State Estimation for Nonlinear Systems with Delayed Output Measurements

Alireza Khosravian, Jochen Trumpf, Robert Mahony

***Abstract*— In this paper, we consider the problem of state estimation for nonlinear systems when the output measurements are delayed. We assume an observer is available that takes the delayed outputs and estimates the delayed states of the system. We propose a novel predictor that takes the delayed estimates from the observer and fuses them with the current input measurements of the system to compensate for the delay. We provide a rigorous stability analysis for globally Lipschitz systems demonstrating that the prediction of the system state converges (asymptotically/exponentially) to the current system trajectory if the observer state converges (asymptotically/exponentially) to the delayed system state. The predictor is computationally simple as it is recursively implementable with a set of delay differential equations. We demonstrate the performance of the proposed predictor via simulation studies.**

I. Introduction

State observers use measurements of the physical outputs of systems to estimate their internal states. A practical challenge arising in state estimation in real world applications is that the sensor measurements are usually delayed in time with respect to the actual physical output of the system. Measurement delays can occur due to various reasons such as physical properties (e.g. slow transients) of sensors or the environment, internal signal processing of sensors, extensive filtering of sensor measurements for noise reduction, and communication delays from sensors to processing units. These delays can degrade the performance of observers and negatively affect their stability and robustness if they are not compensated for properly [1]–[8].

For linear systems, there exist a vast literature for state estimation in the presence of time delays (see for instance [9]–[12] and the references therein) and further extends to classes of nonlinear systems that are linearizable by output injection [13]. Also, for closed loop control of LTI systems with delayed measurement, methods based on Smith predictor [14] and other predictive control approaches have been proposed (see [15] and the references therein).

For nonlinear systems, a classical method to tackle the output delay problem is to employ an observer that has the desired performance for delay free measurements, and modify its innovation term such that it compares each delayed output measurement with its corresponding backward time-shifted estimate. The stability of the modified observer can sometimes be proven using Lyapunov-Krasovskii or Lyapunov-Razumikhin approaches if the delay-free observer has a Lyapunov stability proof (see e.g. [16]–[20]). Although these modified estimators are commonly used in practice (see e.g. [21], [22]), they usually involve complicated stability analysis that require careful and conservative gain tuning and may lead to poor transient behavior of the resulting estimators in practice.

Cascade observer-predictor design is an alternative approach to handle the delay problem for nonlinear systems [1], [6], [23]–[26]. Assuming that an observer is available which has desired stability properties in the presence of delay-free measurements, the observer is fed with the delayed measurements to obtain estimates of the delayed state trajectory. These delayed estimates are then used in predictors to compensate for the effects of the delays such that the prediction of the state converges to the current system trajectory. A crucial challenge is to design the predictors such that they preserve the desired stability properties of their corresponding observer while they compensate for the delays. Such predictors have been proposed in [1], [6], [16], [23] for classes of nonlinear systems on $\mathbb{R}^n$. In addition, [1], [23], [24], [26] cascaded multiple copies of the predictors in which each predictor block only compensates for a portion of the delay increasing the maximum delay that can be compensated for by the whole predictor chain. The authors of this paper have recently proposed a predictor for the particular problem of attitude estimation on the Lie group $\mathsf{SO}(3)$ and proved that it is capable of compensating for arbitrary large delays [7], [8].

All of the above mentioned predictor design methodologies on $\mathbb{R}^n$ build the predictor dynamics such that its internal state directly represents predictions of the current system state. In this paper, we present a new predictor design approach that indirectly predicts the current system state. The proposed approach consist of a dynamics part and a static map. The dynamics part is built to compute a virtual variable that indirectly predicts the difference between the current and the delayed state. This virtual variable together with the delayed estimate of the system state (provided by the observer) are then used in the static map to obtain a prediction of the current state. The predictor dynamics can be implemented recursively using simple delay differential equations. Hence, its computational complexity is low and it is ideal for embedded implementation in practical scenarios. We provide a rigorous *co-stability* analysis for the case of globally Lipschitz systems demonstrating that the state predictions converge asymptotically/exponentially to the current system trajectories if the estimates provided by the corresponding observer converge asymptotically/exponentially to

The authors are with the Research School of Engineering, Australian National University, ACT 2601, Australia (e-mail: alireza.khosravian@anu.edu.au; jochen.trumpf@anu.edu.au; robert.mahony@anu.edu.au).

This work was partially supported by the Australian Research Council through the ARC Discovery Project DP120100316 "Geometric Observer Theory for Mechanical Control Systems".

the delayed system state. Moreover, we provide a lower bound for the convergence rate of the predictor as a function of the Lipschitz constant of the system and the amount of delay. We demonstrate the superior performance of the proposed methodology over a Lyapunov-Krasovskii approach via numerical simulations.

The structure of the paper is as follows. Background and problem formulation is given in Section II. We discuss the predictor-observer approach in Section III where we propose the predictor and prove its co-stability. Performance of the proposed methodology is demonstrated via simulations in Section IV and a brief conclusion is given in Section V.

