Back-and-forth Operation of State Observers and Norm Estimation of Estimation Error

Hyungbo Shim, Aneel Tanwani, and Zhaowu Ping

***Abstract*— In contrast to classical observers operating synchronously with the plant, this paper proposes a state estimation algorithm that executes Luenberger observers in a back-and-forth manner using the stored inputs and output signals. One benefit of this technique is the rapid convergence of state estimation error without relying on high injection gain, so that the amplification of measurement noise is much relieved. Moreover, by operating the observer in the proposed manner, we obtain an upper bound on the estimation error independent of its initial value. Some real-time applications of the proposed idea, and the effect of disturbances, are also discussed.**

I. Introduction

State observers are employed in practice to estimate the internal state variables of a dynamical system (called a 'plant') from the knowledge of the input/output signals and the mathematical model of the plant. Conventionally, state observers (such as the observer by [10]) operate synchronously in time with the plant, meaning that the observer time index is synchronized with the time index of the real plant. On the other hand, a batch operation involves storing the input and output signals for a certain time interval, and then processing all the gathered information at certain time instants to compute the state estimate; see [11] as an example of using an optimization routine to generate state estimates using the signals stored over an interval. As a matter of fact, the synchronous operation has been preferred over the batch operation because the latter is not so easy to implement in real-time whereas the former requires less memory and computing power. It can be recalled that the recursive form of the Kalman filter is preferred over its closed form for the same reasons [9].

However, recent advancements in computing power and memory capacity make the implementation of batch operations quite feasible. This motivates us to present a new method of operation for state observers in this correspondence, that uses the stored input and output signals to process the state estimates. Out of the several benefits emerging from this technique, following are the most notable ones.

1) *Peaking phenomenon:* In order to obtain a good estimate rapidly, one often employs a high injection gain in the observer design, and this incurs amplification of measurement noise and large transients in the estimation error for some initial conditions of the error (known as 'peaking phenomenon [14]'). Researchers in the past have employed time-varying gains to overcome this problem [1], where the basic idea is to obtain rapid convergence initially with high gain and then switch to a low gain around the steady state. Addressing the same issue, our proposed scheme provides an alternate method to avoid these undesirable effects in the transient response without injecting output with high gain, while achieving rapid convergence.
2) *Error estimate:* One drawback of the conventional state observers is that, even if the estimate $\hat{x}(t)$ is guaranteed to converge asymptotically to the actual state $x(t)$ with the passage of time, it is not known how close the current estimate $\hat{x}(t)$ is to the real state $x(t)$ unless the size of initial error is known. The algorithm proposed in this paper is used to obtain an upper bound on state estimation error that is independent of the initial value of the error. Moreover, by successive iterations of the proposed back-and-forth scheme, the estimation error over any finite interval can be made arbitrarily small. The error reduction can be achieved simply at the cost of higher computation, and not necessarily requiring the knowledge of inputs to the observer over larger time period.

The underlying idea of our observer scheme is to generate state estimates using integration forward and backward in time. Specifically, we run a classical Luenberger observer over a certain interval forward in time, and with the estimate obtained at the end of this interval, we run another observer backward in time over the same interval. The poor transient performance of the state estimation error is eliminated by the backward recursion, thus the resulting estimate obtained after one cycle of this back-and-forth operation is close to the actual state on the entire interval. Successive iterations of this routine then yield better and better estimates. This simple idea provides good estimates without necessarily employing the high injection gains and, unlike classical observers, one also obtains the bound on estimation error. Even though these benefits arrive at the cost of some delay in the computation of the estimate $\hat{x}(t)$ for the current state $x(t)$, this delay can be compensated by a 'catch-up' procedure that inherently relies on performing numerical integration faster than the real time frame (this will be clarified in Section IV-A). We envisage the use of proposed technique especially when the knowledge of input and output signals is available for brief amount of time but still the state estimates are required to be arbitrarily close to the actual state.

On an intuitive level, a variant of the proposed scheme, called back-and-forth nudging (BFN) method, has been adopted in several applications [2], [3], [5], [8] to reconstruct the initial state of the system when the output measurements are sparse in time and space. The BFN algorithm was first introduced in [2] for a linear system with convergence analysis. It has been generalized to nonlinear systems in

H. Shim and Z. Ping are with ASRI, School of Electrical Engineering, Seoul National University, Korea.

A. Tanwani is with Team BIPOP, Institut National de Recherche en Informatique et Automatique (INRIA), Rhône-Alpes, France.

[3] without convergence analysis, and the method has been applied to identify the initial state of a quantum system in [5], [8]. The proposed operation is different from the smoothing filter [9] in the stochastic estimation field, in that their objectives are different and that the forward and backward integrations are not simultaneous but sequential, although both of them utilize the integration in forward and backward directions.

Developing further the idea of back-and-forth operation of state observers in this paper, we discuss several utilities of the proposed scheme in an analytical setting. In Section II, we begin with an overview of the back-and-forth operation with convergence analysis of estimation error, and obtain an upper bound on the state estimation error over a finite interval. For linear systems, gains selection method is discussed in Section III. Some real-time applications are discussed in Section IV, and the effect of process and measurement disturbances is studied in Section V.

II. Back-and-forth Process and Estimation of Error Bound

We introduce the back-and-forth operation of an observer for a general nonlinear case in this section. Consider a smooth nonlinear system written as

$$\begin{aligned} \dot{x} &= f(x,u), \\ y &= h(x), \qquad x(0) = x_0, \end{aligned} \tag{1}$$

where $x \in \mathbb{R}^n$ is the state, $u \in \mathbb{R}^m$ is the input, and $y \in \mathbb{R}^p$ is the output. Suppose that, for a given time interval $[0,d]$ with $d > 0$, the input $u(t)$ and the output $y(t)$ are collected and stored in memory for the entire interval.

The solution $x(t)$ obviously satisfies the forward-time description of the system

$$\frac{dx(t)}{dt} = f(x(t),u(t)), \;\; y(t) = h(x(t)), \;\; x(0) = x_0 \tag{2}$$

for $t \in [0,d]$, and the following backward-time description, with $s := d - t$ and $\bar{x}(s) := x(d-s)$,

$$\frac{d\bar{x}(s)}{ds} = -f(\bar{x}(s),u(d-s)), \;\; y(d-s) = h(\bar{x}(s)), \tag{3}$$

with the initial value $\bar{x}(0) = x(d)$ for $s \in [0,d]$.

Now suppose that there are two observers for both system (2) and (3). More specifically, we assume the following.

