

# SINGULAR VALUE ANALYSIS OF HANKEL OPERATORS FOR GENERAL NONLINEAR SYSTEMS

K. Fujimoto<sup>a</sup>, J.M.A. Scherpen<sup>b</sup>

<sup>a</sup>Department of Systems Science  
Kyoto University  
Uji, Kyoto 611-0011, Japan  
fuji@robot.kuass.kyoto-u.ac.jp

<sup>b</sup>Faculty of Info. Tech. and Syst.  
Delft University of Technology  
P.O. Box 5031, 2600 GA Delft, The Netherlands  
J.M.A.Scherpen@its.tudelft.nl

**Keywords** : balanced realization, model reduction, Hankel operators, nonlinear control systems.

## Abstract

This paper discusses singular value analysis of Hankel operators for both continuous-time and discrete-time general nonlinear systems. Singular value analysis clarifies the gain structure of a given operator. Here it is proven that singular value analysis of smooth Hankel operators defined on Hilbert spaces can be characterized by simple equations in terms of their states. A balancing and model reduction procedure is derived based on it. In particular, when the proposed model reduction method is applied to continuous-time nonlinear systems, several gain properties such as Hankel norm, controllability and observability functions are preserved.

## 1 Introduction

A nonlinear extension of the linear state-space concept of balanced realizations has been introduced in [11], mainly based on studying the past input energy and the future output energy. Since then, many results on state-space balancing, modifications, computational issues for model reduction and related minimality considerations for nonlinear systems have appeared in the literature, e.g. [5, 6, 9, 12]. Further, the relation of the state-space notion of balancing for nonlinear systems with the nonlinear Hankel operator has been considered, see e.g. [5, 13, 12]. In particular, *singular value functions* [11] which are nonlinear state-space extension of the Hankel singular values in the linear case play an important role in the nonlinear Hankel theory. It has been shown that singular value analysis of nonlinear Hankel operators can derive natural nonlinear generalization of balanced realization of input-affine nonlinear state-space systems [1, 2]. It also provides some beautiful invariance properties to the corresponding model reduction procedure [3], e.g., the Hankel norm of the target nonlinear system is preserved under model reduction.

On the other hand, recently a balancing method for discrete-time nonlinear systems was investigated [14] and a result on computation of controllability and observability functions of discrete-time nonlinear systems has been reported in [7]. However, typical discrete-time systems are not input-affine and the results developed for continuous-time systems are not directly

applicable to discrete-time systems. The relationship between input-output behavior and the state-space balancing was quite unclear for discrete-time systems so far. Since the authors' former result [1] was strongly dependent on continuous-time input-affine state-space realizations, it cannot handle general nonlinear systems such as discrete-time systems nor input-non-affine ones.

The main objective of this paper is to provide a basic framework for singular value analysis of general nonlinear Hankel operators which does not require any specific state-space realization of the target system. This framework clarifies an algebraic characterization of the singular value structure of the Hankel operators by only using the input-output properties. This characterization derives a concrete balancing procedures based on a more limited technique with our new input-output property which is now applicable to both continuous-time and discrete-time input-non-affine nonlinear state-space systems. Furthermore, we derive a model reduction procedure for continuous-time input-non-affine nonlinear systems where the gain properties such as controllability and observability functions, singular value functions and the Hankel norm of the system are preserved. It is also expected that this framework can be utilized for balancing and model reduction for more general systems such as infinite dimensional systems since the proposed approach can handle any smooth operators on Hilbert spaces.

## 2 Problem setting

This section explains the problem setting for singular value analysis of nonlinear Hankel operators, which is the basic framework for balancing and model reduction of nonlinear control systems.

### 2.1 Hankel operator

Let us consider a nonlinear *Hankel operator*  $\mathcal{H} : U \rightarrow Y$  defined on Hilbert spaces  $U$  and  $Y$ . Here we suppose that  $\mathcal{H}$  can be decomposed as

$$\mathcal{H} = \mathcal{O} \circ \mathcal{C} \quad (1)$$

with the *controllability operator*  $\mathcal{C} : U \rightarrow X$  and the *observability operator*  $\mathcal{O} : X \rightarrow Y$  where  $\mathcal{C}$  is surjective and  $X$  is also a Hilbert space. Typical examples of  $\mathcal{H}$  are related to the following dynamical systems. See [2] for the details.

