2 \mapsto (\epsilon x_1^3 - \epsilon^2 x_2^3, \epsilon x_1 + \epsilon^2 x_2)^T$ is g.i.h. with quadruple $(\mathbf{1}_2, (1, -1)^T, (3, 4)^T, \Phi(x', x''))$ where

$$\Phi(x', x'') := \begin{pmatrix} (x_1')^2 + x_1' x_1'' + (x_1')^2 & (x_2')^2 + x_2' x_1'' + (x_2')^2 \\ 1 & 1 \end{pmatrix}.$$

It is also g.h. with quadruple $(\mathbf{1}_2, (1, -1)^T, (3, 4)^T, \Phi(x')$ where $\Phi(x) := \Phi(x, 0)$.

There are functions, like $\sin x$, which are not g.i.h. but behaves in the upper bound as an g.i.h. function. This motivates the following definition.

Definition *A.2: A parametric function* $\phi(\epsilon) \in \mathbf{C}^0(\mathbb{R}^n, \mathbb{R}^l)$, $\epsilon \in \mathbb{R}_>$*, is said to be incrementally homogeneous in the upper bound (in the generalized sense: g.i.h.u.b.) with quadruple* $(\mathfrak{r}, \mathfrak{d}, \mathfrak{h}, \Phi(x', x''))$ *if there exist* $\mathfrak{d} \in \mathbb{R}^l, \mathfrak{h} \in \mathbb{R}^n, \mathfrak{r} \in \mathbb{R}^n_>$*,* $\Phi \in \mathbf{C}^0(\mathbb{R}^n \times \mathbb{R}^n, \mathbb{R}^{l\times n}_{\geqslant})$ *such that for all* $\epsilon \in (1, +\infty)$ *and* $x', x'' \in \mathbb{R}^n$

$$\left\langle(\Delta\phi(\epsilon))(\epsilon^{\mathfrak{r}} \diamond x', \epsilon^{\mathfrak{r}} \diamond x'')\right\rangle \le \epsilon^{\mathfrak{d}} \diamond \left(\Phi(x', x'')\left\langle\Delta(\epsilon^{\mathfrak{h}} \diamond x', \epsilon^{\mathfrak{h}} \diamond x'')\right\rangle\right)$$

Notice that, in the case of g.i.h.u.b., expanding dilations (i.e. $\epsilon \in (1, +\infty)$) are considered. When the variation Δ of $\phi(\epsilon)$ is computed in between the dilated points $x' := x \in \mathbb{R}^n$ and $x'' := 0$, with $\phi(\epsilon)(0) = 0$, we say $\phi(\epsilon)$ *is homogeneous in the upper bound (in the generalized sense: g.h.u.b.) with quadruple* $(\mathfrak{r}, \mathfrak{d}, \mathfrak{h}, \Phi'(x))$ with $\Phi'(x) = \Phi(x, 0)$. Some properties of incremental homogeneity can be found in [4].

A. Properties of (C, A, B)

For each matrix W, let $\mathrm{Im}\{W\}$ be the span of the columns of W. For any diagonal $G \in \mathbb{R}^{n\times n}$ the matrices A, B and C in (4) have the following properties:

$$CA^T = 0,\ CC^T = I_p,\ B^TA = 0,\ B^TB = I_m, \quad (112)$$

$$C^TC = \mathrm{diag}\{C_1^TC_1, \dots, C_p^TC_p\} \leqslant I_n, \quad (113)$$

$$BB^T = \mathrm{diag}\{B_1B_1^T, \dots, B_mB_m^T\} \leqslant I_n \quad (114)$$

(since $\mathrm{Im}\{C^T\}$ and $\mathrm{Im}\{A^T\}$ are orthogonal subspaces of $\mathbb{R}^n$ and C^T is an orthonormal base of $\mathrm{Im}\{C^T\}$, use duality with $C \Leftrightarrow B^T$ and $A \Leftrightarrow A^T$),

$$(GA^T)^j = 0,\ \forall j \geqslant n, \quad (115)$$

(since A^T is a left-shift operator when acting on the right and, moreover, G is lower triangular),

