WA04 11:00

Proceedings of the 34th
Conference on Decision & Control
New Orleans, LA - December 1995

Asymptotic Observers for Detectable and Poorly Observable Systems

P.E. Moraal* and J.W. Grizzle †

Abstract

A new notion of detectability for nonlinear systems, which generalizes the existing definition of detectability for linear systems, is explored. Based on this notion, a constructive observer design method for detectable, but not necessarily observable, nonlinear systems is given. Preliminary results on the observer design for "poorly observable" systems are also discussed and illustrated with an example.

1 Introduction

When a given system is not completely observable, it is not possible to recover all the state components from the outputs. However, in some cases one may be able to reconstruct at least a part of the state. In linear system theory, the notion of detectability was introduced to deal with systems that are not completely observable, yet for which one can construct an observer whose error decays to zero exponentially.

The notion of detectability has also been introduced for nonlinear systems [1, 2, 7], for the primary purpose of formulating necessary and sufficient conditions for the stability of observer-based feedback control problems. The current literature leaves open the problem of how constructively to check the detectability conditions or actually design the observer. Another important, and thus far unaddressed problem in observer design concerns systems which we will term "poorly observable", i.e., systems with state components that are only weakly coupled in some sense to the outputs.

In this paper, we will address these problems to some extent. Section 2 will provide the necessary preliminaries. We will introduce a new definition of detectability which is shown to be a generalization of the linear notion of detectability; moreover, the observer design for a special class of nonlinear detectable systems will be demonstrated. In Section 4, we will present an exponential observer design method for a more general class of detectable systems which is related to the Newton observers developed in [6]. Finally, in Section 5, some preliminary results on the observer problem for poorly observable systems are discussed, with an example illustrating the ideas.

*Ford Research Laboratories, MD1170 SRL, P.O.Box 2053, Dearborn, MI 48121-2053; e-mail: pmoraal@ford.com

†Department of EECS, University of Michigan, Ann Arbor, MI 48109-2122; e-mail: grizzle@eecs.umich.edu. Research supported in part by NSF Grant ECS-92-13551

2 Background

Consider the following linear system:

$$\Sigma_1 : \begin{cases} x_{k+1} & = & \begin{pmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{pmatrix} x_k & x_k \in \mathbb{R}^n \\ y_k & = & (0 \;\; C_2) x_k & y_k \in \mathbb{R} \end{cases} \tag{2.1}$$

Clearly, this system is not observable: the first component of the state is not measured, nor does it affect the output in an indirect way. However, if the pair (A_{22}, C_2) is observable and A_{11} is asymptotically stable, then it is possible to construct an observer for Σ_1. If these conditions hold, the above system is said to be *detectable* [3]. In general, a system

$$\Sigma_2 : \; x_{k+1} = Ax_k, \; y_k = Cx_k$$

is said to be *detectable* if the unobservable part is asymptotically stable. Recall that for a linear system the unobservable subspace $\mathcal{N}$ is given by $\mathcal{N} = \ker \text{col}(C, CA, \ldots, CA^{n-1})$, so, for a linear system to be detectable, its dynamics restricted to $\mathcal{N}$ must be asymptotically stable.

The task of designing an observer for a detectable linear system in the special form of (2.1) is straightforward. For a linear system in the general form Σ_2, one approach is to compute a state transformation $z_k = Px_k$ such that Σ_2 takes the form of (2.1) in the new coordinates and then proceed, Here, we will give an alternative approach which is related to the recently developed Newton observers [6].

Notation: The pseudo-inverse of a $p \times n$ matrix A $(p < n)$ with rank $A = p$ will be denoted $A^\dagger$.

Suppose that system Σ_2 is detectable and that rank $\text{col}(C, CA, \ldots, CA^{n-1}) = p < n$. Furthermore, for simplicity, assume that A is invertible. Then

$$H := \begin{pmatrix} CA^{-(p-1)} \\ \cdots \\ CA^{-1} \\ C \end{pmatrix}$$

has rank p and the pseudo-inverse of H is given by $H^\dagger = H^T(HH^T)^{-1}$. Define $Y_k = Hx_k = (y_k, \cdots, y_{k-p+1})_T$. Now, an observer for Σ_2 is given by

$$\begin{aligned} \hat{x}_k &= \hat{x}_k^- + H^\dagger(Y_k - H\hat{x}_k^-) \\ \hat{x}_{k+1}^- &= A\hat{x}_k. \end{aligned} \tag{2.2}$$

To see this, consider the error dynamics for $e_k = x_k - \hat{x}_k$:

$$\begin{aligned} e_{k+1} &= Ax_k - A\hat{x}_k^- - AH^\dagger(Y_k - H\hat{x}_k^-) \\ &= A(I - H^T(HH^T)^{-1}H)e_k. \end{aligned} \quad (2.3)$$

Since the matrix $(I - H^T(HH^T)^{-1}H)$ is the projection onto ker H and, by the detectability assumption, the system's dynamics restricted to ker H is asymptotically stable, it follows that (2.3) is asymptotically stable.

