

MULTILAYER STATE PREDICTORS FOR NONLINEAR SYSTEMS WITH TIME-VARYING MEASUREMENT DELAYS

S. BATTILOTTI*

Abstract. In this paper, we propose state predictors for stable genuinely nonlinear systems with time-varying measurement delay, with no restriction on its bound or serious limitations on the growth of the nonlinearities. The measurement delay is assumed to be continuous. A state prediction is generated by chains of nonlinear dynamic observers operating at different layers. On each layer, these observers reconstruct the unmeasurable state vector at different delayed time-instants, which partition the maximal variation interval of the time-varying delay. This partition determines the number of observers in the layer. Transitions from a layer to the next one are triggered by a on-line estimate of the magnitude of the state. Consistently, in passing to the next layer the partition is refined and the number of observers increased. In this sense, our predictor is nonlinear and adaptive.

Key words. Delay systems, time-varying measurement delays, state predictors.

AMS subject classifications. 93A99, 93B07, 93B10

1. Introduction. A challenging problem is the state estimation in the presence of time-varying delayed output measurements. In this case it is important to implement some kind of state prediction based on the delayed measurements and take into account that the measurements are delayed differently at each time.

A first approach, limited to constant delays, is the one which introduces integral terms ([8], [9] see the survey paper [12] and references therein): this approach is 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, this approach cannot be directly extended to the case of time-varying delays.

The second approach is based on dynamic observers that reconstruct the unmeasurable state vector at different delayed time-instants within the time-delay window introduced by the output measurements ([3], [7], [13], [6], [15]). The observers exhibits a chained structure and each observer estimates the unmeasurable state vector at delayed time-instants which differ each other by a sufficiently small amount. Each observer delivers its estimate to the next observer in the chain, until the last observer which gives an asymptotic estimate of the undelayed state. In [3] the innovation process of each observer varies according to the current value of the time-varying delay. The limitation of this approach is that only globally Lipschitz systems are considered and linear observers are used. In the case of constant measurement delays, more recently we have [10] where a modified version of the chained observers, introduced in [5] and [11], is adopted: however, a known compact absorbing set (plus some technical additional assumptions) is assumed for all the system trajectories. In [2] chained nonlinear observers for nonlinear stable systems are studied, removing globally Lipschitz and compact absorbing set assumptions by introducing techniques based on incremental homogeneity properties. The approach of [2] and [10] cannot be extended to the case of time-varying measurement delays. Moreover, approximating the time-varying delay with piecewise constant delays may deteriorate the performances especially in the case of fast-varying delays.

In this paper, we remove globally Lipschitz assumptions by using techniques based on incremental homogeneity ([1]) and consider stable nonlinear systems with time-

*Department of Computer, Control, and Management Engineering “Antonio Ruberti”, Sapienza University of Rome, Via Ariosto 25, Italy. (battilotti@diag.uniroma1.it).

varying measurement delays by using a completely different approach from [2] and being inspired from the observer chain introduced in [3] for globally Lipschitz systems. Basically, in [2] we proposed an observer chain of constant length and, moreover, the structure of each observer in the chain does not depend on time (since the delays are constant). In this paper, a state estimation is generated by chains of nonlinear dynamic observers operating at different layers. At each layer the length of the chain increases. The multi-layer structure is one of the novelty of our approach and it is a specific tool for coping with strong nonlinearities of the system. Another important difference with [2] is that the structure of each observer in the chain changes with time. In particular, according to the position of the time-varying delay inside its maximal variation interval, each observer of the chain uses different innovation processes. This switching structure, together with the variable length of the chain at different layers, stands as a basic difference with [2] and requires also completely different stability analysis tools (Krasovski-Lyapunov or Razumichin-Lyapunov functions and switching system's analysis). The partition of the maximal variation interval of the time-varying delay determines the number of observers in each layer. Transitions from a layer to the next one are triggered by a on-line estimate of the state magnitude. Consistently, during the transition to the next layer the partition is refined and the number of observers increased. Larger the estimated state magnitude is, finer the partition (and the length of the observer chain) of the maximal variation interval of the time-varying delay must be. Therefore, the length of the observer chain varies adaptively according to a on-line estimation of the state magnitude (multilayer structure): this is very natural when considering nonlinear systems with weak growth restrictions. Another novelty of our state predictor in comparison with the literature is the use of nonlinear observers with saturated estimates where the saturation levels are dynamically tuned according to the on-line estimation of the state magnitude.

2. 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}_{\geq} (resp. \mathbb{R}_{\geq}^n , $\mathbb{R}_{\geq}^{n \times s}$) 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). Moreover, for each $v \in \mathbb{R}^n$ $\text{diag}\{v\}$ is the diagonal matrix with diagonal elements v_1, \dots, v_n and for each m -tuple of matrices A_1, \dots, A_m $\text{diag}\{A_1, \dots, A_m\}$ is a block-diagonal matrix with diagonal blocks A_1, \dots, A_m . Also, $|a|$ denotes the absolute value of $a \in \mathbb{R}$, $\|a\|$ denotes the euclidean norm of $a \in \mathbb{R}^n$ with $\|a\|_M := \sqrt{a^T M a}$, $\|A\|$ 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 $\langle v \rangle_i := |v_i|$). For any vectors $x, y \in \mathbb{R}^n$ we write $x \leq y$ if and only if $x_i \leq y_i$ for all $i = 1, \dots, n$. We retain the same notation for matrices $A, B \in \mathbb{R}^{n \times m}$: $A \leq B$ if and only if $A_{ij} \leq B_{ij}$ for all $i = 1, \dots, n$ and $j = 1, \dots, m$. On the other hand $A \leq B$ (resp. $A < B$) for matrices $A, B \in \mathbb{R}^{n \times n}$ if and only if $A - B$ is negative semidefinite (resp. negative definite).

(N3) We denote by $\mathbf{C}^0(\mathcal{X}, \mathcal{Y})$, $\mathcal{X} \subset \mathbb{R}^n$ and $\mathcal{Y} \subset \mathbb{R}^p$, the set of continuous functions $f : \mathcal{X} \rightarrow \mathcal{Y}$, $\mathbf{L}^\infty(\mathcal{X}, \mathcal{Y})$ the set of functions $f \in \mathbf{C}^0(\mathcal{X}, \mathcal{Y})$ such that $\sup_{\theta \in \mathcal{X}} \|f(\theta)\| < +\infty$. For each $d \in \mathbf{L}^\infty(\mathcal{X}, \mathcal{Y})$, we have the sup norm of d defined as $\|d\|_\infty := \sup_{\theta \in \mathcal{X}} \|d(\theta)\|$. Moreover, \mathcal{K} denotes the set of functions $f \in \mathbf{C}^0(\mathbb{R}_{\geq}, \mathbb{R}_{\geq})$ strictly increasing and such that $f(0) = 0$, \mathcal{L} denotes the set of functions $f \in \mathbf{C}^0(\mathbb{R}_{\geq}, \mathbb{R}_{\geq})$ strictly decreasing, and \mathcal{KL} denotes the set of functions $f \in \mathbf{C}^0(\mathbb{R}_{\geq} \times \mathbb{R}_{\geq}, \mathbb{R}_{\geq})$ such that $f(s, \cdot) \in \mathcal{L}$ and $f(\cdot, s) \in \mathcal{K}$ for each $s \in \mathbb{R}_{\geq}$.

(N4) A *saturation function* *sat* with *saturation levels* $l \in \mathbb{R}_{>}^n$ is a function $\text{sat}(x, l) := (\text{sat}(x_1, l_1), \dots, \text{sat}(x_n, l_n))^T$, $x \in \mathbb{R}^n$ and $l \in \mathbb{R}_{>}^n$, such that for each $i = 1, \dots, n$ and $x_i \in \mathbb{R}$:

$$(1) \quad \text{sat}(x_i, l_i) = \begin{cases} x_i & |x_i| \leq l_i \\ \text{sign}(x_i)l_i & \text{otherwise.} \end{cases}$$

It is easy to prove the following inequalities:

$$(2) \quad \langle \text{sat}(x, l) - \text{sat}(y, l) \rangle \leq 2 \langle \text{sat}(x - y, l) \rangle \leq 2l,$$

$$(3) \quad \langle \text{sat}(x, l) \rangle \leq \langle x \rangle$$

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

(N5) For any vectors $x \in \mathbb{R}^n$, $\mathbf{r} \in \mathbb{R}_{>}^n$ and $z \in \mathbb{R}_{>}$, we define

$$(4) \quad z^{\mathbf{r}} := (z^{\mathbf{r}_1}, \dots, z^{\mathbf{r}_n})^T, \quad z^{\mathbf{r}} \diamond x := (z^{\mathbf{r}_1} x_1, \dots, z^{\mathbf{r}_n} x_n)^T$$

viz. $z^{\mathbf{r}} \diamond x$ is the dilation of a vector x with weights \mathbf{r} .

3. The class of systems and problem statement. We consider nonlinear systems with delayed outputs y and each output is measured together with the delay. The measurement delay is time-varying but known and bounded by a known constant δ_{\max} :

$$(5) \quad \dot{x}_t = F(x_t) := Ax_t + \phi(x_t), \quad t \geq -\delta_{\max},$$

$$(6) \quad y_t = H(x_{t-\delta_t}) := Cx_{t-\delta_t}, \quad t \geq 0,$$

with state $x_t \in \mathbb{R}^n$, initial state $x_{-\delta_{\max}}$, output $y_t \in \mathbb{R}^p$ and measurement delay $\delta_t \in \mathbb{R}_{\geq}$. We assume that $A = \text{diag}\{A_1, \dots, A_p\}$, with Brunovskii forms A_i , $C = \text{diag}\{C_1, \dots, C_p\}$, with $C_i = (1, 0, \dots, 0)$, and ϕ is locally Lipschitz. The problem we want to solve in this paper is to design a dynamical estimator for the state of (5) using the delayed measurements (6). For simplicity, we consider only one measurement delay and no output nonlinearities, but the extension to multiple delays and output nonlinearities is straightforward although more complex. Also, it is possible to include the input vector in (5) at no additional technical cost.

3.1. Continuous versus discontinuous delays. The measurement delay δ_t is known and it is known at the same time t at which y_t is measured. This means that the information available for processing at time $t \geq 0$ is the pair $\{y_t, \delta_t\}$. The assumption that the delay is known is realistic in many applications. A common case is that of networked control systems, when the measurements are buffered and then sent over a reliable network that introduces a variable delay. In this case, the delay is typically computed by comparing the time at which the packet is delivered with the time-stamp included in the packet at the sender side.

We assume that δ_t is bounded by δ_{\max} and continuous. A useful property of continuous delays is the following ([3]). Let y_t^* be the undelayed outputs, i.e. $y_{t-\delta_t}^* = y_t$.

LEMMA 1. If δ_t is continuous, when for each $t > \delta_{\max}$ and $\Delta : \delta_t \leq \Delta \leq t$ $\exists \bar{t} \leq t : y_{\bar{t}} = y_{t-\Delta}^*$.

In other words, when the delay δ_t is continuous, all the measurements previous to y_t are available. When the delay is discontinuous, not all the outputs are available. For example, in the case of sampled outputs with sampling period T the outputs in between jT and $(j+1)T$, $j = 0, 1, \dots$, are not available.

4. The structure of the predictor and main result. The state predictor we propose consists of a certain number of layers, each layer consisting of a certain number of chained observers. These observers are chained in the sense that each observer delivers a certain amount of information (like its own estimate) to the next one in each chain. These observer chains run at the same time on different layers and the number of chained observers on each layer varies (and it is increasing) according to the depth of the layer itself. For consistency, the observers which run inside a certain layer with depth k are also running inside any layer with depth $> k$ and, moreover, it is possible to go from layer k to layer $k+1$ (and not viceversa). The observers of the layer with depth k are $2^k m$, where m is the number of observers of the layer with depth 0.

Correspondingly to the layer with depth k , a partition of the interval $[0, \delta_{\max}]$ into a sequence of points $\{\Delta_j^{(k)}\}_{j \in J_k \subset \mathbb{Q}}$ is considered and this partition is refined at each deeper layer, with the consistency condition that $\Delta_j^{(k)} = \Delta_j^{(k-1)}$ and $J_k \subset J_{k+1}$, i.e. the partition at layer with depth k is a sub-partition of the one at preceding layers. We remark that the indexes of the sequence are in general rational (not integer) numbers since at each deeper layer the sequence is refined and its indexes are divided by an integer number (see definition 3). According to this partition, each observer of a layer with depth k computes an estimate of $x_t^{(j)} := x_{t-\Delta_{j-1}^{(k)}}$, denoted by $\hat{x}_t^{(j)}$, with the consistency condition $x_{t-\Delta_{j-1}^{(k)}} = x_{t-\Delta_{j-1}^{(k-1)}}$. Therefore, the observer which computes $\hat{x}_t^{(j)}$ at some layer is the same which computes $\hat{x}_t^{(j)}$ at the next layer. In other words, different (and consecutive) layers correspond to different (and consecutive) time intervals over which the same observer runs and this is consistent with the fact that it is possible to go from layer k to layer $k+1$ (and not viceversa). Therefore, the number of observers increases in time from some layer to the next one but a suitable stabilization mechanism (which will be explained later) will make this number tend to a steady state value.

Each layer is activated by a reference signal \hat{z}_t which provides an estimate of the magnitude of $x_t^{(m)}$. In this sense the chain of observers is *adaptive*. Larger is the magnitude of $x_t^{(m)}$, finer the partition $\{\Delta_j^{(k)}\}$ and deeper the layer must be. In other words, going from one layer to the next deeper one we must increase the number of the observers or in other words we need split the delay in between $x_t^{(j)}$ and $x_t^{(j-1)}$ into smaller pieces and compute estimates of the state with these smaller delays: for this reasons the number of observers increases with the depth of the layer. The reference signal \hat{z}_t generates a signal \hat{w}_t which provides, for each observer j (preceding and following $x_t^{(m)}$ in the chain, an estimate of the magnitude of $x_t^{(j)}$. The reference signal \hat{z}_t reaches a steady state value (which depends on the initial condition of the system) for which no more transitions to deeper layers are possible.