II. Problem formulation

Consider the system

$$\dot{x}(t) = f(x(t), u(t)), \qquad x(0) = x_0 \tag{1}$$

where $x \in \mathbb{R}^n$ is the internal state and $u \in \mathbb{R}^m$ is the input. Denote the outputs of the system by $y(t) = h(x(t)) \in \mathbb{R}^m$. The problem that we consider in this paper is to estimate the current state $x(t)$ when the measurements of the output are delayed such that the output measurement at time t is $y(t-\tau) = h(x(t-\tau))$ for some known constant delay $\tau \geq 0$, but delay-free measurements of the input $u(t)$ are available. We impose the following assumption.

Assumption 1: If the current output measurements $y(t) = h(x(t))$ are available without any delay, a time-invariant[1] state observer with the desired performance is available that provides estimates of $x(t)$. □

Using Assumption 1, we can re-interpret the problem formulation as follows. Assuming that an observer with the desired performance is available when output measurements are delay-free, we aim to propose an estimation algorithm that exhibits the same desired performance when the output measurements are delayed.

III. Observer-predictor approach

The approach that we take to solve the problem discussed in Section II is to combine the observer of Assumption 1 with a predictor in an *observer-predictor* arrangement (see Fig. 1). We inject the delayed output measurements $y(t-\tau)$ together with the delayed input measurements $u(t-\tau)$ into the observer in order to obtain estimates $\hat{x}^\tau(t)$ of $x(t-\tau)$, noting that the convergence of $\hat{x}^\tau(t)$ to $x(t-\tau)$ with the desired performance is ensured due to the assumed time-invariance property of the observer. We inject $\hat{x}^\tau(t)$ into the predictor, whose role is to use the information of the input signal $u(t)$ in order to compute how much the state has changed in the period $t-\tau$ to t and use the result of this computation together with the estimate $\hat{x}^\tau(t)$ to provide a prediction of the current state, denoted by $x^p(t) \in \mathbb{R}^n$. In this paper, we focus only on the predictor part as the observer is assumed to be known *a priori*.

[1]Time-invariant is in the sense of considering the system inputs and outputs altogether as the input of the observer and considering the state estimate as the output of the observer.

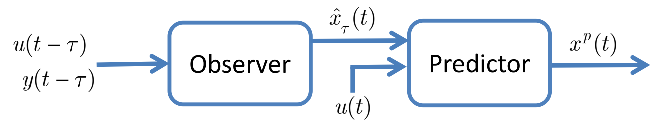

Fig. 1. Illustration of observer-predictor arrangment.

We propose the following predictor

$$\dot{\delta}(t) = f(\hat{x}^\tau(t) + \delta(t) - \delta(t-\tau), u(t)), \quad t \geq \tau \tag{2}$$

$$x^p(t) = \hat{x}^\tau(t) + \delta(t) - \delta(t-\tau), \quad t \geq \tau \tag{3}$$

where $\delta(t) \in \mathbb{R}^n$ is the internal state of the predictor which is obtained using the delay differential equation (2) with an arbitrary initial condition $\delta|_{[0,\tau[} = \delta_0|_{[0,\tau[}$. Although the dynamics of the predictor (2)-(3) is a delay differential equation, it can be easily implemented *recursively* using a memory buffer that stores the trajectory of $\delta(t)$ at least for τ seconds. This makes the predictor ideal for embedded applications where computational load is important.

In the observer-predictor arrangement, the trajectory of the predicted state $x^p(t)$ inherently depends on the trajectory of the observer $\hat{x}^\tau(t)$. Hence, rather than discussing the prediction error $|x^p(t) - x(t)|$ independently, we need to discuss how the prediction error depends on the delayed estimation error of the corresponding observer, i.e. on $\hat{x}^\tau(t) - x(t-\tau)$. The next theorem shows that the state prediction error can be bounded in terms of the delayed estimation error of the observer. We need the following assumption to present the Theorem.

Assumption 2: f is globally Lipschitz with respect to x, that is there exists a constant $0 < L \in \mathbb{R}$ such that $|f(x_1, u) - f(x_2, u)| \leq L|x_1 - x_2|$ for all $x_1, x_2 \in \mathbb{R}^n$ and all $u \in \mathbb{R}^m$ where $|.|$ denotes the Euclidean norm on $\mathbb{R}^n$. □

Theorem 1: Consider the predictor (2)-(3) for system (1) and suppose that Assumption 2 holds. Suppose moreover that a constant $\sigma > 0$ exists such that

$$L\frac{\exp(\sigma\tau) - 1}{\sigma} < 1. \tag{4}$$

Then, for all $t \geq \tau$ we have

$$|x^p(t) - x(t)| \leq \tag{5}$$
$$A \sup_{\tau \leq s \leq t} \big(\exp(\sigma(s-t))|\hat{x}^\tau(s) - x(s-\tau)|\big) + B\exp(-\sigma t)$$

where the constants A and B are $A = \frac{\sigma}{\sigma - L(\exp(\sigma\tau)-1)}$ and $B = \sup_{0 \leq s < \tau} \big(\exp(\sigma s)|x^p(s) - x(s)|\big)$. □

Proof of Theorem 1 is given in the appendix. Note that, using Assumption 2, it is straight-forward to show the solutions of the delay differential equation (2) are unique and exist for all $t \geq \tau$ [27]. Inequality (4) imposes an upper bound on the amount of delay for which the proposed observer-predictor methodology works. This upper bound is reversely related to the Lipschitz constant L of the system. According to the inequality (5), the total prediction error $|x^p(t) - x(t)|$ comprises two terms. The term $A \sup_{\tau \leq s \leq t} \big(\exp(\sigma(s -$

Preprint submitted to 54th IEEE Conference on Decision and Control.
Received March 23, 2015.