In this paper, we explore how such an approach can be extended to observer design for nonlinear detectable systems. To this end, a new definition of detectability for nonlinear systems is introduced. A previous result on Newton observers for nonlinear discrete-time systems will be used throughout this paper and is repeated here for convenience. Consider the following system:

$$\begin{aligned} x_{k+1} &= F(x_k) \\ y_k &= h(x_k) \end{aligned} \quad (2.4)$$

with $x_k \in I\!R^n$ and $y_k \in I\!R^p$, and denote a vector of N consecutive measurements by $Y_k := col(y_{k-N+1}, \cdots, y_k)$

$$Y_k = \begin{pmatrix} h(x_{k-N+1}) \\ h \circ F(x_{k-N+1}) \\ \vdots \\ h \circ F^{N-1}(x_{k-N+1}) \end{pmatrix} =: H(x_{k-N+1}) \quad (2.5)$$

Let $\mathcal{O}$ be a subset of $I\!R^n$, $N \geq 1$ a given integer and $\epsilon > 0$ a positive constant. Denote the complement of $\mathcal{O}$ by $\sim \mathcal{O}$ and define $dist(x, \sim \mathcal{O}) = \inf\{\|x - y\| : y \in \sim \mathcal{O}\}$, and $\mathcal{O}_\epsilon = \{x \in \mathcal{O} | dist(x, \sim \mathcal{O}) \geq \epsilon\}$. Moreover, define constants β, γ, and L by

$$\begin{aligned} \beta &= \sup\{\left\| \left[\frac{\partial H}{\partial x}(x)\right]^{-1} \right\| : x \in \mathcal{O}\} \\ \gamma &= \sup\{\left\| \frac{\partial^2 H}{\partial x^2}(x) \right\| : x \in \mathcal{O}\} \\ L &= \sup\{\left\| \frac{\partial F}{\partial x}(x) \right\| : x \in \mathcal{O}\}. \end{aligned}$$

Then the following result was proven [5, 6]:

Theorem 2.1 *Suppose that F and h in (2.4) are at least twice differentiable with respect to x; that there exists a bounded convex subset $\mathcal{O} \subset I\!R^n$ which is forward invariant for the system (2.4); that is, $F(\mathcal{O}) \subset \mathcal{O}$. Furthermore, suppose that there exists an integer $1 \leq N \leq n$ such that the state-to-measurement map H is square and injective on $\mathcal{O}$ (observability) and has rank n at every point of $\mathcal{O}$ (Observability Rank Condition)* [1].

[1] See [5] for the case that the resulting map H is not square, but does satisfy the two observability conditions.

Then, for every $\epsilon > 0$ and $0 < \mu < 1$, there exists a $\delta > 0$ satisfying

$$\delta \leq \min\left\{\frac{\epsilon}{L}, \frac{\mu}{\beta\gamma L^2}\right\} \quad (2.6)$$

such that

$$\begin{aligned} \hat{x}_k &= \hat{x}_k^- + \left[\frac{\partial H}{\partial x}(\hat{x}_k^-)\right]^{-1}(Y_{k+N-1} - H(\hat{x}_k^-)) \\ \hat{x}_{k+1}^- &= F(\hat{x}_k), \end{aligned}$$

is a quasi-local, exponential observer for (2.4) in the sense that: (A) if $x_1 \in \mathcal{O}$ and $\hat{x}_1^- = x_1$, then $\hat{x}_k = x_k$ for all $k \geq 1$ and (B), if $x_1 \in \mathcal{O}$, $\|\hat{x}_1^- - x_1\| < \delta$ and for all $k \geq 0$, $dist(x_k, \sim \mathcal{O}) \geq \epsilon$, then $\|\hat{x}_{k+1} - x_{k+1}\| \leq \mu\|\hat{x}_k - x_k\|$.

3 Detectability of Nonlinear Systems

We wish to take a constructive approach to the observer design for detectable systems, we will introduce an alternative definition of detectability for nonlinear discrete-time systems, first for systems in a special form: Consider the following system

$$\begin{aligned} x_{k+1}^1 &= F_1(x_k^1, x_k^2) \\ x_{k+1}^2 &= F_2(x_k^2) \\ y_k &= h(x_k^2). \end{aligned} \quad (3.7)$$

Definition 3.1 *System (3.7) is said to be exponentially detectable when the following conditions are satisfied:*
(i) The subsystem $x_{k+1}^2 = F_2(x_k^2)$, $y_k = h(x_k^2)$ is observable and satisfies the Observability Rank Condition.
(ii) Whenever the series $\{\hat{x}_k^2\}$ is such that $\|\hat{x}_k^2 - x_k^2\| \to 0$ exponentially, then $\{\hat{x}_k^1\}$ defined by $\hat{x}_{k+1}^1 = F_1(\hat{x}_k^1, \hat{x}_k^2)$ satisfies $\|\hat{x}_k^1 - x_k^1\| \to 0$ exponentially.