The delayed state $x_t^{(j)}$ is defined for $t \geq \Delta_{j-1}^{(k)}$ and satisfies the differential equation $\dot{x}_t^{(j)} = Ax_t^{(j)} + \phi(x_t^{(j)})$ with $x_0^{(j)} := x_{-\Delta_{j-1}^{(k)}}$. The last element of the observer chain

inside each layer is an observer which provides the desired estimate $\hat{x}_t^{(1)}$ of the current state $x_t^{(1)} := x_t$. Moreover, each observer of the chain, say the j -th observer of the chain, is fed by either one or both of the following type of information:

- a suitable transformation of the available information $\{y_t, \delta_t\}$
- the estimate $\hat{x}_t^{(j+1)}$ provided by the preceding observer in the chain, according to the relative position of the delay δ_t with respect to the interval $[\Delta_{j-1}^{(k)}, \Delta_j^{(k)}]$.

Different transformations of the above type of information determine different innovation processes for each observer to guarantee convergence of the estimate to the delayed state (in conceiving these transformations we are inspired by the algorithm proposed in [3]).

Before plunging into the technical details of the multilayer predictor, we need to make precise what a partition of an interval and any its refinement is.

DEFINITION 2. A sequence $\{\Delta_j\}_{0,1,\dots,m}$ is a m -partition ($m > 1$) of an interval $[0, a]$ if it is strictly increasing, $\Delta_0 := 0$ and $\Delta_m := a$.

DEFINITION 3. A sequence $\{\Delta_j^{(S)}\}_{0, \frac{1}{h}, \frac{2}{h}, \dots, q}$ is a h -refinement ($h \in (1, +\infty)$) of a m -partition $\{\Delta_j\}_{0,1,\dots,q}$ of an interval $[0, a]$ if it is strictly increasing, $\Delta_j = \Delta_j^{(S)}$ for each $j = 0, 1, \dots, m$ and $\Delta_{j+1} - \Delta_j = \sum_{i=0}^{h-1} (\Delta_{j+1-\frac{i}{h}}^{(S)} - \Delta_{j+1-\frac{i+1}{h}}^{(S)})$ for each $j = 0, 1, \dots, m-1$.

Notice that a h -refinement must satisfy a consistency condition (i.e. same points of the partition and its refinement must coincide) and each interval of the partition is divided into h parts in the h -refinement.

Let $\alpha, \nu \in \mathbb{R}_{>}$, $\mathfrak{g} \in \mathbb{R}_{>}^n$ and $Z_0 > 1$ be design parameters and let $\{\Delta_j^{(0)}\}_{0,1,\dots,m}$ be a m -partition of the interval $[0, a]$, with $a > \delta_{max}$, such that

$$(7) \quad \begin{aligned} \Delta_{m-1}^{(0)} &= \delta_{max}, \\ \max_{j=1,\dots,m} (\Delta_j^{(0)} - \Delta_{j-1}^{(0)}) &< \frac{\nu}{\left(2^{\alpha-1}(Z_0^\alpha + \alpha\delta_{max}^\alpha)\right) \frac{|\min_i \mathfrak{g}_i| + 3|\max_i \mathfrak{g}_i|}{\alpha}}. \end{aligned}$$

For notational convenience, let

$$\hat{F}(\chi, \zeta) := A\chi + \phi\left(\text{sat}(\chi, \gamma \diamond \zeta^{\mathfrak{r}})\right)$$

4.1. The structure of the multilayer predictor. The multilayer predictor consists of a certain number of layers, each one showing a chain of observers. Let $\mathfrak{g} \in \mathbb{R}^n$, $\mathfrak{r} \in \mathbb{R}_{>}^n$, $\gamma \in \mathbb{R}_{>}$ and diagonal positive definite $\Gamma \in \mathbb{R}^{n \times n}$ be design parameters.

Depth-0 observer chain. The first observer of the depth-0 chain, which we refer to as depth-0 master observer, is

$$(8) \quad \begin{aligned} \dot{\hat{x}}_t^{(m)} &= \hat{F}(\hat{x}_t^{(m)}, \hat{z}_t) + P^{-1}(\hat{z}_t)C^T R(\hat{z}_t)\epsilon_t^{(m,0)}, \quad t \geq t_0 := 0, \\ \dot{\hat{z}}_t &= \hat{z}_t \text{sat}\left(\left\|\hat{z}_t^{-\mathfrak{r}+\mathfrak{g}} \diamond \left(\hat{x}_t^{(m)} - \text{sat}(\hat{x}_t^{(m)}, \gamma \diamond \hat{z}_t^{\mathfrak{r}})\right)\right\|^2 + \|\hat{z}_t^{-\mathfrak{r}+\mathfrak{g}} \diamond \epsilon_t^{(m,0)}\|^2, \frac{1}{\hat{z}_t^\alpha}\right), \quad t \geq t_0, \end{aligned}$$

with initialization

$$(9) \quad \hat{x}_\theta^{(m)} := 0, \quad \hat{z}_\theta := 1, \quad \theta \in [t_0 - \delta_{max}, t_0],$$

and

$$(10) \quad \begin{aligned} P(\hat{z}) &= (I - G(\hat{z})A^T)^T \text{diag}\{\hat{z}^{-2\mathfrak{r}}\}(I - G(\hat{z})A^T), \\ R(\hat{z}) &= C \text{diag}\{\hat{z}^{-\mathfrak{r}}\}G(\hat{z}) \text{diag}\{\hat{z}^{-\mathfrak{r}}\}C^T, \quad G(\hat{z}) = \text{diag}\{\hat{z}^\mathfrak{g}\}\Gamma \text{diag}\{\hat{z}^\mathfrak{g}\}. \end{aligned}$$

for some innovation process $\epsilon_t^{(m,0)}$ to be defined below.

Remark 4.1. While the signal $\hat{x}_t^{(m)}$ is an estimate of $x_t^{(m)}$ and it is provided by the first equation in (8), the signal \hat{z}_t is an estimate of the magnitude of $x_t^{(m)}$ and it is provided by the second equation in (8) interconnected with the first equation. The initialization of \hat{z}_t requires that $\hat{z}_\theta = 1$ for $\theta \in [t_0 - \delta_{max}, t_0]$. Since $\hat{z}_t \geq 0$ for all $t \geq t_0$, we have $\hat{z}_t \geq 1$ for all $t \geq t_0$ for which \hat{z}_t is defined (actually, as it will be shown \hat{z}_t is bounded in time). The observer gains are tuned and suitably rescaled by the dynamic parameter \hat{z}_t (this is the first adaptive feature of our predictor, while the second is the increase of the chain's length according to the values of \hat{z}_t).

Going down the depth-0 chain, we find the j -th observer ($j = 2, \dots, m$), which we refer to as depth-0 j -th slave observer,

$$(11) \quad \begin{aligned} \dot{\hat{x}}_t^{(m-j+1)} &= \hat{F}(\hat{x}_t^{(m-j+1)}, \hat{w}_t) + P^{-1}(\hat{w}_t)C^T R(\hat{w}_t)\epsilon_t^{(m-j+1,0)}, \quad t \geq t_0, \\ \hat{x}_\theta^{(m-j+1)} &:= 0, \quad \theta \in [t_0 - \delta_{max}, t_0], \end{aligned}$$

for some innovation process $\epsilon_t^{(m-j+1,0)}$ to be defined below and with

$$(12) \quad \hat{w}_t := 2^{\frac{\alpha-1}{\alpha}} \left(\hat{z}_t^\alpha + \alpha \delta_{max}^\alpha \right)^{\frac{1}{\alpha}}.$$

Remark 4.2. The signal \hat{w}_t is defined through the reference signal \hat{z}_t , generated by the master observer, and provides, for each depth-0 j -th slave observer, an estimated upper bound for the magnitude of $x_t^{(m-j+1)}$. The structure of (12) is motivated by the fact that $\dot{\hat{z}}_t \leq \hat{z}_t^{-\alpha}$ in time so that, by this simple differential inequality, an estimation (in excess) \hat{w}_t of the magnitude of $x_t^{(m-j+1)}$ can be obtained as the α -root of the sum of the α -power estimation \hat{z}_t of the magnitude of $x_t^{(m)}$ plus the difference between the two delays $\Delta_{m-1}^{(0)} - \Delta_{m-j}^{(0)} \leq \delta_{max}$.

The innovation process for the j -th observer of the depth-0 chain, $j = 1, \dots, m$, is defined for $j = 1$ (the depth-0 master observer) as

$$(13) \quad \epsilon_t^{(m,0)} = y_{t-\Delta_{m-1}^{(0)}}^* - H(\hat{x}_t^{(m)})$$

(remember y_t^* is the undelayed output) and for $j = 2, \dots, m$ (the depth-0 slave observers) as

$$(14) \quad \epsilon_t^{(m-j+1,0)} = \begin{cases} y_{t-\Delta_{m-j}^{(0)}}^* - H(\hat{x}_{t-\delta_t^{(m-j+1,0)}}^{(m-j+1)}) & \text{if } \delta_t \in [0, \Delta_{m-j}^{(0)}], \\ y_{t-\delta_t}^* - H(\hat{x}_{t-\delta_t^{(m-j+1,0)}}^{(m-j+1)}) & \text{if } \delta_t \in [\Delta_{m-j}^{(0)}, \Delta_{m-j+1}^{(0)}], \\ H(\hat{x}_t^{(m-j+2)}) - H(\hat{x}_{t-\delta_t^{(m-j+1,0)}}^{(m-j+1)}) & \text{if } \delta_t \in [\Delta_{m-j+1}^{(0)}, \Delta_m^{(0)}]. \end{cases}$$

where

$$(15) \quad \delta_t^{(m-j+1,0)} = \begin{cases} 0 & \text{if } \delta_t \in [0, \Delta_{m-j}^{(0)}], \\ \delta_t - \Delta_{m-j}^{(0)} & \text{if } \delta_t \in [\Delta_{m-j}^{(0)}, \Delta_{m-j+1}^{(0)}], \\ \Delta_{m-j+1}^{(0)} - \Delta_{m-j}^{(0)} & \text{if } \delta_t \in [\Delta_{m-j+1}^{(0)}, \Delta_m^{(0)}]. \end{cases}$$

Remark 4.3. Notice that when $\delta_t \in [0, \Delta_{m-j+1}^{(0)})$, $j = 1, \dots, m$, the measured output $y_t = y_{t-\delta_t}^*$ (y_t^* is the undelayed output) is further delayed and $y_{\bar{t}}$ ($\bar{t} < t$)

is used for the innovation $\epsilon_t^{(m-j+1,0)}$ instead of y_t where $\bar{t} = t + \delta_{\bar{t}} - \Delta_{m-j}^{(0)}$, in other words y_t is delayed by an additional $\Delta_{m-j}^{(0)} - \delta_{\bar{t}}$. The estimate $\hat{x}_t^{(m-j+1)}$ is not delayed ($\delta_t^{(m-j+1,0)} = 0$) in the innovation process. This is also the case of the master observer. Notice that for the implementation of this step we need the past outputs $y_{\bar{t}}$ ($\bar{t} < t$) and this requires the continuity of δ_t (see lemma 1).

When $\delta_t \in [\Delta_{m-j}^{(0)}, \Delta_{m-j+1}^{(0)}]$, $j = 1, \dots, m$, the available output y_t is not modified but the estimate $\hat{x}_t^{(m-j+1)}$ is delayed by the amount $\delta_t^{(m-j+1,0)} = \delta_t - \Delta_{m-j}^{(0)}$.

When $\delta_t \in [\Delta_{m-j+1}^{(0)}, \delta_{max}]$, $j = 2, \dots, m$, the available output y_t is replaced by $H(\hat{x}_t^{(m-j+2)})$, where $\hat{x}_t^{(m-j+2)}$, computed by the preceding observer in the chain, is an estimate of $x_t^{(m-j+2)}$. The estimate $\hat{x}_t^{(m-j+1)}$ is delayed by the amount $\delta_t^{(m-j+1,0)} = \Delta_{m-j+1}^{(0)} - \Delta_{m-j}^{(0)}$ in the innovation process.

The reason for which the first observer of the chain is identified as a master observer is accounted for by the definition of $\epsilon_t^{(m,0)}$ and the structure of the observer itself: the first observer receives as information only $\{y_t, \delta_t\}$ and provides the reference signal \hat{z}_t . According to the definition of $\epsilon_t^{(j,0)}$, any subsequent observers j in the chain receives as information not only $\{y_t, \delta_t\}$ but also the estimate $\hat{x}_t^{(j+1)}$ of the preceding observer in the chain and \hat{w}_t which is a suitable transformation of the reference signal.

Depth- k ($k \geq 1$) observer chain. The decision rule for switching to the next layer is the following. Let $t_k > t_{k-1}$ be the time for which

$$(16) \quad \hat{z}_{t_k} = Z_{k-1}$$

(if $\hat{z}_t < Z_{k-1}$ for all $t \geq t_{k-1}$ we will set $t_k := +\infty$). Notice that \hat{z}_t (since $\dot{\hat{z}}_t \geq 0$) is non-decreasing in time so that our definition of t_k is consistent.

Set

$$(17) \quad Z_k := 2^{\lceil \frac{1}{\min_i \vartheta_i} + 3 \rceil \lceil \max_i \vartheta_i \rceil} Z_{k-1}.$$

Let $\{\Delta_j^{(k)}\}_{0, \frac{1}{2^k}, \frac{2}{2^k}, \dots, m}$ be a 2-refinement of the m -partition $\{\Delta_j^{(k-1)}\}_{0, \frac{1}{2^{k-1}}, \frac{2}{2^{k-1}}, \dots, m}$ at layer $k-1$ defined as follows: for each $j = 0, \frac{1}{2^{k-1}}, \frac{2}{2^{k-1}}, \dots, m-1 - \frac{1}{2^{k-1}}, m-1$

$$(18) \quad \Delta_{j+\frac{1}{2}}^{(k)} = \Delta_j^{(k-1)} + \frac{\Delta_{j+1}^{(k-1)} - \Delta_j^{(k-1)}}{2}.$$

(in the refinement we add the mean points of each interval $[\Delta_j^{(k-1)}, \Delta_{j+1}^{(k-1)}]$: this is just one of the possible choices).