$$(I_n - GA^T)^{-1} - I_n = \sum_{j=1}^{n-1}(GA^T)^j \quad (116)$$

(noticing that $I_n - GA^T$ is nonsingular, this follows from (115) and the expansion $(I_n - X)^{-1} = \sum_{j=0}^{n-1} X^j$ for all square matrices X such that $X^n = 0$ and $I_n - X$ is nonsingular),

$$CGA^T = 0,\ C(I_n - GA^T)^{-1} = C(I_n - GA^T) = C, \quad (117)$$

$$B^TGA = 0,\ (I_n - GA^T)^{-1}B = (I_n - A^TG)B = B \quad (118)$$

(since G is diagonal, GA^T is in the span of A^T and on account of (112), (116) and duality) and finally (since A^TA and C^TC all orthogonal projections and duality)

$$GA^TA = A^TAG = A^TAGA^TA,\ , \quad (119)$$

$$C^TC = I_n - A^TA,\ A^TA(I_n - A^TA) = 0, \quad (120)$$

$$BB^T = I_n - AA^T,\ AA^T(I_n - AA^T) = 0. \quad (121)$$

Remark *A.1:* It is important to notice that, since A^T is a down-shift operator when acting on the left, all the above formulas hold true with the following changes: $GA^T \Leftrightarrow A^TG$ and $A^TA \Leftrightarrow AA^T$.

B. Properties of the the left- and right-action $\diamond$

For any diagonal matrix D: (*commutative property of* $\diamond$)

$$D \diamond \epsilon^{\mathfrak{r}} = \epsilon^{\mathfrak{r}} \diamond D. \quad (122)$$

For any matrices R, S with suitable dimensions: (*associative property of* $\diamond$ *and the matrix product* $\cdot$)

$$(RS) \diamond \epsilon^{\mathfrak{r}} = R(S \diamond \epsilon^{\mathfrak{r}}),\ \epsilon^{\mathfrak{r}} \diamond (RS) = (\epsilon^{\mathfrak{r}} \diamond R)S \quad (123)$$

(*commutative property of* $\diamond$ *and the matrix product* $\cdot$)

$$(R \diamond \epsilon^{\mathfrak{r}})S = R(\epsilon^{\mathfrak{r}} \diamond S),\ S(\epsilon^{\mathfrak{r}} \diamond R) = (S \diamond \epsilon^{\mathfrak{r}})R. \quad (124)$$

For any invertible matrix R:

$$(R \diamond \epsilon^{\mathfrak{r}})^{-1} = \epsilon^{-\mathfrak{r}} \diamond R^{-1},\ (\epsilon^{\mathfrak{r}} \diamond R)^{-1} = R^{-1} \diamond \epsilon^{-\mathfrak{r}}. \quad (125)$$

Moreover, $(R \diamond \epsilon^{\mathfrak{r}})^T = \epsilon^{\mathfrak{r}} \diamond R^T$.

C. Auxiliary lemmas

The following two lemmas can be proved by using extensively the definition and properties of incremental homogeneity (the proof is omitted for lack of space). Let

$$\begin{aligned} &S^{(o)} := I_n - G^{(o)}A^T,\ S^{(s)} := (I_n - A^TG^{(s)})^{-1}, \\ &\Sigma^{(o)} := I_n + \Gamma^{(o)}A^T,\ \Sigma^{(o)}_{inv} := (I_n - \Gamma^{(o)}A^T)^{-1}, \\ &\Sigma^{(s)} := (I_n - A^T\Gamma^{(o)})^{-1},\ \Sigma^{(s)}_{inv} := I_n + A^T\Gamma^{(o)}. \end{aligned} \quad (126)$$

Extensively, we say that a matrix F is g.i.h. or g.i.h.u.b. if the associated linear function Fx is g.i.h. or g.i.h.u.b., respectively.

