Proceedings of the
47th IEEE Conference on Decision and Control
Cancun, Mexico, Dec. 9-11, 2008

Observer Design for Discrete-Time Nonlinear Systems

Wei Lin, Jinfeng Wei and Feng Wan

Abstract— **Necessary and sufficient conditions are presented under which a discrete-time autonomous system with outputs is locally diffeomorphic to an output-scaled linear observable system or an output-scaled nonlinear system in the observer form. As a consequence of such characterizations, the nonlinear observer design problem is studied by a time-scaling approach combined with the exact linearization technique, for a broader class of discrete-time nonlinear systems.**

Keywords: Discrete-time systems, linearization, local diffeomorphism, nonlinear observers, time-scaling method

I. Motivation and Discussion

In this paper, we investigate the observer design problem for the discrete-time nonlinear system

$$\sum : \begin{array}{rl} x(k+1) & = f(x(k)) \\ y(k) & = h(x(k)) \end{array} \tag{1.1}$$

where $x \in \mathbb{R}^n$, $f : \mathbb{R}^n \to \mathbb{R}^n$ and $h : \mathbb{R}^n \to \mathbb{R}^m$ are smooth functions, with $f(0) = 0$ and $h(0) = 0$.

The problem has been studied in the literature from various viewpoints. One of the effective and elegant ways has proved to be the use of the differential geometry that provides a powerful tool for the analysis and design of nonlinear observers. In the continuous-time case, the work [4] gave a necessary and sufficient condition for a single-output autonomous system to be locally diffeomorphic to a linear observable system with output injection. Once the system is transformed into the observer form, the nonlinear observer problem can be solved easily by the traditional observer design approach, as illustrated in [4], [1]. The generalizations of [4], [1] to multi-output nonlinear systems were carried out in [5], [10]. Recently, a time-scaling technique, together with coordinate transformation and output injection, has been employed in [9], resulting in a solution to the observer design problem for a larger class of single-output autonomous systems.

For discrete-time nonlinear systems, the observer design problem have also been studied by several researchers; see, for instance, the papers [6], [2], [8], [7], [3], [11], [12] as well as the references therein. Analogous results to those obtained in [4], [5], [10] were reported, for instance, in [6], [8]. The result in [6] was derived under the restricted assumption that a discrete-time autonomous system is invertible, or equivalently, the map $f : \mathbb{R}^n \to \mathbb{R}^n$, is a local diffeomorphism. Such a restriction was removed in [8].

Wei Lin and J.F. Wei are with the Dept. of Electrical Engineering and Computer Science, Case Western Reserve University, Cleveland, Ohio 44106. Their work was supported in part by the NF under grant ESC-0400413 and the Herbold Faculty Fellow Award. F. Wan is with the Dept. of Electrical and Electronic Engineering, University of Macau. His work was supported in part by the Macau Science and Technology Development Fund under Grant FDCT/017/2007/A, and the University of Macau Research Committee under Grant RG062/06-07S/08T/WF/FST.

Following the work [9], in this paper we develop a discrete counterpart of the time-scaling technique to address the observe design problem, for a broader class of discrete-time nonlinear systems than those considered previously in [6], [8]. Specifically, we are interested in the following two questions:

- **Q1:** When is the nonlinear system (1.1) locally diffeomorphic to

$$\left\{ \begin{array}{lcl} z(k+1) & = & s(y(k))Az(k) \\ y(k) & = & Cz(k) \quad ? \end{array} \right. \tag{1.2}$$

  where (C, A) is observable and $s : R^m \to (0, +\infty)$ is a smooth function. (1.2) is referred as the *output-scaled linear observable form.*
- **Q2:** When can the nonlinear system (1.1) be transformed into (by a local diffeomorphism $z = T(x)$)

$$\left\{ \begin{array}{lcl} z(k+1) & = & s(y(k))(Az(k) + \Phi(y(k))) \\ y(k) & = & Cz(k) \quad ? \end{array} \right. \tag{1.3}$$

  where (C, A) is observable, $\Phi(\cdot)$ is a smooth vector field, and $s : R^m \to (0, +\infty)$ is a smooth function. (1.3) is called the *output-scaled observer form with output injection.*

The primary motivation for studying the two issues above comes from the following observation.

Lemma 1.1: For the nonlinear system of the form (1.3), if $s(y)$ is bounded by a constant $l > 0$, then there exists a matrix $L \in R^{n\times m}$ such that

$$\hat{z}(k+1) = s(y(k))[A\hat{z}(k) + \varphi(y(k)) + L(y(k) - C\hat{z}(k)] \tag{1.4}$$

is a global convergent observer for (1.2), satisfying

$$\lim_{k\to\infty} \|z(k) - \hat{z}(k)\| = 0, \forall (z(0), \hat{z}(0)) \in R^n \times R^n. \tag{1.5}$$