Remark 4.4. The signal z_t is the reference signal which activates the k -depth, according if it exceeds $2^{\lceil \frac{1}{\min_i \vartheta_i} + 3 \rceil \lceil \max_i \vartheta_i \rceil}$ times its past value at $t = t_{k-1}$. By the boundedness of $x_t^{(m)}$, as it will be shown, the sequence of times $\{t_k\}$ is always finite and there will be always a k^* (depending on the state initial conditions) such that $\hat{z}_t < Z_{k^*}$ for all $t \geq t_{k^*-1}$ so that $t_{k^*} = +\infty$.

The first 2^k chained observers of the depth- k chain, which we refer to as master observer for the last one and slave observers for the first $2^k - 1$ ones, are

$$(19) \quad \begin{aligned} \dot{\hat{x}}_t^{(m+1-\frac{j}{2^k})} &= \hat{F}(\hat{x}_t^{(m+1-\frac{j}{2^k})}, \hat{z}_t) + P^{-1}(\hat{z}_t)C^T R(\hat{z}_t)\epsilon_t^{(m+1-\frac{j}{2^k}, k)}, \\ t &\geq t_k, j = 1, \dots, 2^k, \end{aligned}$$

with

$$\dot{\hat{z}}_t = \hat{z}_t \text{sat} \left(\left\| \hat{z}_t^{-\tau+\mathfrak{g}} \diamond \left(\hat{x}_t^{(m)} - \text{sat}(\hat{x}_t^{(m)}, \gamma \diamond \hat{z}_t^\tau) \right) \right\|^2 + \left\| \hat{z}_t^{-\tau+\mathfrak{g}} \diamond \epsilon_t^{(m,k)} \right\|^2, \frac{1}{\hat{z}_t^\alpha} \right), \quad t \geq t_k, \quad (20)$$

for some innovation $\epsilon_t^{(m+1-\frac{j}{2^k}, k)}$ to be defined below and with initialization

$$\hat{x}_\theta^{(m+1-\frac{j}{2^k})} = \begin{cases} \hat{x}_\theta^{(m+1-\frac{j/2}{2^{k-1}})}, \theta \in [t_k - \delta_{max}, t_k], & \text{if } j = 2, 4, 6, \dots, 2^k, \\ 0, \theta \in [t_k - \delta_{max}, t_k] & \text{else.} \end{cases}$$

This initialization reflects the fact that an observer which run in the chain of layer k will be running in the chain of the next layer (initialized with its past values).

Remark 4.5. *If an observer is a (slave/master) j -th observer in the $(k-1)$ -depth layer, then it will become a (slave/master) $2j$ -th observer in the k -depth layer and it will be active for $t \geq t_k$ (all subsequent layers) and initialized with past estimates over $[t_k - \delta_{max}, t_k]$ (preceding layer). The other (slave) observers become active from $t = t_k$ and initialized arbitrarily.*

Going down the chain of the layer k we find the remaining $2^k(m-1)$ observers, which we refer to as *slave* observers,

$$\dot{\hat{x}}_t^{(m+1-\frac{j}{2^k})} = \hat{F}(\hat{x}_t^{(m+1-\frac{j}{2^k})}, \hat{w}_t) + P^{-1}(\hat{w}_t) C^T R(\hat{w}_t) \epsilon_t^{(m+1-\frac{j}{2^k}, k)}, \quad t \geq t_k, \quad j = 2^k + 1, \dots, 2^k m, \quad (21)$$

for some innovation $\epsilon_t^{(m+1-\frac{j}{2^k}, k)}$ to be defined below and with initialization

$$\hat{x}_\theta^{(m+1-\frac{j}{2^k})} = \begin{cases} \hat{x}_\theta^{(m+1-\frac{j/2}{2^{k-1}})}, \theta \in [t_k - \delta_{max}, t_k] & \text{if } j = 2^k + 2, 2^k + 4, \dots, 2^k m, \\ 0, \theta \in [t_k - \delta_{max}, t_k] & \text{else.} \end{cases}$$

Remark 4.6. *If an observer is a slave j -th observer in the $(k-1)$ -depth layer, then it will become a slave $2j$ -th observer in the k -depth layer and it will be active for $t \geq t_k$ (all subsequent layers) and initialized with past estimates over $[t_k - \delta_{max}, t_k]$ (preceding layer). The other (slave) observers become active from $t = t_k$ and initialized arbitrarily.*

An illustration of how the 0- and 1-depth levels are organized is given in Fig. 1 if $m = 3$ and $\delta_t \in [\Delta_1, \Delta_{3/2}]$ for all $t \geq t_0$. Notice that the master observer, labeled 1 in layer 0, becomes the master observer labeled 2 in layer 1 (see vertical dotted arrows). The slave observers, labeled 2 and 3 in layer 0, become the slave observers labeled 4 in layer 6 in layer 1. Slave observers, labeled 1, 3 and 5, are the ones which start running from layer 1 and correspond to the 2-refinement of the partition of layer 0.

The innovation for the j -th observer, $j = 1, \dots, 2^k m$, of the depth- k chain is defined as for $j = 1, \dots, 2^k$

$$\epsilon_t^{(m+1-\frac{j}{2^k}, k)} = y_{t-\Delta^{(k)}_{m-\frac{j}{2^k}}}^* - H(\hat{x}_t^{(m+1-\frac{j}{2^k})}) \quad (22)$$

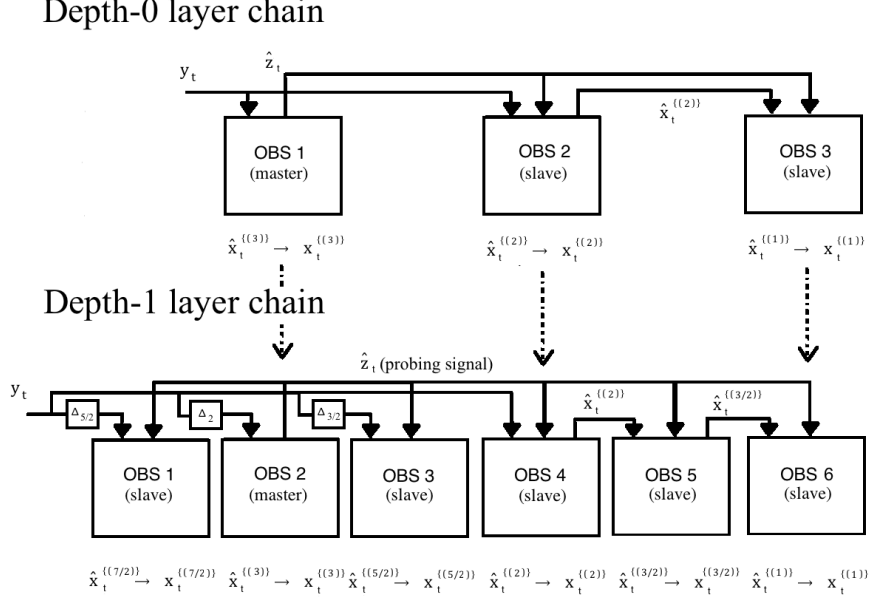


FIG. 1. depth 0 and depth 1 observer chains when $m = 3$ and $\delta_t \in [\Delta_1, \Delta_{3/2}]$.

and for $j = 2^k + 1, \dots, 2^k m$

$$(23) \quad \epsilon_t^{(m+1-\frac{j}{2^k}, k)} = \begin{cases} y_{t-\Delta_{m-\frac{j}{2^k}}^*} - H(\hat{x}_{t-\delta_t}^{(m+1-\frac{j}{2^k}, k)}) & \text{if } \delta_t \in [0, \Delta_{m-\frac{j}{2^k}}^{(k)}], \\ y_{t-\delta_t}^* - H(\hat{x}_{t-\delta_t}^{(m+1-\frac{j}{2^k}, k)}) & \text{if } \delta_t \in [\Delta_{m-\frac{j}{2^k}}^{(k)}, \Delta_{m-\frac{j-1}{2^k}}^{(k)}], \\ H(\hat{x}_t^{(m+1-\frac{j-1}{2^k})}) - H(\hat{x}_{t-\delta_t}^{(m+1-\frac{j}{2^k}, k)}) & \text{if } \delta_t \in [\Delta_{m-\frac{j-1}{2^k}}^{(k)}, \Delta_m^{(k)}], \end{cases}$$

where

$$(24) \quad \delta_t^{(m+1-\frac{j}{2^k}, k)} = \begin{cases} 0 & \text{if } \delta_t \in [0, \Delta_{m-\frac{j}{2^k}}^{(k)}], \\ \delta_t - \Delta_{m-\frac{j}{2^k}}^{(k)} & \text{if } \delta_t \in [\Delta_{m-\frac{j}{2^k}}^{(k)}, \Delta_{m-\frac{j-1}{2^k}}^{(k)}], \\ \Delta_{m-\frac{j-1}{2^k}}^{(k)} - \Delta_{m-\frac{j}{2^k}}^{(k)} & \text{if } \delta_t \in [\Delta_{m-\frac{j-1}{2^k}}^{(k)}, \Delta_m^{(k)}]. \end{cases}$$

Remark 4.7. Similar remarks for the depth k can be made on the innovations of the k -depth observer chain, keeping in mind that the partition of $[0, a]$ at depth k is a two 2-refinement of the partition at depth $k-1$ so that at layer k the value of

the delay δ_t is compared with smaller intervals (actually, half the old ones). It should be worth noting that, by definition of the partition and its subsequent refinements, $\epsilon_t^{(m+1-\frac{j}{2^{k-1}},k)} = \epsilon_t^{(m+1-\frac{j}{2^{k-1}},k-1)}$ for all $k \geq 1$ and $j = 1, \dots, 2^{k-1}$. In other words, the innovation does not switch from layer $k-1$ to the next one for the first 2^{k-1} observers (see also Fig. 1).

In the overall the predictor equations are given by (8)-(9), (11)-(12), (19)-(21) and (21)-(22).

4.2. Main assumptions and results. Our assumptions on the system (5) are the following (see a review of incremental homogeneity in the appendix A).

(H0) (*incremental homogeneity*)

- (i) $A^T \phi$ is incrementally homogeneous in the upper bound (i.h.u.b.) with quadruples $(\mathbf{r}, \mathbf{r} - \mathbf{g}, \mathbf{g}, A^T \Phi(x', x''))$, with lower triangular $\Phi(0, 0)$,
- (ii) $(I - AA^T)\phi$ is i.h.u.b. with quadruple $(\mathbf{r}, (I - AA^T)(\mathbf{r} + \mathbf{g}), \mathbf{g}, (I - AA^T)\Phi(x', x''))$,
- (iii) the degrees \mathbf{g} and weights \mathbf{r} satisfy for each $j = 2, \dots, n$

$$2(\mathbf{g}_j - \mathbf{g}_{j-1}) + \mathbf{g}_{j-1} + \mathbf{r}_{j-1} \leq \mathbf{r}_j - \mathbf{g}_j \leq \mathbf{g}_{j-1} + \mathbf{r}_{j-1},$$

(H1) (*state boundedness*) $x(\cdot, x_0) \in \mathbf{L}^\infty(\mathbb{R}_{\geq}, \mathbb{R})$ for each $x_0 \in \mathbb{R}^n$, where $x(t, x_0)$ is the solution of (5) with initial condition x_0 .

These assumptions are enough general for coping with genuinely nonlinear systems and were adopted in [1], together with an extra observability condition, for designing state estimators for system with constant measurement delays. It is not difficult to check for assumption (H0). In general, it amounts to solve a set of algebraic inequalities in the unknowns $\mathbf{r} \in \mathbb{R}_{>}^n$ and $\mathbf{g} \in \mathbb{R}^n$.

THEOREM 4. *Under assumptions (H0)-(H1) there exist design parameters $\alpha, \nu \in \mathbb{R}_{>}$, $Z_0 > 1$, $\gamma \in \mathbb{R}_{>}$ and $\Gamma \in \mathbb{R}^{n \times n}$ such that*

- (i) *the solution \hat{z}_t of the multilayer predictor (8)-(9), (11)-(12), (19)-(21) and (21)-(22) is bounded for all $t \geq 0$ and there exists $k^* \geq 1$ such that the number of layers is exactly k^* ,*
- (ii) *the solutions $\hat{x}_t^{(j)}$, $j = 1, \dots, 2^{k^*} m$, of the multilayer predictor (8)-(9), (11)-(12), (19)-(21) and (21)-(22) are bounded for all times and*

$$(25) \quad \lim_{t \rightarrow +\infty} (x_t^{(j)} - \hat{x}_t^{(j)}) = 0.$$

In what follows, we prove two results which are instrumental to prove the main theorem 4. First, we need the following lemma (the proof is omitted for reasons of space and follows from simple but lengthy matrix algebra). Recall that $A \leq B$, $A, B \in \mathbb{R}^{m \times l}$, means $A_{ij} \leq B_{ij}$ for all $i = 1, \dots, m$, $j = 1, \dots, l$, and $\max_{\theta \in \mathcal{N}} A(\theta)$, $A(\theta) \in \mathbb{R}^{m \times l}$ for each $\theta \in \mathbb{R}^n$ and compact $\mathcal{N} \subset \mathbb{R}^n$, represents any matrix M such that $A(\theta) \leq M$ for all $\theta \in \mathcal{N}$. If $\mathcal{N}(\gamma)$ is a family of compact sets $\mathcal{N}(\gamma) \subset \mathbb{R}^n$ for each $\gamma \in \mathbb{R}_{>}^n$ and such that $\mathcal{N}(\gamma) \rightarrow \{0\}$ as $\gamma \rightarrow 0$ then $\max_{\theta \in \mathcal{N}(\gamma)} A(\theta)$ is taken in such a way that $\max_{\theta \in \mathcal{N}(\gamma)} A(\theta) \rightarrow A(0)$ as $\gamma \rightarrow 0$.