A somewhat stronger notion of detectability, which is sometimes easier to check is the following:

Definition 3.2 *System (3.7) is said to be strongly detectable when the following conditions are satisfied:*
(i) The subsystem $x_{k+1}^2 = F_2(x_k^2)$, $y_k = h(x_k^2)$ is observable and satisfies the Observability Rank Condition.
(ii) With an appropriate choice of norm $\|\cdot\|$, F_1 satisfies the Lipschitz condition

$$\|F_1(\xi_1, z) - F_1(\xi_2, z)\| \leq L_1\|\xi_1 - \xi_2\|$$

with $L_1 < 1$ for all (ξ_1, z) and (ξ_2, z).

Of course, a given system may not satisfy the strong detectability condition globally. The following definition will then be useful:

Definition 3.3 *System (3.7) is strongly $\mathcal{O}$-detectable when, for a given subset $\mathcal{O} \subset I\!R^n$, the conditions of Definition 3.2 are satisfied for all $x \in \mathcal{O}$.*

The two notions of detectability thus defined are related as shown in the following lemma:

Lemma 3.4 *For systems in the form (3.7) with F_1 globally Lipschitz in its second component with Lipschitz constant $L_2 < \infty$:*

Strong detectability $\Longrightarrow$ Detectability.

The following result establishes that this notion of detectability is a direct extension of the corresponding property for linear systems. The proof is contained in [5].

Lemma 3.5 *For linear systems, Definitions 3.1 and 3.2 are both equivalent to the usual notion of detectability.*

Definition 3.6 *If, for a nonlinear system in general form*

$$\begin{aligned} x_{k+1} &= F(x_k) \\ y_k &= h(x_k) \end{aligned} \tag{3.8}$$

there exists a coordinate transformation $z_k = \Phi(x_k)$ such that, in the new coordinates, (3.8) takes the form of (3.7) and this transformed system satisfies conditions (i) and (ii) of Definition 3.2, then (3.8) is said to be strongly detectable.

In fact, let

$$H_p(x) = \begin{pmatrix} h(x) \\ \vdots \\ h \circ F^{p-1}(x) \end{pmatrix},$$

and suppose that p is such that H_p has rank p and

$$h \circ F^p(x) = \phi_p(h(x), \ldots, h \circ F^{p-1}(x))$$

for some function ϕ_p. Moreover, suppose that there exist functions $\phi_{p+1}(x), \ldots \phi_n(x)$ such that

$$\Phi(x) := (H_p(x), \phi_{p+1}(x), \ldots, \phi_n(x)) =: (H_p(x), \Phi_2(x))$$

is a valid coordinate transformation. Then, in the new coordinates z, system (3.8) takes the form of (3.7) and is given by:

$$\begin{aligned} z^1_{k+1} &= z^2_k \\ &\vdots \\ z^{p-1}_{k+1} &= z^p_k \\ z^p_{k+1} &= \phi_p(z^1_k, \ldots, z^p_k) \\ z^{p+1}_{k+1} &= \phi_{p+1} \circ F \circ \Phi^{-1}(z_k) \\ &\vdots \\ z^n_{k+1} &= \phi_n \circ F \circ \Phi^{-1}(z_k) \\ y_k &= z^1_k \end{aligned} \tag{3.9}$$

A sufficient condition for strong detectability is then given by

$$\sup \| \frac{\partial \Phi_2}{\partial x} \| \cdot \sup \| \frac{\partial F}{\partial x} \| \cdot \sup \| \frac{\partial \Phi}{\partial z_2}^{-1} \| < 1.$$

Example: As a simple example, consider the following system:

$$\begin{aligned} z^1_{k+1} &= (z^1_k)^2 - (z^2_k)^2 + \tfrac{1}{2}(z^1_k - z^2_k) \\ z^2_{k+1} &= (z^1_k)^2 - (z^2_k)^2 - \tfrac{1}{2}(z^1_k - z^2_k) \\ y_k &= z^1_k - z^2_k. \end{aligned} \tag{3.10}$$

Using the transformation $x^1_k = z^1_k - z^2_k$ and $x^2_k = z^1_k$, one obtains the system

$$\begin{aligned} x^1_{k+1} &= x^1_k \\ x^2_{k+1} &= \tfrac{1}{2}x^1_k - (x^1_k)^2 + 2x^1_k x^2_k \end{aligned} \tag{3.11}$$

which, by Definition 3.2, is strongly detectable when $|x^1| = |z^1 - z^2| = |y| \le \frac{1}{2} - \epsilon$ for all $k \ge 0$ and some $\epsilon > 0$. Finally, the following result can be proven:

Lemma 3.7 *If system (3.8) with $F(0) = h(0) = 0$ is strongly detectable in an open set around the origin, then its linearization about the origin is detectable in the usual sense.*

