**Joint 48th IEEE Conference on Decision and Control and
28th Chinese Control Conference
Shanghai, P.R. China, December 16-18, 2009**

FrB07.1

Prediction-based Observation of Nonlinear Systems Non-affine in the Unmeasured States

Yannick Morel

Guidance, Navigation and Control Department
Institut Franco-Allemand de Recherche
de Saint-Louis
Saint-Louis 68300, France
+33 389695058
FAX: +33 389695002
Yannick.Morel@isl.edu

Alexander Leonessa

Mechanical Engineering Department
Virginia Polytechnic Institute
and State University
Blacksburg, VA 24061-0238
(540) 231-3268
FAX: (540) 231-8836
leonessa@vt.edu

Abstract— The presented work addresses the observation problem for a large class of nonlinear systems, including systems which are nonlinear in the unmeasured states. Assuming partial state measurements, the unmeasured states are reconstructed so that a prediction of the measured states converges to a neighborhood of the actual measurements. This prediction-based observer algorithm relies on carefully selected prediction-observation errors, designed using a backstepping technique. Lyapunov's direct method is used to show Lyapunov stability and convergence of these errors to an arbitrarily small neighborhood of the origin. The technique is applied to two different nonlinear systems. Results of numerical simulations are presented for both cases and illustrate the efficacy of the algorithm. Experimental results are also provided for one of the examples.

Keywords: observer, predictor, nonlinear systems, nonlinear observer, systems non-affine in the unmeasured states.

I. INTRODUCTION

The observation problem for linear systems has received considerable attention from the research community over the years, and a wide range of results can be found in the literature, with outstanding contributions from Kalman ([1], [2]) and Luenberger ([3], [4]), to name but a few. For nonlinear systems however, the observation problem still constitutes a challenging problem which is largely open.

The strength of the observation techniques developed for linear systems naturally led to attempts at extending them to nonlinear systems. Extended Kalman filtering ([5], [6]), for instance, is the result of one such tangent, and has shown its utility in a number of cases ([7]). Alternately, when only the output appears nonlinearly in the system's dynamics, it is possible to construct an observer for which the observation error dynamics are linear ([8], [9]). Linear techniques can thus be used to solve the observation problem for this particular type of systems, which are sometimes referred to as being in output feedback form ([8]).

The observation problem becomes more difficult if the considered nonlinear system is not in output feedback form, that is, if unmeasured states appear nonlinearly in the system's dynamics. High gain observers ([9], [10]) can be used to address the problem for such systems. Such techniques essentially consist in selecting observer gains sufficiently large such that the negative influence of nonlinearities on the observation errors is overcome. The approach can however lead to large transient observation errors, due to an issue referred to as peaking phenomenon ([9]). In [11], the authors present an alternate design methodology for nonlinear observers applicable to systems which are not in output feedback form, with an approach taking advantage of the notions of immersion and invariance ([12]). Although the technique is of great interest, it requires solving a set of partial differential equations, which, quoting the authors, can be "extremely difficult." Nevertheless, observers for a number of nonlinear systems can be constructed using the approach, a pair of them being presented in [11]. This immersion and invariance based procedure constitutes a fairly open-ended approach.

In the following, we introduce a novel observer design which, while concerned with a smaller class of systems, has the advantage of being systematic and relatively simple to apply. In a departure from high gain observation techniques such as those in [9], [10], the presented approach relaxes high gain requirements by using derivative estimates ([13], [14]). The observer relies on a partial state predictor, which compares predictions on the measured states with actual measurements. The prediction error is then used to estimate the unmeasured states. The design of this predictor-observer relies on specific prediction-observation errors, designed using a backstepping technique ([8]). A Lyapunov stability analysis of the errors' dynamics shows that they are Lyapunov stable, and converge to a neighborhood of the origin. The size of this neighborhood depends upon the accuracy of the derivative estimate, as well as on a number of design constants. Accordingly, the errors can be made arbitrarily small by using a high performance derivative estimator, and/or selecting appropriate gains.

The paper is structured as follows. Section II describes the class of systems considered and presents the observation strategy. The predictor-observer is then introduced in Section III. Performance of the obtained observer was tested through numerical simulations for two different systems, a nonlinear

oscillator and a three degree of freedom helicopter model. Simulation results are presented in Section IV and illustrate the efficacy of the proposed observer. The algorithm was also tested experimentally, on a three degree of freedom helicopter setup. Implementation results are given in Section V. Section VI concludes this paper.

II. PROBLEM STATEMENT AND OBSERVATION STRATEGY

Consider a system of the form

$$
\begin{aligned}
\dot{x}(t) &= f(x(t),t), \quad x(0)=x_0, \quad t \geq 0, \quad &(1)\\
y(t) &= Cx(t), &(2)
\end{aligned}
$$

where $x(t) \in \mathbb{R}^n$, $t \geq 0$, is the system's state vector, $y(t) \in \mathbb{R}^m$, $t \geq 0$, the measured output vector, with $m \geq n/2$, and $C = \begin{bmatrix} I_m & 0_{m \times p} \end{bmatrix}$, with $p = n - m \leq m$.

A sufficient condition for system (1)–(2) to be observable is that $\mathrm{rank}[J(x(t),t)] = n$, $t \geq 0$, ([15], [16]), where

$$
J(x(t),t) \triangleq \begin{bmatrix} \nabla y(t) \\ \nabla y^{(1)}(t) \\ \vdots \\ \nabla y^{(n-1)}(t) \end{bmatrix}, \quad t \geq 0, \quad (3)
$$

where $\nabla g(x) \triangleq \frac{\mathrm{d}g(x)}{\mathrm{d}x}$, and $(\cdot)^{(i)}$ denotes the ith derivative with respect to time. In the following, we focus on a subset of the set of observable systems as defined by the above condition. More specifically, we consider systems for which $\mathrm{rank}[J_{\mathrm{s}}(x(t),t)] = n$, $t \geq 0$, where

$$
J_{\mathrm{s}}(x(t),t) \triangleq \begin{bmatrix} \nabla y(t) \\ \nabla y^{(1)}(t) \end{bmatrix}, \quad t \geq 0. \quad (4)
$$