Assumption 1: There exist the forward observer for system (2) given by

$$\frac{d\hat{x}_f(t)}{dt} = \hat{f}_f(\hat{x}_f(t),u(t),y(t)) \tag{4}$$

and the backward observer for (3) given by

$$\frac{d\hat{x}_b(s)}{ds} = \hat{f}_b(\hat{x}_b(s),u(d-s),y(d-s)) \tag{5}$$

such that, for $t, s \in [\frac{d}{2}, d]$,

$$|\hat{x}_f(t) - x(t)| \le \alpha |\hat{x}_f(0) - x(0)| \tag{6}$$

$$|\hat{x}_b(s) - \bar{x}(s)| \le \alpha |\hat{x}_b(0) - \bar{x}(0)| \tag{7}$$

where $0 < \alpha < 1$. ◊

Remark 1: In the literature, there are many design methods for nonlinear observers which guarantee asymptotic error convergence. These designs are suitable for Assumption 1 by taking d sufficiently large. If d is not large, then several nonlinear observer designs, that assign arbitrary convergence rate, can also be employed. For example, the high-gain observer design in [6] yields the error inequality $|\hat{x}_f(t) - x(t)| \le \lambda(\theta)\exp(-\theta t)|\hat{x}_f(0) - x(0)|$ for a given constant $\theta > 0$, and the polynomial λ in θ. Here, by choosing observer gains appropriately, θ can be made arbitrarily large. Although, the value of $\lambda(\theta)$ increases with θ in general, the growth of λ is suppressed by the exponential decrease of $\exp(-\theta t)$ at a positive time t. Hence, Assumption 1 can still be met by increasing θ. Once the forward observer satisfying (6) is obtained, the backward observer is often derived by the same method such that (7) holds. ◊

We now explain the back-and-forth operation of the observers. Suppose that, at time $t = d$, the following operation is performed. Let $\hat{x}_0(0)$ be an arbitrary initial guess for $x(0)$. Setting $\hat{x}_f(0) = \hat{x}_0(0)$, the forward observer (4) is integrated first, with the stored $u(t)$ and $y(t)$, over the interval $[0,d]$. The corresponding solution $\hat{x}_f(t), t \in [0,d]$, is then stored in the memory. At time $t = d$, the backward observer (5) is integrated over the same interval $[0,d]$ with $\hat{x}_f(d)$ as the initial condition of $\hat{x}_b$, i.e., $\hat{x}_b(0) = \hat{x}_f(d)$, the solution obtained is then stored as $\hat{x}_b(s), s \in [0,d]$. Finally, we take

$$\hat{x}_1(t) := \begin{cases} \hat{x}_b(d-t), & t \in [0, \frac{d}{2}) \\ \hat{x}_f(t), & t \in [\frac{d}{2}, d] \end{cases} \tag{8}$$

as the estimate of $x(t)$ on the interval $[0,d]$. Then it is trivially seen that $\hat{x}_1(t)$ is quite a rich estimate of $x(t)$ on the interval without transients (but with a possible discontinuity at $t = d/2$). Indeed, by Assumption 1, we have

$$\begin{aligned} \sup_{t \in [\frac{d}{2},d]} |\hat{x}_1(t) - x(t)| &= \sup_{t \in [\frac{d}{2},d]} |\hat{x}_f(t) - x(t)| \\ &\le \alpha |\hat{x}_f(0) - x(0)| = \alpha |\hat{x}_0(0) - x(0)| \end{aligned}$$

and

$$\begin{aligned} \sup_{t \in [0,\frac{d}{2})} |\hat{x}_1(t) - x(t)| &= \sup_{s \in (\frac{d}{2},d]} |\hat{x}_b(s) - \bar{x}(s)| \\ &\le \alpha |\hat{x}_b(0) - \bar{x}(0)| = \alpha |\hat{x}_f(d) - x(d)| \\ &\le \alpha^2 |\hat{x}_0(0) - x(0)| \le \alpha |\hat{x}_0(0) - x(0)|. \end{aligned}$$

Therefore, we obtain that

$$\sup_{t \in [0,d]} |\hat{x}_1(t) - x(t)| \le \alpha |\hat{x}_0(0) - x(0)|, \tag{9}$$

$$\text{and} \qquad |\hat{x}_1(0) - x(0)| \le \alpha^2 |\hat{x}_0(0) - x(0)|, \tag{10}$$

in which, inequality (9) indicates that $\hat{x}_1(t)$ is an estimate of $x(t)$ on the *entire* interval $[0,d]$ without any transients.

As a matter of fact, the back-and-forth process can be repeated by setting $\hat{x}_f(0) = \hat{x}_i(0)$ in order to obtain an improved estimate $\hat{x}_{i+1}(t)$ from another round-trip excursion. If the process is repeated R times, then

$$\sup_{t \in [0,d]} |\hat{x}_R(t) - x(t)| \le \alpha^R |\hat{x}_0(0) - x(0)|, \tag{11}$$

$$|\hat{x}_i(0) - x(0)| \le \alpha^2 |\hat{x}_{i-1}(0) - x(0)|, \quad 1 \le i \le R, \tag{12}$$

that is, the estimation error decreases by the factor of α^R times the initial error $|\hat{x}_0(0) - x(0)|$, and thus, converges to zero as R tends to infinity.

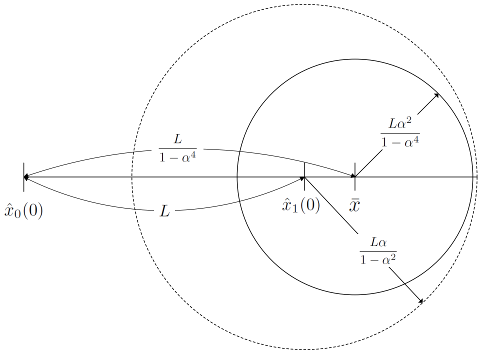

Fig. 1. With $L = |\hat{x}_1(0) - \hat{x}_0(0)|$, solid circle shows the region containing $x(0)$ according to Remark 3. The dashed circle displays the set $\{x : |x - \hat{x}_1(0)| \le \frac{L\alpha}{1-\alpha^2}\}$ related to the coarse bound in (14).

Another benefit of the back-and-forth operation comes from inequality (10). From (10) and the triangular inequality

$$|\hat{x}_0(0) - x(0)| - |\hat{x}_1(0) - \hat{x}_0(0)| \le |\hat{x}_1(0) - x(0)|,$$

it follows that

$$|\hat{x}_0(0) - x(0)| \le \frac{1}{1-\alpha^2}|\hat{x}_1(0) - \hat{x}_0(0)|. \quad (13)$$

Using (9), this in turn leads to

$$\sup_{t\in[0,d]} |\hat{x}_1(t) - x(t)| \le \frac{\alpha}{1-\alpha^2}|\hat{x}_1(0) - \hat{x}_0(0)|. \quad (14)$$

Note that the right-hand sides of (13) and (14) become known after one excursion, which constitute the guaranteed upper bound on the estimation error. If the computing power is strong so that the back-and-forth operation is quickly performed on-line, then the upper bound of estimation error, as well as the estimate $\hat{x}_1(t)$, can be obtained soon after $t = d$ when the computation begins. These discussions are summed up in the following result:

Theorem 1: Suppose that Assumption 1 holds for system (1). For each $i = 1, 2, \ldots$, let

$$\hat{x}_i(t) := \begin{cases} \hat{x}_b(d-t), & t \in [0, \frac{d}{2}) \\ \hat{x}_f(t), & t \in [\frac{d}{2}, d] \end{cases} \quad (15)$$

where $\hat{x}_f(\cdot)$ and $\hat{x}_b(\cdot)$ are obtained from the integration of (4) and (5) by setting $\hat{x}_f(0) = \hat{x}_{i-1}(0)$, and $\hat{x}_b(0) = \hat{x}_f(d)$, respectively, with $\hat{x}_0(0)$ picked arbitrarily. Then, it holds that,

$$\sup_{t\in[0,d]} |\hat{x}_{i+1}(t) - x(t)| \le \frac{\alpha}{1-\alpha^2}|\hat{x}_{i+1}(0) - \hat{x}_i(0)|, \quad (16)$$

where $|\hat{x}_{i+1}(0) - \hat{x}_i(0)| = O(\alpha^i)$, that is, the state-estimation error is bounded by a known quantity and converges to zero as i tends to infinity. ◊

Note that, by (11), the right-hand side of (16) goes to zero as i increases, so Theorem 1 states that one can perform the back-and-forth operation repeatedly until the required precision of the estimation is obtained, and stop. This is simply done by monitoring $|\hat{x}_{i+1}(0) - \hat{x}_i(0)|$.