Lemma *A.1: Assume (H2) and let* $G^{(o)}$ *and* $\Gamma^{(o)}$ *be as in (8).*

(i) $A^TAG^{(o)}A^TA$ *is g.i.h. with quadruple (* $\mathfrak{r}, \mathfrak{r} + \mathfrak{f}^{(o)}, \mathfrak{f}^{(o)}, A^TA\Gamma^{(o)}A^TA)$*,*

(ii) $S^{(o)}f$ *(resp.* f*) is g.i.h.u.b. with quadruple* $(\mathfrak{r}, \mathfrak{r} + \mathfrak{f}^{(o)}, \mathfrak{f}^{(o)}, \Sigma^{(o)}\Phi^{(o)}(x', x'')$ *(resp.* $(\mathfrak{r}, \mathfrak{r} + \mathfrak{f}^{(o)}, \mathfrak{f}^{(o)}, \Phi^{(o)}(x', x''))$*,*

(iii) $A + S^{(o)}A\sum_{j=1}^{n-1}(G^{(o)}A^T)^j$ *is g.i.h.u.b. with quadruple* $(\mathfrak{r}, \mathfrak{r} + \mathfrak{f}^{(o)}, \mathfrak{f}^{(o)}, A + \Sigma^{(o)}A\sum_{j=1}^{n-1}(\Gamma^{(o)}A^T)^j)$*,*

(iv) ${S^{(o)}}^{-1}$ *(resp.* $S^{(o)}$*) is g.i.h.u.b. with quadruple* $(\mathfrak{r}, \mathfrak{r} - \mathfrak{f}^{(o)}, \mathfrak{f}^{(o)}, \Sigma^{(o)}_{inv})$ *(resp.* $(\mathfrak{r}, \mathfrak{r} - \mathfrak{f}^{(o)}, \mathfrak{f}^{(o)}, \Sigma^{(o)})$*).*

Lemma *A.2: Assume (H1) and let $G^{(s)}$ and $\Gamma^{(s)}$ be as in (15).*

(i) $AA^TG^{(s)}AA^T$ is g.i.h. with quadruple $(\mathfrak{r}, \mathfrak{r}+\mathfrak{f}^{(s)}, \mathfrak{f}^{(s)}, AA^T\Gamma^{(s)}AA^T)$,
(ii) $S^{(s)}\phi$ is g.i.h.u.b. with quadruple $(\mathfrak{r}, \mathfrak{r}-\mathfrak{f}^{(s)}, \widehat{\mathfrak{f}}^{(s)}, \Sigma^{(s)}\Phi^{(s)}(x))$,
(iii) $A+\sum_{j=1}^{n-1}(A^TG^{(s)})^jAS^{(s)^{-1}}$ is g.i.h.u.b. with quadruple $(\mathfrak{r}, \mathfrak{r}+\mathfrak{f}^{(s)}, \mathfrak{f}^{(s)}, A+\sum_{j=1}^{n-1}(A^T\Gamma^{(s)})^jA\Sigma_{inv}^{(s)})$,
(v) $S^{(s)}$ (resp. $S^{(s)^{-1}}$) is g.i.h.u.b. with quadruple $(\mathfrak{r}, \mathfrak{r}, 0, \Sigma^{(s)})$ (resp. $(\mathfrak{r}, \mathfrak{r}, 0, \Sigma_{inv}^{(s)})$).

In what follows, we give a sketchy proof (simple but lengthy matrix algebra is needed) of a couple of auxiliary results which we need to prove theorem 4.1. Recall that $A \preceq B$, $A, B \in \mathbb{R}^{m\times l}$, means $A_{ij} \leqslant B_{ij}$ for all $i = 1, \dots n$, $j = 1, \dots, l$, and $\sup_{\theta\in\mathcal{N}} A(\theta)$, $A(\theta) \in \mathbb{R}^{n\times l}$ for each $\theta \in \mathbb{R}^n$ and compact $\mathcal{N} \subset \mathbb{R}^n$, represents any matrix M such that $A(\theta) \preceq M$ for all $\theta \in \mathcal{N}$. If $\{\mathcal{N}(c)\}_{c\in\mathbb{R}^n_{>}}$ is a family of compact sets $\mathcal{N}(c) \subset \mathbb{R}^n$ continuously depending on c and such that $\mathcal{N}(c) \to \{0\}$ as $c \to 0$ then $\sup_{\theta\in\mathcal{N}(c)} A(\theta)$ is assumed to be such that $\sup_{\theta\in\mathcal{N}(c)} A(\theta) \to \Phi(0)$ as $c \to 0$.