The goal of the presented work is to reconstruct the entire state vector $x(t)$, $t \geq 0$, using the information provided by the output measurement $y(t)$, $t \geq 0$, and partial knowledge of the form of the right-hand-side of (1). To reach this goal, we define $x_1(t) \triangleq y(t)$, $t \geq 0$, and decompose the state vector as $x(t) = \begin{bmatrix} x_1^{\mathrm{T}}(t) & x_2^{\mathrm{T}}(t) \end{bmatrix}^{\mathrm{T}}$, $t \geq 0$, where $x_2(t) \in \mathbb{R}^p$, $t \geq 0$, denotes the unmeasured states which we attempt to reconstruct. Furthermore, the dynamics of the measured states are described by

$$
\dot{x}_1(t) = f_1(x_1(t),x_2(t),t), \quad x_1(0) = Cx_0 \triangleq x_{10} = y(0),
$$
$$
t \geq 0, \quad (5)
$$

where $f_1(x_1(t),x_2(t),t) \triangleq Cf(x(t),t)$, $t \geq 0$. Note that, using (2) and (5), we can rewrite (4) as

$$
J_{\mathrm{s}}(x,t) \triangleq \begin{bmatrix} I_m & 0_{m \times p} \\ \frac{\partial f_1(x_1,x_2,t)}{\partial x_1} & \frac{\partial f_1(x_1,x_2,t)}{\partial x_2} \end{bmatrix}, \quad (6)
$$

and the above observability condition reduces to the condition that $\mathrm{rank}[\partial f_1(x_1,x_2,t)/\partial x_2] = p$. In addition, from (2) and (5), we have that

$$
\dot{y}(t) - f_1(y(t),x_2(t),t) = 0, \quad t \geq 0. \quad (7)
$$

According to the implicit function theorem ([17]), if $\mathrm{rank}[\partial f_1(x_1,x_2,t)/\partial x_2] = p$, there exists a unique function $g : \mathbb{R}^m \times \mathbb{R}^m \times \mathbb{R} \to \mathbb{R}^p$, such that $g(y,\dot{y},t) = x_2$. However, obtaining a closed form expression of this function can prove

challenging, and its existence can only be guaranteed locally. In addition, while $y(t)$, $t \geq 0$, is measured and hence known, $\dot{y}(t)$, $t \geq 0$, is not necessarily available. Thus, directly working in terms of the above $g(\cdot)$ does not constitute a viable option.

Instead, we will *indirectly* work in terms of $g(\cdot)$. In particular, we will construct an estimate $\hat{x}_2(t)$ of $x_2(t)$, $t \geq 0$, which uniformly converges to a neighborhood of $g(y(t),\dot{y}(t),t)$, $t \geq 0$. In addition, since $\dot{y}(t)$, $t \geq 0$, is not measured, we will estimate its value, using measured information (i.e. $y(t)$, $t \geq 0$) and a derivative estimator (such as those in [13], [14]). The resulting estimate $\hat{x}_2(t)$, $t \geq 0$, is then further refined by adjusting it according to a prediction error. The latter is computed as the difference between the measured output and a predicted $\hat{x}_1(t)$, $t \geq 0$, obtained by integrating a partial state predictor whose form is similar to (5), with $x_2(t)$ replaced by $\hat{x}_2(t)$, $t \geq 0$.

III. NONLINEAR OBSERVER DESIGN

Following the approach delineated in Section II, it is possible to construct a partial state predictor of the form

$$
\dot{\hat{x}}_1(t) = f_1(x_1(t),\hat{x}_2(t),t), \quad \hat{x}_1(0) = x_{10}, \quad t \geq 0, \quad (8)
$$

where $\hat{x}_2(t) \in \mathbb{R}^p$, $t \geq 0$, is designed so that $\hat{x}_1(t) \in \mathbb{R}^m$, $t \geq 0$, provides a predicted value of $x_1(t)$, $t \geq 0$. However, to provide more flexibility in the design procedure, the algorithm will rely on predictions of an augmented output $x_{1\mathrm{a}}(t)$, $t \geq 0$, obtained from

$$
\dot{x}_{1\mathrm{a}}(t) = f_1(x_1(t),x_2(t),t) + \dot{x}_{\mathrm{a}}(t), \quad x_{1\mathrm{a}}(0) = x_{10}, \quad t \geq 0, (9)
$$

where $\dot{x}_{\mathrm{a}}(t) \in \mathbb{R}^m$, $t \geq 0$, is the rate of change of a known augmenting signal, generated by

$$
\ddot{x}_{\mathrm{a}}(t) = W_{\mathrm{a}}(x_1(t),\hat{x}_2(t),t)v(t) + u(t), \quad \dot{x}_{\mathrm{a}}(0) = x_{\mathrm{a}}(0) = 0_m,
$$
$$
t \geq 0, (10)
$$

where the matrix $W_{\mathrm{a}}(\cdot) \in \mathbb{R}^{m \times (m-p)}$ is designed so that $W_2(x_1,\hat{x}_2,t) \triangleq \begin{bmatrix} \frac{\partial f_1(x_1,\hat{x}_2,t)}{\partial \hat{x}_2} & W_{\mathrm{a}}(x_1,\hat{x}_2,t) \end{bmatrix} \in \mathbb{R}^{m \times m}$ is nonsingular, and $v(t) \in \mathbb{R}^{m-p}$, $u(t) \in \mathbb{R}^m$, $t \geq 0$, are signals to be designed at a later stage of the proceedings. In particular, $u(t)$, $t \geq 0$, will be constructed to ensure boundedness of $x_{\mathrm{a}}(t)$.

Next, we construct the following partial state predictor,

$$
\dot{\hat{x}}_{1\mathrm{a}}(t) = f_1(x_1(t),\hat{x}_2(t),t) + \dot{x}_{\mathrm{a}}(t), \quad \hat{x}_{1\mathrm{a}}(0) = x_{10}, \quad t \geq 0, (11)
$$

where $\hat{x}_{1\mathrm{a}}(t) \in \mathbb{R}^m$ is a prediction of $x_{1\mathrm{a}}(t)$, $t \geq 0$, and $\hat{x}_2(t) \in \mathbb{R}^p$ is an estimate of $x_2(t)$, $t \geq 0$.