LEMMA 5. *Let Φ be as in (H0). For each $\gamma \in \mathbb{R}_{>}^n$ and positive definite diagonal*

$\Gamma \in \mathbb{R}^{n \times n}$ define

$$(26) \quad N(\Gamma) := A + (I + \Gamma A^T)A \sum_{j=1}^{n-1} (\Gamma A^T)^j$$

$$(27) \quad L(\Gamma) := 2^\alpha \max_i |g_i| \Gamma A^T \sum_{j=0}^{n-1} (\Gamma A^T)^j$$

$$(28) \quad M_\gamma(\Gamma) := 2(I + \Gamma A^T) \max_{-\gamma \leq x', x'' \leq \gamma} \Phi(x', x'')(I - \Gamma A^T)^{-1}$$

and

$$(29) \quad K_\gamma(\Gamma) := I - 2\Gamma + N(\Gamma) + M_\gamma(\Gamma) + L(\Gamma) + (N(\Gamma) + M_\gamma(\Gamma) + L(\Gamma))^T.$$

There exists a choice of γ and Γ such that

$$(30) \quad K_\gamma(\Gamma) \leq -I$$

For the stated purposes, consider the “auxiliary” system

$$(31) \quad \dot{x}_t = Ax_t + \phi(x_t), \quad t \geq T - a,$$

$$(32) \quad y_t^{aux} = H(x_t) + d_t, \quad t \geq T,$$

where $a > \delta_{max}$ and $d_t, t \geq T - a$, represents a bounded disturbance or uncertainty and $T \geq 0$. For this auxiliary system we consider two types of observers which correspond in our multilayer predictor to master and, respectively, slave observers and we prove convergence of the observation error under different assumptions. In our context, the disturbance in (32) models the mismatch between the switching innovation ϵ of each (master or slave) observer and a fixed innovation equal to the undelayed estimation error $H(x_t) - H(\hat{x}_t)$. This mismatch varies with the relative position of δ_t with respect to the current partition of $[0, a]$ (see the definition of the innovation processes). In the case of master observers we will have $d_t = 0$ and in the case of slave observers a more complicated term, which however may depend only on the following estimation errors: delayed and undelayed estimation error of the observer itself and delayed estimation error of the preceding observer in the chain (see assumption (H2) in the following proposition). This last error is modeled as an exogenous input for the slave observer ($\hat{\chi}_t$ in assumption (H2)).

Let Φ be as in (H0) and $\gamma \in \mathbb{R}_{>}^n$ and $\Gamma \in \mathbb{R}^{n \times n}$ be selected as in lemma 5 and define

$$(33) \quad \Theta := \|C\Gamma C^T\|(\|Q\| + \|C\Gamma C^T\|)$$

where

$$(34) \quad Q := C \left(A + \max_{-\gamma \leq x', x'' \leq \gamma} \Phi(x', x'') \right) (I - \Gamma A^T)^{-1}.$$

4.2.1. Error convergence for slave observers. The first kind of observer we consider for (31) is a slave-type observer

$$(35) \quad \dot{\hat{x}}_t = \hat{F}(\hat{x}_t, \hat{w}_t) + P^{-1}(\hat{w}_t)C^T R(\hat{w}_t) \left(y_t^{aux} - H(\hat{x}_t) \right), \quad t \geq T,$$

with initialization

$$(36) \quad \hat{x}_\theta = \chi_\theta \in \mathbf{C}^0([T - a, T], \mathbb{R}^n), \quad \theta \in [T - a, T],$$

where \hat{w}_t , $t \geq T - a$, is a bounded exogenous signal. The initialization (36) reflects the fact that the slave-type observer may be running already at a previous layer. We state and prove the following convergence result. Recall that $\langle\langle v \rangle\rangle$, $v \in \mathbb{R}^n$, means $(|v_1|, \dots, |v_n|)^T$.

PROPOSITION 6. Under assumptions (H0)-(H1), if $\gamma \in \mathbb{R}_>^n$ and $\Gamma \in \mathbb{R}^{n \times n}$ are selected as in lemma 5 and if, in addition,

(H2) for all $t \geq T$

$$\langle\langle d_t \rangle\rangle \leq \langle\langle \hat{\chi}_t \rangle\rangle + \langle\langle H(x_t) - H(\hat{x}_t) - (H(x_{t-\tau_t}) - H(\hat{x}_{t-\tau_t})) \rangle\rangle$$

where $\hat{\chi}_t$, $t \geq T$, is a bounded exogenous signal such that

$$(37) \quad \lim_{t \rightarrow +\infty} \hat{\chi}_t = 0$$

and τ_t is a bounded delay such that

$$(H3) \quad 0 \leq \tau_t < \frac{1}{8\Theta \hat{w}_t^{\min_i g_i} + 3|\max_i g_i|} \text{ for all } t \geq T,$$

$$(H4) \quad 0 \leq \frac{\dot{\hat{w}}_t}{\hat{w}_t} \leq \frac{2^{\alpha-1}}{\hat{w}_t^\alpha} \text{ for all } t \geq T \text{ and } \hat{w}_t \geq 1 \text{ for all } t \in [T-a, T], \text{ with } \alpha \geq 2|\max_i g_i| + 1,$$

$$(H5) \quad \lim_{t \rightarrow +\infty} (x_t - \text{sat}(x_t, \gamma \diamond \hat{w}_t^\dagger)) = 0,$$

the solutions (x_t, \hat{x}_t) of (31)-(32)-(35) are bounded for all $t \geq T$ and

$$(38) \quad \lim_{t \rightarrow +\infty} (x_t - \hat{x}_t) = 0.$$

Proof. For simplicity of notation, we write $\text{sat}(\cdot)$ instead of $\text{sat}(\cdot, \gamma \diamond \hat{w}_t^\dagger)$. Let

$$(39) \quad e = x - \hat{x}, \quad \eta := (I - G(\hat{w})A^T)e,$$

$$(40) \quad V(e, w) := e^T P(w)e.$$

We also write V_t instead of $V(e_t, w_t)$, P_t instead of $P(\hat{w}_t)$, R_t instead of $R(\hat{w}_t)$ and G_t instead of $G(\hat{w}_t)$. Moreover, for notational convenience we use the following operator ∇ acting on functions $f \in \mathbf{C}^0(\mathbb{R}^n, \mathbb{R}^n)$

$$(41) \quad \nabla f(\chi_1, \chi_2) := f(\chi_1) - f(\chi_2).$$

All the inequalities involving time functions are meant to hold for $t \geq T$ (if not otherwise stated). We prove proposition 6 by successive steps.

(A) An upper bound for the time derivative of V_t along the trajectories of (31)-(32)-(35). We have using properties B.1 and B.2 of Appendix B

$$\begin{aligned} & \dot{V}_t|_{(31)-(32)-(35)} \\ &= \overbrace{2(\eta_t \diamond \hat{w}_t^{-\mathbf{r}})^T \left\{ \text{diag}\{\hat{w}_t^{-\mathbf{r}}\} (I - G_t A^T) A e_t - C^T C G_t \text{diag}\{\hat{w}_t^{-\mathbf{r}}\} C^T C e_t \right\}}^{(1)} \\ &+ \overbrace{2(\eta_t \diamond \hat{w}_t^{-\mathbf{r}})^T \text{diag}\{\hat{w}_t^{-\mathbf{r}}\} (I - G_t A^T) \nabla(\phi \circ \text{sat})(x_t, x_t - e_t)}^{(2)} \\ &- \overbrace{2(\eta_t \diamond \hat{w}_t^{-\mathbf{r}})^T C^T C G_t \text{diag}\{\hat{w}_t^{-\mathbf{r}}\} C^T d_t}^{(3)} + \overbrace{e_t^T \dot{P}_t e_t}^{(4)} \\ &+ \overbrace{2(\eta_t \diamond \hat{w}_t^{-\mathbf{r}})^T \text{diag}\{\hat{w}_t^{-\mathbf{r}}\} (I - G_t A^T) \nabla \phi(x_t, \text{sat}(x_t))}^{(5)} \end{aligned} \tag{42}$$

We majorize the terms under graphs in the above inequality. Let's begin with the bracketed term (1) in (42). Using properties B.1 and B.2 of Appendix B with (i), (iii) of lemma 10

$$\begin{aligned}
& \overbrace{2(\eta_t \diamond \hat{w}_t^{-\mathfrak{r}})^T \left\{ \text{diag}\{\hat{w}_t^{-\mathfrak{r}}\}(I - G_t A^T) A e_t - C^T C G_t \text{diag}\{\hat{w}_t^{-\mathfrak{r}}\} C^T C e_t \right\}}^{(1)} \\
& \leq 2 \langle \langle \eta_t \diamond \hat{w}_t^{-\mathfrak{r}+\mathfrak{g}} \rangle \rangle^T \left\{ -\Gamma + A + (I + \Gamma A^T) A \sum_{j=1}^{n-1} (\Gamma A^T)^j \right\} \langle \langle \eta_t \diamond \hat{w}_t^{-\mathfrak{r}+\mathfrak{g}} \rangle \rangle \\
(43) \quad & = \langle \langle \eta_t \diamond \hat{w}_t^{-\mathfrak{r}+\mathfrak{g}} \rangle \rangle^T \left\{ -2\Gamma + N(\Gamma) + N^T(\Gamma) \right\} \langle \langle \eta_t \diamond \hat{w}_t^{-\mathfrak{r}+\mathfrak{g}} \rangle \rangle,
\end{aligned}$$

with $N(\Gamma)$ as in (26). Let's consider the bracketed term (2) in (42). Notice that for all $x \in \mathbb{R}^n$ and $\hat{w} \geq 1$

$$(44) \quad -\gamma \leq \hat{w}^{-\mathfrak{r}} \diamond \text{sat}(x) \leq \gamma.$$

On account of (2), using properties B.2 of Appendix B and (ii), (iv) and (v) of lemma 10

$$\begin{aligned}
& \overbrace{2(\eta_t \diamond \hat{w}_t^{-\mathfrak{r}})^T \text{diag}\{\hat{w}_t^{-\mathfrak{r}}\}(I - G_t A^T) \nabla(\phi \circ \text{sat})(x_t, x_t - e_t)}^{(2)} \\
& \leq 4 \langle \langle \eta_t \diamond \hat{w}_t^{-\mathfrak{r}+\mathfrak{g}} \rangle \rangle^T (I + \Gamma A^T) \Phi \Big|_{\substack{x' = \hat{w}_t^{-\mathfrak{r}} \diamond \text{sat}(x_t) \\ x'' = \hat{w}_t^{-\mathfrak{r}} \diamond \text{sat}(\hat{x}_t)}} (I - \Gamma A^T)^{-1} \langle \langle \eta_t \diamond \hat{w}_t^{-\mathfrak{r}+\mathfrak{g}} \rangle \rangle \\
(45) \quad & \leq \langle \langle \eta_t \diamond \hat{w}_t^{-\mathfrak{r}+\mathfrak{g}} \rangle \rangle^T (M_\gamma(\Gamma) + M_\gamma(\Gamma)^T) \langle \langle \eta_t \diamond \hat{w}_t^{-\mathfrak{r}+\mathfrak{g}} \rangle \rangle
\end{aligned}$$

with $M_\gamma(\Gamma)$ as in (28). Next, we majorize the bracketed term (3) in (42). Using Young's inequality

$$\begin{aligned}
& \overbrace{-2(\eta_t \diamond \hat{w}_t^{-\mathfrak{r}})^T C^T C G_t \text{diag}\{\hat{w}_t^{-\mathfrak{r}}\} C^T d_t}^{(3)} \leq \frac{1}{2} \|\eta_t \diamond \hat{w}_t^{-\mathfrak{r}+\mathfrak{g}}\|^2 \\
& + 2 \|C \Gamma \text{diag}\{\hat{w}_t^{-\mathfrak{r}+\mathfrak{g}}\} C^T d_t\|^2.
\end{aligned}$$

Now, we majorize the bracketed term (4) in (42). We have from (vi) of lemma 10 and (H4)

$$\begin{aligned}
& \overbrace{e_t^T \dot{P}_t e_t}^{(4)} \leq -4 \frac{\dot{\hat{w}}_t}{\hat{w}_t} (\eta_t \diamond \hat{w}_t^{-\mathfrak{r}+\mathfrak{g}})^T \text{diag}\{\mathfrak{g} \diamond \hat{w}_t^{-2\mathfrak{g}}\} \text{diag}\{\hat{w}_t^{-\mathfrak{r}+\mathfrak{g}}\} G_t A^T (I - G_t A^T)^{-1} \eta_t \\
& \leq 4 \frac{\dot{\hat{w}}_t}{\hat{w}_t} \max_i |\mathfrak{g}_i| \hat{w}_t^{2|\max_i \mathfrak{g}_i|} \langle \langle \eta_t \diamond \hat{w}_t^{-\mathfrak{r}+\mathfrak{g}} \rangle \rangle^T \left(\Gamma A^T \sum_{j=0}^{n-1} (\Gamma A^T)^j \right) \langle \langle \eta_t \diamond \hat{w}_t^{-\mathfrak{r}+\mathfrak{g}} \rangle \rangle \\
(46) \quad & \leq \langle \langle \eta_t \diamond \hat{w}_t^{-\mathfrak{r}+\mathfrak{g}} \rangle \rangle^T (L^T(\Gamma) + L(\Gamma)) \langle \langle \eta_t \diamond \hat{w}_t^{-\mathfrak{r}+\mathfrak{g}} \rangle \rangle
\end{aligned}$$

with $L(\Gamma)$ as in (27). Finally, the bracketed term (5) in (42) is upper bounded via similar passages used for the bracketed term (2) and using Young's inequality:

$$\begin{aligned}
& \overbrace{2(\eta_t \diamond \hat{w}_t^{-\mathfrak{r}})^T \text{diag}\{\hat{w}_t^{-\mathfrak{r}}\}(I - G_t A^T) \left\{ \phi(x_t) - \phi(\text{sat}(x_t)) \right\}}^{(5)} \leq (1/2) \|\eta_t \diamond \hat{w}_t^{-\mathfrak{r}+\mathfrak{g}}\|^2 \\
& + 2 \left\| (I + \Gamma A^T) \Phi \Big|_{\substack{x' = \hat{w}_t^{-\mathfrak{r}} \diamond x_t \\ x'' = \hat{w}_t^{-\mathfrak{r}} \diamond \text{sat}(x_t)}} \sum_{j=0}^{n-1} (\Gamma A^T)^j \right\} \langle \langle \hat{w}_t^{-\mathfrak{r}+\mathfrak{g}} \diamond (x_t - \text{sat}(x_t)) \rangle \rangle \|^2.
\end{aligned}$$