Remark 2: The norm of $x(t)$ can also be estimated from (16) because $\hat{x}_i$ and $\hat{x}_{i+1}$ are known. An advantage over the norm-estimator of [13] is that the estimated norm is a guaranteed upper bound obtained at time $t = d + D$ where D is the time elapsed for computation while the information from the norm-estimator of [13] converges to the actual norm of the state as time tends to infinity. On the other hand, the norm-estimator of [13] requires much weaker condition of 'output-to-state stability' than Assumption 1. ◊

Remark 3: (Tighter bound on the estimation error) In fact, without using the triangular inequality, tighter bound than (13) can be obtained. Note that $\{x : |z_1 - x| \le \alpha^2|z_0 - x|\} = \{x : |x - (\frac{1}{1-\alpha^4}z_1 - \frac{\alpha^4}{1-\alpha^4}z_0)| \le \frac{\alpha^2}{1-\alpha^4}|z_1 - z_0|\}$. Therefore, it follows from (10) that $x(0)$ is located within the circle of radius $\frac{\alpha^2}{1-\alpha^4}L$, where L is the distance $|\hat{x}_1(0) - \hat{x}_0(0)|$, centered at $\bar{x} = \frac{1}{1-\alpha^4}\hat{x}_1(0) - \frac{\alpha^4}{1-\alpha^4}\hat{x}_0(0) = \frac{1}{1-\alpha^4}(\hat{x}_1(0) - \hat{x}_0(0)) + \hat{x}_0(0)$. See Fig. 1 for illustration. ◊

The back-and-forth operation is intrinsically a batch process. However, with strong computing power and large memory, it can be used on-line. The details of on-line implementation will be given in Section IV.

III. Linear Systems Case

For linear systems, more concrete method for observer construction can be addressed, under which Assumption 1 holds. Consider the linear system

$$\begin{aligned} \dot{x} &= Ax + Bu \\ y &= Cx \end{aligned} \quad (17)$$

with $x(0) = x_0$, where (A, C) is an observable pair.

Proposition 1: For a given $d > 0$ and a given α such that $0 < \alpha < 1$, there exist gain matrices L_f and L_b such that

$$\|\exp((A - L_f C)t)\| \le \alpha, \qquad {}^\forall t \in [d/2, d], \quad (18)$$

and

$$\|\exp(-(A - L_b C)t)\| \le \alpha, \qquad {}^\forall t \in [d/2, d]. \quad (19)$$

See the Appendix for a constructive proof of Proposition 1.

Then, the forward observer is given by

$$\frac{d}{dt}\hat{x}_f = A\hat{x}_f + Bu(t) + L_f(y(t) - C\hat{x}_f) \quad (20)$$

while the backward observer is

$$\frac{d}{ds}\hat{x}_b = -A\hat{x}_b - Bu(d-s) - L_b(y(d-s) - C\hat{x}_b). \quad (21)$$

In fact, the backward observer is based on the backward-time description of the system (17) written as

$$\frac{d}{ds}\bar{x} = -A\bar{x} - Bu(d-s), \quad y(d-s) = C\bar{x}, \quad \bar{x}(0) = x(d),$$

with $\bar{x}(s) = x(d-s)$ for $s = d - t \in [0, d]$. With Proposition 1, it is clear that Assumption 1 holds.

IV. Real-time Applications

In this section, we discuss some applications of the proposed back-and-forth state estimation algorithm.

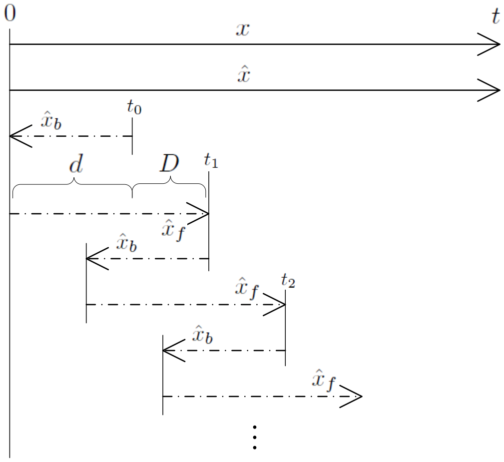

Fig. 2. Operation time chart. Solid arrow implies real time frame, and solid-dot means numerical integration time frame that is faster than real time.

A. Enhancing Convergence Rate of Conventional Observers

In order to enhance the convergence rate without using unnecessarily high injection gain, a conventional observer given by

$$\dot{\hat{x}}(t) = A\hat{x}(t) + Bu(t) + L(y(t) - C\hat{x}(t)) \tag{22}$$

can be equipped with the estimation update law

$$\hat{x}(t_i) = \hat{x}_f(t_i) \tag{23}$$

in which the time t_i and $\hat{x}_f(t_i)$ will be specified shortly as an outcome of the back-and-forth operation. Since the update law (23) introduces discontinuity in the state trajectory, the overall observer becomes hybrid-type.

1) Choosing the gains: Suppose that L is designed as $L = P^{-1}C^T$ where P is the symmetric positive definite solution of

$$PA + A^TP - 2C^TC = -\theta P \tag{24}$$

with some $\theta \geq 0$, so that $(A - LC)$ becomes Hurwitz. (Note that θ is not necessarily high.) Let $L_f = L$ and find d such that $\|\exp((A - L_fC)t)\| < 1$ for $t > d$. Then, find L_b that satisfies $\|\exp(-(A - L_bC)t)\| < 1$ for $t > d$ (one may refer to the Appendix).

2) Computing t_i and $\hat{x}_f(t_i)$: With the gains L_f and L_b, prepare two additional observers (20) and (21). The state observer given in (22) is started at $t = 0$ with some initial condition $\hat{x}(0)$, and it runs synchronously with the plant in the regular time frame. At time $t = t_0 := d$, the backward observer (21) starts with the history of inputs and outputs over the interval $[t_0 - d, t_0]$ and with the initial condition $\hat{x}_b(0) = \hat{x}(t_0)$. After the backward integration reaches the time $t_0 - d$, the forward observer (20) takes over with $\hat{x}_f(t_0 - d) = \hat{x}_b(d)$. It is integrated with the input and output history for $[t_0 - d, t_0 + D]$, where positive D indicates the elapsed time for going backward and forward. At time $t = t_1 := t_0 + D$, the state $\hat{x}(t_1)$ is reset as in (23) (with $i = 1$). Then, with $\hat{x}_b(0) = \hat{x}(t_1)$ at the time $t = t_1$, this back-and-forth operation repeats as illustrated in Fig. 2.