Proof. Let $e(k) = z(k) - \hat{z}(k)$ be the estimate state error. From (1.3) and (1.4), it is clear that the error dynamics is $e(k+1) = s(y(k))(A - LC)e(k)$. Consequently,

$$e(k+1) = (A - LC)^{k+1} e(0) \prod_{i=0}^{k} s(y(i))$$

from which it follows that

$$\|e(k+1)\| \quad \le \quad \|A - LC\|^{k+1} \|e(0)\| \; l^{k+1}$$

Since (A, C) is observable pair, one can choose L to assign the eigenvalues of $(A - LC)$ arbitrarily, so that $\|A - LC\|^{k+1} \le \beta \frac{1}{(2l)^{k+1}}$, where $\beta > 0$ is a real number. As a consequence, $\|e(k+1)\| \le \frac{\beta}{2^{k+1}} \|e(0)\|$. This, in turn, yields (1.5). ∎

978-1-4244-3124-3/08/$25.00 ©2008 IEEE

With the aid of Lemma 1.1, the observer design problem for the nonlinear system (1.1) can be solved straightforwardly. In fact, for the nonlinear system (1.1) that is diffeomorphic to the output-scaled linear observable form (1.2) or the output-scaled observer form with output injection (1.3), one can design the observer (1.4) so that the estimate error $e(k) = \hat{z}(k) - z(k)$ tends to zero as $k \to \infty$. This, in turn, implies that $\hat{x} = T^{-1}(\hat{z})$ is an estimate state of the nonlinear system (1.1) and $\hat{x}$ eventually approaches to the state x. In other words, the dynamic system

$$\hat{x}_{k+1} = T^{-1} \circ (AT(\hat{x}_k) + \varphi(y_k) + L[y_k - CT(\hat{x}_k)])$$

is a nonlinear observer for the discrete-time system (1.1).

In the next two sections, necessary and sufficient conditions will be derived for system (1.1) to be locally equivalent to, via a change of coordinates $z = T(x)$, the output-scaled linear observable system (1.2) and the output-scaled nonlinear system with output injection (1.3), respectively. The results of this paper can be regarded as a discrete analogous of the ones presented in [9]. They extend the previous work [8] and enlarge the class of nonlinear systems under consideration, due to the introduction of the output-scaling factor $s(y)$. In the case when $s(y) = 1$, the results of this paper recover the previous ones given in [8]

We end this section by introducing some notations to be frequently used in the rest of the paper.

- Given a mapping $f : R^n \to R^n$, define $f^0(x) = x$ and $f^i(x) = f(f^{i-1}(x)) = f \circ f^{i-1}(x)$, for $i = 1, 2, \cdots$
- Let $f : R^n \to R^m, g : R^m \to R^p$ and $h : R^p \to R^q$ be the continuous functions. Then, the composite function $h(g(f(x)))$ is denoted as $h \circ g \circ f(x)$.
- Let $T : R^n \to R^n$ be a smooth mapping, the Jacobian matrix of T is denoted as $\frac{\partial T}{\partial x} = (T)_*$.
- Given a real-valued function $\lambda : R^n \to R$, its differential is defined as $d\lambda = \frac{\partial \lambda}{\partial x} = \left(\frac{\partial \lambda}{\partial x_1} \ \frac{\partial \lambda}{\partial x_2} \ \cdots \frac{\partial \lambda}{\partial x_n} \right)$.

II. Output-Scaled Linear Observable Form

In this section, we focus on the question **Q1** and characterize a necessary and sufficient condition for the discrete-time nonlinear system (1.1) to be locally diffeomorphic to the output-scaled linear observable form (1.2). For simplicity, we first consider the single-output case.

A. The Single-Output Case

In the case when $m = 1$, the following result answers the question of when system (1.1) can be transformed into (1.2) by a local diffeomorphism $z = T(x)$.

Theorem 2.1: The single-output nonlinear system (1.1) is locally equivalent to the output-scaled linear observable form (1.2) via a change of coordinates $z = T(x)$ if, and only if

(i) the pair $\left(\frac{\partial f}{\partial x}|_{x=0}, \frac{\partial h}{\partial x}|_{x=0} \right)$ is observable;

(ii) there are real constants $a_i, i = 1, 2, \cdots, n$ such that

$$h \circ f^n(x) = \sum_{i=1}^{n} a_i \Big(\prod_{j=i}^{n} [s \circ h \circ f^{j-1}(x)] \Big) \ h \circ f^{i-1}(x)$$

for all x in a neighborhood of $x = 0$.

Proof. *Necessity:* If there exists a local diffeomorphism $z = T(x)$ that transforms system (1.1) into the output-scaled linear observable form (1.2), the following relations hold in the neighborhood of the origin.

$$\begin{aligned} T \circ f \circ T^{-1}(z) &= s(Cz)Az \\ h \circ T^{-1}(z) &= Cz \end{aligned} \quad (2.1)$$

Thus, at $z = x = 0$,

$$\begin{aligned} s(0)A &= (T)_*|_{x=0} \big(\frac{\partial f}{\partial x} \big)_{x=0} (T^{-1})_*|_{z=0} \\ C &= dh|_{x=0} (T^{-1})_*|_{z=0} \end{aligned} \quad (2.2)$$

Note that (A, C) is observable and $s(0) \neq 0$. By the nonsingularity of the mappings T and T^{-1}, it is concluded from (2.2) that the pair $\left(\frac{\partial f}{\partial x}|_{x=0}, \frac{\partial h}{\partial x}|_{x=0} \right)$ is observable.