Collecting all the above majorizations

$$(47) \quad \begin{aligned} \dot{V}|_{(31)-(32)-(35)} &\leq \langle\langle \eta_t \diamond \hat{w}_t^{-\tau+\mathfrak{g}} \rangle\rangle^T K_\gamma(\Gamma) \langle\langle \eta_t \diamond \hat{w}_t^{-\tau+\mathfrak{g}} \rangle\rangle \\ &\quad + 2\|C\Gamma \text{diag}\{\hat{w}_t^{-\tau+\mathfrak{g}}\}C^T d_t\|^2 + \rho_t^{(0)} \end{aligned}$$

with $K_\gamma(\Gamma)$ as in (30) and

$$\rho_t^{(0)} := 2\left\| (I + \Gamma A^T) \Phi \Big|_{\substack{x' = \hat{w}_t^{-\tau} \circ x_t \\ x'' = \hat{w}_t^{-\tau} \circ \text{sat}(x_t)}} \sum_{j=0}^{n-1} (\Gamma A^T)^j \langle\langle \hat{w}_t^{-\tau+\mathfrak{g}} \diamond (x_t - \text{sat}(x_t)) \rangle\rangle \right\|^2$$

bounded and such that $\lim_{t \rightarrow +\infty} \rho_t^{(0)} = 0$ (on account of (H4)-(H5) and (H0)). But $\gamma \in \mathbb{R}_{>}^n$ and $\Gamma \in \mathbb{R}^{n \times n}$ are selected as in lemma 5 so that $K_\gamma(\Gamma) \leq -I$. As a result of (47)

$$(48) \quad \dot{V}|_{(31)-(32)-(35)} \leq -\|\eta_t \diamond \hat{w}_t^{-\tau+\mathfrak{g}}\|^2 + 2\hat{w}_t^{2|\max_i \mathfrak{g}_i|} \|C\Gamma \text{diag}\{\hat{w}_t^{-\tau}\}C^T d_t\|^2 + \rho_t^{(0)}.$$

Now, we compute an upper bound for the second term on the left of (48). Let $\Pi_t := C\Gamma \text{diag}\{\hat{w}_t^{-\tau}\}C^T$ and $\tau_{max} := \frac{\beta}{6\Theta}$. Notice that Π_t is diagonal and positive definite for each t . By (H2)

$$(49) \quad \langle\langle \Pi_t d_t \rangle\rangle \leq \Pi_t \langle\langle \hat{\chi}_t \rangle\rangle + \Pi_t \int_{t-\tau_t}^t \left\langle\left\langle \frac{d}{ds} (H(x_s) - H(\hat{x}_s)) \right\rangle\right\rangle ds$$

for all $t \geq \max\{T, \tau_{max}\}$ (τ_{max} is an upper bound for τ_t by (H4)), recalling that $\int_{t''}^{t'} \frac{d}{ds} v_s ds = v_{t'} - v_{t''}$. On account of properties B.3 of Appendix B, for all $s \leq t$

$$(50) \quad \Pi_s C P_s^{-1} C^T R \Pi_s^{-1} \leq \hat{w}_t^{2|\max_i \mathfrak{g}_i|} C \Gamma C^T, \quad \Pi_t \Pi_s^{-1} \leq I.$$

Continuing from (49) and applying once again (H2) it follows that

$$(51) \quad \begin{aligned} \langle\langle \Pi_t d_t \rangle\rangle &\leq \int_{t-\tau_t}^t \overbrace{\Pi_s C \left\langle\left\langle \frac{de_s}{ds} + P_s^{-1} C^T R_s d_s \right\rangle\right\rangle}^{(7)} ds \\ &\quad + \hat{w}_t^{2|\max_i \mathfrak{g}_i|} C \Gamma C^T \int_{t-\tau_t}^t \overbrace{\Pi_s \left(\langle\langle C e_s \rangle\rangle + \langle\langle C e_{s-\tau_s} \rangle\rangle \right)}^{(8)} ds + \rho_t^{(1)} \end{aligned}$$

where

$$(52) \quad \rho_t^{(1)} := C \Gamma C^T \left(\langle\langle \hat{\chi}_t \rangle\rangle + \hat{w}_t^{2|\max_i \mathfrak{g}_i|} C \Gamma C^T \int_{t-\tau_t}^t \langle\langle \hat{\chi}_s \rangle\rangle ds \right)$$

is a bounded signal such that $\lim_{t \rightarrow +\infty} \rho_t^{(1)} = 0$ (on account of (H5), boundedness from above and below of \hat{w}_t by (H4) and boundedness of τ_t by (H3)).

Next, we obtain a majorant for the bracketed terms (7) and (8) in (51). For the term (7) for all $s \in [t - \tau_t, t]$ we have, on using properties B.3 of Appendix B,

$$(53) \quad \begin{aligned} &\Pi_s C \left\langle\left\langle \frac{de_s}{ds} + P_s^{-1} C^T R_s d_s \right\rangle\right\rangle \\ &\leq \Pi_s C \left\langle\left\langle A(I - G_s A^T)^{-1} \eta_s + \phi(\text{sat}(x_s)) - \phi\left(\text{sat}\left(x_s - (I - G_s A^T)^{-1} \eta_s\right)\right)\right\rangle\right\rangle \\ &+ \Pi_s C \Gamma \text{diag}\{\hat{w}_s^{2\mathfrak{g}}\} C^T \langle\langle C \eta_s \rangle\rangle + \Pi_s C \langle\langle \phi(x_s) - \phi(\text{sat}(x_s)) \rangle\rangle \end{aligned}$$

With (44) and (ii) and (vi) of lemma 10 we obtain for all $s \in [t - \tau_t, t]$

$$(54) \quad \Pi_s C \left\langle \left\langle \frac{de_s}{ds} + P_s^{-1} C^T R_s d_s \right\rangle \right\rangle \leq \hat{w}_s^{2|\max_i \mathfrak{g}_i|} C \Gamma C^T Q \langle \eta_s \diamond \hat{w}_s^{-\tau} \rangle + \rho_s^{(2)}$$

with Q as in (34) and bounded

$$\rho_s^{(2)} := \hat{w}_s^{|\max_i \mathfrak{g}_i|} C \Gamma C^T C \Phi \Big|_{\substack{x' = \hat{w}_s^{-\tau} \diamond x_s \\ x'' = \hat{w}_s^{-\tau} \diamond \text{sat}(x_s)}} (I - \Gamma A^T)^{-1} \langle \langle (x_s - \text{sat}(x_s)) \diamond \hat{w}_s^{-\tau+\mathfrak{g}} \rangle \rangle$$

such that $\lim_{t \rightarrow +\infty} \rho_t^{(2)} = 0$ (on account of (H5), boundedness of x_t and boundedness (from above and below) of \hat{w}_t by (H4)).

As to the bracketed term (8) in (51), for all $s \in [t - \tau_t, t]$

$$(55) \quad \Pi_s \langle \langle C e_s \rangle \rangle + \Pi_s \langle \langle C e_{s-\tau_s} \rangle \rangle \leq C \Gamma C^T \left(\langle \langle \eta_t \diamond \hat{w}_t^{-\tau} \rangle \rangle + \langle \langle \eta_{s-\tau_s} \diamond \hat{w}_{s-\tau_s}^{-\tau} \rangle \rangle \right).$$

Finally, using (51), (54) and (55) and recalling the definition of $V_t = \|\eta_t \diamond \hat{w}_t^{-\tau}\|^2$, we get the desired majorant for $\Pi_t d_t$

$$(56) \quad \|\Pi_t d_t\| \leq \hat{w}_t^{2|\max_i \mathfrak{g}_i|} \Theta \int_{t-\tau_t}^t (\sqrt{V_s} + \sqrt{V_{s-\tau_s}}) ds + \|\rho_t^{(1)}\| + \int_{t-\tau_t}^t \|\rho_s^{(2)}\| ds$$

with Θ as in (33) and for all $t \geq \max\{T, \tau_{max}\}$.

(B) Asymptotic convergence of η_t and, therefore, e_t

By Jensen's and Young's inequalities

$$(57) \quad \left(\int_{t-\tau_t}^t (\sqrt{V_s} + \sqrt{V_{s-\tau_s}}) ds \right)^2 \leq 2\tau_t \int_{t-\tau_t}^t (V_s + V_{s-\tau_s}) ds.$$

From (48), (56) and (107), boundedness of \hat{w}_t and Young's inequality it follows that for $t \geq \max\{T, \tau_{max}\}$

$$(58) \quad \dot{V}_t \leq -\hat{w}_t^{-2|\min_i \mathfrak{g}_i|} V_t + 8\hat{w}_t^{6|\max_i \mathfrak{g}_i|} \Theta^2 \tau_t \int_{t-\tau_t}^t (V_s + V_{s-\tau_s}) ds + \rho_t$$

with

$$\rho_t := 4\hat{w}_t^{2|\max_i \mathfrak{g}_i|} \left(\|\rho_t^{(1)}\| + \int_{t-\tau_t}^t \|\rho_s^{(2)}\| ds \right)^2 + \rho_t^{(0)}$$

such that

$$(59) \quad \lim_{t \rightarrow +\infty} \rho_t = 0.$$

We apply Razumikhin theorem to (58) (theorem 2.4 of [4] or theorem 1 of [14]) and for proving asymptotic convergence of η_t we can assume that $t \geq \max\{T, \tau_{max}\}$. To this aim, set

$$(60) \quad \alpha_1(s) := \beta^2 s, \quad \alpha_2(s) := 4(\sup_{t \geq T} \hat{w}_t)^{2|\min_i \mathfrak{g}_i|} s, \quad \alpha_3(s) = \frac{1}{2} (\sup_{t \geq T} \hat{w}_t)^{-2|\min_i \mathfrak{g}_i|} s,$$

with $\beta \in (0, 1)$ such that

$$(61) \quad \tau_t < \frac{\beta}{8\Theta \hat{w}_t^{3|\max_i \mathfrak{g}_i| + |\min_i \mathfrak{g}_i|}}, \quad \forall t \geq T$$

(the choice of β is feasible by (H3)). All the above functions are \mathcal{K} -class functions. Moreover, $\alpha_1(s) < s$ for $s > 0$. If for each $t \geq \max\{T, \tau_{max}\}$

$$V_t \geq \alpha_1(V_{t+\theta}) + \alpha_2(\rho_t), \quad \theta \in [-2\tau_{max}, 0],$$

then on account of (H3)

$$\dot{V}_t \leq -\alpha_3(V_t).$$

By virtue of theorem 2.4 of [4] (or theorem 1 of [14]) it follows

$$(62) \quad \begin{aligned} V_t &\leq \chi_1(\|V\|_{[\max\{T, \tau_{max}\}-2\tau_{max}, \max\{T, \tau_{max}\}], t - \max\{T, \tau_{max}\}}) \\ &+ \chi_2(\|\rho\|_{[\max\{T, \tau_{max}\}, +\infty)}), \quad t \geq \max\{T, \tau_{max}\} \end{aligned}$$

for some $\chi_1 \in \mathcal{KL}$ and $\chi_2 \in \mathcal{K}$ ($\|v\|_I$ denotes $\sup_{t \in I} \|v_t\|$). Actually, we claim that $\lim_{t \rightarrow +\infty} V_t = 0$, which proves our theorem since z_t is bounded from below and above and

$$(63) \quad \lim_{t \rightarrow +\infty} V_t = \lim_{t \rightarrow +\infty} \|\eta_t\| = \lim_{t \rightarrow +\infty} \|e_t\| = 0.$$

Indeed, by passing to the limit for $t \rightarrow +\infty$ in (62)

$$(64) \quad \limsup_{t \rightarrow +\infty} V_t \leq \chi_2(\|\rho\|_{[\max\{T, \tau_{max}\}, +\infty)})$$

Fix $\epsilon > 0$ and let $h > 0$ be such that $\chi_2(h) < \epsilon$. On account of (59) there exists $\bar{T} \geq 0$ such that $\rho_{t+\bar{T}} \leq h$ for all $t \geq \max\{T, \tau_{max}\}$.

$$(65) \quad \limsup_{t \rightarrow +\infty} V_t = \limsup_{t \rightarrow +\infty} V_{t+\bar{T}} \leq \chi_2(\|\rho\|_{[\bar{T}+\max\{T, \tau_{max}\}, +\infty)}) \leq \chi_2(h) < \epsilon$$

which by letting $\epsilon \rightarrow 0$ gives $0 \leq \limsup_{t \rightarrow +\infty} V_t \leq \chi_2(\limsup_{t \rightarrow +\infty} \rho_t) = 0$. \square

4.2.2. Error convergence for master observers. The second kind of observer we consider for (31) is a master-type observer

$$(66) \quad \begin{aligned} \dot{\hat{x}}_t &= A\hat{x}_t + \phi(\text{sat}(\hat{x}_t, \gamma \diamond \hat{w}_t^r)) + P^{-1}(\hat{w}_t)C^T R(\hat{w}_t)(y_t^{aux} - H(\hat{x}_t)), \quad t \geq T, \\ \dot{\hat{z}}_t &= \hat{z}_t \text{sat}\left(\|\hat{z}_t^{-r+\mathfrak{g}} \diamond (\hat{x}_t - \text{sat}(\hat{x}_t, \gamma \diamond \hat{z}_t^r))\|^2 + \|\hat{z}_t^{-r+\mathfrak{g}} \diamond (y_t^{aux} - H(\hat{x}_t))\|^2, \frac{1}{\hat{z}_t^\alpha}\right), \quad t \geq T, \end{aligned}$$

with initialization

$$(67) \quad \hat{x}_\theta = \chi_\theta \in \mathbf{C}^0([T-a, T], \mathbb{R}^n), \quad \hat{z}_\theta = 1, \quad \theta \in [T-a, T],$$

where $a > \delta_{max}$. . The initialization (67) reflects the fact that the master-type observer may be running already at a previous layer. We state and prove the following convergence result.