We assume that the backward and forward observers can be integrated at a much faster time scale σt where $\sigma \gg 1$. Thus, the integration of the backward observer takes d/σ seconds, and the time of integration for both observers is $(2d + D)/\sigma$ seconds, which should be the same D seconds measured in the regular time frame (see Fig. 2). This implies that $D = 2d/(\sigma - 1)$, and let the time $t_i = t_{i-1} + D$.

3) Analysis of state estimation error: We now analyze the reduction in estimation error due to (23) compared to the observer without such reset map. If $\hat{x}_c(t)$ denotes the continuous solution of (22) without the reset map (23), then the size of estimation error $\tilde{x}_c := \hat{x}_c - x$ can be evaluated in terms of the error Lyapunov function $V = \tilde{x}_c(t)^TP\tilde{x}_c(t)$. It is observed that $\dot{V} = -\theta V$ along the observer (22) and the plant (17) when there is no reset (23). Considering the interval $[0, t_1]$ for the sake of simplicity, we have

$$\tilde{x}_c(t_1)^TP\tilde{x}_c(t_1) = \exp\left(-\theta\frac{2}{\sigma - 1}d\right)\tilde{x}_c(t_0)^TP\tilde{x}_c(t_0).$$

On the other hand, to see the error reduction with the reset map (23), let $\tilde{x} := \hat{x} - x$, $\tilde{x}_f := \hat{x}_f - x$ and $A_b := -A + L_bC$. By the fact that $A - LC = A - L_fC$ and $d + D = ((\sigma + 1)/(\sigma - 1))d$, we obtain that

$$\begin{aligned}
&\tilde{x}(t_1)^TP\tilde{x}(t_1) = \tilde{x}_f(t_1)^TP\tilde{x}_f(t_1) \\
&= \exp\left(-\theta\frac{\sigma + 1}{\sigma - 1}d\right)\tilde{x}_f(t_0 - d)^TP\tilde{x}_f(t_0 - d) \\
&= \exp\left(-\theta\frac{\sigma + 1}{\sigma - 1}d\right)\tilde{x}_c(t_0)^T\exp(A_b^Td)P\exp(A_bd)\tilde{x}_c(t_0).
\end{aligned}$$

Therefore, the use of state update (23) provides more reduction in estimation error within the same real time interval by the amount

$$\begin{aligned}
&\tilde{x}_c(t_1)^TP\tilde{x}_c(t_1) - \tilde{x}(t_1)^TP\tilde{x}(t_1) \\
&\quad = e^{-\theta\frac{2}{\sigma-1}d}\tilde{x}(t_0)^T(P - e^{-\theta d}e^{A_b^Td}Pe^{A_bd})\tilde{x}(t_0)
\end{aligned}$$

where the matrix $P - e^{-\theta d}e^{A_b^Td}Pe^{A_bd}$ is positive definite. This argument can be applied at each t_i, $i \geq 1$, which suggests that the update law (23) indeed diminishes the size of estimation error.

Example 1: A simulation is performed for a linear system with $A = [0, 1, 0; 0, 0, 1; -0.03, -0.5, -0.2]$, $B = [0.5; 0.5; 1]$, $C = [1, 0, 0]$, and $u(t) = \sin(t)$. Both injection gains L_f and L_b are chosen with $\theta = 1.5$ in (24). For comparison, another gain L_h is chosen with $\theta = 2.5$, which will be used for a conventional observer, e.g., like (22). The value $\theta = 2.5$ is selected such that the error reduces by more than 1/8-th of its initial value within 3 seconds, which is verified in Fig. 3. On the other hand, we suppose that the back-and-forth operation begins at $t_0 = 3$ like in Fig. 2. Since $\max\{\|e^{3(A-L_fC)}\|, \|e^{-3(A-L_bC)}\|\} < 1/2$, the reduction in error after one back-and-forth operation is more than $1/8 = (1/2)^3$. This can be clearly observed in Fig. 4, in which the estimate is obtained immediately after $t = 3$, and after that, a conventional observer with L_f is used. However, with a measurement disturbance (analyzed formally in Section V), $y = Cx + \sin(3t)$, it can be seen in Fig. 5 that the performance of the high-gain observer becomes degraded while the back-and-forth observer gives better results because the relatively small gain. $\Diamond$

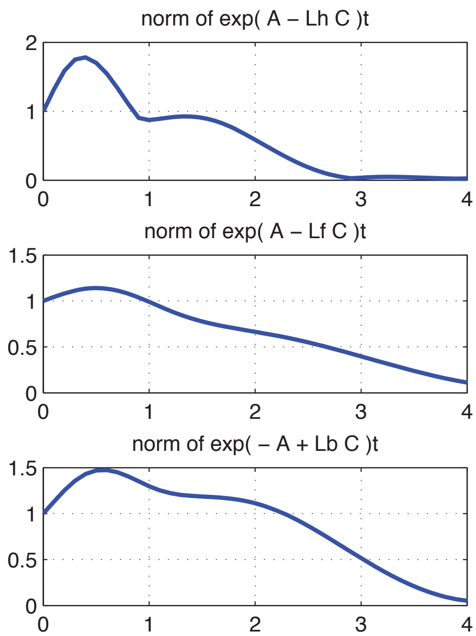

Fig. 3. Plot of $\|e^{(A-L_hC)t}\|$, $\|e^{(A-L_fC)t}\|$, and $\|e^{-(A-L_bC)t}\|$, in which $L_h = [3.6; 8.2; 4.4]$, $L_f = [2.1; 2.7; 0.25]$, $L_b = [-2.5; 4.2; -1.9]$.

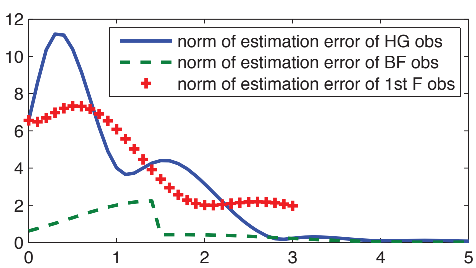

Fig. 4. Plot of the estimation error $\|\hat{x}(t) - x(t)\|$ from the conventional observer with high gain L_h (blue solid), the back-and-forth operation (green dashed), and the first run of the forward observer (red +). It is seen that, even though the first estimate of the forward observer is not satisfactory, the back-and-forth operation results in the estimate as good as that of the high gain observer at time $t = 3$. Initial condition of the plant is $[5; -3; -3]$ while initial conditions of both observers are set to zero.

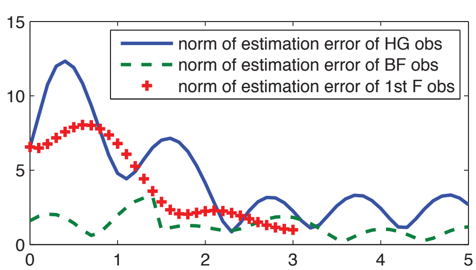

Fig. 5. Plot of the estimation error $\|\hat{x}(t) - x(t)\|$ as in Fig. 4, but with a measurement disturbance $\sin(3t)$.