To prove (ii), observe that $h(x) = Cz = CT(x)$. Moreover, it is deduced from (2.1) that

$$\begin{aligned} h \circ f(x) &= (h \circ T^{-1}) \circ (T \circ f(x)) = C \circ (s(Cz)Az) \\ &= s(h(x))CAT(x) \\ h \circ f^2(x) &= h \circ f(f(x)) = s(h(f(x)))CAT(f(x)) \\ &= [s \circ h \circ f(x)][CAT \circ f \circ T^{-1}(z)] \\ &= [s \circ h \circ f(x)]CA \circ s(h(x))Az \\ &= [s \circ h \circ f(x)] \cdot [s \circ h(x)] \cdot CA^2T(x) \end{aligned}$$

By induction, for $i = 1, 2, \cdots, n$

$$h \circ f^i(x) = \Big(\prod_{j=1}^{i} [s \circ h \circ f^{j-1}(x)] \Big) CA^i T(x) \quad (2.3)$$

By the observability of (A, C), there exist $a_1, a_2, \cdots, a_n$ such that $CA^n = \sum_{i=1}^{n} a_i CA^{i-1}$. Hence,

$$\begin{aligned} h \circ f^n(x) &= \Big(\prod_{j=1}^{n} [s \circ h \circ f^{j-1}(x)] \Big) \sum_{i=1}^{n} a_i CA^{i-1} T(x) \\ &= \sum_{i=1}^{n} a_i \Big[\prod_{j=i+1}^{n} [s \circ h \circ f^{j-1}(x)] \Big] \Big[\prod_{j=1}^{i} [s \circ h \circ f^{j-1}(x)] \Big] CA^i T(x) \\ &= \sum_{i=1}^{n} a_i \Big(\prod_{j=i}^{n} [s \circ h \circ f^{j-1}(x)] \Big) [h \circ f^{i-1}(x)]. \end{aligned}$$

This completes the proof of necessity.

Sufficiency: Construct the change of coordinates

$$z = \begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_i \\ \vdots \\ z_n \end{bmatrix} = \begin{bmatrix} h(x) \\ \frac{h \circ f(x)}{s \circ h(x)} \\ \vdots \\ \frac{h \circ f^{i-1}(x)}{\prod_{j=1}^{i-1} s \circ h \circ f^{j-1}(x)} \\ \vdots \\ \frac{h \circ f^{n-1}(x)}{\prod_{j=1}^{n-1} s \circ h \circ f^{j-1}(x)} \end{bmatrix} := T(x). \quad (2.4)$$

By the condition (i) and the property of $s(y) \neq 0$, the Jacobian matrix $(\frac{\partial T}{\partial x})_{x=0} = (T)_*|_{x=0}$ is nonsingular. Hence, $z = T(x)$ is a local diffeomorphism.

By construction, $h \circ f^i(x) \equiv z_{i+1}\prod_{j=1}^{i} s \circ h \circ f^{j-1}(x)$. Moreover,

$$\begin{aligned} z_i(k+1) &= \frac{h \circ f^{i-1}(x(k+1))}{\prod_{j=1}^{i-1} s \circ h \circ f^{j-1}(x(k+1))} \\ &= \frac{z_{i+1}(k)\prod_{j=1}^{i} s \circ h \circ f^{j-1}(x(k))}{\prod_{j=1}^{i-1} s \circ h \circ f^{j}(x(k))} \\ &= s \circ h(x(k))\ z_{i+1}(k) \\ &= s(y(k))z_{i+1}(k), \quad i = 1, \cdots, n-1 \end{aligned} \tag{2.5}$$

and

$$\begin{aligned} z_n(k+1) &= \frac{h \circ f^{n-1}(x(k+1))}{\prod_{j=1}^{n-1} s \circ h \circ f^{j-1}(x(k+1))} \\ &= \frac{h \circ f^{n}(x(k))}{\prod_{j=1}^{n-1} s \circ h \circ f^{j}(x(k))} \end{aligned} \tag{2.6}$$

Using the condition (ii), it is deduced from (2.6) that

$$\begin{aligned} z_n(k+1) &= \frac{\sum_{i=1}^{n} a_i \left[\prod_{j=i}^{n} s \circ h \circ f^{j-1}(x_k)\right] h \circ f^{i-1}(x_k)}{\prod_{j=1}^{n-1} s \circ h \circ f^{j}(x(k))} \\ &= s \circ h(x(k)) \sum_{i=1}^{n} a_i \left(\frac{h \circ f^{i-1}(x(k))}{\prod_{j=1}^{i-1} s \circ h \circ f^{j-1}(x(k))} \right) \\ &= s(y(k)) \sum_{i=1}^{n} a_i z_i(k) \end{aligned} \tag{2.7}$$

Putting (2.5) and (2.7) together, it is clear that the local diffeomorphism $z = T(x)$ defined by (2.4), transforms the nonlinear system (1.1) into

$$\begin{aligned} z(k+1) &= s(y(k))Az(k) \\ y(k) &= Cz(k) \end{aligned}$$

where

$$A = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \\ a_1 & a_2 & a_3 & \dots & a_n \end{bmatrix} \text{ and } C = [1 \quad 0 \dots \quad 0]. \tag{2.8}$$