In the following, we will construct the vector $\dot{\hat{x}}_{2\mathrm{a}}(t) \triangleq \begin{bmatrix} \dot{\hat{x}}_2^{\mathrm{T}}(t) & v^{\mathrm{T}}(t) \end{bmatrix}^{\mathrm{T}} \in \mathbb{R}^m$, $t \geq 0$, so that the prediction error

$$
e_1(t) \triangleq x_{1\mathrm{a}}(t) - \hat{x}_{1\mathrm{a}}(t) = x_1(t) - \hat{x}_1(t), \quad t \geq 0, \quad (12)
$$

is Lyapunov stable and asymptotically converges to a neighborhood of the origin, uniformly in time. Note that, from (8) and (11), we have that $\hat{x}_1(t) = \hat{x}_{1\mathrm{a}}(t) - x_{\mathrm{a}}(t)$, $t \geq 0$, provides a prediction of the measured output $x_1(t)$, $t \geq 0$.

For the statement of the following theorem, we will use the notations

$$W_1(x_1, \hat{x}_2, t) \triangleq \frac{\partial f_1(x_1, \hat{x}_2, t)}{\partial x_1}, \qquad (13)$$

$$W_3(x_1, \hat{x}_2, t) \triangleq \frac{\partial f_1(x_1, \hat{x}_2, t)}{\partial t}, \qquad (14)$$

as well as the following lemma,

Lemma 3.1: [18] Let $A, \Sigma \in \mathbb{R}^{m \times m}$ be Hurwitz, and $Q \triangleq R^T R$, with $R \in \mathbb{R}^{m \times m}$ full rank. If

$$H \triangleq \begin{bmatrix} A & \Sigma\Sigma^T \\ -Q & -A^T \end{bmatrix}, \qquad (15)$$

has no eigenvalues on the imaginary axis, then there exists $P \geq 0$ such that

$$A^T P + PA + Q + P\Sigma\Sigma^T P = 0. \qquad (16)$$

In addition, if (A, R) is observable, $P > 0$.

Theorem 3.1: Consider the system given by (1)–(2), and the partial state predictor (11). Assume that a continuously differentiable signal $z(t)$ is available such that

$$\dot{y}(t) = z(t) - \dot{x}_a(t) + \epsilon(t), \quad t \geq 0, \qquad (17)$$

with $\|\epsilon(t)\| \leq \sqrt{\varepsilon/3}$, $\varepsilon \geq 0$, $t \geq 0$. Then, consider the estimated trajectory $\hat{x}_2(t)$, $t \geq 0$, generated by

$$\begin{aligned}
\dot{\hat{x}}_{2a}(t) = & W_2^{-1}(x_1(t), \hat{x}_2(t), t)\Big[P_2^{-1}\Big(P_1 e_1(t) - \frac{1}{2}Q_3 e_2(t)\Big) \\
& - A_1(A_1 e_1(t) - e_2(t)) - W_3(x_1(t), \hat{x}_2(t), t) + \dot{z}(t) \\
& - u(t) + A_2 e_2(t) - W_1(x_1(t), \hat{x}_2(t), t)\Big(z(t) - \dot{x}_a(t) \\
& + \frac{1}{2}W_1^T(x_1(t), \hat{x}_2(t), t)P_2 e_2(t)\Big)\Big], \quad \hat{x}_{2a}(0) = 0_m, \\
& \qquad\qquad\qquad\qquad\qquad\qquad t \geq 0, \quad (18)
\end{aligned}$$

$$\hat{x}_2(t) = \begin{bmatrix} I_p & 0_{p \times (m-p)} \end{bmatrix} \hat{x}_{2a}(t), \qquad (19)$$

where A_1, A_2 are chosen Hurwitz, the observation error is defined as

$$e_2(t) \triangleq f_1(x_1(t), \hat{x}_2(t), t) + \dot{x}_a(t) + A_1 e_1(t) - z(t), \ t \geq 0. \qquad (20)$$

In addition, $u(t) \triangleq -\omega_0^2 x_a(t) - 2\zeta\omega_0 \dot{x}_a(t) - P_2^{-1}Q_3 e_2(t)/2$, $t \geq 0$, with $\zeta, \omega_0, Q_3 \in \mathbb{R}^{m \times m}$ chosen positive definite, and

$$v(t) \triangleq \begin{bmatrix} 0_p & I_{p \times (m-p)} \end{bmatrix} \hat{x}_{2a}(t), \quad t \geq 0. \qquad (21)$$

The matrices P_1, $P_2 \in \mathbb{R}^{m \times m}$ are obtained from the following Riccati equations,

$$A_1^T P_1 + P_1 A_1 + Q_1 + P_1^2 = 0, \qquad (22)$$

$$A_2^T P_2 + P_2 A_2 + Q_2 + P_2 A_1 A_1^T P_2 = 0, \qquad (23)$$

where $Q_1 \triangleq R_1^T R_1 > 0$, $Q_2 \triangleq R_2^T R_2 > 0$, are chosen such that

$$H_1 \triangleq \begin{bmatrix} A_1 & I_m \\ -Q_1 & -A_1^T \end{bmatrix}, \quad H_2 \triangleq \begin{bmatrix} A_2 & A_1 A_1^T \\ -Q_2 & -A_2^T \end{bmatrix}, \quad (24)$$

have no eigenvalues on the imaginary axis and (A_1, R_1), (A_2, R_2) are both observable. Finally, define $\bar{Q}_2 \triangleq Q_2 + Q_3$,

$$M \triangleq \begin{bmatrix} Q_1^{-1} P_1 & 0_{m \times m} \\ 0_{m \times m} & \bar{Q}_2^{-1} P_2^T \end{bmatrix}, \qquad (25)$$

and let $\lambda_{\max}(M)$ denote the maximum eigenvalue of M. The solution $\hat{x}_{2a}(t)$, $t \geq 0$, to (18) guarantees convergence of the prediction-observation errors (e_1, e_2) to $\mathcal{D} \triangleq \{(e_1, e_2) : e_1^T P_1 e_1 + e_2^T P_2 e_2 \leq \varepsilon\lambda_{\max}(M)\}$.