Lemma *A.3: Let $\Phi^{(o)}(x', x'')$ be as in* (H2). *For each $l^{(o)} \in \mathbb{R}_{>}$ and positive definite diagonal $\Gamma^{(o)} \in \mathbb{R}^{n\times n}$ define*

$$N^{(o)} := A + \Sigma^{(o)}A\sum_{j=1}^{n-1}(\Gamma^{(o)}A^T)^j$$
$$M^{(o)} := 2\Sigma^{(o)}F^{(o)}\Sigma_{\rm inv}^{(o)},\ F^{(o)} := \sup_{\substack{x',x''\in\mathbb{R}^n:\\ \|x'\|,\|x''\|\leqslant nl^{(o)}}} \Phi^{(o)}(x', x'') \quad (127)$$

For each $a \in (0,1)$ there exist $l_\infty^{(o)} \in \mathbb{R}_{>}$ and $\Gamma^{(o)}$ such that for all $l^{(o)} \leqslant l_\infty^{(o)}$

$$K^{(o)} := aI_n + N^{(o)} + M^{(o)} + N^{(o)^T} + M^{(o)^T} - 2a\Gamma^{(o)} \leqslant 0. \quad (128)$$

Proof: (*Sketch*). Recall that $\sup_{x',x''\in\mathcal{N}(l^{(o)})}\Phi^{(o)}(x', x'') \to \Phi^{(o)}(0,0)$ as $l^{(o)} \to 0$, with $\mathcal{N}(l^{(o)}) := \{x \in \mathbb{R}^n : \|x\| \leqslant nl^{(o)}\}$ so that $F^{(o)}|_{l^{(o)}=0} = \Phi^{(o)}(0,0)$. Using assumption (H2) for which $\Phi^{(o)}(0,0)$ is lower triangular, for each $a \in (0,1)$ find positive definite diagonal $\Gamma^{(o)} \in \mathbb{R}^{n\times n}$ such that $K^{(o)}|_{l^{(o)}=0} \leqslant -I_n$. Finally, pick $l_\infty^{(o)} \in \mathbb{R}_{>}$ (sufficiently small) such that $K^{(o)}|_{l^{(o)}=l_\infty^{(o)}} - K^{(o)}|_{l^{(o)}=0} \leqslant I$, taking into account that $F^{(o)} \preceq F^{(o)}|_{l^{(o)}=l_\infty^{(o)}}$ (and therefore $M^{(o)} \preceq M^{(o)}|_{l^{(o)}=l_\infty^{(o)}}$) for all $l^{(o)} \leqslant l_\infty^{(o)}$. ■

The following lemma is dual to lemma A.4 and the proof goes exactly in the same way.

Lemma *A.4: Let $\Phi^{(s)}$ be as in* (H1). *For each $l^{(s)} \in \mathbb{R}_{>}$ and positive definite diagonal $\Gamma^{(s)} \in \mathbb{R}^{n\times n}$ define*

$$N^{(s)} := A + [\sum_{j=1}^{n-1}(A^T\Gamma^{(s)})^j]A\Sigma_{inv}^{(s)} \quad (129)$$
$$M^{(s)} := \Sigma^{(s)}F^{(s)}\Sigma_{inv}^{(s)},\ F^{(s)} := \sup_{\substack{x'\in\mathbb{R}^n:\\ \|x'\|\leqslant\|\Sigma_{\rm inv}^{(s)}\mathbf{1}_n\|l^{(s)}}} \Phi^{(s)}. \quad (130)$$