Obviously, (A, C) is observable. ∎

Remark 2.1: It is not difficult to show that the condition (i) in Theorem 2.1 (i.e., $(\frac{\partial f}{\partial x}|_{x=0}, \frac{\partial h}{\partial x}|_{x=0})$ is observable) is equivalent to the statement that for all x in the neighborhood U of $x = 0 \in R^n$,

$$\dim\left(\text{span}\left\{dh(x), dh \circ f(x), \cdots, dh \circ f^{n-1}(x)\right\}\right) = n, \tag{2.9}$$

which is a discrete analogue of the condition

$$\dim(\text{span}\{dh(x), dL_f h(x), \cdots, dL_f^{n-1}h(x)\}) = n$$

in the continuous-time case.

B. The Multi-Output Case

We now consider the nonlinear system (1.1) with $m > 1$, i.e. the multi-output case.

Theorem 2.2: The multi-output nonlinear system (1.1) is locally equivalent to the output-scaled linear observable form (1.2) if, and only if there exist observability indices $k_1 \geq k_2 \geq \cdots \geq k_m \geq 1$ with $\sum_{i=1}^{m} k_i = n$, such that for all x in a neighborhood of $x = 0$,

(i) dim $(span\{dh_1(x), \ldots, dh \circ f^{k_1-1}(x); \ldots; dh_m(x), \ldots, dh_m \circ f^{k_m-1}(x)\}) = n$;

(ii) there exist real constants $a^i_{11}, a^i_{12}, \cdots, a^i_{1k_1}, \cdots, a^i_{m1}, a^i_{m2}, \cdots, a^i_{mk_m}$, such that for $i = 1, 2, \cdots, m$,

$$\begin{aligned} h_i \circ f^{k_i}(x) &= \sum_{j=i}^{m} \sum_{l=1}^{k_j} a^i_{jl}\Big(h_i \circ f^{l-1}(x) \cdot \prod_{r=l}^{k_i-1} s \circ h \circ f^r(x)\Big) \\ &+ \sum_{j=1}^{i-1} \sum_{l=1}^{k_i} a^i_{jl}\Big(h_i \circ f^{l-1}(x) \cdot \prod_{r=l}^{k_i-1} s \circ h \circ f^r(x)\Big) \\ &+ \sum_{j=1}^{i-1} \sum_{l=k_i+1}^{k_j} a^i_{jl}\Big(\frac{h_i \circ f^{l-1}(x)}{\prod_{r=k_i+1}^{k_j-1} s \circ h \circ f^r(x)}\Big) \end{aligned} \tag{2.10}$$

Proof. *Necessity:* Similar to the single-input case, system (1.1) being locally equivalent to the output-scaled linear observable system (1.2) implies that the pair $\left(\frac{\partial f}{\partial x}|_{x=0}, \frac{\partial h}{\partial x}|_{x=0}\right)$ is observable. This, in view of Remark 2.1, yields the condition (i) immediately.

The Condition (ii) of Theorem 2.2 can be proven in a manner similar to the proof of Theorem 2.1. Indeed, it is easy to show that the relationship (2.3) still holds in the multi-output case. By the observability of (A, C), there exist indices $k_1 \geq k_2 \geq \cdots \geq k_m \geq 0$ with $\sum_{i=1}^{m} k_i = n$, such that after reordering C_i's, $\{C_iA^{j-1} : i = 1, 2, \cdots, m; j = 1, \cdots, k_i\}$ are linearly independent and $C_iA^{k_i} \in \text{span}\{C_iA^{j-1} : i = 1, 2, \cdots, m; j = 1, \cdots, k_i\}$. Since $s(y) \neq 0$ and $T(0)$ is nonsingular, (2.3) implies that for all x in a neighborhood of $x = 0$ and $i = 1, \cdots, m$,

$$\begin{aligned} h_i \circ f^{k_i}(x) \in span\Big\{&\{\Big(h_j \circ f^l(x)\textstyle\prod_{r=l}^{k_i-1} s \circ h \circ f^r(x)\Big) : \\ &0 \leq l \leq k_j - 1, \text{if } j \geq i;\ 0 \leq l \leq k_i - 1, \text{if } j \leq i\} \\ &\bigcup\{\Big(h_j \circ f^l(x)/\textstyle\prod_{r=k_i}^{k_j-1} s \circ h \circ f^r(x)\Big) : j \leq i, \\ &k_j \geq l \geq k_i\}\Big\}, \end{aligned}$$

which leads to (2.10), i.e. the condition (ii).

Sufficiency: Consider the state transformation

$$z_i = \begin{bmatrix} z_{i1} \\ z_{12} \\ \vdots \\ z_{1k_i} \end{bmatrix} = T_i(x) = \begin{bmatrix} h_i(x) \\ \frac{h_i \circ f(x)}{s \circ h(x)} \\ \vdots \\ \frac{h_i \circ f^{k_i-1}(x)}{\prod_{j=1}^{k_i-1} s \circ h \circ f^{j-1}(x)} \end{bmatrix}$$

$$z = (z_1^T, \cdots, z_m^T)^T = (T_1^T(x), \cdots, T_m^T(x))^T = T(x) \tag{2.11}$$

By the observability condition (i), together with the fact that $s(y) \neq 0$, it is easy to see that the Jacobian $\frac{\partial T}{\partial x}|_{x=0} =$

$(T)_*|_{x=0}$ is nonsingular. Therefore, the transformation $z = T(x)$ defined by (2.11) is a local diffeomorphism. The rest of the proof is similar to the single-input case.