$t))|\hat{x}^\tau(s) - x(s-\tau)|\big)$ represents the effect of the observer error and the term $B\exp(-\sigma t)$ represents the error due to the mismatch between the initial condition of the predictor and the system. The error due to the initial condition vanishes as time goes to infinity. Ideally, we would like that $|x^p(t) - x(t)|$ remains bounded and converges asymptotically/exponentially to zero if $|\hat{x}^\tau(t) - x(t-\tau)|$ is bounded and converges asymptotically/exponentially to zero. The following definitions formalize this notion.

Definition 1: A predictor is called co-stable if $|x^p(t) - x(t)|$ is bounded provided that $|\hat{x}^\tau(t) - x(t-\tau)|$ is bounded. The predictor is called asymptotically (resp. exponentially) co-stable if it is co-stable and $|x^p(t) - x(t)|$ converges asymptotically (resp. exponentially) to zero provided that $|\hat{x}^\tau(t) - x(t-\tau)|$ converges asymptotically (resp. exponentially) to zero. □

Note that exponential co-stability of a predictor *does not* generally imply its asymptotic co-stability. The following corollary shows that the predictor (2)-(3) satisfies the above defined co-stability properties.

Corollary 1: Under the assumptions of Theorem 1, the predictor (2)-(3) is both asymptotically and exponentially co-stable. Moreover, if $\hat{x}^\tau(t)$ converges to $x(t-\tau)$ exponentially with convergence rate $\alpha > 0$, then a lower bound for the convergence rate of $x^p(t)$ to $x(t)$ is given by $\min(\sigma, \alpha)$ where σ is given in Theorem 1. □

Proof of Corollary 1 is given in the appendix. The convergence rate of $x^p(t)$ to $x(t)$ depends on both the convergence rate of the observer and also on the variable σ defined in Theorem 1. The variable σ roughly represents the convergence rate due to the predictor part in the observer-predictor arrangement. Using (4), one obtains $\tau \le \frac{1}{\sigma}\ln(1 + \frac{\sigma}{L})$. Given a system with a Lipschitz constant L, this inequality provides a lower bound for the maximum delay for which the convergence rate σ can be achieved. Defining $\bar{\tau} := \frac{1}{\sigma}\ln(1 + \frac{\sigma}{L})$, $\bar{\tau}$ versus σ is illustrated in Fig. 2 for different values of the Lipschitz constant L. According to Fig. 2, larger delays yield slower convergence rates of the predictor. For $\bar{\tau} \to 0$ we have $\sigma \to \infty$. This is because for zero delay the term $\delta(t) - \delta(t-\tau)$ in (3) equals to zero and hence the predictor does not contribute to the trajectory of $x^p(t)$. The maximum amount of delay (determined by Theorem 1) for which the predictor remains co-stable is obtained by computing $\bar{\tau}$ for σ close to zero. This maximum delay is given by $\bar{\tau}_{\max} = \lim_{\sigma \to 0^+} \ln(1 + \frac{\sigma}{L}) = \frac{1}{L}$. In practice, this is a lower bound for the maximum delay since the inequality (4) is just a sufficient condition for the co-stability of the predictor (see Section IV).

Remark 1: Defining $\eta(t) := \delta(t) - \delta(t-\tau)$, the predictor dynamics (2)-(3) can be rewritten as

$$\begin{aligned} \dot{\eta}(t) &= f(\hat{x}^\tau(t) + \eta(t), u(t)), \qquad (6)\\ &\quad - f(\hat{x}^\tau(t-\tau) + \eta(t-\tau), u(t-\tau)), \quad t \ge \tau \\ x^p(t) &= \hat{x}^\tau(t) + \eta(t), \quad t \ge \tau. \qquad (7) \end{aligned}$$

In an ideal condition where there is no input noise, there is no important difference between implementing the predictor either based on (2)-(3) or (6)-(7). However, when the input $u(t)$ is disturbed with measurements noise, this noise is accumulated in the variable $\eta(t)$ as we continue integrating the dynamics (6). A practical advantage of (2)-(3) is that the effect of the measurement noise is canceled via the term $\delta(t) - \delta(t-\tau)$ and does not accumulate in the variable $\delta(t)$ during the integration of dynamics (2). Hence, the quality of the prediction $x^p(t)$ of (2)-(3) degrades less than the corresponding prediction provided by (6)-(7) in the presence of input noise.

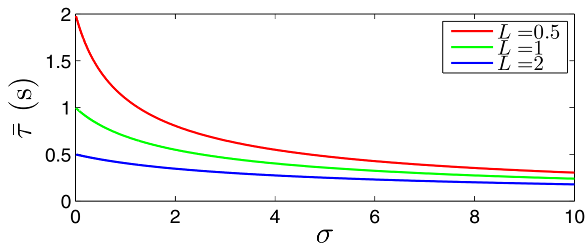

Fig. 2. Maximum delay ($\bar{\tau}$) versus the convergence rate of predictor (σ).