PROPOSITION 7. *Under assumptions (H0)-(H1) and if $\gamma \in \mathbb{R}_>^n$ and $\Gamma \in \mathbb{R}^{n \times n}$ are selected as in lemma 5, the solutions $(x_t, \hat{x}_t, \hat{z}_t)$ of (31)-(32)-(66), with $d_t = 0$, are bounded for all $t \geq T$ and*

$$(68) \quad \lim_{t \rightarrow +\infty} (x_t - \hat{x}_t) = 0,$$

$$(69) \quad \lim_{t \rightarrow +\infty} (\hat{x}_t - \text{sat}(\hat{x}_t, \gamma \diamond \hat{z}_t^r)) = 0.$$

Proof. We adopt the same notations and simplifications of the proof of proposition 6. Let e , η and V be as in the proof of proposition 6, with \hat{w} replaced by \hat{z} . Properties in the appendices B.2 and B.3 remain valid (with \hat{w} replaced by \hat{z}). The estimates \hat{x}_t and \hat{z}_t are defined over $[T, +\infty)$. In particular, for all $t \geq T$ from (66)

$$(70) \quad 0 \leq \dot{\hat{z}}_t \leq \frac{1}{\hat{z}_t^{\alpha-1}} \leq 1$$

(since $\text{sat}(v, \frac{1}{\hat{z}_t^\alpha}) \leq \frac{1}{\hat{z}_t^\alpha}$ for all $v \in \mathbb{R}$). Following the same lines for the proof of (58) in proposition 6

$$(71) \quad \dot{V}_t \leq -\|\hat{z}_t^{-\tau+\mathfrak{g}} \diamond \eta_t\|^2 + \rho_t$$

with

$$\rho_t := 2 \left\| (I + \Gamma A^T) \Phi \left| \begin{array}{c} x' = \hat{z}_t^{-\tau} \diamond x_t \\ x'' = \hat{z}_t^{-\tau} \diamond \text{sat}(x_t) \end{array} \right. \sum_{j=0}^{n-1} (\Gamma A^T)^j \right\} \left\langle \left\langle \hat{z}_t^{-\tau+\mathfrak{g}} \diamond (x_t - \text{sat}(x_t)) \right\rangle \right\rangle^2$$

Since \hat{z}_t is non-decreasing, we have either $\lim_{t \rightarrow +\infty} \hat{z}_t < +\infty$ or $\lim_{t \rightarrow +\infty} \hat{z}_t = +\infty$. Assume by absurd that $\lim_{t \rightarrow +\infty} \hat{z}_t = +\infty$. By (H0) it follows the existence of $\bar{T} \geq T$ such that

$$(72) \quad x_t - \text{sat}(x_t) = 0, \quad \forall t \geq \bar{T}.$$

and, consequently, $\rho_t = 0, \forall t \geq \bar{T}$. Following (71), for all $t \geq \bar{T}$

$$(73) \quad \dot{V}_t \leq -\|\hat{z}_t^{-\tau+\mathfrak{g}} \diamond \eta_t\|^2.$$

This implies that for all $t \geq \bar{T}$

$$(74) \quad \int_{\bar{T}}^t \|\hat{z}_s^{-\tau+\mathfrak{g}} \diamond \eta_s\|^2 ds \leq V_{\bar{T}}.$$

Next, we find a majorant for \hat{z}_t . To this aim, using once again (72) with (iv) of lemma 10

$$(75) \quad \|\hat{z}_t^{-\tau+\mathfrak{g}} \diamond (\hat{x}_t - \text{sat}(\hat{x}_t))\| \leq 2\|(I - \Gamma A^T)^{-1}\| \|\hat{z}_t^{-\tau+\mathfrak{g}} \diamond \eta_t\|.$$

In the same way,

$$(76) \quad \|\hat{z}_t^{-\tau+\mathfrak{g}} \diamond (y_t^{aux} - H(\hat{x}_t))\| \leq 2\|(I - \Gamma A^T)^{-1}\| \|\hat{z}_t^{-\tau+\mathfrak{g}} \diamond \eta_t\|.$$

Finally, using (74) and (75)-(76) and integrating $\dot{\hat{z}}_t$, for all $t \geq \bar{T}$

$$\hat{z}_t \leq \hat{z}_{\bar{T}} + 8\|(I - \Gamma A^T)^{-1}\|^2 \int_{\bar{T}}^t \|\hat{z}_s^{-\tau+\mathfrak{g}} \diamond \eta_s\| ds \leq \hat{z}_{\bar{T}} + 8\|(I - \Gamma A^T)^{-1}\|^2 V_{\bar{T}}.$$

But this contradicts $\lim_{t \rightarrow +\infty} \hat{z}_t = +\infty$. Therefore, $\lim_{t \rightarrow +\infty} \hat{z}_t < +\infty$.

Next, from (73) with the boundedness of x_t and \hat{z}_t it follows that \hat{x}_t is bounded and uniformly continuous. By integration of $\dot{\hat{x}}_t$ and Barbalat lemma we obtain that $\lim_{t \rightarrow +\infty} (\hat{x}_t - \text{sat}(\hat{x}_t)) = 0$ which proves the first claim of our proposition and gives that $\lim_{t \rightarrow +\infty} \rho_t = 0$. According to the same lines of the proof of (63) in proposition 6

$$\lim_{t \rightarrow +\infty} V_t = \lim_{t \rightarrow +\infty} \eta_t = \lim_{t \rightarrow +\infty} e_t = 0$$

which proves the second part of our proposition. \square

4.2.3. Proof of theorem 4. We will use throughout the notation $e_t^{(j)} := x_t^{(j)} - \hat{x}_t^{(j)}$. Let $\gamma \in \mathbb{R}_>^n$ and $\Gamma \in \mathbb{R}^{n \times n}$ be selected as in lemma 5. Choose the design parameters Z_0 and ν (see (7)) as follows: set $Z_0 := 2$ (the initial guess for Z_0 may be also greater) and $\nu := \frac{1}{8\Theta}$ with Θ as in (33) and pick any $\alpha \geq 2 \max_i \{\mathbf{g}_i\} + 1$.

Consider the first (master) observer of the 0-depth layer, given by equations (8) for $t \geq [0, t_1]$, t_1 the time at which $\hat{z}_{t_1} = Z_0$. This observer becomes the 2^k -observer in the k -depth layer, $k \geq 1$, with equations given by (19) for $t \geq [t_k, t_{k+1}]$, where t_k is the time at which $\hat{z}_{t_k} = Z_{k-1}$. Moreover, it can be directly checked from (14), (15) that

$$(77) \quad \epsilon_t^{(m,0)} = \epsilon_t^{(m,k)} = H(x_t^{(m)}) - H(\hat{x}_t^{(m)}) = C e_t^{(m)}$$

for all $k > 0$ (see also remark 4.7). In the overall, we can consider all these observers running over the intervals $[t_{k-1}, t_k]$, $k = 1, \dots$, as the same observer running over $[0, +\infty)$.

Taking into account (77), we notice that (8) and (19) (which are the same observer at different layers) have the form (66) if in (66) we let \hat{x}_t correspond to $\hat{x}_t^{(m)}$ and in (31)-(32) we let x_t and d_t correspond, respectively, to $x_t^{(m)}$ and 0. From proposition 7 we conclude that

$$(78) \quad \begin{aligned} \lim_{t \rightarrow +\infty} e_t^{(m)} &= 0, \\ \lim_{t \rightarrow +\infty} (\hat{x}_t^{(m)} - \text{sat}(\hat{x}_t^{(m)}, \gamma \diamond \hat{z}_t^{\mathbf{r}})) &= 0 \end{aligned}$$

and \hat{z}_t is bounded. From (78) it also follows that

$$(79) \quad \lim_{t \rightarrow +\infty} (x_t^{(m)} - \text{sat}(x_t^{(m)}, \gamma \diamond \hat{z}_t^{\mathbf{r}})) = 0.$$

On account of the algorithm which activates each subsequent layer, \hat{z}_t bounded implies that the sequence $\{t_j\}$ of layer activation times is finite and the number of activated layers is finite. Let's denote this number by k^* and set $t_{k^*+1} = +\infty$.

Consider the first $2^k - 1$ observers of each k -depth layer, given by equations (19) for $t \geq [t_k, t_{k+1}]$, $1 \leq k \leq k^*$. The j -th one of these observers, $j = 1, \dots, 2^k - 1$, becomes the $(2^{h-k}j)$ -th observer in the h -depth layer, $h > k$, with equations given by (19) for $t \geq [t_h, t_{h+1}]$. Moreover, it can be directly checked from (23), (24) that

$$(80) \quad \epsilon_t^{(m+1-\frac{j}{2^k}, k)} = \epsilon_t^{(m+1-\frac{j}{2^k}, h)} = H(x_t^{(m+1-\frac{j}{2^k})}) - H(\hat{x}_t^{(m+1-\frac{j}{2^k})}) = C e_t^{(m+1-\frac{j}{2^k})}$$

for all $j = 1, \dots, 2^k - 1$ and $1 \leq k \leq k^*$ (see also remark 4.7). In the overall, we can consider all these observers running over the intervals $[t_h, t_{h+1}]$, $h = k, \dots, k^*$, as the same observer running over $[t_k, +\infty)$.

We notice that (19) has the form (35) if in (35) we let \hat{x}_t and \hat{w}_t correspond, respectively, to $\hat{x}_t^{(m+1-\frac{j}{2^k})}$ and \hat{z}_t and in (31)-(32) x_t and d_t correspond, respectively, to $x_t^{(m+1-\frac{j}{2^k})}$ and 0. Moreover, since $x_t^{(m+1-\frac{j}{2^k})} = x_{t-\Delta_{m-\frac{j}{2^k}}+\Delta_{m-1}}^{(m)}$, from (78) we have

$$(81) \quad \lim_{t \rightarrow +\infty} (\hat{x}_t^{(m+1-\frac{j}{2^k})} - \text{sat}(\hat{x}_t^{(m+1-\frac{j}{2^k})}, \gamma \diamond \hat{z}_{t-\Delta_{m-\frac{j}{2^k}}+\Delta_{m-1}}^{\mathbf{r}})) = 0.$$

Since \hat{z}_t and the sequence $\{\Delta_j\}$ are increasing and $-\Delta_{m-\frac{j}{2^k}} + \Delta_{m-1} < 0$ for $j = 1, \dots, 2^k - 1$,

$$(82) \quad \hat{z}_{t-\Delta_{m-\frac{j}{2^k}}+\Delta_{m-1}} \leq \hat{z}_t$$

for all $t \geq t_k$ and $j = 1, \dots, 2^k - 1$. Therefore,

$$(83) \quad \lim_{t \rightarrow +\infty} (\hat{x}_t^{(m+1-\frac{j}{2^k})} - \text{sat}(\hat{x}_t^{(m+1-\frac{j}{2^k})}, \gamma \diamond \hat{z}_t^r)) = 0.$$

for $j = 1, \dots, 2^k - 1$. Moreover, (see (70) and recall the choice of α at the beginning of the proof) for all $t \geq t_k$

$$(84) \quad 0 \leq \frac{\dot{\hat{z}}_t}{\hat{z}_t} \leq \frac{1}{\hat{z}_t^\alpha}$$

We see that (H2)-(H5) of proposition 6 are met with $\tau_t = 0$ so that we conclude that

$$(85) \quad \lim_{t \rightarrow +\infty} e_t^{(m+1-\frac{j}{2^k})} = 0, \quad j = 1, \dots, 2^k - 1.$$

Finally, consider the last $2^k m - 2^k$ observers of each k -depth layer, given by equations (21) for $t \geq [t_k, t_{k+1})$, $0 \leq k \leq k^*$. The j -th one of these observers, $j = 2^k + 1, \dots, 2^k m$, becomes the $(2^{h-k}j)$ -th observer in the h -depth layer, $h > k$, with equations given by (21) for $t \geq [t_h, t_{h+1})$. In the overall, we can consider all these observers running over the intervals $[t_h, t_{h+1})$, $h = k, \dots, k^*$, as the same observer running over $[t_k, +\infty)$ with switching innovations $\epsilon_t^{(m+1-\frac{j}{2^k}, h)}$, $h = k, \dots, k^*$. Again, we notice that (21) has the form (35) if in (35) we let \hat{x}_t correspond, respectively, to $\hat{x}_t^{(m+1-\frac{j}{2^k})}$ and in (31)-(32) x_t , d_t and τ_t correspond, respectively, to $x_t^{(m+1-\frac{j}{2^k})}$,

$$d_t = C \begin{cases} 0 & \text{if } \delta_t \in [0, \Delta_{m-\frac{j}{2^k}}^{(h)}], \quad j = 2^k + 1, \dots, 2^k m - 1, \\ -e_t^{(m+1-\frac{j}{2^k})} + e_{t-\delta_t}^{(m+1-\frac{j}{2^k}, h)} & \text{if } \delta_t \in [\Delta_{m-\frac{j}{2^k}}^{(h)}, \Delta_{m-\frac{j-1}{2^k}}^{(h)}], \quad j = 2^k + 1, \dots, 2^k m, \\ -e_t^{(m+1-\frac{j}{2^k})} + e_{t-\delta_t}^{(m+1-\frac{j}{2^k}, h)} + e_t^{(m+1-\frac{j-1}{2^k})} & \text{if } \delta_t \in [\Delta_{m-\frac{j-1}{2^k}}^{(h)}, \Delta_m^{(h)}], \quad j = 2^k + 1, \dots, 2^k m, \end{cases}$$

and

$$(86) \quad \tau_t = \begin{cases} 0 & \text{if } \delta_t \in [0, \Delta_{m-\frac{j}{2^k}}^{(h)}], \quad j = 2^k + 1, \dots, 2^k m - 1, \\ \delta_t - \Delta_{m-\frac{j}{2^k}}^{(h)} & \text{if } \delta_t \in [\Delta_{m-\frac{j}{2^k}}^{(h)}, \Delta_{m-\frac{j-1}{2^k}}^{(h)}], \quad j = 2^k + 1, \dots, 2^k m, \\ \Delta_{m-\frac{j-1}{2^k}}^{(h)} - \Delta_{m-\frac{j}{2^k}}^{(h)} & \text{if } \delta_t \in [\Delta_{m-\frac{j-1}{2^k}}^{(h)}, \Delta_m^{(h)}], \quad j = 2^k + 1, \dots, 2^k m. \end{cases}$$