As a matter of fact, for $i = 1, 2, \cdots, m$ and $j = 1, 2, \cdots, k_{i-1}$, it follows from (2.11) that

$$\begin{aligned} z_{ij}(k+1) &= \frac{h_i \circ f^{j-1}(x(k+1))}{\prod_{l=1}^{j-1} s \circ h \circ f^{l-1}(x(k+1))} \\ &= \frac{h_i \circ f^{j}(x(k))}{\prod_{l=1}^{j-1} s \circ h \circ f^{l}(x(k))} \end{aligned} \quad (2.12)$$

Observe that by construction,

$$h_i \circ f^j(x(k)) = z_{i,j+1}(k) \prod_{l=1}^{j} s \circ h \circ f^{l-1}(x(k)) \quad (2.13)$$

Substituting (2.13) into (2.12) yields immediately for $i = 1, 2, \cdots, m, j = 1, 2, \cdots, k_i - 1$,

$$z_{ij}(k+1) = s \circ h(x(k)) z_{i,j+1}(k) = s(y(k)) z_{i,j+1}(k) \quad (2.14)$$

Finally, it is deduced from (2.11) that

$$\begin{aligned} z_{ik_i}(k+1) &= \frac{h_i \circ f^{k_i-1}(x(k+1))}{\prod_{l=1}^{k_i-1} s \circ h \circ f^{l-1}(x(k+1))} \\ &= \frac{h_i \circ f^{k_i}(x(k))}{\prod_{l=1}^{k_i-1} s \circ h \circ f^{l}(x(k))} \end{aligned} \quad (2.15)$$

Substituting (2.10) into (2.15), we have

$$\begin{aligned} z_{ik_i}(k+1) &= s \circ h(x(k)) \sum_{j=1}^{m} \sum_{l=1}^{k_j} a_{jl}^i \\ &\quad \times \left(\frac{h_i \circ f^{l-1}(x(k))}{\prod_{r=1}^{l-1} s \circ h \circ f^{r}(x(k))} \right) \\ &= s(y(k)) \sum_{j=1}^{m} \sum_{l=1}^{k_j} a_{jl}^i z_{jl} \end{aligned} \quad (2.16)$$

Combining (2.14) and (2.10), together with $y(k) = h(x(k)) = (z_{11}(k), z_{21}(k), \cdots, z_{m1}(k))^T$, we conclude that the local diffeomorphism (2.11) transforms system (1.1) to the output-scaled linear observable form (1.2) with $C = diag(C_1, \ldots, C_m) \in \mathbb{R}^{m \times n}$ and $A \in \mathbb{R}^{n \times n}$ whose diagonal blocks are $A_1, \ldots, A_m$, where the matrices $A_i \in \mathbb{R}^{k_i \times k_i}$ and $C_i \in \mathbb{R}^{1 \times k_i}$ are of the form (2.8), for $i = 1, \cdots, m$. ∎

Remark 2.2: It is interesting to note that in the case when $s(y) = 1$, the results of this section recover the previous work [8].

III. Output-Scaled Observer Form with Output Injection

In this section, we turn our attention to the question **Q2**. The purpose of this section is to present a necessary and sufficient condition under which the discrete-time autonomous system (1.1) is locally equivalent, via a change of coordinates $z = T(x)$, to the output-scaled observer form with output injection (1.3).

For the sake of a technical convenience and the readability, we first investigate, as proceeded in the previous section, the single-output case.

A. The Single-Output Case: $m = 1$

Given an observable pair (A, C), with $C \in R^{1 \times n}$ and $A \in R^{n \times n}$, it is well-known that after a linear change of coordinates, the pair can always be put into the form

$$A = \begin{bmatrix} 0 & \ldots & 0 & a_1 \\ 1 & \ldots & 0 & a_2 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \ldots & 1 & a_n \end{bmatrix} \quad \text{and} \quad C = [0 \quad \ldots \quad 0 \quad 1].$$

Let $L = (a_1, a_2 \cdots, a_n)^T$. Then,

$$\bar{A} = A - LC = \begin{bmatrix} 0 & \ldots & 0 & 0 \\ 1 & \ldots & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \ldots & 1 & 0 \end{bmatrix} \quad (3.1)$$

Accordingly, the output-scaled observer form (1.3) can be rewritten as

$$\begin{aligned} z(k+1) &= s(y(k))(Az(k) + \Phi(y(k))) \\ &= s(y(k))(\bar{A}z(k) + \bar{\Phi}(y(k))) \\ y(k) &= Cz(k) \end{aligned}$$

where $\bar{\Phi}(y) := Ly + \Phi(y)$.

The discussion above suggests that without loss of generality, one can assume that the observable pair (A, C) in (1.3) is of the form $(\bar{A}, C)$ given in (3.1).