The above discussion can be repeated by differentiating $x^p(t)$ of (2) to obtain the following predictor.

$$\dot{x}^p(t) = \dot{\hat{x}}^\tau(t) + f(x^p(t), u(t)) - f(x^p(t-\tau), u(t-\tau)), \tag{8}$$

for all $t \ge \tau$. The difference here is that this dynamics depends on $\dot{\hat{x}}^\tau(t)$ (rather than $\hat{x}^\tau(t)$ directly) and hence the information of the initial condition of the dynamics is lost during differentiation. One can add an innovation term to the dynamics (8) in order to stabilize it to the trajectory $x^p(t)$ of (2)-(3). This is in fact what the predictor of [1, Eq. (3.8)] does. Consequently, it is evident that the predictions of the current state by [1, Eq. (3.8)] exponentially converge to the predictions provided by our proposed predictor (2)-(3). Note that [1, Eq. (3.8)] uses an integral innovation term that is not immediately implementable recursively unless one implements this term based on a differential equation. Similar to the discussions provided earlier, the noise characteristics of the predictor (2)-(3) is superior to [1, Eq. (3.8)] because of the noise cancellation effect due to the term $\delta(t) - \delta(t-\tau)$. □

Remark 2: In ideal conditions where the input measurements are noise-free, it is possible to show that the solution $\delta(t)$ of the delay differential equation (2) is bounded for all $t \ge \tau$. In practice, when measurement noise exists, the amplitude of $\delta(t)$ might grow larger and larger due to the integration of the input noise. This might cause numerical errors when computing the prediction $x^p(t)$ using (3) since the term $\delta(t) - \delta(t-\tau)$ might remain bounded while $\delta(t)$ and $\delta(t-\tau)$ both grow very large. For the particular case where the system dynamics is invariant, a periodic resetting technique has been proposed in [8, Lemma 1] to keep the trajectories of $\delta(t)$ bounded while maintaining the co-stability of the predictor. An alternative method is to use feedback in the predictor dynamics (2) to keep its trajectories bounded in the presence of input noise. In the general case where the underlying system does not have any particular invariance, this remains an open question for future research. Note that a similar boundedness issue also occurs in the predictor proposed in [1, Eq. (3.8)] if the corresponding integral terms are recursively computed using differential equations. □

IV. SIMULATIONS

In this section, we provide numerical simulations to compare the performance of the proposed observer-predictor approach with the Lyapunov-Krasovskii method proposed in [18]. As in [18, Section 4], we consider the following system.

$$\dot{x}_1(t) = c_1 x_2(t) - l_1 x_1(t), \tag{9a}$$

$$\dot{x}_2(t) = c_2 \sin(x_2(t)) + c_3 \cos(x_2(t)) + c_4 u(t), \tag{9b}$$

$$y(t-\tau) = x_1(t-\tau) \tag{9c}$$

When the output measurement delay τ is zero, the system (9) belongs to a class of uniformly observable systems for which various observer design methods are available in the literature (see e.g. [28]–[31]). We employ the observer design methodology of [28][2] to obtain an observer and then feed the resulting observer with the delayed measurement (9c) and the delayed input to estimate the delayed state $x(t-\tau) = [x_1(t-\tau), x_2(t-\tau)]^\top$ as follows.

$$\begin{bmatrix} \dot{\hat{x}}_1^\tau(t) \\ \dot{\hat{x}}_2^\tau(t) \end{bmatrix} = \begin{bmatrix} c_1 \hat{x}_2^\tau(t) - l \hat{x}_1^\tau(t) \\ c_2 \sin(\hat{x}_2^\tau(t)) + c_3 \cos(\hat{x}_2^\tau(t)) + c_4 u(t-\tau) \end{bmatrix} - \mathrm{diag}(\theta, \theta^2) S^{-1} C^\top (\hat{x}_1^\tau(t) - y(t-\tau)) \tag{10}$$

where $\theta > 1$ and $\mathbb{R}^{2\times 2} \ni S = S^\top > 0$ are observer gains and $C = [1,\ 0]$. With appropriate choices of observer gains, globally exponential convergence of $\hat{x}^\tau(t)$ to $x(t-\tau)$ is proved in [28]. We feed the estimate $\hat{x}^\tau(t)$ into the predictor (2)-(3) to obtain the prediction of the current state $x(t)$. For the simulation, we choose the same system parameters and observer gains as [18, Section 4], that is $c_1 = 1$, $c_2 = c_3 = 0.02$, $c_4 = 8$, $l = 0.04$, $\theta = 1.55$, and