B. Intermittent Monitoring

As mentioned earlier, the conventional observers only guarantee the convergence of state estimation error to zero but don't provide any information about the size of error. However, the use of the back-and-forth operation, in particular the use of backward observer in conjunction with a conventional observer, allows us to monitor the quality of estimation error. With a gain L for the conventional observer, one can find d such that $\|\exp((A - LC)d)\| =: \alpha < 1$, and then design L_b such that $\|\exp((-A + L_bC)d)\| \leq \alpha$. From time to time when the information about the estimation error is required, say, at $t = t^*$, the backward observer is employed with $\hat{x}_b(0) = \hat{x}(t^*)$ over the interval $[t^* - d, t^*]$ to obtain the estimate

$$|\hat{x}(t^*) - x(t^*)| \leq (\alpha/(1-\alpha^2))|\hat{x}_b(d) - \hat{x}(t^* - d)|$$

which is derived in the same manner as one arrives at (14) using (10).

C. State Estimation of a Switched System

Consider a switched system given in Fig. 6, which has two modes of operation described by

$$\Sigma_1 : \begin{cases} \dot{x}_1 = A_1x_1 + B_1u, & y = C_1x_1, \\ \dot{x}_2 = A_2x_2 + A_{21}x_1, & \end{cases} \tag{25}$$

for mode 1, and

$$\Sigma_2 : \begin{cases} \dot{x}_1 = A_1x_1 + A_{12}x_2 & \\ \dot{x}_2 = A_2x_2 + B_2u, & y = C_2x_2, \end{cases} \tag{26}$$

for mode 2, where the pair (A_i, C_i), $i = 1, 2$, is observable. Now suppose that the system configuration switches between modes 1 and 2 (i.e., between (25) and (26)) after every T seconds, and we want to estimate the states x_1 and x_2 completely. Note that, at each mode, the system is not completely observable. For example, at mode 1, the state x_2 is unobservable.

One can design a separate observer for each mode of operation as follows:

$$\dot{\hat{x}}_1 = A_1\hat{x}_1 + B_1u - L_1C_1\hat{x}_1 + L_1y \tag{27a}$$
$$\dot{\hat{x}}_2 = A_2\hat{x}_2 + A_{21}\hat{x}_1 \tag{27b}$$

for Σ_1, and

$$\dot{\hat{x}}_1 = A_1\hat{x}_1 + A_{12}\hat{x}_2 \tag{28a}$$
$$\dot{\hat{x}}_2 = A_2\hat{x}_2 + B_2u - L_2C_2\hat{x}_2 + L_2y \tag{28b}$$

for Σ_2, where L_1 and L_2 are large enough so that some meaningful estimates $\hat{x}_1$ and $\hat{x}_2$ are obtained over the interval of length T. In fact, at the end of the first period $[0, T)$, one can obtain the estimate $\hat{x}_1(T)$ by (27a) for mode 1. For the second interval $[T, 2T)$, this estimate serves as the initial condition of (28a), and the observer (28b) starts to estimate $x_2(t)$. However, the observer (28b) will exhibit some transients during the initial period of the interval $[T, 2T)$, which may corrupt the estimate $\hat{x}_1(t)$ being obtained through the observer (28a) because of initially large error between $\hat{x}_2(t)$ and $x_2(t)$.

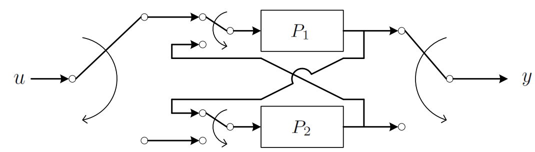

Fig. 6. All switches are synchronized.

To overcome this problem, the following hybrid-type observer may be utilized instead of (27) and (28)

$$\hat{\Sigma}_1 : \begin{cases} \dot{\hat{x}}_1 = A_1\hat{x}_1 + B_1 u \\ \dot{\hat{x}}_2 = A_2\hat{x}_2 + A_{21}\hat{x}_1 \end{cases} \tag{29}$$

$$\hat{\Sigma}_2 : \begin{cases} \dot{\hat{x}}_1 = A_1\hat{x}_1 + A_{12}\hat{x}_2 \\ \dot{\hat{x}}_2 = A_2\hat{x}_2 + B_2 u \end{cases} \tag{30}$$

$$\begin{pmatrix} \hat{x}_1(kT) \\ \hat{x}_2(kT) \end{pmatrix} = \begin{pmatrix} \hat{\xi}_1(kT^-) \\ \hat{\xi}_2(kT^-) \end{pmatrix}, \qquad k \geq 1, \tag{31}$$

where the variables $\hat{\xi}_1$ and $\hat{\xi}_2$ will be obtained shortly using the back-and-forth operation such that the inequality,

$$|\hat{x}((k+2)T) - x((k+2)T)| \leq \gamma|\hat{x}(kT) - x(kT)|, \tag{32}$$

holds for all $k \geq 1$, $x := (x_1^T, x_2^T)^T$, and a desired parameter $\gamma < 1$. The inequality (32) guarantees the convergence of estimation error to zero due to the fact that $\sup_{t\in[kT,(k+1)T)} |\hat{x}(t) - x(t)| \leq M|\hat{x}(kT) - x(kT)|$ where M is a constant. The latter inequality holds because the dynamics for $\hat{x} - x$ are linear and their growth is bounded over a finite interval.

In order to design $\hat{\xi}_1$ and $\hat{\xi}_2$, we prepare the back-and-forth observer (20) and (21) for the x_1-subsystem of (25) and for the x_2-subsystem of (26), respectively. For each subsystem, the injection gains L_f and L_b are designed such that (18) and (19) hold with $d = T/2$ and

$$\alpha = \left(\frac{\gamma}{\sqrt{2}\max\{\alpha_1, \alpha_2, \alpha_3, \alpha_4\}} \right)^{\frac{1}{R}}$$

where $\alpha_1 = M_1 + L_1L_2 + L_2$, $\alpha_2 = L_1M_2 + M_2$, $\alpha_3 = M_2 + L_1L_2 + L_1$, $\alpha_4 = L_2M_1 + M_1$ and

$$M_i := \|e^{A_iT}\|, \;\; L_i := \left\| \int_0^T e^{A_is} ds A_{ij} \right\|, \;\; i, j = 1, 2, i \neq j,$$

and let $R+1$ be the number of round-trips of numerical back-and-forth integrations that are possible within the interval of length $T/2$. Clearly, the number R relies on the computation power.