3.1 Observer design for detectable systems

It will now be shown how to construct an observer for the unobservable but detectable system in canonical form (3.7). First, define

$$Y_k := \begin{pmatrix} h(x^2_{k-N+1}) \\ \vdots \\ h \circ F_2^{N-1}(x^2_{k-N+1}) \end{pmatrix} =: H(x^2_{k-N+1}). \tag{3.12}$$

and, for a given compact subset $\mathcal{O} \subset \mathbb{R}^n$

$$L_3 := \sup_{x \in \mathcal{O}} \left\| \frac{\partial F_2}{\partial x}(x) \right\|$$

$$\beta := \sup_{x \in \mathcal{O}} \left\| \left(\frac{\partial H}{\partial x}(x) \right)^{-1} \right\|$$

$$\gamma := \sup_{x \in \mathcal{O}} \left\| \frac{\partial^2 H}{\partial x^2}(x) \right\|$$

Theorem 3.8
(Exponential observer for strongly detectable systems I) *Consider the nonlinear system (3.7) and its associated extended output map H, describing N consecutive outputs as a function of the system's state as defined in (3.12). Suppose there exists a compact, convex subset $\mathcal{O} \subset \mathbb{R}^n$ such that the following hold:*

1. *The system (3.7) is strongly $\mathcal{O}$-detectable in the sense of Definition 3.3 with Lipschitz constant $L_1 < 1$.*
2. *F_1 is Lipschitz continuous in $\mathcal{O}$ with respect to its second component with constant L_2*
3. *$\mathcal{O}$ is F-invariant, i.e., $F(\mathcal{O}) \subset \mathcal{O}$.*

Then, for every $\epsilon > 0$ and $\delta > 0$ satisfying

$$\delta \le \min \left\{ \frac{\epsilon}{L_3}, \frac{L_1}{2\beta\gamma L_3^2}, \frac{\epsilon L_1}{8 L_2 L_3^N}, \frac{\epsilon}{2 L_3^N} \right\} \tag{3.13}$$

the system

$$\begin{aligned} z_{k+1} &= F_2(z_k) + \left[\frac{\partial H}{\partial z}(F_2(z_k))\right]^{-1}(Y_k - H(F_2(z_k))) \\ \hat{x}_k^2 &= F_2^{N-1}(z_k) \\ \hat{x}_{k+1}^1 &= F_1(\hat{x}_k^1, \hat{x}_k^2) \end{aligned} \tag{3.14}$$

is a quasi-local exponential observer for (3.7) in the sense that:
(A) if $(x_1^1, x_1^2) \in \mathcal{O}$ *and* $z_{N+1} = x_1^2$ *and* $\hat{x}_1^1 = x_1^1$, *then* $\hat{x}_k = x_k$ *for all* $k \geq N+1$
(B), if $(x_1^1, x_1^2) \in \mathcal{O}$, $\|z_{N+1} - x_1^2\| < \delta$, *and* $\|x_N^1 - \hat{x}_N^1\| < \frac{\epsilon}{4}$, *and for all* $k \geq 0$, $dist(x_k, \sim \mathcal{O}) \geq \epsilon$, *then* $\|\widehat{x}_{k+N} - x_{k+N}\| \leq L_1^k \epsilon$ *for all* $k \geq 0$.

The idea of the proof is to apply Theorem 2.1 to the subsystem involving x_k^2 in order to obtain a sufficiently fast exponential observer for x_k^2. Along with the detectability assumption (i.e., F_1 is a contraction mapping with respect to its first variable), this will ensure exponential convergence of $\|x_k^1 - \hat{x}_k^1\|$.

4 Asymptotic observers for strongly detectable systems

In this section, it will be shown how to construct a quasi-local exponential observer for a more general nonlinear system in the original coordinates *without* explicitly performing a state coordinate transformation. We will need the following result:

Theorem 4.1 Newton-like method for underdetermined systems *Let* $R : \mathbb{R}^n \to \mathbb{R}^p$ *with* $p < n$ *be twice continuously differentiable in an open convex set* $\mathcal{O} \subset \mathbb{R}^n$ *and let* $J(x) := \frac{\partial R}{\partial x}(x)$. *Assume that there exist* $r_1 > 0$ *and* $x_\star \in \mathcal{O}$ *such that* $R(x_\star) = 0$ *and* $\|\frac{\partial J}{\partial x}(x)\| \leq \gamma$, $\|J(x)\| \leq \alpha$ *and* $\|J^\dagger(x)\| \leq \beta$ *for all* $x \in B(x_\star, r_1) \subset \mathcal{O}$. *Then there exists* $0 < r_2 < r_1$ *such that* $\forall x_0 \in B(x_\star, r_2)$, *the sequence generated by*