For each $a \in (0,1)$ there exist $l_\infty^{(s)} \in \mathbb{R}_{>}$ and $\Gamma^{(s)}$ such that for all $l^{(s)} \leqslant l_\infty^{(s)}$

$$K^{(s)} := aI_n + N^{(s)} + M^{(s)} + N^{(s)^T} + M^{(s)^T} - a\Gamma^{(s)} \leqslant 0. \quad (131)$$

Lemma *A.5: Consider the function $\xi : \mathbb{R}_{\geqslant} \times \mathbb{R}_{\geqslant} \to \mathbb{R}$ defined in (87) with $d \in \mathbb{R}_{>}$, $\rho, \pi : \mathbb{R}_{\geqslant} \to \mathbb{R}$ and $\chi : \mathbb{R}_{\geqslant} \to \mathbb{R}_{>}$ such that*

$$\lim_{\epsilon\to+\infty}\rho(\epsilon) = \rho_\infty > 0,\ \lim_{\epsilon\to+\infty}\frac{\pi(\epsilon)}{\chi(\epsilon)} = +\infty. \quad (132)$$

There exists $\epsilon_\infty > 0$ such that for all $\epsilon \geqslant \epsilon_\infty$: $\xi(\mathbf{v}, \mathbf{w}) \geqslant 0$ for all $\mathbf{v}, \mathbf{w} \in \mathbb{R}_{\geqslant}$.

Proof: We distinguish two cases. If $0 \leqslant \mathbf{w} \leqslant {l^{(s)}}^2/d$:

$$\xi(\mathbf{v}, \mathbf{w}) \geqslant \rho(\epsilon)\mathbf{v} + \chi(\epsilon)\Big[\frac{\pi(\epsilon)}{\chi(\epsilon)}\frac{1}{d+{l^{(s)}}^2} - 1\Big]d\mathbf{v}. \quad (133)$$

From (132) and since $\chi(\epsilon) > 0$ for all $\epsilon \geqslant 0$, it follows the existence of $\epsilon_\infty > 1$ such that for all $\epsilon \geqslant \epsilon_\infty$: $\xi(\mathbf{v}, \mathbf{w}) \geqslant 0$ for all $\mathbf{v}, \mathbf{w} \in \mathbb{R}_{\geqslant} : \mathbf{w} \leqslant {l^{(s)}}^2/d$. On the other hand, if $\mathbf{w} \geqslant {l^{(s)}}^2/d$:

$$\xi(\mathbf{v}, \mathbf{w}) \geqslant \rho(\epsilon)\mathbf{v} + \chi(\epsilon)\Big[\frac{\pi(\epsilon)}{\chi(\epsilon)}\frac{1}{d+{l^{(s)}}^2} - 1\Big]{l^{(s)}}^2. \quad (134)$$

From (132) it follows the existence of $\epsilon_\infty > 1$ (possibly larger) such that for all $\epsilon \geqslant \epsilon_\infty$: $\xi(\mathbf{v}, \mathbf{w}) \geqslant 0$ for all $\mathbf{v}, \mathbf{w} \in \mathbb{R}_{\geqslant} : \mathbf{w} \geqslant {l^{(s)}}^2/d$. The above facts prove the claim of the lemma. ■

References

[1] T. Ahmed-Ali, E. Cherrier, F. Lamnabhi-Lagarrigue, Cascade high-gain predictor for a class of nonlinear systems, *IEEE Trans. Autom. Contr.*, bf 57, 2012, pp. 221-226

[2] T. Ahmed-Ali , F. Lamnabhi-Lagarrigue, Global exponential sampled-data observers for nonlinear systems with delayed measurements, *Syst. Control Lett.*, **67**, no. 7, 2013, pp. 539-549.

[3] T. Ahmed-Ali, I. Karafyllis, M. Krstic, F. Lamnabhi-Lagarrigue, Robust stabilization of nonlinear globally Lipschitz delay systems, *Advances in Delays and Dynamics*, **4**, New York: Springer, 2016, pp. 43-60.