Proof: From (10), (11) and (12), we obtain

$$\begin{aligned}
\dot{e}_1(t) = & f_1(x_1(t), x_2(t), t) + \dot{x}_a(t) - (f_1(x_1(t), \hat{x}_2, t) \\
& + \dot{x}_a(t)), \quad t \geq 0, \qquad (26)
\end{aligned}$$

which, using (20), can be rewritten

$$\begin{aligned}
\dot{e}_1(t) = & f_1(x_1(t), x_2(t), t) + \dot{x}_a(t) - z(t) + A_1 e_1(t) - e_2(t) \\
= & A_1 e_1(t) - e_2(t) + \epsilon(t), \quad t \geq 0. \qquad (27)
\end{aligned}$$

In addition, using (10), (13), (14), (20), and the definition of $W_2(\cdot)$, we obtain

$$\begin{aligned}
\dot{e}_2(t) = & A_1 \dot{e}_1(t) + W_1(x_1(t), \hat{x}_2(t), t)\dot{x}_1(t) - \dot{z}(t) + u(t) \\
& + W_2(x_1(t), \hat{x}_2(t), t)\dot{\hat{x}}_{2a}(t) + W_3(x_1(t), \hat{x}_2(t), t) \\
& \qquad\qquad\qquad\qquad\qquad\qquad t \geq 0. \qquad (28)
\end{aligned}$$

Substituting (17) and (27) into (28) yields

$$\begin{aligned}
\dot{e}_2(t) = & A_1 \epsilon(t) + A_1(A_1 e_1(t) - e_2(t)) + W_3(x_1(t), \hat{x}_2(t), t) \\
& + W_1(x_1(t), \hat{x}_2(t), t)(z(t) - \dot{x}_a(t) + \epsilon(t)) - \dot{z}(t) \\
& + u(t) + W_2(x_1(t), \hat{x}_2(t), t)\dot{\hat{x}}_{2a}, \quad t \geq 0. \qquad (29)
\end{aligned}$$

Finally, substituting (18) into (29), we obtain

$$\begin{aligned}
\dot{e}_2(t) = & P_2^{-1}(P_1 e_1(t) - Q_3 e_2(t)/2) + A_2 e_2(t) + A_1 \epsilon(t) \\
& - W_1(x_1(t), \hat{x}_2(t), t)(W_1^T(x_1(t), \hat{x}_2(t), t)P_2 e_2(t)/2 \\
& - \epsilon(t)), \quad t \geq 0. \qquad (30)
\end{aligned}$$

Then, consider the following Lyapunov function candidate,

$$V(e_1, e_2) = e_1^T P_1 e_1 + e_2^T P_2 e_2, \qquad (31)$$

where P_1, $P_2 > 0$ are obtained from (22) and (23), respectively. The time derivative of (31) along the trajectories of (27) and (30) is given by

$$\begin{aligned}
\dot{V}(t) = & 2e_2^T(t)P_2 A_1 \epsilon(t) + e_2^T(t)P_2 W_1(x_1(t), \hat{x}_2(t), t)(2\epsilon(t) \\
& - W_1^T(x_1(t), \hat{x}_2(t), t)P_2 e_2(t)) - e_1^T(t)(Q_1 + P_1^2)e_1(t) \\
& - e_2^T(t)(Q_2 + Q_3 + P_2 A_1 A_1^T P_2)e_2(t) + 2e_1^T(t)P_1 \epsilon(t), \\
& \qquad\qquad\qquad\qquad\qquad\qquad t \geq 0, \quad (32)
\end{aligned}$$

where we have used (22)–(23). Next, using the completion of the square rule, we obtain

$$2e_1^T P_1 \epsilon = -(P_1 e_1 - \epsilon)^T(P_1 e_1 - \epsilon) + e_1^T P_1^2 e_1 + \epsilon^T \epsilon, \ (33)$$

$$\begin{aligned}
2e_2^T P_2 A_1 \epsilon = & -(A_1^T P_2 e_2 - \epsilon)^T(A_1^T P_2 e_2 - \epsilon) + e_2^T P_2 A_1 A_1^T P_2 e_2 \\
& + \epsilon^T \epsilon, \qquad (34)
\end{aligned}$$

$$\begin{aligned}
2e_2^T P_2 W_1(x_1, \hat{x}_2, t)\epsilon = & e_2^T P_2 W_1(x_1, \hat{x}_2, t)W_1^T(x_1, \hat{x}_2, t)P_2 e_2 \\
& + \epsilon^T \epsilon - (\epsilon - W_1^T(x_1, \hat{x}_2, t)P_2 e_2)^T(\epsilon \\
& - W_1^T(x_1, \hat{x}_2, t)P_2 e_2). \qquad (35)
\end{aligned}$$

Substituting (33)–(35) into (32), we obtain

$$\dot{V}(t) \leq -e_1^T(t)Q_1 e_1(t) - e_2^T(t)\bar{Q}_2 e_2(t) + \varepsilon, \quad t \geq 0. \ (36)$$

Hence, $\dot{V}(t)$, $t \geq 0$, is strictly negative outside of the set $\{(e_1, e_2) : e_1^T Q_1 e_1 + e_2^T \bar{Q}_2 e_2 \leq \varepsilon\}$, which allows to con-

clude ultimate boundedness of $(e_1(t), e_2(t))$, $t \geq 0$ ([9], [19]). In addition, the ultimate bound can be characterized by $\alpha \triangleq \min(e_1^{\mathrm{T}}P_1e_1 + e_2^{\mathrm{T}}P_2e_2)$, subject to the constraint $e_1^{\mathrm{T}}Q_1e_1 + e_2^{\mathrm{T}}Q_2e_2 = \varepsilon$. This constrained minimization problem is easily solved using Lagrange multipliers, yielding $\alpha = \varepsilon\lambda_{\max}(M)$, which proves convergence of the error trajectories to $\mathcal{D}$ and concludes this proof. ∎

Remark 3.1: Note that, in the ideal case that $\epsilon(t) \equiv 0$, $t \geq 0$, we would obtain that $(e_1(t), e_2(t)) \to 0$ as $t \to \infty$. Hence, from the definition of $e_2(t)$, $t \geq 0$, we would obtain convergence of $\hat{x}_2(t)$ to the implicit function $g(y(t), \dot{y}(t), t)$, $t \geq 0$, defined by (7).