$$S = \begin{bmatrix} 1 & -1 \\ -1 & 2 \end{bmatrix}.$$

Using these system parameters, one can show that a lower bound for the Lipschitz constant of the system is $L = 1$. According to Fig. 2, the proposed predictor (2)-(3) is proven to be co-stable for delays smaller than $\bar{\tau}_{\max} = \frac{1}{L} = 1$ (s). On the other hand, as reported in [18, Section 4], the observer of [18, Eq. (7)] is stable for delays smaller than 0.01 (s). Hence, the predictor presented in this paper is provably able to compensate for much larger delays. We chose the initial condition $x(0) = [-50, -50]^\top$ for the system (9) and we initialize both the observer (10), the predictor dynamics (2), and the observer of [18, Eq. (7)] with zero initial conditions. For the small delay of $\tau = 0.01$ (s), Fig. 3 shows the system trajectory $x(t)$ along with the prediction $x^p(t)$ provided by (2)-(3) and the estimate $\hat{x}(t)$ provided by [18, Eq. (7)]. As demonstrated by Fig. 3, estimates of the current state provided by either the method proposed in this paper or the approach presented in [18] converge to the true system trajectory with similar convergence rates (note that the estimated trajectories are not available for $t < \tau$ since $y(t-\tau)$ is unavailable for that period). Fig. 4 shows the plot of the same variables when the delay is increased to $\tau = 0.45$ (s). According to Fig. 4, the approach presented in this paper still provides exponentially convergent predictions of the current state while the observer of [18] becomes unstable. Even though the co-stability of the proposed predictor is only proved for delays of up to $\bar{\tau}_{\max} = 1$ (s), our simulations show that the observer-predictor approach remains stable for much larger delays. Fig. 5 shows the state predictions for the very large delay of $\tau = 10$ (s). As is evident from Fig. 5, the state predictions still converge to the true system trajectories, though their convergence rate is decreased compared to when the delay is small. This observation is compatible with Fig. 2 which suggests that the convergence rate decreases as the delay becomes larger.

[2]We have used a slightly different notation from [28] to be more compatible with the notations used in [18].

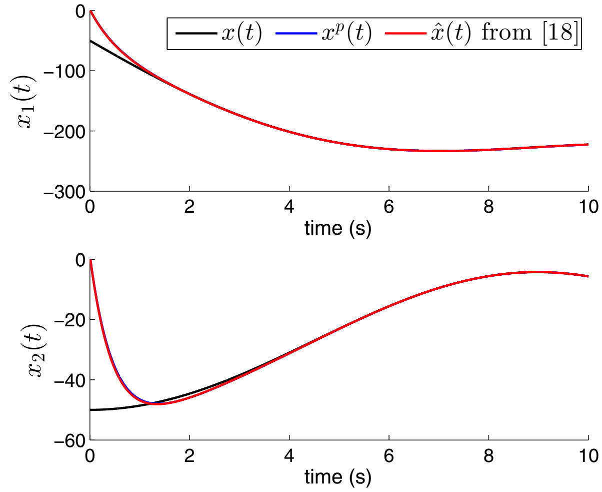

Fig. 3. Comparison of the estimation methode of this paper vs [18] for $\tau = 0.01$ (s).

V. CONCLUSION

We propose a novel predictor to compensate for output measurement delays for state estimation of nonlinear systems. The predictor is employed in a cascade observer-predictor arrangement where estimates of the delayed system state are provided by an observer and the predictor's role is to compensate for the delay. We provide rigorous co-stability results demonstrating that the predictions of the state asymptotically/exponentially converge to the current system trajectory if the observer's state estimates converge asymptotically/exponentially to the delayed system state. The computational complexity of the predictor is low as it is recursively implemented using simple delay differential equations. Superior performance of the proposed predictor over a Lyapunov-Krasowskii method and its robustness

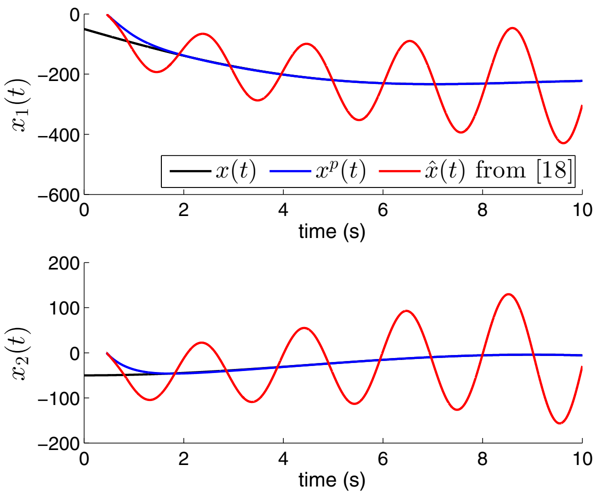

Fig. 4. Comparison of the estimation methode of this paper vs [18] for $\tau = 0.45$ (s).

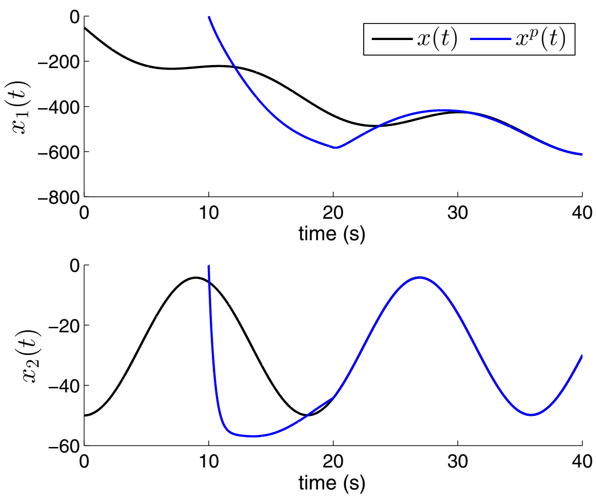

Fig. 5. Performance of the estimation methode of this paper for $\tau = 10$ (s).

against large measurement delays are demonstrated using numerical simulations. Methods to ensure boundedness of the internal state of the predictor in the presence of input noise as well as chaining multiple copies of the predictor to compensate for larger delays potential topics for future research.