Now, we are ready to present the main result of the section. In what follows, we denote $\theta_i = h \circ f^{i-1}$, $i = 1, 2, \cdots$, and $\Phi = (\phi_1, \phi_2, \cdots, \phi_n)^T$.

Lemma 3.1: The nonlinear system (1.1) is locally diffeomorphic to the output-scaled observer form with output injection (1.3) if, and only if

(i) the pair $(\frac{\partial f}{\partial x}|_{x=0}, \frac{\partial h}{\partial x}|x=0)$ is observable;

(ii) for all x in a neighborhood of $x = 0$,

$$\theta_{n+1}(x) = \sum_{i=1}^{n} \Big(\phi_i(\theta_i(x)) \prod_{j=i}^{n} s(\theta_j(x)) \Big) \quad (3.2)$$

Proof. *Necessity:* By assumption, there exists a local diffeomorphism $z = T(x)$ such that

$$\begin{aligned} s(Cz)Az &= T \circ f \circ T^{-1}(z) - s(Cz)\Phi(Cz) \\ Cz &= h \circ T^{-1}(z) \end{aligned} \quad (3.3)$$

As a consequence,

$$\begin{aligned} s(0)A &= (T)_*|_{x=0} \left(\tfrac{\partial f}{\partial x}\right)_{x=0} (T^{-1})_*|_{z=0} \\ &\quad -s(0)\tfrac{\partial \Phi}{\partial y}|_{y=0}\, C \\ C &= \left(\tfrac{\partial h}{\partial x}\right)_{x=0} (T^{-1})_*|_{z=0} \end{aligned}$$

Since (A, C) is observable and $s(0) \neq 0$, so is the pair

$$\left((T)_*|_{x=0} \left(\frac{\partial f}{\partial x} \right)_{x=0} (T^{-1})_*|_{z=0},\ \left(\frac{\partial h}{\partial x} \right)_{x=0} (T^{-1})_*|_{z=0} \right).$$

This, together with the nonsingularity of $(T)_*|_{x=0}$ and $(T^{-1})_*|_{z=0}$, yields the condition (i).

To prove the condition (ii), note that A and C are of the canonical form (3.1). With this in mind, it is deduced from (3.3) that

$$\begin{aligned} h(x) &= h\circ T^{-1}(z) = Cz = z_n \\ z_1(k+1) &= s(y(k))\phi_1(z_n(k)) \\ z_2(k+1) &= s(y(k))(z_1(k)+\phi_2(z_n(k))) \\ &\vdots \\ z_n(k+1) &= s(y(k)))(z_{n-1}(k)+\phi_n(z_n(k))) \end{aligned}$$

where z_i and ϕ_i are the i-th component of the vectors $z = T(x)$ and $\Phi(y)$, respectively.

Using the relationship above, we arrive at

$$\begin{aligned} z_n(k) &= h(x(k)) \\ z_{n-1}(k) &= \frac{z_n(k+1)}{s(h(x(k)))} - \phi_n(z_n(k)) \\ &= \frac{h\circ f(x(k))}{s\circ h(x(k))} - \phi_n(h(x(k))) \\ &= \frac{\theta_2(x(k))}{s(\theta_1(x(k)))} - \phi_n(\theta_1(x(k)) \\ &\vdots \\ z_1(k) &= s(y(k))(z_2(k+1)-\phi_2(z_n(k))) \\ &= \frac{\theta_n(x(k))}{\prod_{j=1}^{n-1} s(\theta_j(x(k)))} - \phi_2(\theta_1(x(k))) \\ &\quad -\sum_{i=3}^{n} \frac{\phi_i(\theta_{i-1}(x(k)))}{\prod_{j=1}^{i-2} s(\theta_j(x(k)))} \end{aligned} \quad (3.4)$$

Therefore,

$$\begin{aligned} z_1(k+1) &= s(y(k))\phi_1(z_n(k)) = s(\theta_1(x(k)))\phi_1(\theta_1(x(k))) \\ &= \frac{\theta_{n+1}(x(k))}{\prod_{j=2}^{n} s(\theta_j(x(k)))} - \phi_2(\theta_2(x(k))) \\ &\quad -\sum_{i=3}^{n} \frac{\phi_i(\theta_i(x(k)))}{\prod_{j=2}^{i} s(\theta_j(x(k)))} \end{aligned}$$

from which it follows that

$$\theta_{n+1}(x) = \sum_{i=1}^{n} \Big(\phi_i(\theta_i(x)) \prod_{j=i}^{n} s(\theta_j(x))\Big). \quad (3.5)$$

This completes the proof of necessity.

Sufficiency: Suppose the conditions (i) and (ii) hold. In view of the condition (i) and Remark 2.1, $\{d\theta_1(x), d\theta_2(x), \cdots, d\theta_n(x)\}$ are linearly independent in the neighborhood of $x = 0$. Consequently, a straightforward calculation shows that the mapping $z = T(x)$ defined in (3.4) is a local diffeomorphism, and transforms the nonlinear system (1.1) into the output-scaled observer form with output injection (1.3). ∎

Although Lemma 3.1 has provided a characterization on when the nonlinear system (1.1) is locally equivalent to the output-scaled observer form (1.3), it is not easy to be used. This is because the condition (ii) of Lemma 3.1 relies on the information of $\phi_i(y), i = 1, \cdots, n$, and $s(y)$, thus making the application of Lemma 3.1 a non-trivial job.