Indeed, from (23) when $\delta_t \in [0, \Delta_{m-\frac{j}{2^k}}^{(h)})$

$$(87) \quad \epsilon_t^{(m+1-\frac{j}{2^k}, h)} = H(x_t^{(m+1-\frac{j}{2^k})}) - H(\hat{x}_t^{(m+1-\frac{j}{2^k})}) = Ce_t^{(m+1-\frac{j}{2^k})}.$$

When $\delta_t \in [\Delta_{m+1-\frac{j}{2^k}}^{(h)}, \Delta_{m-\frac{j-1}{2^k}}^{(h)}]$

$$(88) \quad \epsilon_t^{(m+1-\frac{j}{2^k}, h)} = H(x_t^{(m+1-\frac{j}{2^k})}) - H(\hat{x}_{t-\delta_t}^{(m+1-\frac{j}{2^k}, h)}) = Ce_t^{(m+1-\frac{j}{2^k})}$$

and when $\delta_t \in [\Delta_{m-\frac{j-1}{2^k}}^{(h)}, \Delta_m^{(h)}]$

$$(89) \quad \begin{aligned} \epsilon_t^{(m+1-\frac{j}{2^k}, h)} &= H(\hat{x}_t^{(m+1-\frac{j-1}{2^k})}) - H(\hat{x}_{t-\delta_t}^{(m+1-\frac{j}{2^k}, h)}) \\ &= -Ce_t^{(m+1-\frac{j-1}{2^k})} + Ce_t^{(m+1-\frac{j}{2^k}, h)}. \end{aligned}$$

We can see that (H2)-(H5) of proposition 6 are met. Indeed, by definition of \hat{w}_t (see (12)) and (70)

$$(90) \quad \frac{\dot{\hat{w}}_t}{\hat{w}_t} = \frac{2^{\frac{\alpha-1}{\alpha}}}{\hat{w}_t} \left(\hat{z}_t^\alpha + \alpha \delta_{max}^\alpha \right)^{\frac{1-\alpha}{\alpha}} \hat{z}_t^{\alpha-1} \dot{\hat{z}}_t = \frac{2^{\alpha-1}}{\hat{w}_t^\alpha} \hat{z}_t^{\alpha-1} \dot{\hat{z}}_t \leq \frac{2^{\alpha-1}}{\hat{w}_t^\alpha}$$

for all $t \geq t_k$, which is (H3). Also, with τ_t as in (86)

$$(91) \quad 0 \leq \tau_t < \frac{1}{8\Theta \hat{w}_t^{|\min_i \mathfrak{g}_i|+3} |\max_i \mathfrak{g}_i|}$$

for all $t \geq t_k$. Indeed, for each $j = 3, \dots, m$ (layer 0) and for all $t \in [0, t_1)$

$$(92) \quad \begin{aligned} \tau_t &\leq \Delta_{m-j+1}^{(0)} - \Delta_{m-j}^{(0)} \leq \max_{i=1, \dots, m} (\Delta_i^{(0)} - \Delta_{i-1}^{(0)}) \\ &< \frac{\nu}{\left[2^{\frac{\alpha-1}{\alpha}} \left(Z_0^\alpha + \alpha \delta_{max}^\alpha \right)^{\frac{1}{\alpha}} \right]^{|\min_i \mathfrak{g}_i|+3} |\max_i \mathfrak{g}_i|} \\ &< \frac{\nu}{\left[2^{\frac{\alpha-1}{\alpha}} \left(\hat{z}_t^\alpha + \alpha \delta_{max}^\alpha \right)^{\frac{1}{\alpha}} \right]^{|\min_i \mathfrak{g}_i|+3} |\max_i \mathfrak{g}_i|} = \frac{1}{8\Theta \hat{w}_t^{|\min_i \mathfrak{g}_i|+3} |\max_i \mathfrak{g}_i|} \end{aligned}$$

on account of (7), definition of ν and τ_t in (86) and since $\hat{z}_t < Z_0$ for $t \in [0, t_1)$ (by definition of t_1). Moreover, for each $k \geq 1$ and $j = 2^k + 1, \dots, 2^k m$ (layer k) and for all $t \in [t_k, t_{k+1})$

$$(93) \quad \begin{aligned} \tau_t &\leq \Delta_{m-\frac{j-1}{2^k}}^{(k)} - \Delta_{m-\frac{j}{2^k}}^{(k)} \leq \frac{1}{2^k} \max_{i=1, \dots, m} (\Delta_i^{(0)} - \Delta_{i-1}^{(0)}) \\ &< \frac{1}{2^k} \frac{\nu}{\left[2^{\frac{\alpha-1}{\alpha}} \left(Z_0^\alpha + \alpha \delta_{max}^\alpha \right)^{\frac{1}{\alpha}} \right]^{|\min_i \mathfrak{g}_i|+3} |\max_i \mathfrak{g}_i|} \\ &< \frac{\nu}{\left[2^{\frac{\alpha-1}{\alpha}} \left(Z_k^\alpha + \alpha \delta_{max}^\alpha \right)^{\frac{1}{\alpha}} \right]^{|\min_i \mathfrak{g}_i|+3} |\max_i \mathfrak{g}_i|} \\ &< \frac{\nu}{\left[2^{\frac{\alpha-1}{\alpha}} \left(\hat{z}_t^\alpha + \alpha \delta_{max}^\alpha \right)^{\frac{1}{\alpha}} \right]^{|\min_i \mathfrak{g}_i|+3} |\max_i \mathfrak{g}_i|} = \frac{1}{8\Theta \hat{w}_t^{|\min_i \mathfrak{g}_i|+3} |\max_i \mathfrak{g}_i|} \end{aligned}$$

on account of (7), (18), definition of ν and τ_t and since $\hat{z}_t < Z_k$ for $t \in [t_k, t_{k+1})$ (by definition of t_{k+1}) where $Z_k := 2^{\frac{k}{|\min_i \mathfrak{g}_i| + 3} |\max_i \mathfrak{g}_i|} Z_0$ by (17). This proves (H4).

Finally, we must prove (H5). Notice that by (70) and since $-\Delta_{m-\frac{j}{2^k}} + \Delta_{m-1} > 0$ for $j = 2^k + 1, \dots, 2^k m$,

$$\hat{z}_{t-\Delta_{m-\frac{j}{2^k}}+\Delta_{m-1}} \leq \left(\hat{z}_t^\alpha + \alpha(-\Delta_{m-\frac{j}{2^k}} + \Delta_{m-1}) \right)^{\frac{1}{\alpha}} \leq 2^{\frac{\alpha-1}{\alpha}} \left(\hat{z}_t^\alpha + \alpha \delta_{max}^\alpha \right)^{\frac{1}{\alpha}} = \hat{w}_t$$

Using once again (81)

$$\begin{aligned} 0 &= \lim_{t \rightarrow +\infty} \left(\hat{x}_t^{(m+1-\frac{j}{2^k})} - \text{sat}(\hat{x}_t^{(m+1-\frac{j}{2^k})}, \gamma \diamond \hat{z}_{t-\Delta_{m-\frac{j}{2^k}}+\Delta_{m-1}}^f) \right) \\ (94) \quad &= \lim_{t \rightarrow +\infty} \left(\hat{x}_t^{(m+1-\frac{j}{2^k})} - \text{sat}(\hat{x}_t^{(m+1-\frac{j}{2^k})}, \gamma \diamond \hat{w}_t^f) \right) \end{aligned}$$

for $j = 2^k + 1, \dots, 2^k m$, which proves (H5).

5. Example and simulations. For testing our predictor we consider a Van Der Pol oscillator

$$(95) \quad \begin{aligned} \dot{x}_{1,t} &= x_{2,t} \\ \dot{x}_{2,t} &= -x_{1,t} + (1 - x_{1,t}^2)x_{2,t}, \quad y = x_{1,t-\delta_t} \end{aligned}$$

where the measurement delay δ_t (Fig. 2) models the following situation: the measurements are taken over intervals of the form $[1.1k, 1.1k + 1]$ for $k = 0, 1, \dots$ at low rate 1 and are supplied at a high rate 10 during the time interval $[1.1k + 1, 1.1(k + 1)]$. By approximating δ_t with a (mean) constant delay (using techniques in [2]) we would have a significant error in the state prediction, especially during the time intervals with high rate delivery. The maximum delay is 1 so that $\delta_{max} = 1$ and $a = 1.3$. System (95) satisfies assumptions (H0) and (H1) of theorem 4 with $\mathfrak{r}_1 = 1/4$, $\mathfrak{r}_2 = 3/4$, $\mathfrak{g}_1 = \mathfrak{g}_2 = 1/4$.

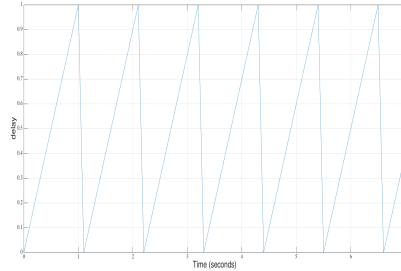


FIG. 2. Time-varying delay δ_t .

A state predictor has been designed according to our procedure and a first simulation has been worked out with initial conditions $x(0) = (-5, -4)^T$. With these state initial conditions an observer chain with three observers ($m = 3$) is sufficient to estimate the state $x(t)$. The interval $[0, a]$ has been partitioned in four parts ($\Delta_3^{(0)} = 1.5, \Delta_2^{(0)} = 1, \Delta_1^{(0)} = 0.5$ and $\Delta_0^{(0)} = 0$), corresponding to the estimates $\hat{\xi}_t^{(4)}, \hat{\xi}_t^{(3)}, \hat{\xi}_t^{(2)}$ and $\hat{\xi}_t^{(1)}$. The saturation levels of the estimates are set with $\gamma = 0.1$,

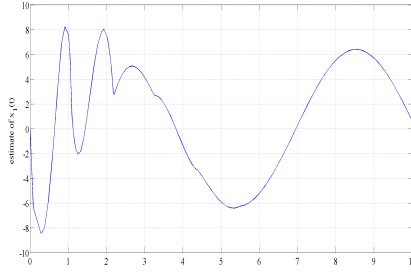
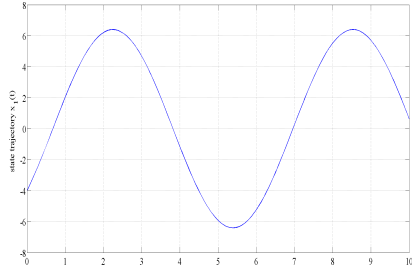


FIG. 3. State trajectory $x_{1,t}$ and estimate $\xi_{1,t}$ ($x_{1,-\delta_{max}} = -5$ and $x_{2,-\delta_{max}} = -4$).

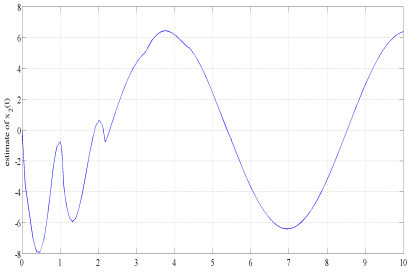
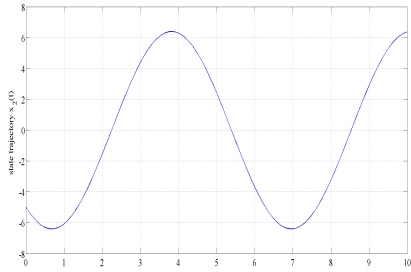


FIG. 4. State trajectory $x_{2,t}$ and estimate $\xi_{2,t}$ ($x_{1,-\delta_{max}} = -5$ and $x_{2,-\delta_{max}} = -4$).

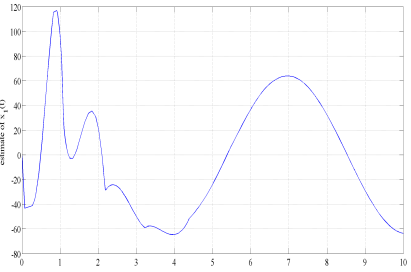
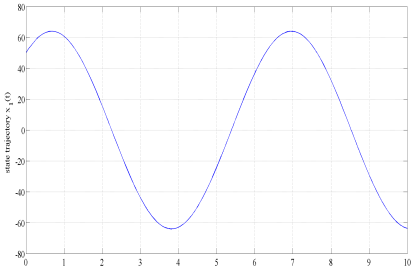


FIG. 5. State trajectory $x_{1,t}$ and estimate $\xi_{1,t}$ ($x_{1,-\delta_{max}} = -40$ and $x_{2,-\delta_{max}} = 50$).

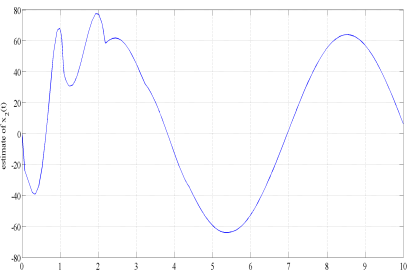
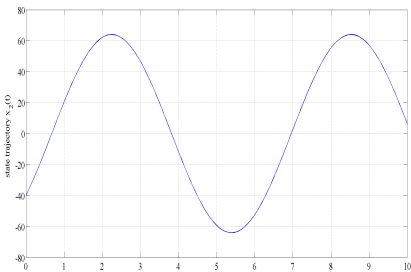


FIG. 6. State trajectory $x_{2,t}$ and estimate $\xi_{2,t}$ ($x_{1,-\delta_{max}} = -40$ and $x_{2,-\delta_{max}} = 50$).

the diagonal elements of Γ are respectively 30 and 100 (see lemma 5). Moreover, we set $Z_0 = 50$. The reference signal \hat{z} tends to the constant value ≈ 8.09 and there is no switching from depth 0 to 1 since $\hat{z}_t < Z_0$. The states $x_{1,t}, x_{2,t}$ together with their estimates $\hat{\xi}_{1,t}^{(1)}, \hat{\xi}_{2,t}^{(1)}$, are shown versus time in Figs. 3 and 4.