APPENDIX

Proof of Theorem 1:

The proof is inspired by the proof of [1, Lemma 3.3]. Integrating the sides of (1) from $t-\tau$ to t yields

$$x(t) = x(t-\tau) + \int_{t-\tau}^{t} f(x(s), u(s))ds. \qquad (11)$$

Similarly, integrating the sides of (2) from $t-\tau$ to t and replacing into (3), it is easy to see that the predictor (2)-(3) can be rewritten into the following integral formulation

$$x^p(t) = \hat{x}^\tau(t) + \int_{t-\tau}^{t} f(x^p(s), u(s))ds, \quad t \geq \tau \qquad (12)$$

with the initial condition $x^p|_{[0,\tau[} = \hat{x}^p_0|_{[0,\tau[} := \hat{x}^\tau(\tau)+\delta(\tau)-\delta(0)$. Using (12) and (11) and resorting to Assumption 2 we have

$$\begin{aligned} |x^p(t) - x(t)| &= |\hat{x}^\tau(t) - x(t-\tau)| \\ &\quad + \int_{t-\tau}^{t} \big(f(x^p(s), u(s)) - f(x(s), u(s))\big)ds| \\ &\leq |\hat{x}^\tau(t) - x(t-\tau)| \\ &\quad + \int_{t-\tau}^{t} |f(x^p(s), u(s)) - f(x(s), u(s))|ds \\ &\leq |\hat{x}^\tau(t) - x(t-\tau)| + L\int_{t-\tau}^{t} |x^p(s) - x(s)|ds \end{aligned}$$

Multiplying by $\exp(\sigma t)$ on both sides we have

$$\begin{aligned} \exp(\sigma t)|x^p(t) - x(t)| \leq\ &\exp(\sigma t)|\hat{x}^\tau(t) - x(t-\tau)| \qquad (13) \\ &+ L\exp(\sigma t)\int_{t-\tau}^{t} |x^p(s) - x(s)|ds. \end{aligned}$$

On the other hand, we have

$$\begin{aligned} &\int_{t-\tau}^{t} |x^p(s) - x(s)|ds = \int_{t-\tau}^{t} \exp(\sigma(s-s))|x^p(s) - x(s)|ds \\ &\leq \sup_{t-\tau\leq s\leq t} \big(\exp(\sigma s)|x^p(s) - x(s)|\big) \int_{t-\tau}^{t} \exp(-\sigma s)ds \\ &= \exp(-\sigma t)\frac{\exp(\sigma\tau) - 1}{\sigma} \sup_{t-\tau\leq s\leq t} \big(\exp(\sigma s)|x^p(s) - x(s)|\big). \end{aligned} \qquad (14)$$

Using (13) and (14) we have

$$\begin{aligned} &\exp(\sigma t)|x^p(t) - x(t)| \leq \qquad (15) \\ &\quad \exp(\sigma t)|\hat{x}^\tau(t) - x(t-\tau)| \\ &\quad + L\frac{\exp(\sigma\tau) - 1}{\sigma} \sup_{t-\tau\leq s\leq t} \big(\exp(\sigma s)|x^p(s) - x(s)|\big) \end{aligned}$$

Assuming $t \geq \tau$ and taking the supremum over t in (15) we have

$$\begin{aligned} &\sup_{\tau\leq s\leq t} \big(\exp(\sigma s)|x^p(s) - x(s)|\big) \leq \qquad (16) \\ &\quad \sup_{\tau\leq s\leq t} \big(\exp(\sigma s)|\hat{x}^\tau(s) - x(s-\tau)|\big) \\ &\quad + L\frac{\exp(\sigma\tau) - 1}{\sigma} \sup_{0\leq s\leq t} \big(\exp(\sigma s)|x^p(s) - x(s)|\big) \end{aligned}$$

Now, we have two cases depending on the time at which the supremum occurs.

Case 1 (supremum occurs in $[0, \tau]$*):* If $\sup_{0\leq s\leq t} \big(\exp(\sigma s)|x^p(s) - x(s)|\big) = \sup_{0\leq s\leq \tau} \big(\exp(\sigma s)|x^p(s) -$

$x(s)|\big)$ then (16) simplifies to

$$\sup_{\tau\le s\le t}\big(\exp(\sigma s)|x^p(s)-x(s)|\big)\le \sup_{\tau\le s\le t}\big(\exp(\sigma s)|\hat{x}^\tau(s)-x(s-\tau)|\big) + L\frac{\exp(\sigma\tau)-1}{\sigma}\sup_{0\le s\le\tau}\big(\exp(\sigma s)|x^p(s)-x(s)|\big)\big) \le \frac{\sigma}{\sigma-L(\exp(\sigma\tau)-1)}\sup_{\tau\le s\le t}\big(\exp(\sigma s)|\hat{x}^\tau(s)-x(s-\tau)|\big) + \sup_{0\le s\le\tau}\big(\exp(\sigma s)|x^p(s)-x(s)|\big)\big) \quad (17)$$

where the last inequality is derived using (4) which implies $\frac{\sigma}{\sigma-L(\exp(\sigma\tau)-1)}>1$.