$$x_{k+1} = x_k - J^\dagger(x_k)R(x_k) \tag{4.15}$$

is well-defined, satisfies

$$\|R(x_{k+1})\| \leq \frac{1}{2}\gamma\beta^2\|R(x_k)\|^2 \tag{4.16}$$

and $R(x_k)$ *converges to zero quadratically.*

Note: the theorem does not state that the sequence $\{x_k\}$ converges to $x_\star$, however, it does guarantee that $\{x_k\}$ is such that the sequence $\{R(x_k)\}$ converges to zero quadratically. The proof may be found in [5].
Now consider a system in general form

$$\begin{aligned} x_{k+1} &= F(x_k) \\ y_k &= h(x_k) \ , \end{aligned} \tag{4.17}$$

which is detectable in the sense that there exists a coordinate transform $z = \Phi(x)$ such that, in the new coordinates, the system is in the form of (3.7) and satisfies Definition 3.2. Furthermore, define

$$\begin{aligned} H(x) &= \begin{pmatrix} h(x) \\ h \circ F(x) \\ \vdots \\ h \circ F^{(N-1)}(x) \end{pmatrix} \\ A(z) &= \Phi \circ F \circ \Phi^{-1}(z) \\ c(z) &= h \circ \Phi^{-1}(z) \\ C(z) &= \begin{pmatrix} c(z) \\ c \circ A(z) \\ \vdots \\ c \circ A^{(N-1)}(z) \end{pmatrix} = H \circ \Phi^{-1}(z) \end{aligned}$$

where N is such that $C(z)$, which in the z-coordinates is a function of z_2 only, is square, injective and has full rank. The following theorem can be stated:

Theorem 4.2 (Exponential observer for strongly detectable systems II) *Suppose that the system (4.17) is strongly detectable in the sense that there exists a coordinate transform* $z = \Phi(x)$ *such that, in the new coordinates, the system is in the form of (3.7) and satisfies Definition 3.2, then the system*

$$\begin{aligned} \hat{x}_k &= \hat{x}_k^- + \left(\frac{\partial H}{\partial x}(\hat{x}_k^-)\right)^\dagger (Y_{k+N-1} - H(\hat{x}_k^-)) \\ \hat{x}_{k+1}^- &= F(\hat{x}_k) \end{aligned} \tag{4.18}$$

is a quasi-local exponential observer for (4.17); i.e., the observer error $\|\hat{x}_k - x_k\| \to 0$ *exponentially, provided that* $\|\hat{x}_0 - x_0\|$ *is sufficiently small.*

Note that for this theorem to apply, one doesn't need to know the coordinate transformation explicitly, as long as it is known that there does exist one.

A sketch of the proof is outlined below; again, the complete proof can be found in [5]. Notation: let $\hat{z}_k := \Phi(\hat{x}_k)$ and $\hat{z}_k^- := \Phi(\hat{x}_k^-)$. The structure of the proof is as follows: (1) Using Theorem 4.1 it is established that

$$\|H(\hat{x}_{k+1}^-) - H(x_{k+1})\| \leq \frac{\mu}{2}\|H(\hat{x}_k^-) - H(x_k)\|$$

for any $0 < \mu < 1$ provided that $\|x_0 - \hat{x}_0\|$ is sufficiently small.
(2) Then, it is shown that this implies $\|\hat{z}_k^2 - z_k^2\| \leq Q_1 \left(\frac{\mu}{2}\right)^k \|\hat{z}_0^2 - z_0^2\|$ for some finite Q_1, and hence,
(3) using the detectability assumption, the special structure of the system equations in the z-coordinates, and Theorem 3.8: $\|\hat{z}_k - z_k\| \leq Q_2\mu^k\|\hat{z}_0 - z_0\|$ for some finite Q_2.
(4) Finally, this then implies that $\|\hat{x}_k - x_k\| \leq Q_3\mu^k$ for some finite Q_3, and $\mu < 1$, which proves the theorem.

5 Poorly observable systems

In practice, the distinction between observable and unobservable systems is often not so clear. A system may be found to be technically observable; however, if, for example, one of the states is only very weakly coupled to the outputs, it might well be found to be unobservable for all practical purposes. This section explores some of the issues and possible ways of dealing with such "poorly observable" systems.

Consider the following system and assume for simplicity $x_k^1, x_k^2, y_k \in I\!R$:

$$\begin{array}{rcl} x_{k+1}^1 &=& F_1(x_k^1, x_k^2) \\ x_{k+1}^2 &=& F_2(x_k^2) + \epsilon x_k^1 \\ y_k &=& h(x_k^2) \end{array} \tag{5.19}$$

The state-to-measurement map is given by

$$Y_{k+1} = H(x_k) = \begin{pmatrix} h(x_k^2) \\ h(F_2(x_k^2) + \epsilon x_k^1) \end{pmatrix} \tag{5.20}$$

The Jacobian matrix $\frac{\partial H}{\partial x}(x_k)$ of this map is given by

$$\begin{pmatrix} 0 & \frac{\partial h}{\partial x}(x_k^2) \\ \epsilon \frac{\partial h}{\partial x}(F_2(x_k^2) + \epsilon x_k^1) & \frac{\partial h}{\partial x}(F_2(x_k^2) + \epsilon x_k^1) \frac{\partial F_2}{\partial x}(x_k^2) \end{pmatrix}$$