Remark 3.2: In Theorem 3.1, it is assumed that $m \geq n/2$. In the particular case that $m = n/2$, the design procedure is simpler. Indeed, in such a situation, the augmenting signal $x_{\mathrm{a}}(t)$, $t \geq 0$, becomes superfluous. Hence, in such a case, $x_{1\mathrm{a}}(t) = x_1(t)$, $\hat{x}_{1\mathrm{a}}(t) = \hat{x}_1(t)$, and $\hat{x}_{2\mathrm{a}}(t) = \hat{x}_2(t)$, $t \geq 0$.

IV. ILLUSTRATIVE NUMERICAL EXAMPLES

To obtain an estimate $z(t)$ of $\dot{y}(t) + \dot{x}_{\mathrm{a}}(t)$, $t \geq 0$, we will use the Adaptive Integral Variable Structure Derivative Estimator (AIVSDE) described in [14]. The object of the technique is, using measurements of a signal $r(t)$, $t \geq 0$, to obtain an estimate of its time derivative $\dot{r}(t)$, $t \geq 0$. This is achieved by constructing a signal $\gamma(t)$, $t \geq 0$, which converges to $r(t)$, $t \geq 0$. The signal $\gamma(t)$, $t \geq 0$, is obtained from

$$\dot{\gamma}(t) = k_r(t)\mathrm{sign}(\sigma(t)) + k_b\sigma(t) - k_1k_2^{-1}\gamma(t), \quad \gamma(0) = \gamma_0,$$
$$t \geq 0, \quad (37)$$

$$\sigma(t) = k_2(r(t) - \gamma(t)) + k_1\int_0^t(r(\tau) - \gamma(\tau))\mathrm{d}\tau, \quad (38)$$

where k_b, k_1, $k_2 \geq 0$, the adaptive gain $k_r(t)$, $t \geq 0$, is obtained from

$$\dot{k}_r(t) = \begin{cases} \alpha_r & \text{for} \|\sigma(t)\| \geq \mu, \\ 0 & \text{for} \|\sigma(t)\| < \mu, \end{cases} \quad k_r(0) = k_{r0}, \ t \geq 0, (39)$$

where $\alpha_r \geq 0$ and $\mu \geq 0$. The estimate of $\dot{r}(t)$ is given by $\dot{\gamma}(t)$, $t \geq 0$. In particular, setting $r(t) = y(t) + x_{\mathrm{a}}(t)$, and $z(t) = \dot{\gamma}(t)$, $t \geq 0$, we obtain an estimate $z(t)$ of $\dot{y}(t) + \dot{x}_{\mathrm{a}}(t)$, $t \geq 0$.

In the following, we will apply the predictor-observer described in Section III, in conjunction with the above AIVSDE technique, to address the observation problem for two distinct systems. In a first example, we consider a system with state $[x_1^{\mathrm{T}}(t)\, x_2(t)]^{\mathrm{T}} \in \mathbb{R}^3$, $t \geq 0$, where $x_1(t) \in \mathbb{R}^2$, $t \geq 0$, is measured, and we attempt to construct an estimate $\hat{x}_2(t)$ of the unmeasured $x_2(t) \in \mathbb{R}$, $t \geq 0$. In a second example, we apply the technique to the same observation problem, for a system with state $[\eta^{\mathrm{T}}(t)\, \nu^{\mathrm{T}}(t)]^{\mathrm{T}} \in \mathbb{R}^6$, $t \geq 0$, with output $\eta(t) \in \mathbb{R}^3$, $t \geq 0$. The estimate $\hat{\nu}(t)$ is shown to converge to neighborhood of the unmeasured $\nu(t)$, $t \geq 0$.

Example 4.1: Consider the system

$$\dot{x}_{11}(t) = -x_{11}^3(t) + \tanh(x_{12}(t)) + 3\sin(x_2(t)),$$
$$x_{11}(0) = x_{110}, \quad t \geq 0, (40)$$

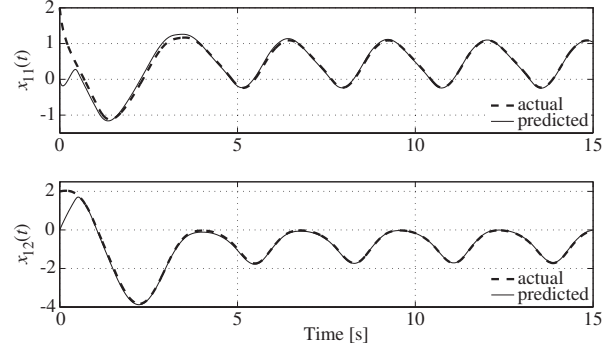
Fig. 1.   Actual and predicted trajectories.

$$\dot{x}_{12}(t) = 4\tanh(x_{11}(t)) - 2\tanh(x_{12}(t)) - 3\cos(x_2(t)),$$
$$x_{12}(0) = x_{120}, \quad (41)$$
$$\dot{x}_2(t) = -x_2^3(t) - 2\tanh((x_{11}(t) + x_{12}(t))^2),$$
$$x_2(0) = x_{20}, \quad (42)$$
$$y(t) = x_1(t), \quad (43)$$

where $x_1(t) \triangleq [\ x_{11}(t) \quad x_{12}(t)\ ]^{\mathrm{T}} \in \mathbb{R}^2$, and $x_2(t) \in \mathbb{R}$, $t \geq 0$. Applying the technique described in Theorem 3.1, we construct the following predictor-observer,

$$\dot{\hat{x}}_{1\mathrm{a}}(t) = \begin{bmatrix} -x_{11}^3(t) + \tanh(x_{12}(t)) + 3\sin(\hat{x}_2(t)) \\ 4\tanh(x_{11}(t)) - 2\tanh(x_{12}(t)) - 3\cos(\hat{x}_2(t)) \end{bmatrix}$$
$$+ \dot{x}_{\mathrm{a}}(t), \quad \hat{x}_{1\mathrm{a}}(0) = [\ x_{110} \quad x_{120}\ ]^{\mathrm{T}}, \quad t \geq 0, \quad (44)$$
$$\dot{\hat{x}}_{2\mathrm{a}}(t) = W_2^{-1}(\hat{x}_2(t))\Big(P_2^{-1}(P_1e_1(t) - Q_3e_2(t)/2) + A_2e_2(t)$$
$$- A_1(A_1e_1(t) - e_2(t)) + \dot{z}(t) - u(t) - W_1(x_1(t))z(t)$$
$$+ W_1(x_1(t))(\dot{x}_{\mathrm{a}}(t) - W_1^{\mathrm{T}}(x_1(t))P_2e_2(t)/2)\Big),$$
$$\hat{x}_{2\mathrm{a}}(0) = \hat{x}_{2\mathrm{a}0}, \quad (45)$$