**Example 1** Consider an asymptotically stable finite dimensional continuous-time linear system

$$\begin{cases} \dot{x} &= Ax + Bu \\ y &= Cx + Du \end{cases}.$$

Its controllability operator  $\mathcal{C} : L_2^m(\mathbb{R}_+) \rightarrow \mathbb{R}^n$  and observability operator  $\mathcal{O} : \mathbb{R}^n \rightarrow L_2^p(\mathbb{R}_+)$  are defined by

$$\begin{aligned} x^0 &= \mathcal{C}(u) := \int_0^\infty e^{A\tau} Bu(\tau) d\tau \\ y &= \mathcal{O}(x^0) := Ce^{At} x^0. \end{aligned}$$

Its Hankel operator is given by the composition (1) with  $U = L_2^m(\mathbb{R}_+)$ ,  $X = \mathbb{R}^n$  and  $Y = L_2^p(\mathbb{R}_+)$ .

**Example 2** Consider an  $L_2$ -stable finite dimensional continuous-time nonlinear system

$$\begin{cases} \dot{x} &= f(x, u, t) \\ y &= h(x, u, t) \end{cases}.$$

Its controllability operator  $\mathcal{C} : L_2^m(\mathbb{R}_+) \rightarrow \mathbb{R}^n$  and observability operator  $\mathcal{O} : \mathbb{R}^n \rightarrow L_2^p(\mathbb{R}_+)$  are defined by

$$\begin{aligned} x^0 = \mathcal{C}(u) &: \begin{cases} \dot{x} &= -f(x, u, t) & x(\infty) = 0 \\ x^0 &= x(0) \end{cases} \quad (2) \\ y = \mathcal{O}(x^0) &: \begin{cases} \dot{x}(t) &= f(x, 0, t) & x(0) = x^0 \\ y &= h(x, 0, t) \end{cases}. \quad (3) \end{aligned}$$

Its Hankel operator is given by the composition (1) with  $U = L_2^m(\mathbb{R}_+)$ ,  $X = \mathbb{R}^n$  and  $Y = L_2^p(\mathbb{R}_+)$ .

**Example 3** Consider an  $\ell_2$ -stable finite dimensional discrete-time nonlinear system

$$\begin{cases} x(t+1) &= f(x(t), u(t), t) \\ y(t) &= h(x(t), u(t), t) \end{cases}.$$

Its controllability operator  $\mathcal{C} : \ell_2^m(\mathbb{Z}_+) \rightarrow \mathbb{R}^n$  and observability operator  $\mathcal{O} : \mathbb{R}^n \rightarrow \ell_2^p(\mathbb{Z}_+)$  are defined by

$$\begin{aligned} x^0 = \mathcal{C}(u) &: \begin{cases} x(t-1) &= f(x(t), u(t), t) & x(\infty) = 0 \\ x^0 &= x(0) \end{cases} \\ y = \mathcal{O}(x^0) &: \begin{cases} x(t+1) &= f(x(t), 0, t) & x(0) = x^0 \\ y(t) &= h(x(t), 0, t) \end{cases}. \end{aligned}$$

Its Hankel operator is given by the composition (1) with  $U = L_2^m(\mathbb{Z}_+)$ ,  $X = \mathbb{R}^n$  and  $Y = L_2^p(\mathbb{Z}_+)$ .

Here we investigate the singular value structure of Hankel operators which is a generalized version of the results in [1]. This investigation will derive a balancing and model reduction procedure which are applicable to a much wider class of nonlinear systems such as time-varying systems, input-non-affine systems, discrete-time systems.

The controllability and observability functions  $L_c : X \rightarrow \mathbb{R}_+$  and  $L_o : X \rightarrow \mathbb{R}_+$  with respect to the Hankel operator  $\mathcal{H}$  given in (1) are defined by

$$L_c(x^0) := \inf_{\substack{\mathcal{C}(u)=x^0 \\ u \in U}} \frac{1}{2} \|u\|^2 \quad (4)$$

$$L_o(x^0) := \frac{1}{2} \|\mathcal{O}(x^0)\|^2. \quad (5)$$

If the pseudo-inverse  $\mathcal{C}^\dagger : X \rightarrow U$  of  $\mathcal{C} : U \rightarrow X$  defined by