For this system, consider the following observer:

$$\begin{array}{rcl} \hat{x}_k &=& \hat{x}_k^- + G(\hat{x}_k^-)(Y_{k+1} - H(\hat{x}_k^-)) \\ \hat{x}_{k+1}^- &=& F(\hat{x}_k) \end{array} \tag{5.21}$$

When $\epsilon \neq 0$, one could, in principle, choose $G(\hat{x}_k) = \frac{\partial H}{\partial x}(\hat{x}_k^-)^{-1}$ to obtain the standard Newton observer. On the other hand, when $\epsilon = 0$, one would reconstruct only x_k^2 from Y_{k+1}, which is accomplished by choosing $G(x) = \begin{pmatrix} 0 \\ G_2(x) \end{pmatrix}$ where $G_2(x) = \left(\frac{\partial H}{\partial x_2}(x)^T \frac{\partial H}{\partial x_2}(x)\right)^{-1} \frac{\partial H}{\partial x_2}(x)^T$. With this choice of G, one obtains a Newton observer with overdetermined state-to-measurement map (because two measurements are used to reconstruct a single state variable).

The approach just outlined is, of course, not acceptable in practice. Clearly, for ϵ very small, $\frac{\partial H}{\partial x}$ is almost singular. The x_k^1 component of the state is almost unobservable and the problem of determining x_k^1 from the outputs y_k becomes numerically ill-conditioned. Rather than attempting to solve this ill-conditioned problem (recovering x_k^1 from the measurements), one would want to reconstruct x_k^2 only and estimate x_k^1 using the system's dynamics, provided that the system is detectable in some sense.

In general, if a system is poorly observable (e.g. large condition number for $\frac{\partial H}{\partial x}$), it will not be so straightforward to separate the observable from the poorly observable states, unless the system happens to be in a special form like (5.19). To get an idea as to how one could deal with the more general case, it is useful to recall the idea that lies at the heart of the Newton observer design method developed in [5, 6]. Namely, at a given time k, one wants to find an approximate solution x_+ of $Y_k - H(x) = 0$ given an initial guess x_c. Newtons's method proposes to do so by solving x_+ from the following system of equations, which is linear once x_c is given:

$$Y_k - H(x_c) = -\frac{\partial H}{\partial x}(x_c)(x_+ - x_c) \tag{5.22}$$

For convenience, denote the Jacobian by $J := \frac{\partial H}{\partial x}(x_c)$. When J is known to be potentially ill-conditioned, a numerically robust way of computing its rank is to use a singular value decomposition (SVD) [4]. Let an SVD of J be given by $J = USV^T$, that is, U and V are unitary matrices and $S = \text{diag}(\sigma_1, \ldots, \sigma_n)$ with $\sigma_1 \geq \cdots \geq \sigma_n \geq 0$. When J is (nearly) rank deficient, one or more singular values will be (almost) zero. Theoretically, the rank of J is equal to the number of nonzero singular values σ_i of J. The "numerical" rank of J is determined similarly, *after* replacing all singular values that are less than a certain threshold δ by zero. In the context of observability, the "numerical" rank of J could, loosely speaking, be interpreted as the dimension of the "robustly observable" part of the system's state.

Suppose that $S = \text{diag}(\sigma_1, \ldots, \sigma_n)$ is such that $\sigma_{p+1}, \ldots, \sigma_n < \delta$ for a given threshold δ. A pseudo-inverse J_δ^+ of J using SVD's is then defined as [4]

$$J_\delta^+ := VS^+U^T \tag{5.23}$$

where $S_\delta^+ := \text{diag}(1/\sigma_1, \ldots, 1/\sigma_p, 0, \ldots, 0)$ The pseudo-inverse is subscripted with a δ to indicate the dependence on the choice of threshold. It can be easily verified that if a matrix P of size $p \times n$ has full rank and usual pseudo-inverse $P^\dagger = P^T(PP^T)^{-1}$, then, for sufficiently small threshold δ, the above defined pseudo-inverse for $\tilde{P} = \begin{pmatrix} P \\ 0 \end{pmatrix}$, an $n \times n$ matrix, satisfies $\tilde{P}_\delta^+ = (P^\dagger \quad 0)$. Moreover, when a square matrix A is invertible, $A_\delta^+ = A^{-1}$, again, provided that δ is chosen sufficiently small. Using this special pseudo-inverse, we can now construct an observer for a class of poorly observable linear systems.

Consider the following family of linear discrete-time systems parametrized by ϵ and assume for simplicity that the output is scalar valued:

$$\begin{array}{rcl} x_{k+1} &=& A(\epsilon)x_k \\ y_k &=& C(\epsilon)x_k \end{array} \tag{5.24}$$

Define $H(\epsilon) := \begin{pmatrix} C(\epsilon) \\ \vdots \\ C(\epsilon)A(\epsilon)^{n-1} \end{pmatrix}$ and $Y_k = H(\epsilon)x_{k-n+1}$, and assume that $A(\epsilon)$ and $C(\epsilon)$ are differentiable in ϵ.