where $e_1(t) = x_1(t) + x_{\mathrm{a}}(t) - \hat{x}_{1\mathrm{a}}(t)$, $e_2(t) = \dot{\hat{x}}_{1\mathrm{a}}(t) + \dot{x}_{\mathrm{a}}(t) - z(t) + A_1e_1(t)$, $t \geq 0$, and the augmenting signal $x_{\mathrm{a}}(t)$, $t \geq 0$, is obtained from (9), with $u(t) = -2.5/1000\,x_{\mathrm{a}}(t) - \sqrt{2}/20\,\dot{x}_{\mathrm{a}}(t) - 20e_2(t)$, $t \geq 0$. In addition,

$$W_1(x_1) = \begin{bmatrix} -3x_{11}^2 & \mathrm{sech}^2(x_{12}) \\ 4\mathrm{sech}^2(x_{11}) & -2\mathrm{sech}^2(x_{12}) \end{bmatrix}, \quad (46)$$

$$W_2(\hat{x}_2) = 3\begin{bmatrix} \cos(\hat{x}_2) & -\sin(\hat{x}_2) \\ \sin(\hat{x}_2) & \cos(\hat{x}_2) \end{bmatrix}. \quad (47)$$

Choosing $A_1 = -10I_2$, $Q_1 = 10I_2$, $A_2 = -100I_2$, and $Q_2 = 80I_2$, we obtain, from (22)–(23), $P_1 = 0.5132I_2$, $P_2 = 0.5528I_2$. The parameters of the AIVSDE are chosen as follows, $k_1 = k_b = 10$, $k_2 = 20$, $\alpha_r = 0$, $\mu = 1$, $\gamma_0 = x_{10}$ and $k_{r0} = 145$. The initial conditions for the system are chosen as $x_{10} = [\ 2 \quad 2\ ]^{\mathrm{T}}$, and $x_{20} = -1$. Finally, the initial prediction is $\hat{x}_{1\mathrm{a}0} = [\ 0 \quad 0\ ]^{\mathrm{T}}$, while the initial estimate is chosen as $\hat{x}_{20} = [\ -4 \quad 0\ ]^{\mathrm{T}}$.

As seen in Figure 1, the predicted trajectory $\hat{x}_1(t) = \hat{x}_{1\mathrm{a}}(t) - x_{\mathrm{a}}(t)$, $t \geq 0$, smoothly converges to the actual trajectory $x_1(t)$, $t \geq 0$. Similarly, after a short transient, the observed trajectory $\hat{x}_2(t)$, $t \geq 0$, matches the unmeasured state $x_2(t)$, $t \geq 0$, very closely, as observed from Figure 2.
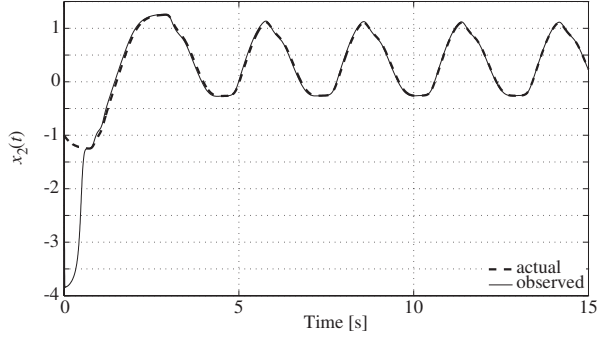
Fig. 2. Actual and observed trajectories.

Fig. 3. Actual and predicted trajectories of $\eta(t)$, $t \geq 0$.

Fig. 4. Actual and observed trajectories of $\nu(t)$, $t \geq 0$.

Example 4.2: Consider the system

$$\dot{\eta}(t) = J(\eta(t))\nu(t), \qquad \eta(0) = \eta_0, \quad t \geq 0, \quad (48)$$
$$\dot{\nu}(t) = \Theta_1^* \varphi(\eta(t)) + \Theta_2^* \tau(t), \quad \nu(0) = \nu_0, \qquad (49)$$
$$y(t) = \eta(t), \qquad (50)$$

where $\eta(t) \triangleq \begin{bmatrix} \phi(t) & \theta(t) & \psi(t) \end{bmatrix}^{\mathrm{T}} \in \mathbb{R}^3$, $t \geq 0$, is the measured output, $\nu(t) \in \mathbb{R}^3$, $t \geq 0$, is not measured, and $\tau(t) \in \mathbb{R}^2$, $t \geq 0$, is the control input applied to the system. In addition, $\Theta_1^* \in \mathbb{R}^{3\times 2}$ and $\Theta_2^* \in \mathbb{R}^{3\times 2}$ are unknown constant matrices, and

$$J(\eta) = \begin{bmatrix} 1 & \tan(\theta)\sin(\phi) & \tan(\theta)\cos(\phi) \\ 0 & \cos(\phi) & -\sin(\phi) \\ 0 & \sin(\phi)/\cos(\theta) & \cos(\phi)/\cos(\theta) \end{bmatrix}, \quad (51)$$
$$\varphi(\eta) = \begin{bmatrix} \cos(\theta)\cos(\phi) & -\cos(\theta)\sin(\phi) \end{bmatrix}^{\mathrm{T}}. \qquad (52)$$

Applying the technique described in Theorem 3.1, we construct the following predictor-observer,