[4] S. Battilotti, Incremental generalized homogeneity, observer design and semiglobal stabilization, *Asian Journ. Contr.*, **16**, 2014, pp. 498-508.

[5] S. Battilotti, Nonlinear predictors for systems with bounded trajectories and delayed measurements, *Automatica*, **59**, 2015, pp. 127-138.

[6] D. Bresch-Pietri, N. Petit, Robust compensation of a chattering time-varying input delay, in *Proc. IEEE Conf. Decision and Control*, 2014, pp. 457-462, Los Angeles, CA.

[7] G. Besancon, D. Georges, Z. Benayache, Asymptotic state prediction for continuous-time systems with delayed input output and application to control, *Europ. Contr. Conf.*, 2007, pp. 552-557.

[8] F. Cacace, A. Germani, C. Manes, An observer for a class of nonlinear systems with time varying observation delays, *Syst. Contr. Lett.*, **59**, 2010, pp. 305-312.

[9] F. Cacace, A. Germani, C. Manes, A chain observer for nonlinear systems with multiple time varying measurement delays, SIAM Journ. Opt. Contr., **52**, 2014, pp. 1862-1885.

[10] F. Cacace, F. Conte, A.Germani, Memoryless approach to the LQ and LQG problems with variable input delay, *IEEE Trans. Autom. Control*, **61**, no. 1, 2016, pp. 216-221.

[11] F. Cacace, A. Germani, C. Manes, Predictor-based control of linear systems with large and variable measurement delays, *Int. J. Control*, **87**, no. 4, 2014, pp. 704-714.

[12] E. Fridman and S.-I. Niculescu, On complete Lyapunov-Krasovskii functional techniques for uncertain systems with fast-varying delays, *Int. J. Robust and Nonlin. Control*, **18**, no. 3, 2008, pp. 364-374, .

[13] A. Germani, C. Manes, P. Pepe, A new approach to state observation of nonlinear systems with delayed output, *IEEE Trans. Autom. Control*, **47**, no. 1, pp. 2002, pp. 96-101.

[14] M. Ghanes, J. De Leon, J.-P. Barbot, Observer design for nonlinear systems observer design for nonlinear systems under unknown time-varying delays, *IEEE Trans. Autom. Contr.*, **58**, 2013, pp. 1529-1534.

[15] S. Ibrir, Observer based control of a class of time delay nonlinear systems, *Automatica*, **47**, 2011, pp. 388-394.

[16] I. Karafyllis, Stabilization by means of approximate predictors for systems with delayed input, *SIAM Journ. Opt. Contr.*, **49**, 2012, pp. 1141-1154.

[17] I. Karafyllis, M. Krstic, Nonlinear stabilization under sampled and delayed measurements and with inputs subject to delay and zero-order hold, *IEEE Trans. Autom. Contr.* **57**, 2012, pp. 1141-1154.

[18] I. Karafyllis, M. Krstic, Predictor-based output feedback for nonlinear delay systems, arXiv:1108.4499.

[19] I. Karafyllis, M. Krstic, T. Ahmed-Ali, F. Lamnabhi-Lagarrigue, Global stabilization of nonlinear delay systems with a compact absorbing set, *Int. Journ. Contr.*, **87**, 2013, pp. 1010-1027.

[20] J. Lei, H. K. Khalil, High-gain-predictor-based output feedback control for time-delay nonlinear systems, *Automatica*, **71**, 2016, pp. 324-333.

[21] N. Kazantzis, R. A. Wright, Nonlinear observer design in the presence of delayed output measurements, *Syst. Contr. Lett.*, **54**, 2001, pp. 1100-1123.

[22] M. Krstic, *Delay compensation for nonlinear adaptive and PDE systems*, Birkhauser, Boston, 2009.

[23] L. Marquez-Martinez, C. Moog, M. Velasco-Villa, Observability and observers with time-delay, *2nd IFAC Workshop Time Delay Syst.*, 2000, pp. 52-57.

[24] F. Mazenc, M. Malisoff, Local stabilization of nonlinear systems through the reduction model approach, *IEEE Trans. Autom. Control*, **59**, no. 11, 2014, pp. 3033-3039.