Case 2 (supremum occurs in $[\tau,t]$): If $\sup_{0\le s\le t}\big(\exp(\sigma s)|x^p(s)-x(s)|\big)=\sup_{\tau\le s\le t}\big(\exp(\sigma s)|x^p(s)-x(s)|\big)$ then (16) is simplifies to

$$\sup_{\tau\le s\le t}\big(\exp(\sigma s)|x^p(s)-x(s)|\big)\le \frac{\sigma}{\sigma-L(\exp(\sigma\tau)-1)}\sup_{\tau\le s\le t}\big(\exp(\sigma s)|\hat{x}^\tau(s)-x(s-\tau)|\big) \le \frac{\sigma}{\sigma-L(\exp(\sigma\tau)-1)}\sup_{\tau\le s\le t}\big(\exp(\sigma s)|\hat{x}^\tau(s)-x(s-\tau)|\big) + \sup_{0\le s\le\tau}\big(\exp(\sigma s)|x^p(s)-x(s)|\big)\big) \quad (18)$$

Combining (17) and (18) and noting that $\exp(\sigma t)|x^p(t)-x(t)|\le\sup_{\tau\le s\le t}\big(\exp(\sigma s)|x^p(s)-x(s)|\big)$ yields

$$|x^p(t)-x(t)|\le \frac{\sigma}{\sigma-L(\exp(\sigma\tau)-1)}\sup_{\tau\le s\le t}\big(\exp(\sigma(s-t))|\hat{x}^\tau(s)-x(s-\tau)|\big) + \exp(-\sigma t)\sup_{0\le s\le\tau}\big(\exp(\sigma s)|x^p(s)-x(s)|\big)\big)$$

for all $t\ge\tau$ as claimed in (5). ∎

Proof of Corollary 1:

Assume that $|\hat{x}^\tau(t)-x(t-\tau)|\le c$ for some $c\ge 0$ and all $t\ge\tau$. By (5) we have $|x^p(t)-x(t)|\le Ac+B$ for all $t\ge\tau$. This proves that the predictor is co-stable.

Assume $\lim_{t\to\infty}|\hat{x}^\tau(t)-x(t-\tau)|=0$. Then for all $\epsilon_\tau>0$, there exists a T_τ such that for all $t\ge T_\tau$ we have $|\hat{x}^\tau(t)-x(t-\tau)|<\epsilon_\tau$. By (5) we have $|x^p(t)-x(t)|\le A\sup_{\tau\le s\le t}\big(\exp(\sigma(s-t))|\hat{x}^\tau(s)-x(s-\tau)|\big)+B\exp(-\sigma t)<A\epsilon_\tau\sup_{\tau\le s\le t}\big(\exp(\sigma(s-t))\big)+B\exp(-\sigma t)<A\epsilon_\tau+B\exp(-\sigma t)$ for all $t\ge T_\tau$. For a given $\epsilon>0$, choose ϵ_τ small enough and T^* large enough such that $A\epsilon_\tau+B\exp(-\sigma T^*)<\epsilon$ (such ϵ_τ and T^* always exist). We have $|x^p(t)-x(t)|<\epsilon$ for all $t\ge\max(T^*,T_\tau)$ which means that $\lim_{t\to\infty}|x^p(t)-x(t)|=0$. This proves that the predictor is asymptotically co-stable.

If the estimation error $\hat{x}^\tau(t)-x(t-\tau)$ converges exponentially to zero then there exist $A_0,\alpha>0$ such that $|\hat{x}^\tau(t)-x(t-\tau)|\le A_0|\hat{x}^\tau(\tau)-x(0)|\exp(-\alpha t)$ for all $t\ge\tau$. This yields

$$\sup_{\tau\le s\le t}\big(\exp(\sigma(s-t))|\hat{x}^\tau(s)-x(s-\tau)|\big) \le A_0|\hat{x}^\tau(\tau)-x(0)|\exp(-(\sigma+\alpha)t)\sup_{\tau\le s\le t}\big(\exp(\sigma(s))\big) \le A_0|\hat{x}^\tau(\tau)-x(0)|\exp(-\alpha t). \quad (19)$$

for all $t\ge\tau$. Substituting (19) into (5) implies that $|x^p(t)-x(t)|\le\bar{A}\exp(-\bar{\alpha}t)$ for all $t\ge\tau$ where $\bar{\alpha}=\min(\sigma,\alpha)$ and $\bar{A}=AA_0|\hat{x}^\tau(\tau)-x(0)|\exp((\bar{\alpha}-\alpha)\tau)+B\exp((\bar{\alpha}-\sigma)\tau)$. This shows that the predictor is exponentially co-stable and completes the proof. ∎

References

[1] T. Ahmed-Ali, I. Karafyllis, and F. Lamnabhi-Lagarrigue, "Global exponential sampled-data observers for nonlinear systems with delayed measurements," *Systems & Control Letters*, vol. 62, no. 7, pp. 539–549, 2013.

[2] B. Khaleghi, A. Khamis, F. O. Karray, and S. N. Razavi, "Multisensor data fusion: A review of the state-of-the-art," *Information Fusion*, vol. 14, no. 1, pp. 28–44, 2013.

[3] F. Deza, E. Busvelle, J. Gauthier, and D. Rakotopara, "High gain estimation for nonlinear systems," *Systems & Control letters*, vol. 18, no. 4, pp. 295–299, 1992.