Theorem 5.1 *Suppose that system (5.24) is detectable for $\epsilon = 0$ and rank $H(0) = p < n$. Then $\exists \epsilon^\star > 0$ and $\delta(\epsilon^\star) > 0$ such that for any fixed ϵ satisfying $0 < |\epsilon| < \epsilon^\star$, the following is an exponential observer for (5.24):*

$$\begin{aligned} \hat{x}_k &= \hat{x}_k^- + H_\delta^+(\epsilon)(Y_{k+n-1} - H(\epsilon)\hat{x}_k^-) \\ \hat{x}_{k+1}^- &= A(\epsilon)\hat{x}_k \end{aligned} \tag{5.25}$$

For a proof, see [5].

As of yet, we have not been able to generalize the above result to nonlinear systems, however, given this result and recalling the statement of Theorem 4.2 (Observer for strongly detectable systems in general form), the following Conjecture seems plausible:

Conjecture 5.2 *Consider a system Σ_ϵ : $x_{k+1} = F_\epsilon(x_k)$, $Y_k = H_\epsilon(x_{k-N+1})$, where F_ϵ and H_ϵ are such that, for $\epsilon = 0$, the system is strongly detectable. Then $\exists \epsilon^\star > 0$ and $\delta(\epsilon^\star) > 0$ such that $\forall 0 < |\epsilon| < \epsilon^\star$, a quasi-local exponential observer for Σ_ϵ is given by*

$$\begin{aligned} \hat{x}_k &= \hat{x}_k^- + \left(\frac{\partial H_\epsilon}{\partial x}(\hat{x}_k^-)\right)_\delta^+ (Y_{k+N-1} - H_\epsilon(\hat{x}_k^-)) \\ \hat{x}_{k+1}^- &= F_\epsilon(\hat{x}_k) \end{aligned} \tag{5.26}$$

where $\left(\frac{\partial H}{\partial x}\right)_\delta^+$ is as defined in (5.23).

Remark 5.3 Note that if the system is poorly observable only in a certain part of the state space, but "robustly" observable otherwise, the above observer structure naturally exploits the available information to the full extent possible: In the poorly observable region, the system is implicitly interpreted as a detectable system and only a part of the state estimate is updated via the measurements, whereas otherwise the observer is exactly the standard Newton observer developed in [5, 6].

Remark 5.4 Consider again system Σ_ϵ given in Conjecture 5.2 and assume that $F_\epsilon = 0$ and $h(0) = 0$. Using Lemma 3.7, it follows that the linearization of Σ_ϵ about the origin is detectable. Theorem 5.1 can then directly be applied to prove the above conjecture for the special case of this linearized system.

As an illustration in further support of the above Conjecture, consider the following example:

Example: Consider the following system:

$$\begin{aligned} x_{k+1}^1 &= 1.08x_k^1 - 0.03x_k^1x_k^2 + 0.04x_k^3 \arctan x_k^2 \\ x_{k+1}^2 &= 0.7x_k^2 + 0.05x_k^1x_k^2 \\ x_{k+1}^3 &= 0.999x_k^3 \\ y_k &= x_k^1 \end{aligned} \tag{5.27}$$

This system has the structure of a predator-prey model, with x_k^1 and x_k^2 representing the prey and predator populations at time k, respectively. The additional state variable x_k^3 can be interpreted as a very slowly varying component of the growth rate for the prey species, whose initial value is unknown. Furthermore, at every sampling instant, a reliable count can be made of the prey species only, i.e., the output is $y_k = x_k^1$. A first look at the model structure reveals that the observability of the third state variable is dependent on the size of x_k^2: if x_k^2 is close to zero, the third state is practically unobservable; if $x_k^2 = 0$, then the above system is actually unobservable, however, it *is* detectable. Figure 1 illustrates the

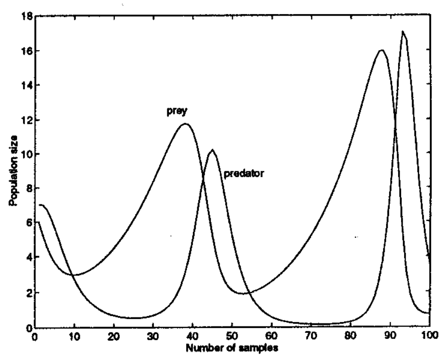

Figure 1: *Characteristic evolution of predator and prey species.*

evolution of the two species over time: Initially, the predator species diminishes the prey species, until the latter is so low that, due to higher mortatility rate of the predator population (reflected in the 0.7 growth rate), the predators become almost extinct. This allows the prey population to grow again, which in turn feeds the predator species after some time. During the period that the predator population is very low, the state x_k^3 is practically unobservable, whereas the system exhibits better observability properties when x_k^2 is high. The observer structure proposed in Conjecture 5.2 is applied to this system, with a threshold $\delta = 0.0005$ for the computation of the pseudo inverse. Moreover, as in [5, 6], the observer update equation (5.26) is slightly modified by including a gain $\lambda = 0.1$ (equivalent to modifying the step size in the conventional Newton algorithm):

$$\hat{x}_k = \hat{x}_k^- + \lambda \left(\frac{\partial H_\epsilon}{\partial x}(\hat{x}_k^-)\right)_\delta^+ (Y_{k+N-1} - H_\epsilon(\hat{x}_k^-)). \tag{5.28}$$