Next, a state predictor has been designed according to our procedure and a first simulation has been worked out with initial conditions $x(0) = (-5, -4)^T$. The state initial conditions are very large now and the reference signal \hat{z} tends to the constant value ≈ 80 so that we have a switching from the depth-0 chain to the depth-1 chain since \hat{z}_t crosses Z_0 . Therefore, the 3-partition of $[0, a]$ is refined into a 6-partition. The interval $[0, a]$ is next partitioned in 6 parts ($\Delta_3^{(1)} = 1.5, \Delta_{5/2}^{(1)} = 1.25, \Delta_2^{(1)} = 1, \Delta_{3/2}^{(1)} = 0.75, \Delta_1^{(1)} = 0.5, \Delta_{1/2}^{(1)} = 0.25$ and $\Delta_0 = 0$), corresponding to the estimates $\hat{\xi}_t^{(4)}, \hat{\xi}_t^{(7/2)}, \hat{\xi}_t^{(3)}, \hat{\xi}_t^{(5/2)}, \hat{\xi}_t^{(2)}, \hat{\xi}_t^{(3/2)}$ and $\hat{\xi}_t^{(1)}$. The states $x_{1,t}, x_{2,t}$ together with their estimates $\hat{\xi}_{1,t}^{(1)}, \hat{\xi}_{2,t}^{(1)}$, are shown versus time in Figs. 5 and 6.

6. Conclusions. We have proposed a state predictor for stable nonlinear systems with time-varying measurement delay and weak limitations on the growth of the nonlinearities. The measurement delay is assumed to be continuous. The state prediction consists of chains of nonlinear dynamic observers operating at different layers. On each layer, these observers reconstruct the unmeasurable state vector at different delayed time-instants, which partition the maximal variation interval of the time-varying delay. Transitions from a layer to the next one are triggered by a on-line non-decreasing estimate of the magnitude of the state: as this on-line estimate doubles its value the partition of the maximal variation interval of the time-varying delay is refined by the same factor and, correspondingly, the number of observers is doubled. Further research will be devoted to the case of discontinuous time-varying delay.

Appendix A. Incremental homogeneity.

The notion of (incremental) homogeneity has been introduced in [1] in the context of semi-global stabilization and observer design problems. Here we recall this notion in a slightly more general form.

DEFINITION 8. A parametrized function $\phi_z \in \mathbf{C}^0(\mathbb{R}^n, \mathbb{R}^l)$, $z \in \mathbb{R}_{>}$, is said to be incrementally homogeneous (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 $z > 0$ and $x', x'' \in \mathbb{R}^n$

$$(96) \quad \phi_z(z^{\mathfrak{r}} \diamond x') - \phi_z(z^{\mathfrak{r}} \diamond x'') = z^{\mathfrak{d}} \diamond \left(\Phi(x', x'') \left(z^{\mathfrak{h}} \diamond (x' - x'') \right) \right)$$

In few words, the increment of ϕ_z between two dilated points $z^{\mathfrak{r}} \diamond x'$ and $z^{\mathfrak{r}} \diamond x''$ behaves “homogeneously” in the sense that it is equal to the image of a linear operator $\phi(x', x'') \in \mathbb{R}^{l \times n}$ under the increment between the two dilated points $z^{\mathfrak{h}} \diamond w'$ and $z^{\mathfrak{h}} \diamond w''$, followed by a componentwise dilation by $z^{\mathfrak{d}}$. The vector $\mathfrak{d} \in \mathbb{R}^l$ describes the “vertical” degrees and the vector $\mathfrak{h} \in \mathbb{R}^n$ describes the “horizontal” degrees.

There are functions, like $\sin x$, which are not i.h. but behaves in the upper bound as an i.h. function. This motivates the following definition ($\langle\langle a \rangle\rangle$ denotes the column vector of the absolute values of the elements of $a \in \mathbb{R}^n$).

DEFINITION 9. A parametrized function $\phi_z \in \mathbf{C}^0(\mathbb{R}^n, \mathbb{R}^l)$, $z \in \mathbb{R}_{>}$, is said to be incrementally homogeneous in the upper bound (i.h.u.b.) with quadruple $(\mathfrak{r}, \mathfrak{d}, \mathfrak{h}, \Phi(x', x''))$ if there exist $\mathfrak{d} \in \mathbb{R}^l, \mathfrak{h} \in \mathbb{R}^n, \mathfrak{r} \in \mathbb{R}_{>}^n, \Phi \in \mathbf{C}^0(\mathbb{R}^n \times \mathbb{R}^n, \mathbb{R}_{\geq}^{l \times n})$ such that for all $z \geq 1$

and $x', x'' \in \mathbb{R}^n$

$$(97) \quad \langle\langle \phi_z(z^\tau \diamond x') - \phi_z(z^\tau \diamond x'') \rangle\rangle \leq z^\mathfrak{d} \diamond \left(\Phi(x', x'') \langle\langle z^\mathfrak{h} \diamond (x' - x'') \rangle\rangle \right)$$

When in (97) only the pairs (x', x'') such that $x'' = 0$ are considered and $\phi_z(0) = 0$ we will simply say that ϕ_z is homogeneous in the upper bound (h.u.b.) with quadruple $(\mathfrak{r}, \mathfrak{d}, \mathfrak{h}, \Phi'(x))$, where $\Phi'(x) = \Phi(x, 0)$.

The function $\phi_z(x) = z(x_2 \ x_2^3 g(x_1))^T$, $g \in \mathbf{C}^0(\mathbb{R}, \mathbb{R})$ any bounded and globally Lipschitz function, is i.h.u.b. with triple $(\mathfrak{r}, \mathfrak{d}, \mathfrak{h}, \Phi(x', x''))$, where $\mathfrak{r} := (1, 2)^T$, $\mathfrak{d} := (3, 7)^T$, $\mathfrak{h} := (1, 0)^T$ and the matrix $\Phi(x', x'')$ defined as

$$\begin{aligned} [\Phi(x', x'')]_{11} &:= 0, [\Phi(x', x'')]_{12} := 1, \\ [\Phi(x', x'')]_{21} &:= (x_2'')^3 \frac{|g(x_1') - g(x_1'')|}{|x_1' - x_1''|}, \\ [\Phi(x', x'')]_{22} &:= |(x_2')^2 + (x_2'')^2 + x_2' x_2''| |g(x_1')|. \end{aligned}$$

Clearly, ϕ_z is also h.u.b. with triple $(\mathfrak{r}, \mathfrak{d}, \mathfrak{h}, \Phi'(x))$, where $\Phi'(x) = \Phi(x, 0)$.

Properties of incremental homogeneity can be found in [1].

The following facts can be proved by using extensively the definition and properties of incremental homogeneity and (H0). Let A and ϕ as in (5) with $G \in \mathbb{R}^{n \times n}$ a positive definite diagonal matrix which is i.h. with quadruple $(\mathfrak{r}, \mathfrak{r} + \mathfrak{g}, \mathfrak{g}, \Gamma)$, Γ diagonal positive definite (we say that a matrix D is i.h. or i.h.u.b. in the sense that the associated linear function $\phi(x) = Dx$ is i.h. or i.h.u.b.).

LEMMA 10. Assume (H0).

- (i) $A^T A G A^T A$ is i.h.u.b. with quadruple $(\mathfrak{r}, \mathfrak{r} + \mathfrak{g}, \mathfrak{g}, A^T A G A^T A)$,
- (ii) $A^T \phi$ is i.h.u.b. with quadruple $(\mathfrak{r}, \mathfrak{r} - \mathfrak{g}, \mathfrak{g}, A^T \Phi(x', x''))$ and ϕ is i.h.u.b. with quadruple $(\mathfrak{r}, \mathfrak{r} + \mathfrak{g}, \mathfrak{g}, \Phi(x', x''))$,
- (iii) $A + (I - G A^T) A \sum_{j=1}^{n-1} (G A^T)^j$ is i.h.u.b. with quadruple $(\mathfrak{r}, \mathfrak{r} + \mathfrak{g}, \mathfrak{g}, A + (I + \Gamma A^T) A \sum_{j=1}^{n-1} (\Gamma A^T)^j)$,
- (iv) $(I - G A^T)^{-1}$ is i.h.u.b. with quadruple $(\mathfrak{r}, \mathfrak{r} - \mathfrak{g}, \mathfrak{g}, (I - \Gamma A^T)^{-1})$,
- (v) $G A^T (I - G A^T)^{-1}$ is i.h.u.b. with quadruple $(\mathfrak{r}, \mathfrak{r} - \mathfrak{g}, \mathfrak{g}, \Gamma A^T (I - \Gamma A^T)^{-1})$,
- (vi) $A(I - G A^T)^{-1}$ is i.h.u.b. with quadruple $(\mathfrak{r}, \mathfrak{r} + \mathfrak{g}, \mathfrak{g}, A(I - \Gamma A^T)^{-1})$.

Appendix B. Auxiliary results.

B.1. Properties of (C, A) . For any diagonal $G \in \mathbb{R}^{n \times n}$ and C, A as in (5)-(6):

$$(98) \quad C A^T = 0, C C^T = I$$

$$(99) \quad (G A^T)^j = 0, \forall j \geq n, (I - G A^T)^{-1} = \sum_{j=0}^{n-1} (G A^T)^j$$

$$(100) \quad C G A^T = 0, C(I - G A^T)^{-1} = C, G A^T A = A^T A G A^T A$$

$$(101) \quad C^T C = I - A^T A, A^T A (I - A^T A) = 0, C^T C A^T A = 0$$

B.2. Properties of dilations and \diamond operator. Dilations and operator \diamond have the following useful properties: for all $\eta, w \in \mathbb{R}^n$ and $\mathfrak{a}, \mathfrak{b} \in \mathbb{R}^n$

$$(102) \quad \text{diag}\{w^{\mathfrak{a}+\mathfrak{b}}\} = \text{diag}\{w^{\mathfrak{a}}\} \text{diag}\{w^{\mathfrak{b}}\}$$

$$(103) \quad \eta \diamond w^{\mathfrak{a}+\mathfrak{b}} = (\eta \diamond w^{\mathfrak{a}}) \diamond w^{\mathfrak{b}} = \eta \diamond (w^{\mathfrak{a}} \diamond w^{\mathfrak{b}}) = (\eta \diamond w^{\mathfrak{b}}) \diamond w^{\mathfrak{a}}$$

$$(104) \quad \eta^T \text{diag}\{w^{\mathfrak{a}}\} = (\eta \diamond w^{\mathfrak{a}})^T$$

$$(105) \quad \eta^T \text{diag}\{w^{\mathfrak{a}+\mathfrak{b}}\} = (\eta \diamond w^{\mathfrak{a}+\mathfrak{b}})^T = (\eta \diamond w^{\mathfrak{a}})^T \text{diag}\{w^{\mathfrak{b}}\}.$$

B.3. Properties of monotone signals and dilations. Let $\hat{w}_t, t \geq T \geq 0$, be monotonically increasing and such that $\hat{w}_t \geq 1$ for all $t \geq T$. We have for all $t \geq T$

$$(106) \quad 1 \leq w_s^{\mathbf{a}} \leq w_t^{\mathbf{a}}, \quad \forall \mathbf{a} \in \mathbb{R}_{\geq}^n, \quad s \in [T, t],$$

$$(107) \quad w_t^{-|\min_i a_i|} \leq w_t^{\min_i a_i} \leq w_t^{\mathbf{a}} \leq w_t^{\max_i a_i} \leq w_t^{|\max_i a_i|}, \quad \forall \mathbf{a} \in \mathbb{R}^n.$$

References.

- [1] S. BATTILOTTI, *Incremental generalized homogeneity, observer design and semiglobal stabilization*, Asian Journal of Control, 16 (2014), pp. 498–508.
- [2] S. BATTILOTTI, *Nonlinear predictors for systems with bounded trajectories and delayed measurements*, Automatica, (2015), pp. 127–138.
- [3] F. CACACE, A. GERMANI, AND C. MANES, *A chain observer for nonlinear systems with multiple time varying measurement delays*, SIAM Journal on Optimization and Control, 52 (2014), pp. 1862–1885.
- [4] S. DASHKOVSKIY AND L. NAUJOK, *Lyapunov-razumikhin and lyapunov-krasovskii theorems for interconnected iss time-delay systems*, in Proceedings of the 19th International Symposium on Mathematical Theory of Networks and Systems, 2010, pp. 1179–1184.
- [5] A. GERMANI, C. MANES, AND P. PEPE, *A new approach to state observation of nonlinear systems with delayed output*, IEEE Transactions on Automatic Control, 47 (2002), pp. 96–101.
- [6] M. GHANES, J. DE LEON, , AND J.-P. BARBOT, *Observer design for nonlinear systems under unknown time-varying delays*, IEEE Transactions on Automatic Control, 58 (2013), pp. 1529–1534.
- [7] S. IBRIR, *Observer based control of a class of time delay nonlinear systems*, Automatica, 47 (2011), pp. 388–394.
- [8] I. KARAFYLLIS AND M. KRSTIC, *Nonlinear stabilization under sampled and delayed measurements and with inputs subject to delay and zero-order hold*, IEEE Transactions on Automatic Control, 57 (2012), pp. 1141–1154.
- [9] I. KARAFYLLIS AND M. KRSTIC, *Predictor-based output feedback for nonlinear delay systems*, Preprint, arXiv:1108.4499 (2012).
- [10] I. KARAFYLLIS, M. KRSTIC, T. AHMED-ALI, AND F. LAMNABHI-LAGARRIGUE, *Global stabilization of nonlinear delay systems with a compact absorbing set*, Journal International Journal of Control, 87 (2013), pp. 1010–1027.
- [11] N. KAZANTZIS AND R. A. WRIGHT, *Nonlinear observer design in the presence of delayed output measurements*, Systems and Control Letters, 54 (2005), pp. 877–886.
- [12] M. KRSTIC, *Delay compensation for nonlinear adaptive and PDE systems*, Birkhauser Boston, 2009.
- [13] K. SUBBARAO AND P. MURALIDHAR, *State observer for linear systems with piecewise constant output delays*, IET Control Theory and Applications, (2009), pp. 1017–1022.
- [14] A. R. TEEL, *Connections between razumikhin-type theorems and the iss nonlinear small gain theorem*, IEEE Trans. Automat. Control, 43 (1998), pp. 960–964.
- [15] B. ZHOU, Z. LIN, AND G.-R. DUAN, *Truncated predictor feedback for linear systems with long time-varying delay*, Automatica, 48 (2006), pp. 2387–2399.