[4] W. H. Heemels, A. R. Teel, N. van de Wouw, and D. Nesic, "Networked control systems with communication constraints: Tradeoffs between transmission intervals, delays and performance," *IEEE Trans. Automatic Control*, vol. 55, no. 8, pp. 1781–1796, 2010.

[5] H. Hammouri, M. Nadri, and R. Mota, "Constant gain observer for continuous-discrete time uniformly observable systems," in *Proc. IEEE Conf. Decision and Control*, 2006, pp. 5406–5411.

[6] I. Karafyllis and C. Kravaris, "From continuous-time design to sampled-data design of observers," *IEEE Trans. Automatic Control*, vol. 54, no. 9, pp. 2169–2174, 2009.

[7] A. Khosravian, J. Trumpf, R. Mahony, and T. Hamel, "Recursive attitude estimation in the presence of multi-rate and multi-delay vector measurements," in *Proc. American Control Conference*, July 2015, (to appear).

[8] ——, "Velocity aided attitude estimation on SO(3) with sensor delay," in *Proc. IEEE Conf. on Decision and Control*, December 2014.

[9] A. W. Olbrot, "Observability and observers for a class of linear systems with delays," *IEEE Trans. Automatic Control*, vol. 26, no. 2, pp. 513–517, 1981.

[10] M. Boutayeb, "Observers design for linear time-delay systems," *Systems & Control Letters*, vol. 44, no. 2, pp. 103–109, 2001.

[11] M. Fliess and H. Mounier, "Controllability and observability of linear delay systems: an algebraic approach," *ESAIM: Control, Optimisation and Calculus of Variations*, vol. 3, pp. 301–314, 1998.

[12] E. Lee and A. Olbrot, "Observability and related structural results for linear hereditary systems," *International Journal of Control*, vol. 34, no. 6, pp. 1061–1078, 1981.

[13] L. A. Márquez-Martínez, C. H. Moog, and M. Velasco-Villa, "Observability and observers for nonlinear systems with time delays," *Kybernetika*, vol. 38, no. 4, pp. 445–456, 2002.

[14] O. J. Smith, "Closer control of loops with dead time," *Chemical Engineering Progress*, vol. 53, no. 5, pp. 217–219, 1957.

[15] D. W. Clarke, C. Mohtadi, and P. Tuffs, "Generalized predictive controlpart i and ii," *Automatica*, vol. 23, no. 2, pp. 137–148, 1987.

[16] T. Ahmed-Ali, E. Cherrier, and F. Lamnabhi-Lagarrigue, "Cascade high gain predictors for a class of nonlinear systems," *IEEE Trans. Automatic Control*, vol. 57, no. 1, pp. 221–226, 2012.

[17] S. Bahrami and M. Namvar, "Delay compensation in global estimation of rigid-body attitude under biased velocity measurement," in *Proc. IEEE Conf. on Decision and Control*, December 2014.

[18] V. Van Assche, T. Ahmed-Ali, C. Hann, and F. Lamnabhi-Lagarrigue, "High gain observer design for nonlinear systems with time varying delayed measurements," in *IFAC World congress*, 2011, pp. 692–696.

[19] W. Aggoune, M. Boutayeb, and M. Darouach, "Observers design for a class of nonlinear systems with time-varying delay," in *Proc. IEEE Conf. on Decision and Control*, vol. 3, 1999, pp. 2912–2913.

[20] F. Cacace, A. Germani, and C. Manes, "An observer for a class of nonlinear systems with time varying observation delay," *Systems & Control Letters*, vol. 59, no. 5, pp. 305–312, 2010.

[21] P. Riseborough. [Online]. Available: https://github.com/priseborough/InertialNav (visited on September 2014)

[22] "APM: Navigation extended Kalman filter overview," ArduPilot/APM. [Online]. Available: http://copter.ardupilot.com/wiki/common-apm-navigation-extended-kalman-filter-overview/ (visited on September 2014)

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

[24] S. Battilotti, "Nonlinear predictors for systems with bounded trajectories and delayed measurements¡," in *Proc. IFAC World Congr.*, 2014, pp. 6812–9817.

[25] K. Subbarao and P. C. Muralidhar, "A state observer for iti systems with delayed outputs: Time-varying delay," in *American Control Conference*, 2008, pp. 3029–3033.

[26] N. Kazantzis and R. A. Wright, "Nonlinear observer design in the presence of delayed output measurements," *Systems & Control letters*, vol. 54, no. 9, pp. 877–886, 2005.

[27] A. Bellen and M. Zennaro, *Numerical methods for delay differential equations*. Oxford university press, 2013.

[28] G. Bornard and H. Hammouri, "A high gain observer for a class of uniformly observable systems," in *Proc. IEEE Conf. on Decision and Control*, 1991, pp. 1494–1496.

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

[30] H. Shim, Y. I. Son, and J. H. Seo, "Semi-global observer for multi-output nonlinear systems," *Systems & Control Letters*, vol. 42, no. 3, pp. 233–244, 2001.

[31] M. Farza, M. M'Saad, and L. Rossignol, "Observer design for a class of mimo nonlinear systems," *Automatica*, vol. 40, no. 1, pp. 135–143, 2004.

Preprint submitted to 54th IEEE Conference on Decision and Control.
Received March 23, 2015.