In what follows, it is illustrated that in the case when $s(y) = constant = s$, by using Lemma 3.1 one can obtain a necessary and sufficient condition that depends only on the information of the controlled plant (1.1), i.e., the vector fields $f(x)$ and $h(x)$.

Proposition 3.1: The nonlinear system (1.1) is locally diffeomorphic to the output-scaled observer form (1.3) with a constant scaling factor if, and only if

(a) the pair $(\frac{\partial f}{\partial x}|_{x=0}, \frac{\partial h}{\partial x}|x=0)$ is observable;

(b) for all x in a neighborhood of $x = 0$, $\frac{\partial^2 h\circ f^n\circ\psi^{-1}(q)}{\partial q^2}$ is a diagonal matrix, where $x = \psi^{-1}(q)$ is the inverse mapping of

$$q = \psi(x) = \left[h(x), h\circ f(x), \ldots, h\circ f^{n-1}(x)\right]^T. \quad (3.6)$$

Proof. *Necessity:* The condition (a) is obvious. Moreover, it follows from Lemma 3.1 that for all x in a neighborhood of $x = 0$,

$$h\circ f^n(x) = \sum_{i=1}^{n} s^{n-1+i}\phi_i(h\circ f^{i-1}(x)) \quad (3.7)$$

By Remark 2.1, the inverse mapping of $q = \psi(x)$ defined in (3.6) exists and is locally smooth. Thus, it follows from (3.7) and (3.6) that

$$h\circ f^n\circ\psi^{-1}(q) = \sum_{i=1}^{n} s^{n-1+i}\phi_i(q_i) \quad (3.8)$$

This, in turn, implies that the $n\times n$ matrix $\frac{\partial^2 h\circ f^n\circ\psi^{-1}(q)}{\partial q^2}$ is diagonal. Thus, the condition (b) holds.

Sufficiency: Using the condition (b), it is not difficult to prove the existence of a set of smooth functions $\phi_i, 1\le i\le n$, satisfying (3.8). In fact, without loss of generality, assume $\phi_i(0) = 0$. Then, it is easy to check that

$$\phi_i(q_i) = \frac{1}{s^{n-1+i}} h\circ f^n\circ\psi^{-1}(0,\ldots,0,q_i,0,\ldots,0). \quad (3.9)$$

From (3.8) and (3.6), we arrive at (3.7) immediately. Using the smooth functions $\phi_i (1\le i\le n)$ thus obtained, one can construct a state transformation $z = T(x)$ as defined in (3.4), with $s > 0$ being a constant.

Now, a straightforward calculation shows that

$$\frac{\partial T}{\partial x}|_{x=0} = \begin{bmatrix} * & \cdots & * & 1 \\ \vdots & & 1/s & 0 \\ & \cdot & & \\ * & \cdot & \vdots & \vdots \\ & \cdot & & \\ 1/s^{n-1} & \cdots & 0 & 0 \end{bmatrix} \left(\frac{\partial\psi}{\partial x}\right)_{x=0}.$$

In view of the condition (a) and Remark 2.1, it is clear that $\frac{\partial T}{\partial x}|_{x=0}$ is nonsingular, and hence $z = T(x)$ defined by (3.4) is a local diffeomorphism. Moreover, it is easy to verify that the change of coordinates $z = T(x)$ transforms the nonlinear system (1.1) into the observer form (1.3) with $C = [0\ \cdots\ 0\ 1]$ and A being in the form (3.1). ∎

It is worth mentioning that to check the condition (ii), one needs to compute the inverse map ψ^{-1}, which may be a tedious task.

B. The Multi-Output Case: $m > 1$

For the sake of a technical convenience, we assume that A and C are in the following canonical forms when $m > 1$.

$$A = diag(A_1, \dots, A_m), A_i = \begin{bmatrix} 0 & 0 & \dots & 0 & 0 \\ 1 & 0 & \dots & 0 & 0 \\ \vdots & \ddots & & \vdots & \vdots \\ & & \ddots & & \\ 0 & 0 & \dots & 1 & 0 \end{bmatrix}_{k_i \times k_i}$$

$$C = \begin{bmatrix} c_1 & & 0 \\ & c_2 & \\ & & \ddots \\ 0 & & & c_m \end{bmatrix}_{m \times n} \quad c_i = [0 \quad 0 \dots 0 \quad 1]_{1 \times k_i} \tag{3.10}$$

with $k_1 \geq k_2 \geq \cdots \geq k_m \geq 1$ and $\sum_{i=1}^{m} k_i = n$.