$$\mathcal{C}^\dagger(x^0) := \arg \inf_{\substack{\mathcal{C}(u)=x^0 \\ u \in U}} \|u\| \quad (6)$$

exists, then  $L_c$  can be written as

$$L_c(x^0) = \frac{1}{2} \|\mathcal{C}^\dagger(x^0)\|^2.$$

## 2.2 Singular value analysis of nonlinear operators

The main objective of this paper is to investigate the gain structure of the Hankel operator  $\mathcal{H}$ , i.e., to examine

$$\begin{aligned} \rho_{\max}(c) &:= \sup_{\|u\|=c} \frac{\|\mathcal{H}(u)\|}{\|u\|} \\ v_{\max}(c) &:= \arg \sup_{\|u\|=c} \frac{\|\mathcal{H}(u)\|}{\|u\|}. \end{aligned}$$

We add the constraint  $\|u\| = c$  because we are interested in the maximizing input for each input magnitude.

Here we suppose the Hankel operator  $\mathcal{H}$  is (Fréchet) differentiable. Since  $v_{\max}$  is the stationary (critical) points of  $(\|\mathcal{H}(u)\|/\|u\|)$ ,  $v = v_{\max}$  needs to satisfy

$$d\left(\frac{\|\mathcal{H}(u)\|}{\|u\|}\right)(v) = 0 \quad \text{subj. to} \quad \|v\| = c. \quad (7)$$

Actually calculating the derivative of the first equation in (7), we obtain

$$\begin{aligned} 0 &= d\left(\frac{\|\mathcal{H}(u)\|}{\|u\|}\right)(v) \\ &= \frac{\|u\| \cdot d(\|\mathcal{H}(u)\|)(du) - \|\mathcal{H}(u)\| \cdot d(\|u\|)(du)}{\|u\|^2} \\ &= \frac{\langle \|u\|/\|\mathcal{H}(u)\| \langle \mathcal{H}(u), d\mathcal{H}(u)(du) \rangle - (\|\mathcal{H}(u)\|/\|u\|) \langle u, du \rangle}{\|u\|^2} \\ &= \frac{\langle \|u\|/\|\mathcal{H}(u)\| \langle (d\mathcal{H}(u))^* \circ \mathcal{H}(u), du \rangle - (\|\mathcal{H}(u)\|/\|u\|) \langle u, du \rangle}{\|u\|^2} \\ &= \frac{\langle (d\mathcal{H}(u))^* \circ \mathcal{H}(u) - (\|\mathcal{H}(u)\|/\|u\|)u, du \rangle}{\|u\|^4 \cdot \|\mathcal{H}(u)\|}. \end{aligned} \quad (8)$$

On the other hand, the constraint in (7) reduces to

$$\langle u, du \rangle = 0.$$

Hence we can rewrite the problem (7) into

$$\begin{aligned} \langle (d\mathcal{H}(v))^* \circ \mathcal{H}(v) - (\|\mathcal{H}(v)\|/\|v\|)v, du \rangle &= 0 \\ \text{for } \forall du \text{ s.t. } \langle v, du \rangle &= 0. \end{aligned}$$

Finally, we can obtain an alternative formulation of (7) as follows. The proof is straightforwardly obtained from the above discussion.

**Proposition 1** Suppose the Hankel operator  $\mathcal{H} : U \rightarrow Y$  is Fréchet differentiable. Then a  $v \in U$  satisfies (7) if and only if it satisfies  $\|v\| = c$  and

$$(\mathrm{d}\mathcal{H}(v))^* \circ \mathcal{H}(v) = \lambda v, \quad \lambda \in \mathbb{R}. \quad (9)$$

Here Equation (9) characterizes all stationary (critical) inputs as well as the maximizing input  $v_{\max}$ . For the objective of Hankel theory in control, we are only interested in such stationary inputs in the image space of  $\mathcal{C}^\dagger$ , see [1] for a detailed discussion on this matter. Therefore what we have to solve here is Equation (9) and

$$v \in \mathrm{Im} \mathcal{C}^\dagger. \quad (10)$$

We call investigation of the solution of the above equation ‘‘singular value analysis of  $\mathcal{H}$ ’’. Here the solution  $v$  is ‘‘a singular vector’’ and the corresponding scalar  $\rho$  defined by

$$\rho = \frac{\|\mathcal{H}(v)\|}{\|v\|} \quad (11)$$

is called a ‘‘singular value’’ of  $\mathcal{H}$ .