Let the initial condition $\hat{\xi}_1(0^-)$ and $\hat{\xi}_2(0^-)$ be arbitrary. Fig. 7 illustrates the strategy to obtain $\hat{\xi}_1(kT^-)$ and $\xi_2(kT^-)$, for $k = 2$, over the interval $[T, 2T)$ when mode 2 is active. At time $t = T$, the estimate $\hat{x}_1(T)$ and $\hat{x}_2(T)$ are set to $\hat{\xi}_1(T^-)$ and $\hat{\xi}_2(T^-)$ respectively, and they are integrated in the real time by (30). At the same time, the initial condition of the forward observer (for estimating x_2) is set by $\hat{\xi}_2(T^-) = \hat{x}_2(T)$, and this forward observer runs in the real time first until $T + T/2$. At $T + T/2$, the backward observer is employed with the terminal state of the forward observer as its initial condition. The round-trip of back-and-forth operation continues R times with the input-output data of the interval $[T, T + T/2]$, after which the forward observer is finally integrated from T to $2T$. Since the time elapsed by the back-and-forth operation and the last forward operation does not exceed $T/2$, the last forward integration will 'catch up' with the real time, as indicated in Fig. 7. While these operations are performed, the information about $\hat{\xi}_2(t)$ is collected as illustrated in the figure. At the same

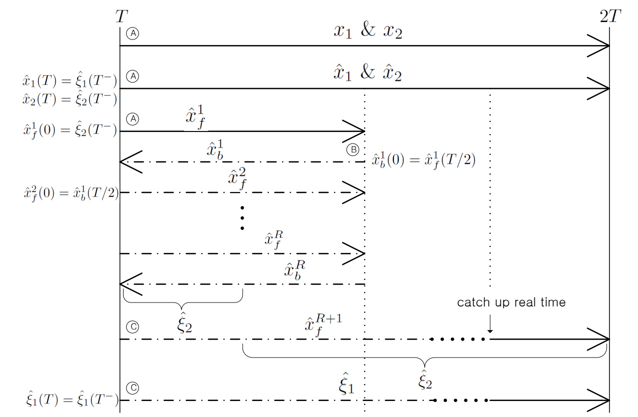

Fig. 7. Operation time chart for mode 2 in the interval $[T, 2T)$. Solid arrow implies real time frame, and solid-dot means numerical integration time frame that is faster than real time. Circles with letters indicate the synchronized time.

time with the start of the last forward observer, we begin the integration of

$$\dot{\hat{\xi}}_1 = A_1\hat{\xi}_1 + A_{12}\hat{\xi}_2 \tag{33}$$

with the initial condition $\hat{\xi}_1(T) = \hat{\xi}_1(T^-)$ and with the signal $\hat{\xi}_2(t)$ obtained by back-and-forth operation over the interval $[T, 2T)$. By this procedure, we obtain $\hat{\xi}_1(2T^-)$ and $\hat{\xi}_2(2T^-)$. This procedure repeats in the next interval, with the role for x_1 and x_2 being switched, and instead of (33), following equation is used to compute $\hat{\xi}_2$:

$$\dot{\hat{\xi}}_2 = A_2\hat{\xi}_2 + A_{21}\hat{\xi}_1 \tag{34}$$

instead of (33).

We now proceed with the error analysis. Let $\epsilon := \hat{x} - x$. Then, $\epsilon(kT) = \hat{x}(kT) - x(kT) = \hat{\xi}(kT^-) - x(kT)$. We note that, for the $(2k+1)$-th interval, $k \in \mathbb{N}$, $\hat{\xi}_1$ is the estimate from the back-and-forth observer while $\hat{\xi}_2$ is the state of (34), and for the $2k$-th interval, their roles are reversed. Therefore, in the first interval $[0, T)$, we have that

$$|\epsilon_1(T)| \leq \alpha^R|\epsilon_1(0)|$$

$$|\epsilon_2(T)| \leq \|e^{A_2T}\||\epsilon_2(0)| + \left\| \int_0^T e^{A_2s} ds A_{21} \right\| \left[\alpha^R|\epsilon_1(0)|\right].$$

Similarly, we can derive the following expressions for the second interval $[T, 2T)$,

$$\begin{aligned} |\epsilon_1(2T)| &\leq \|e^{A_1T}\||\epsilon_1(T)| + \left\| \int_0^T e^{A_1s} ds A_{12} \right\| \left[\alpha^R|\epsilon_2(T)|\right] \\ &\leq \alpha^R\|e^{A_1T}\||\epsilon_1(0)| + \alpha^R \left\| \int_0^T e^{A_1s} ds A_{12} \right\| \|e^{A_2T}\||\epsilon_2(0)| \\ &\quad + \alpha^{2R} \left\| \int_0^T e^{A_1s} ds A_{12} \right\| \left\| \int_0^T e^{A_2s} ds A_{21} \right\| |\epsilon_1(0)| \end{aligned}$$

and

$$\begin{aligned} |\epsilon_2(2T)| &\leq \alpha^R|\epsilon_2(T)| \\ &\leq \alpha^R\|e^{A_2T}\||\epsilon_2(0)| + \alpha^{2R} \left\| \int_0^T e^{A_2s} ds A_{21} \right\| |\epsilon_1(0)|. \end{aligned}$$

The terms within the brackets, $[\cdot]$, are due to the back-and-forth observer, which yield a rich estimation on the entire interval of length T, including the initial period. Finally, it is seen that

$$\begin{aligned}|\epsilon(2T)| &\leq |\epsilon_1(2T)| + |\epsilon_2(2T)| \\ &\leq \alpha^R(M_1 + L_1L_2 + L_2)|\epsilon_1(0)| + \alpha^R(L_1M_2 + M_2) \\ &\quad \times |\epsilon_2(0)| \\ &\leq \frac{\gamma}{\sqrt{2}}(|\epsilon_1(0)| + |\epsilon_2(0)|) \leq \gamma|\epsilon(0)|.\end{aligned}$$

This proves the claim (32) for even k. For odd k, the proof is similar and thus omitted.

Thus we have shown that the use of the back-and-forth observer has improved the transient response of the state estimation error, thus leading to quality estimates of the state variable over the entire interval.

V. Effects of Disturbances

In this section we are interested in the effects of external disturbances on the back-and-forth operation. In particular, we consider an observable linear system given by

$$\begin{aligned}\dot{x} &= Ax + Bu + d_1(t) \\ y &= Cx + d_2(t)\end{aligned} \tag{35}$$

where d_1 and d_2 denote the process disturbance and the measurement disturbance, respectively. Then, the following result shows that, even under the disturbances, the upper bound of the estimation error is obtained.