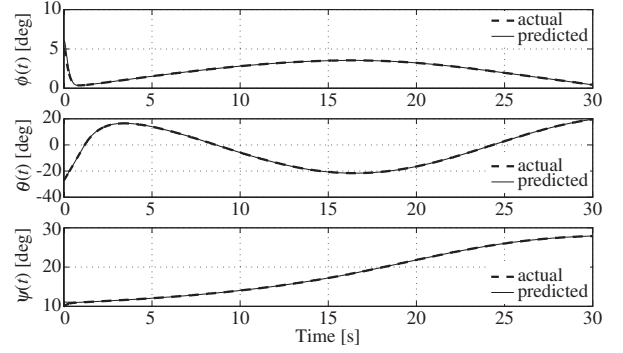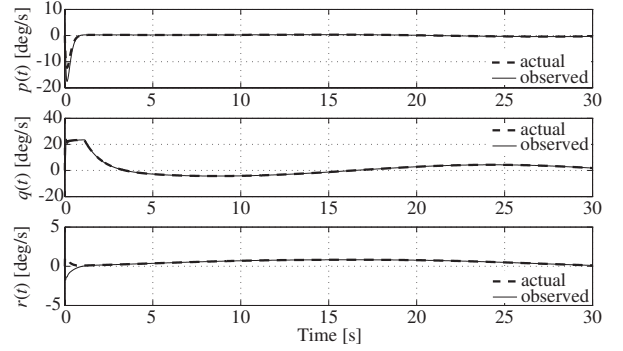$$\dot{\hat{\eta}}(t) = J(\eta(t))\hat{\nu}(t), \qquad \hat{\eta}(0) = \eta_0, \quad t \geq 0, \qquad (53)$$
$$\dot{\hat{\nu}}(t) = J^{-1}(\eta(t))\Big(P_2^{-1} P_1 e_1(t) - A_1(A_1 e_1(t) - e_2(t))$$
$$- W_1(\eta(t), \hat{\nu}(t))(z(t) - W_1^{\mathrm{T}}(\eta(t), \hat{\nu}(t))P_2 e_2(t)/2)$$
$$+ \dot{z}(t) + A_2 e_2(t) \Big), \quad \hat{\nu}(0) = \hat{\nu}_0, \qquad (54)$$

where $e_1(t) \triangleq \eta(t) - \hat{\eta}(t)$, $e_2(t) \triangleq A_1 e_1(t) - z(t) + J(\eta(t))\hat{\nu}(t)$, $t \geq 0$, and $W_1(\eta, \hat{\nu}) \triangleq \frac{\partial J(\eta)}{\partial \eta}\hat{\nu}$. We choose $A_1 = -4I_3$, $A_2 = -40I_3$, $Q_1 = 12I_3$, $Q_2 = 40I_3$, and obtain, from (22) and (23), $P_1 = 6I_3$, $P_2 = (2.5+\sqrt{15}/2)I_3$. The AIVSDE parameters are chosen as follows, $k_1 = 10$, $k_2 = 15$, $k_b = 1/2$, $\alpha_r = 0$, $\mu = 1$, $\gamma_0 = x_{10}$ and $k_{r0} = 1$.

In addition, we use the following for plant parameters,

$$\Theta_1^* = \begin{bmatrix} -2.6828 & 3.2966 \\ 9.8298 & 9.9455 \\ 0 & -20 \end{bmatrix}, \; \Theta_2^* = \begin{bmatrix} 0.25 & 0 \\ 0 & -0.575 \\ 0 & 0 \end{bmatrix}. \; (55)$$

The initial conditions are chosen as $\eta_0 = \begin{bmatrix} 5 & -27 & 10 \end{bmatrix}^{\mathrm{T}}$, $\nu_0 = 0_3$, $\hat{\eta}_0 = \eta_0 + \begin{bmatrix} 1 & 1 & 1 \end{bmatrix}^{\mathrm{T}}$, $\hat{\nu}_0 = \begin{bmatrix} 5 & 5 & 5 \end{bmatrix}^{\mathrm{T}}$, and $z(0) = 0_3$. The trajectories $\eta(t)$ and $\nu(t)$, $t \geq 0$, are shown in Figure 3 and Figure 4, respectively. As was the case with Example 4.1, the algorithm performs admirably. The predicted and observed trajectories closely match the real ones.

V. EXPERIMENTAL TEST

The previous section presented numerical simulation results, illustrating the performance of the algorithm. Performance was further tested by implementing the presented predictor-observer algorithm on the Quanser "3DOF helicopter" setup. Equations (48)–(50) provide a model of the setup's dynamics. Hence, we will apply the algorithm as designed for Example 4.2. However, a significant hurdle in the way of implementation is the fact that the measurements provided by the experimental setup are discrete and quantized, with quantization intervals of $\pi/2048$.

The observation of systems with quantized outputs constitutes a challenging problem (as discussed in [20]), which remains essentially open. The presented predictor-observer technique does not account for quantization of the measurements, and application of the technique as is to such a system yields poor results. This poor performance is mainly due to the difficulty in obtaining a $z(t)$, $t \geq 0$, providing a reasonable approximation of $\dot{y}(t) + \dot{x}_{\mathrm{a}}(t)$, $t \geq 0$. In particular, when the measurements change from a level of quantization to the next, most numerical derivation techniques lead to large spikes in the estimate of the derivative. In an attempt to compensate for this issue, we used a second order low-pass filter to smooth-out the measurements. This filter has a damping coefficient $\zeta = \sqrt{2}/2$ and a natural frequency $\omega_{\mathrm{n}} = 10$. This simple addition to the algorithm significantly improved experimental results, to an extent that performance of the algorithm when subjected to quantized measurements
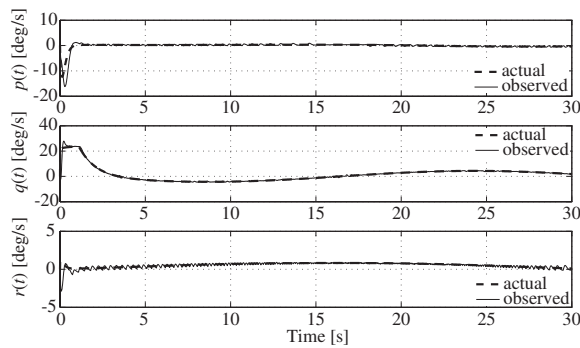
Fig. 5. Numerical simulation with measurement quantization, actual and observed trajectories of $\nu(t)$, $t \geq 0$.
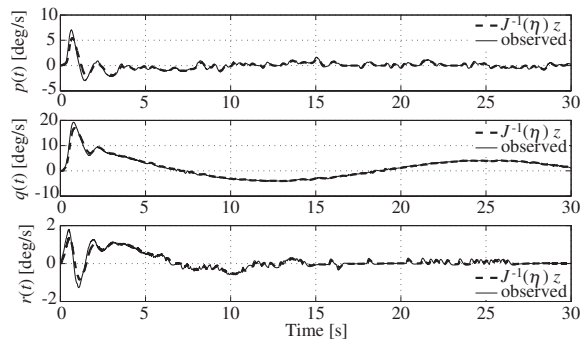

Fig. 6. Experimental results, observed trajectory of $\nu(t)$, $t \geq 0$, compared to a numerical derivative.

became comparable to that without quantization (see Figure 5 for simulation results showing actual and observed $\nu(t)$, $t \geq 0$, with quantized measurements), and allowed implementation of the presented algorithm.