In addition,

$$\Phi(y) = \begin{bmatrix} \phi_1(y) \\ \phi_2(y) \\ \vdots \\ \phi_m(y) \end{bmatrix}, \ \phi_i(y) = \begin{bmatrix} \phi_{i1}(y) \\ \phi_{i2}(y) \\ \vdots \\ \phi_{ik_i}(y) \end{bmatrix}, \ 1 \leq i \leq m. \tag{3.11}$$

Similar to the single-input case, the following notations are used: $\theta_{ij} = h_i \circ f^{j-1}, \ i = 1, 2, \cdots, m; \ j = 1, 2, \cdots, k_i$ with $\theta_{i0} = h_i \circ f^0 = h_i$.

Lemma 3.2: The multi-output nonlinear system (1.1) is locally equivalent, via a state transformation $z = T(x)$, to the output-scaled observer form (1.3) with A and C being of the form (3.10) if, and only if, for all x in a neighborhood of $x = 0$,

(1) dim (span$\{d\theta_{ij}(x) : i = 1, 2, \cdots, m; \ j = 1, 2, \cdots, k_i\}) = n$;

(2) There exist $s : R^m \to (0, +\infty)$ and $\Phi : R^m \to R^n$ of the form (3.11), such that for $i = 1, \cdots, m$

$$\theta_{i,k_i+1}(x) = \sum_{j=1}^{k_i} \Big(\phi_{ij}(\theta_{1j}(x), \theta_{2j}(x), \cdots, \theta_{mj}(x)) \times \prod_{l=j}^{k_i} s(\theta_{1l}(x), \theta_{2l}(x), \cdots, \theta_{ml}(x)) \Big). \tag{3.12}$$

This result can be shown by combining the arguments used in the proofs of Lemma 3.1 and Theorem 2.2, and therefore is omitted for the sake of spaces.

Remark 3.1: Notably, it is not an easy job to check the condition (ii) of Lemma 3.2 as it is involved with $s(\cdot)$ and $\phi_{ij}(\cdot)$. Nevertheless, using Lemma 3.2 as a starting point, one can show that if the scaling factor is a constant, a checkable condition can be derived and a result similar to Proposition 3.1 can be established. Due to the limited space, we left the derivation to the reader as an exercise.

IV. Conclusions

In this paper, we have investigated the question of when a discrete-time autonomous system with outputs is locally diffeomorphic to either an output-scaled linear observable form or an output-scaled observer form with output injection. Necessary and sufficient conditions were derived for the existence of a local change of coordinates, by using the differential geometric and functional analysis techniques. Once the local diffeomorphism is found and the nonlinear system is transformed into either one of the observer forms, observer design can be carried out straightforwardly. Indeed, the conventional linear observer design approach can still be employed, resulting in a solution to the observe design problem for a class of discrete-time nonlinear systems.

It should be pointed out that one of the conditions obtained in this paper, namely, the condition (2) of Lemma 3.2 is not easy to check. Although Lemma 3.2 has provided some insights on what kind of discrete-time nonlinear systems can be transformed into the output-scaled nonlinear observer form, it is still not convenient for the observer design. A more friendly-user condition needs to be developed and the class of nonlinear systems should be identified for which the condition (ii) can be simplified or becomes easily checkable. These topics are currently under investigations.

References

[1] D. Bestle and M. Zeitz, Canonical form observer design for nonlinear time variable systems, *Int. J. Control,* Vol. 38, pp. 419-431 (1983).

[2] G. Ciccarela, M. Dalla Mora, and A. Germani, "Observers for discrete-time nonlinear systems," *Syst. Control Lett.,* Vol. 20, p. 373 (1993).

[3] N. Kazantzis and C. Kravaris, "Discrete-time nonlinear observer design using functional equations," *Syst. Control Lett.,* Vol. 42, pp. 81-94 (2001).

[4] A.J. Krener and A. Isidori, Linearization by output injection and nonlinear observers, *Syst. Contr. Lett.,* Vol. 3, pp. 47-52 (1983).

[5] A.J. Krener and W. Respondek, Nonlinear observers with linearizable error dynamics, *SIAM J. Contr. Optimiz.,* Vol. 23, 197-216 (1985).

[6] W. Lee and K. Nam, Observer design for autonomous discrete-time nonlinear systems, *Syst. Contr. Lett.,* Vol. 17, pp. 49-58 (1991).

[7] T. Lilge, "On observer design for nonlinear discrete-time systems," *European J. Control,* Vol. 4, pp. 306-319 (1998).

[8] W. Lin and C. Byrnes, Remarks on linearization of discrete-time automous systems and nonlinear observer design, *Systems & Control Letters,* Vol. 25, pp. 31-40 (1995).

[9] W. Respondek, A. Pogromsky, H. Nijmeijer, Time scaling for observer design with linearizable error dynamics, *Automatica*, **40**, pp. 277-285, 2004.

[10] X. Xia and W. B. Gao, Nonlinear observer design by observer error linearization, *SIAM J. Contr. Optimiz.*, Vol. 27, 199-216 (1989).

[11] M. Xiao, N. Kazantzis, C. Kravaris, and A. Krener, Nonlinear discrete-time observer design with linearizable error dynamics, *IEEE Trans. Automat. Contr.,* Vol. 48, 622-626 (2003).

[12] M. Xiao, "A direct method for the construction of nonlinear discrete-time observer with linearizable error dynamics," *IEEE Trans. Automat. Contr.,* Vol. 51, pp. 128-135 (2006).