Proposition 2: Pick $d > 0$ and $0 < \alpha < 1$, and suppose that two gains L_f and L_b are designed by Proposition 1. The repeated back-and-forth operation of the forward observer (20) and the backward observer (21) for system (35) results in

$$\begin{aligned}\sup_{t\in[0,d]} |\hat{x}_{i+1}(t) - x(t)| &\leq \frac{\alpha}{1-\alpha^2}|\hat{x}_{i+1}(0) - \hat{x}_i(0)| \\ &+ \left(\frac{\alpha(\alpha M_f + M_b)}{1-\alpha^2} + (M_f + M_b)\right) \\ &\times \max\{\|d_1(t) - L_f d_2(t)\|_{[0,d]}, \|d_1(t) - L_b d_2(t)\|_{[0,d]}\}\end{aligned} \tag{36}$$

for $i = 0, 1, \cdots$, where $M_f := \int_0^d \|\exp((A - L_fC)t)\|dt$, $M_b := \int_0^d \|\exp(-(A - L_bC)t)\|dt$, and $\|x(t)\|_{[0,d]} := \sup_{t\in[0,d]}|x(t)|$, with $\hat{x}_0(0) = x_0$ being the initial guess. Moreover,

$$\lim_{i\to\infty} |\hat{x}_{i+1}(0) - \hat{x}_i(0)| = 0. \tag{37}$$

Proof: From the repeated back-and-forth operation, we get

$$\hat{x}_i(t) = \begin{cases}\hat{x}_b^i(d-t), & t \in [0, \frac{d}{2}) \\ \hat{x}_f^i(t), & t \in [\frac{d}{2}, d]\end{cases} \tag{38}$$

where $\hat{x}_f^i$ and $\hat{x}_b^i$ are obtained from the i-th iteration of equations (20) and (21), respectively. Let $\tilde{x}_i(t) = \hat{x}_i(t) - x(t)$. The initial conditions for each iteration are given as $\hat{x}_b^i(0) = \hat{x}_f^i(d)$ and $\hat{x}_f^{i+1}(0) = \hat{x}_b^i(d)$ with $\hat{x}_f^1(0) = \hat{x}_0(0)$. Now with the error variables $\tilde{x}_f^i(t) = \hat{x}_f^i(t) - x(t)$ and $\tilde{x}_b^i(s) = \hat{x}_b^i(s) - \bar{x}(s) = \hat{x}_b^i(s) - x(d-s)$, two error dynamics become

$$\begin{aligned}\frac{d\tilde{x}_f^i(t)}{dt} &= (A - L_fC)\tilde{x}_f^i + (L_f d_2(t) - d_1(t)), \\ \frac{d\tilde{x}_b^i(s)}{ds} &= -(A - L_bC)\tilde{x}_b^i - (L_b d_2(d-s) - d_1(d-s)).\end{aligned}$$

For convenience, let $D_f(t) = L_f d_2(t) - d_1(t)$ and $D_b(d-s) = L_b d_2(d-s) - d_1(d-s)$. Then it is seen that

$$\begin{aligned}\sup_{t\in[\frac{d}{2},d]} |\tilde{x}_{i+1}(t)| &= \sup_{t\in[\frac{d}{2},d]} |\tilde{x}_f^{i+1}(t)| \\ &= \sup_{t\in[\frac{d}{2},d]} \Big| \exp((A - L_fC)t)\tilde{x}_f^{i+1}(0) \\ &\quad + \int_0^t \exp((A - L_fC)(t-\tau))D_f(\tau)d\tau \Big| \\ &\leq \alpha|\tilde{x}_i(0)| \\ &\quad + \sup_{t\in[\frac{d}{2},d]} \left| \int_0^t \exp((A - L_fC)\gamma)D_f(t-\gamma)d\gamma \right| \\ &\leq \alpha|\tilde{x}_i(0)| + M_f\|D_f(t)\|_{[0,d]}.\end{aligned}$$

Similarly,

$$\begin{aligned}\sup_{t\in[0,\frac{d}{2})} |\tilde{x}_{i+1}(t)| &= \sup_{t\in(\frac{d}{2},d]} |\tilde{x}_b^{i+1}(t)| \\ &= \sup_{t\in(\frac{d}{2},d]} \Big| \exp(-(A - L_bC)t)\tilde{x}_b^{i+1}(0) \\ &\quad - \int_0^t \exp(-(A - L_bC)(t-\tau))D_b(d-\tau)d\tau \Big| \\ &\leq \alpha|\tilde{x}_f^{i+1}(d)| \\ &\quad + \sup_{t\in(\frac{d}{2},d]} \left| \int_0^t \exp(-(A - L_bC)\gamma)D_b(d-t+\gamma)d\gamma \right| \\ &\leq \alpha(\alpha|\tilde{x}_i(0)| + M_f\|D_f(t)\|_{[0,d]}) + M_b\|D_b(t)\|_{[0,d]}.\end{aligned}$$

Therefore we have that

$$\sup_{t\in[0,d]} |\tilde{x}_{i+1}(t)| \leq \alpha|\tilde{x}_i(0)| + (M_f + M_b)\Delta \tag{39}$$

$$|\tilde{x}_{i+1}(0)| \leq \alpha^2|\tilde{x}_i(0)| + (\alpha M_f + M_b)\Delta \tag{40}$$

where $\Delta = \max\{\|D_f(t)\|_{[0,d]}, \|D_b(t)\|_{[0,d]}\}$. From this and the triangular inequality $|\hat{x}_i(0) - x(0)| \leq |\hat{x}_{i+1}(0) - \hat{x}_i(0)| + |\hat{x}_{i+1}(0) - x(0)|$, it follows that

$$|\hat{x}_i(0) - x(0)| \leq \frac{1}{1-\alpha^2}\left(|\hat{x}_{i+1}(0) - \hat{x}_i(0)| + (\alpha M_f + M_b)\Delta\right).$$

Using (39), we obtain

$$\begin{aligned}\sup_{t\in[0,d]} |\hat{x}_{i+1}(t) - x(t)| &\leq \frac{\alpha}{1-\alpha^2}|\hat{x}_{i+1}(0) - \hat{x}_i(0)| \\ &+ \frac{\alpha}{1-\alpha^2}(\alpha M_f + M_b)\Delta + (M_f + M_b)\Delta\end{aligned}$$

which corresponds to (36).

In order to prove (37), we first observe from (20) that

$$\begin{aligned}\dot{\hat{x}}_f^{i+1} - \dot{\hat{x}}_f^i &= A\hat{x}_f^{i+1} + Bu + L_f(y - C\hat{x}_f^{i+1}) \\ &\quad - [A\hat{x}_f^i + Bu + L_f(y - C\hat{x}_f^i)] \\ &= (A - L_fC)(\hat{x}_f^{i+1} - \hat{x}_f^i).\end{aligned}$$

Hence, from (18),

$$\|\hat{x}_{i+1}(t) - \hat{x}_i(t)\|_{[\frac{d}{2},d]} = \|\hat{x}_f^{i+1}(t) - \hat{x}_f^i(t)\|_{[\frac{d}{2},d]} \\ \leq \alpha|\hat{x}_i(0) - \hat{x}_{i-1}(0)|.$$

Similarly, from (19),

$$\|\hat{x}_{i+1}(t) - \hat{x}_i(t)\|_{[0,\frac{d}{2})} = \|\hat{x}_b^{i+1}(d-t) - \hat{x}_b^i(d-t)\|_{[0,\frac{d}{2})} \\ \leq \alpha|\hat{x}_f^{i+1}(d) - \hat{x}_f^i(d)| \leq \alpha^2|\hat{x}_i(0) - \hat{x}_{i-1}(0)|.$$

This implies

$$|\hat{x}_{i+1}(0) - \hat{x}_i(0)| \leq \alpha^{2i}|\hat{x}_1(0) - \hat{x}_0(0)| \qquad (41)$$

which proves (37). ■

VI. Conclusion

In this paper, we have studied a new technique for estimation of state variables in dynamical systems. The scheme is based on running conventional observers forward and backward in time to obtain not only better state estimates, but also the bounds on state estimation error. Further, we presented three notable applications of the proposed observer, where each of them illustrates one of the three utilities of the back-and-forth operation. These are: a) improving convergence rate without high-gain, b) monitoring of estimation error, and c) obtaining quality estimates on the whole interval including the initial period. We remark that even though the computation may take time, the catch-up procedure outlined in Section IV-A and IV-C allows for the compensation of such delays. The effect of disturbances on the estimation bounds was analyzed which revealed why the choice of low gain is favorable.