[25] F. Mazenc, M. Malisoff, Stabilization and robustness analysis for time-varying systems with time-varying delays using a sequential subpredictors approach, **82**, 2017, pp. 118-127.

[26] F. Mazenc, M. Malisoff, T. Dinh, Robustness of nonlinear systems with respect to delay and sampling of the controls, *Automatica*, **49**, no. 6, 2013, pp. 1925-1931.

[27] F. Mazenc, M. Malisoff, Z. Lin, Further results on input-to-state stability for nonlinear systems with delayed feedbacks, *Automatica*, **44**, no. 9, 2008, pp. 2415-2421.

[28] F. Mazenc, M. Malisoff, S.-I. Niculescu, Reduction model approach for linear time-varying systems with delays, *IEEE Trans. Autom. Control*, **59**, no. 8, 2014, pp. 2068-2082.

[29] F. Mazenc, S.-I. Niculescu, M. Krstic, Lyapunov-Krasovskii functionals and application to input delay compensation for linear time invariant systems, *Automatica*, **48**, no. 7, 2012, pp. 1317-1323.

[30] F. Mazenc, M. Malisoff, S.-I. Niculescu, Stabilization of Nonlinear Time-Varying Systems Through a New Prediction Based Approach, *IEEE Trans. Autom. Control*, **62**, no. 8, 2017, pp. 2908-2915.

[31] W. Michiels, S.-I. Niculescu, *Stability and Stabilization of Time-Delay Systems*, Philadelphia, PA: SIAM, 2007.

[32] M. Najafi, S. Hosseinnia, F. Sheikholeslam, M. Karimadin, Closed-loop control of dead time systems via sequential sub-predictors,*Int. J. Control*, **86**, no. 4, 2013, pp. 599-609.

[33] H. Omran, L. Hetel, J.-P. Richard, and F. Lamnabhi-Lagarrigue, Stability analysis of bilinear systems under aperiodic sampled-data control, *Automatica*, **50**, no. 4, 2014, pp. 1288-1295.

[34] P. Richard, Time delay systems: an overview of some recent advances and open problems, *Automatica*, **29**, 2003, pp. 388-394.

[35] H. Shim, A. R. Teel, Asymptotic controllability and observability imply semiglobal practical asymptotic stabilizability by sampled-data output feedback, *Automatica*, **39**,2003, pp. 441-453.

[36] H. Smith, *An introduction to delay differential equations*, vol. 57, Springer, 2011.

[37] K. Subbarao and P. Muralidhar, State observer for linear systems with piecewise constant output delays, *IET Contr. Theory and Appl.*, 2009, pp. 1017-1022.

[38] C. Ning, Y. He, M. Wu, J. She, Improved Razumichin-Type Theorem for Input-To-State Stability of Nonlinear Time-Delay Systems, *IEEE Trans. Autom. Contr.*, **59**, 2014, pp. 1983-1988.

[39] V. Van Assche, T. Ahmed-Ali, C. Hann, F. Lamnahbi-Lagarrigue, High gain observer design for nonlinear systemswith time varying delayed mea- surement, *18th IFAC World Congr.*, 2011, pp. 692-696.

[40] B. Zhou, Pseudo-predictor feedback stabilization of linear systems with time-varying input delays, *Automatica*, **50**, 2014, pp. 2861-2871.

[41] B. Zhou, Z. Lin, G.-R. Duan, Truncated predictor feedback for linear systems with long time-varying input delays, *Automatica*, **48**, 2012, pp. 2387-2399.

Stefano Battilotti was born in Rome, Italy, in 1962. In 1992 he joined the Dipartimento di Ingegneria Informatica, Automatica e Gestionale Antonio Ruberti (DIAG) of Sapienza in Rome (Italy), where since 2005 he is a Professor of Automatic Control. He authored more than 50 journal papers and the monograph "Noninteracting control with stability for nonlinear systems", LNCIS Springer series. His research interests are currently focused on state estimation and control of nonlinear systems and networked systems.