Experimental results are shown in Figure 6. Note that the body fixed velocity vector $\nu(t)$, $t \geq 0$, is not measured. Hence, to assess performance of the presented algorithm, we compare the observed $\hat{\nu}(t)$, $t \geq 0$, to an estimate of $\nu(t)$, $t \geq 0$, obtained from the output of the AIVSDE, $z(t)$, $t \geq 0$. Since $z(\cdot)$ is an estimate of $\dot{\eta}(t)$, $t \geq 0$, it can be pre-multiplied by $J^{-1}(\eta(t))$, $t \geq 0$, to obtain an estimate of $\nu(t)$, $t \geq 0$. As seen in Figure 6, the observed trajectories are in practice significantly noisier than in simulation. The algorithm however yields promising results. In particular, the transient shows that the observer reacts more quickly than the output of the AIVSDE to large changes in velocity.

VI. CONCLUSION

This paper presents a prediction-based observer, relevant to a wide class of nonlinear systems, including systems whose unmeasured states appear nonlinearly in the dynamics. The algorithm relies on a pair of prediction-observation errors, constructed using a backstepping technique. Lyapunov's second method was used to prove Lyapunov stability of these errors, as well as their convergence to a neighborhood of the origin. The algorithm makes use of a derivative estimator, which allows to relax high gain requirements commonly found in observation techniques for nonlinear systems. In addition, the relationship between the ultimate bound on the prediction-observation errors and the derivative estimator's level of performance, as well as a number of relevant design constants, was made explicit. The presented technique could be extended to a larger class of observable systems, by working in terms of the p first derivatives of the output $y(t)$, $t \geq 0$, where $p \leq n - 1$ is the smallest integer such that $\mathrm{rank}(\begin{bmatrix} \nabla y^{\mathrm{T}}(t) & \nabla \dot{y}^{\mathrm{T}}(t) & \ldots & \nabla y^{(p)\mathrm{T}}(t) \end{bmatrix}^{\mathrm{T}}) = n$. Another possible extension of the presented work could consist in using the information provided by knowledge of the entire right-hand-side of (1), as opposed to only that of (5).

REFERENCES

[1] R. E. Kalman, "A new approach to linear filtering and prediction problems," *Trans. of the ASME Journal of Basic Engineering*, vol. 82, no. Series D, pp. 35–45, 1960.

[2] R. E. Kalman and R. Bucy, "New results in linear filtering and prediction theory," *Trans. of the ASME Journal of Basic Engineering*, vol. 83, no. Series D, pp. 95–108, 1961.

[3] D. G. Luenberger, "Observer for multivariable systems," *IEEE Trans. on Automatic Control*, vol. AC-11, no. 2, pp. 190–197, 1966.

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

[5] M. S. Grewal and A. P. Andrews, *Kalman Filtering*. Upper Saddle River, NJ: Prentice-Hall, 1993.

[6] S. J. Julier and J. K. Uhlmann, "A new extension of the kalman filter to nonlinear systems," in *Int. Symp. Aerospace/Defense Sensing, Simul. and Controls*, pp. 182–193, 1997.

[7] F. L. Lewis, L. Xie, and D. Popa, *Optimal and Robust Estimation*. Boca Raton, FL: Taylor and Francis, 2008.

[8] M. Krstić, I. Kanellakopoullos, and P. Kokotović, *Nonlinear and Adaptive Control Design*. New York, NY: John Wiley and Sons, Inc, 1995.

[9] H. K. Khalil, *Nonlinear Systems*. Upper Saddle River, NJ: Prentice Hall, third ed., 2002.

[10] S. Nazrulla and H. K. Khalil, "Robust stabilization of non-minimum phase nonlinear systems using extended high gain observers," in *Proc. 2008 Amer. Contr. Conf.*, (Seattle, WA), pp. 1734–1739, 2008.

[11] D. Karagiannis and A. Astolfi, "Nonlinear observer design using invariant manifolds and applications," in *Proc. 44th IEEE Conf. Dec. Contr.*, (Seville, Spain), pp. 7775–7780, 2005.

[12] A. Astolfi and R. Ortega, "Immersion and invariance: a new tool for stabilization and adaptive control of nonlinear systems," *IEEE Trans. on Automatic Control*, vol. 48, no. 4, pp. 590–606, 2003.

[13] J. X. Xu, Q. W. Jia, and T. H. Lee, "On the design of a nonlinear adaptive variable structure derivative estimator," *IEEE Trans. on Automatic Control*, vol. 45, no. 5, pp. 1028–1033, 2000.

[14] C.-C. Cheng and M.-W. Chang, "Design of derivative estimator using adaptive sliding mode," in *Proc. 2006 IEEE Amer. Contr. Conf.*, (Minneapolis, MN), pp. 2611–2615, 2006.

[15] A. M. Kovalev, "Criteria for the observability of nonlinear dynamic systems in multidiensional measurements," *Prikladnaya Mekhanika*, vol. 13, no. 12, pp. 95–101, 1977.

[16] R. Hermann and A. J. Krener, "Nonlinear controllability and observability," *IEEE Trans. on Automatic Control*, vol. AC-22, no. 5, pp. 728–740, 1977.

[17] S. G. Krantz and H. R. Parks, *The Implicit Function Theorem: History, Theory, and Applications*. Boston, MA: Birkhäuser, 2008.

[18] W. M. Haddad and V. Chellaboina, *Nonlinear Dynamical Systems and Control*. Princeton, NJ: Princeton, 2008.

[19] J. LaSalle and S. Lefschets, *Stability by Liapunov's Direct Method with Applications*. Academic Press, 1961.

[20] D. Liberzon, "Observer-based quantized output feedback control of nonlinear systems," in *Proc. 2007 Mediterranean Conf. on Contr. and Automation*, (Athens, Greece), 2007.