The result are shown in Figures 2 and 3. Figure 2

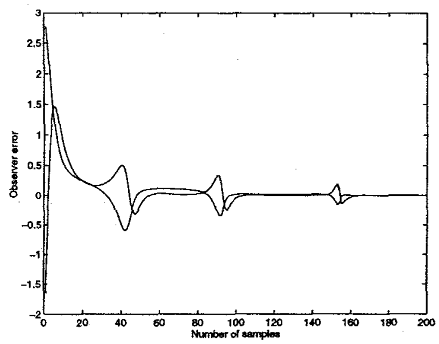

Figure 2: *Observer error for* x_k^1 *and* x_k^2.

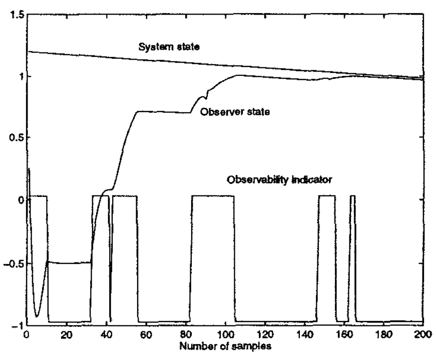

Figure 3: *The poorly observable state:* x_k^3 *and its estimate* $\hat{x}_k^3$. *The observability indicator shows when the system is observable (indicator is* 0*) and when it is practically unobservable, but detectable (indicator* -1*).*

shows the observer error for x_k^1 and x_k^2. Figure 3 illustrates that the observer does exactly what it was intended to do. In order to show what happens, we have included a signal (labeled "Observability indicator") which is 0 when the system is judged to be observable and -1 otherwise. It is immediately apparent that the observer is able to reduce the observer error during periods of observability (except for some "transient" behavior during the first few samples), whereas it only propagates the estimated $\hat{x}_k^3$ through the system's dynamics when x_k^3 is judged to be unobservable.

6 Conclusions

In this paper we have explored a new definition of detectability for nonlinear discrete-time systems, which is more amenable to the construction of observers for nonlinear detectable systems. Furthermore, we have proven the exponential convergence of an observer for such systems, which can be constructed without performing state coordinate transformations. This observer was based on the Newton observers described in [5, 6]. Finally, we presented some preliminary results on the observer design problem for systems that are "poorly observable", meaning that the state-to-measurement map is nearly rank deficient. In the case of linear systems, it was shown that a Newton observer with an appropriately defined pseudo inverse yields an exponential observer. For nonlinear systems, we suspect that the same is true, although a conclusive proof has, sofar, eluded us. In one particular example, namely a predator-prey model, our conjecture appeared to hold.

The most severe drawback to this latter type of observer is that, at every iteration of the observer, it has to be checked whether the system is observable or merely detectable. Since this amounts to a rank test, a singular value decomposition is required at each iteration. Certainly when the sampling periods are short, this places a heavy burden on the computational unit. On the other hand, to the best of our knowledge, the proposed approach is the only one known to date that can deal, in a systematic way, with such poorly observable systems. Moreover, it may only be a matter of time until the computation of SVD's is considered to be a computationally low cost operation.

References

[1] C.I. Byrnes and A. Isidori, "Steady state response, separation principle and the output regulation of nonlinear systems", *Proceedings of the 28th Conference on Decision and Control*, Tampa, 1989, pp. 2247–2251.

[2] C.I. Byrnes and W. Lin, "On nonlinear discrete-time control," *Proc. 32nd IEEE CDC* (1993), pp.2992–2996.

[3] C.T. Chen, *"Linear System Theory and Design,"* 1984.

[4] G. H. Golub and C. F. VanLoan, *Matrix computations*, The Johns Hopkins University Press, 1990.

[5] P.E. Moraal, *Nonlinear observer design: theory and applications to automotive control,* Doctoral Dissertation, University of Michigan, June 1994.

[6] P.E. Moraal and J.W. Grizzle, "Observer design for nonlinear systems with discrete-time measurements", *IEEE Transactions on Automatic Control*, March 1995.

[7] M. Vidyasagar (1981), "On the stabilization of nonlinear systems using state detection", *IEEE Trans. Autom. Control*, Vol. 25, No.3, pp. 504-509.