It was proven in our former paper [1] that the singular value structure (9) can be characterized by an algebraic equation using  $L_c$  and  $L_o$  if the target system is an input-affine continuous-time nonlinear system. However, this result were not directly applicable to general (neither non-affine nor discrete-time) nonlinear systems so far.

### 3 Singular value analysis of Hankel operators

The objective of this section is to provide the solution of Equations (9) and (10) for singular value analysis of the Hankel operator  $\mathcal{H}$ . Here we assume the smoothness of the operators  $\mathcal{C}$ ,  $\mathcal{O}$  and  $\mathcal{C}^\dagger$ .

**Assumption A1** the operators  $\mathcal{C} : U \rightarrow X$ ,  $\mathcal{O} : X \rightarrow Y$  and  $\mathcal{C}^\dagger : X \rightarrow U$  exist and are differentiable.

Under Assumption A1 we can obtain an alternative characterization of singular value analysis of the Hankel operator on the signal space  $X$  which is much simpler than (9).

First of all, in order to characterize the signal space satisfying the constraint (10), let us consider the properties of the pseudo-inverse operator  $\mathcal{C}^\dagger$ . By Assumption A1, both  $\mathcal{C}$  and  $\mathcal{C}^\dagger$  exist and are smooth. Hence the constraint (10) can be characterized by singular value analysis of  $\mathcal{C}^\dagger \circ \mathcal{C}$ . That is,

$$\arg \sup_{u \in U} \frac{\|\mathcal{C}^\dagger \circ \mathcal{C}(u)\|}{\|u\|}$$

with the maximum singular value 1 characterizes the elements of  $\mathrm{Im} \mathcal{C}^\dagger$ , because

$$\begin{aligned} \frac{\|\mathcal{C}^\dagger \circ \mathcal{C}(u)\|}{\|u\|} &= 1 & u \in \mathrm{Im} \mathcal{C}^\dagger \\ \frac{\|\mathcal{C}^\dagger \circ \mathcal{C}(u)\|}{\|u\|} &< 1 & \text{otherwise} \end{aligned}$$

hold for the definition of  $\mathcal{C}^\dagger$  in (6).

By the argument similar to Equation (8) we know that Equation (10) reduces to singular value analysis

$$(\mathrm{d}(\mathcal{C}^\dagger \circ \mathcal{C})(v))^* \circ \mathcal{C}^\dagger \circ \mathcal{C}(v) = \frac{\|\mathcal{C}^\dagger \circ \mathcal{C}(v)\|}{\|v\|} v = v$$

since the maximum singular value is 1. This turns out to be

$$(\mathrm{d}\mathcal{C}(v))^* \circ (\mathrm{d}\mathcal{C}^\dagger(\mathcal{C}(v)))^* \circ \mathcal{C}^\dagger \circ \mathcal{C}(v) = v. \quad (12)$$

On the other hand, the decomposition of  $\mathcal{H}$  in (1) implies that the singular value analysis equation (9) can be written as

$$(\mathrm{d}\mathcal{C}(v))^* \circ (\mathrm{d}\mathcal{O}(\mathcal{C}(v)))^* \circ \mathcal{O} \circ \mathcal{C}(v) = \lambda v. \quad (13)$$

Comparing (12) and (13), we obtain a sufficient condition to characterize the singular structure of  $\mathcal{H}$  as

$$(\mathrm{d}\mathcal{O}(\mathcal{C}(v)))^* \circ \mathcal{O} \circ \mathcal{C}(v) = \lambda (\mathrm{d}\mathcal{C}^\dagger(\mathcal{C}(v)))^* \circ \mathcal{C}^\dagger \circ \mathcal{C}(v) \quad (14)$$

using the linearity of the operator  $(\mathrm{d}\mathcal{C}(v))^*$ . Defining the intermediate signal  $\xi := \mathcal{C}(v)$ , we can obtain a simpler expression

$$(\mathrm{d}\mathcal{O}(\xi))^* \circ (\mathcal{O}(\xi)) = \lambda (\mathrm{d}\mathcal{C}^\dagger(\xi))^* \circ (\mathcal{C}^\dagger(\xi)). \quad (15)$$