We conclude this article by mentioning that the back-and-forth scheme may have several utilities because of its simple structure and useful implications. This has already led to its usage in observers for nonlinear switched systems by [12]. Another important utility could be in the estimation of output derivatives for certain class of systems. This may prove to be a useful alternative to the existing numerical differentiation schemes since additional computation leads to accurate estimates on entire interval while alleviating the effect of noise.

Appendix

We prove Proposition 1 by constructing L_f and L_b here. In fact, this task could be achieved in several different ways, but we discuss one possible method.

In order to obtain L_f, find a symmetric positive definite matrix P_f such that

$$P_f A + A^T P_f - 2C^T C = -\theta_f P_f \qquad (42)$$

or, equivalently,

$$(-A - (\theta_f/2)I)^T P_f + P_f(-A - (\theta_f/2)I) = -2C^T C \qquad (43)$$

with a sufficiently large positive constant θ_f. Then, set $L_f = P_f^{-1}C^T$. In fact, there exists a constant $\theta_f^* > 0$, dependent on the given $\alpha < 1$ and $d > 0$, such that, for each L_f obtained from (43) with $\theta_f \geq \theta_f^*$, the inequality (18) holds. This is because, for sufficiently large θ_f, the matrix $(-A-(\theta_f/2)I)$ becomes Hurwitz while the pair $(-A-(\theta_f/2)I, \sqrt{2}C)$ is observable (since (A, C) is observable), which guarantees the existence of the unique positive definite solution P_f to (43) [4]. Moreover, each element of the matrix P_f is a rational function of θ_f that is continuous for $\theta_f \geq \theta_f^*$ (because it is a solution of linear equation (43)). Hence, the ratio of its maximum and minimum eigenvalues $\lambda_{\max}(P_f)/\lambda_{\min}(P_f)$ are bounded by a polynomial of θ_f.[1] Let $V_f = \tilde{x}_f^T P_f \tilde{x}_f$ and $\tilde{x}_f = \hat{x}_f - x$. Since $\dot{V}_f = -\theta_f V_f$, it follows that $\lambda_{\min}(P_f)|\tilde{x}_f(t)|^2 \leq V_f(t) = \exp(-\theta_f t)V_f(0) \leq \exp(-\theta_f t)\lambda_{\max}(P_f)|\tilde{x}_f(0)|^2$. Hence, it holds that

$$|\tilde{x}_f(t)| = |\exp((A - L_f C)t)\tilde{x}_f(0)| \\ \leq \sqrt{\frac{\lambda_{\max}(P_f)}{\lambda_{\min}(P_f)}} \exp\left(-\frac{\theta_f}{2}t\right)|\tilde{x}_f(0)|$$

for all $\tilde{x}_f(0)$. Since the induced matrix norm is a tight bound [7], it follows that

$$\|\exp((A - L_f C)t)\| \leq \sqrt{\frac{\lambda_{\max}(P_f)}{\lambda_{\min}(P_f)}} \exp\left(-\frac{\theta_f}{2}t\right).$$

Then, since the quantity $\sqrt{\lambda_{\max}(P_f)/\lambda_{\min}(P_f)}$ as a function of θ_f is bounded by a polynomial of θ_f, the right-hand side can be made arbitrarily small on the interval $t \in [d/2, d]$ by increasing θ_f.

The gain L_b can be designed in a similar manner, but instead of (42), one solves the following Lyapunov equation $P_b A + A^T P_b - 2C^T C = \theta_b P_b$ for a sufficiently large constant $\theta_b > 0$, and set $L_b = P_b^{-1}C^T$. Rest of the proof proceeds identically.

References

[1] J. H. Ahrens and H. K. Khalil, "High-gain observers in the presence of measurement noise: A switched-gain approach," *Automatica*, vol. 45, no. 4, pp. 936-943, 2009.

[2] D. Auroux and J. Blum, "Back and forth nudging algorithm for data assimilation problems," *Comptes Rendus Mathematique*, vol. 340, no. 12, pp. 873-878, 2005.

[3] D. Auroux and J. Blum, "A nudging-based data assimilation method: the Back and Forth Nudging (BFN) algorithm," *Nonlinear Processes in Geophysics*, vol. 15, pp. 305-319, 2008.

[4] C. T. Chen, *Linear System Theory and Design*, New York: Holt, Rinehart, and Winston, 1984.

[5] A. Donovan, M. Mirrahimi, and P. Rouchon, "Back and forth nudging for quantum state reconstruction," in *4th International Symposium on Communications, Control and Signal Processing*, Limassol, Cyprus, March 2010, pp. 1-5.

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

[7] P. Lancaster and M. Tismenetsky, *The Theory of Matrices: with Application*, Orlando: Academic Press, 1985.

[8] Z. Leghtas, M. Mirrahimi, and P. Rouchon, "Back and forth nudging for quantum state estimation by continuous weak measurement," in *Proceedings of American Control Conference*, San Francisco, USA, June 2011, pp. 4334-4339.

[9] F. L. Lewis, *Optimal Estimation*, New York: Wiley, 1986.

[10] D. G. Luenberger, "An introduction to observers," *IEEE Transactions on Automatic Control*, vol. 16, no. 6, pp. 596-602, 1971.

[11] H. Michalska and D. Q. Mayne, "Moving horizon observers and observer-based control," *IEEE Transactions on Automatic Control*, vol. 40, no. 6, pp. 995-1006, 1995.

[12] H. Shim and A. Tanwani, "On a sufficient condition for observability of nonlinear switched systems and observer design strategy," in *Proceedings of American Control Conference*, San Francisco, USA, June 2011, pp. 1206-1211.

[13] E. Sontag and Y. Wang, "Output-to-state stability and detectability of nonlinear systems," *Systems & Control Letters*, vol. 29, no. 5, pp. 279-290, 1997.

[14] H. Sussmann and P. V. Kokotovic, "The Peaking phenomenon and the global stabilization of nonlinear systems," *IEEE Transactions on Automatic Control*, vol. 36, no. 4, pp. 424-440, 1991.

[1] Since each element of P_f is a continuous rational function for $\theta_f \geq \theta_f^*$, its maximum eigenvalue $\lambda_{\max}(P_f)$ is bounded by a continuous rational function of θ_f from the Gershgorin disc theorem [7]. Similarly, each element of P_f^{-1} is again a rational function of θ_f so that its maximum eigenvalue, i.e., $1/\lambda_{\min}(P_f)$ is bounded by a rational function of θ_f. It in turn implies that $\lambda_{\max}(P_f)/\lambda_{\min}(P_f)$ is bounded by a polynomial function of θ_f.