Recall that the derivative of the controllability and observability functions  $L_c$  and  $L_o$  defined in (4) and (5) are given by

$$\mathrm{d}L_c(x)(\mathrm{d}x) = \langle \mathcal{C}^\dagger(x), \mathrm{d}\mathcal{C}^\dagger(x)(\mathrm{d}x) \rangle = \langle (\mathrm{d}\mathcal{C}^\dagger(x))^* \circ \mathcal{C}^\dagger(x), \mathrm{d}x \rangle \quad (16)$$

$$\mathrm{d}L_o(x)(\mathrm{d}x) = \langle \mathcal{O}(x), \mathrm{d}\mathcal{O}(x)(\mathrm{d}x) \rangle = \langle (\mathrm{d}\mathcal{O}(x))^* \circ \mathcal{O}(x), \mathrm{d}x \rangle. \quad (17)$$

Therefore Equation (9) reduces down to

$$\mathrm{d}L_o(\xi) = \lambda \mathrm{d}L_c(\xi).$$

Finally we can obtain the following result which is the generalized version of the result in [1] in the sense that it is applicable to a larger class of input-state-output systems.

**Theorem 1** Suppose that Assumption A1 holds. Assume moreover that there exist  $\lambda \in \mathbb{R}$  and  $\xi \in X$  satisfying

$$\mathrm{d}L_o(\xi) = \lambda \mathrm{d}L_c(\xi). \quad (18)$$

Then  $v \in U$  defined by

$$v := \mathcal{C}^\dagger(\xi) \quad (19)$$

satisfies the equation for singular value analysis of  $\mathcal{H}$

$$(\mathrm{d}\mathcal{H}(v))^* \circ \mathcal{H}(v) = \lambda v. \quad (20)$$

**Proof.** Suppose  $\xi \in X$  is the solution of Equation (18). Then obviously it satisfies the condition (15) by the relations (16) and (17). Note that we can define the corresponding input  $v \in U$  by (19). Then  $v$  has to be the solution of (12) because it is an element of the signal space  $\mathrm{Im} \mathcal{C}^\dagger$ . Equations (12) and (14)

imply it also satisfies (13). That is, Equation (20) holds. This completes the proof.  $\square$

Note that the corresponding singular value  $\rho$  defined in (11) is given by

$$\begin{aligned}\rho &= \frac{\|\mathcal{H}(v)\|}{\|v\|} = \frac{\|\mathcal{O} \circ \mathcal{C} \circ \mathcal{C}^\dagger(\xi)\|}{\|\mathcal{C}^\dagger(\xi)\|} = \frac{\|\mathcal{O}(\xi)\|}{\|\mathcal{C}^\dagger(\xi)\|} \\ &= \sqrt{\frac{(1/2)\|\mathcal{O}(\xi)\|^2}{(1/2)\|\mathcal{C}^\dagger(\xi)\|^2}} = \sqrt{\frac{L_o(\xi)}{L_c(\xi)}}.\end{aligned}$$

In particular, if we can characterize all the solutions  $\xi_i$ 's of (18) and let  $\rho_i$ 's denote the corresponding singular values, then clearly we obtain the property that the Hankel norm, which is the gain of the Hankel operator, coincides with the maximum singular value. That is,

$$\sup_{\substack{u \in U \\ u \neq 0}} \frac{\|\mathcal{H}(u)\|}{\|u\|} = \sup_{\xi_i} \rho_i|_{\xi=\xi_i}.$$

**Example 4** Suppose that our target system is the linear dynamical system given in Example 1. Then the solution of singular value analysis of the corresponding Hankel operator can be characterized by

$$\xi^T Q = \lambda \xi^T P^{-1}$$

with the controllability and observability Gramians  $P$  and  $Q$ , which is equivalent to

$$PQ \xi = \lambda \xi.$$

That is,  $\xi$  is the eigenvector of  $PQ$  and the solution set of  $\xi$  plays the role of the coordinate axes of the balanced realization. Furthermore, the singular value  $\rho$  coincides with the Hankel singular values (square root of the eigenvalues of  $PQ$ ).

**Example 5** Suppose that our target system is the dynamical system given in Example 2 or 3. Then the solution of singular value analysis of the corresponding Hankel operator can be characterized by an algebraic equation

$$\frac{\partial L_o}{\partial x}(\xi) = \lambda \frac{\partial L_c}{\partial x}(\xi). \quad (21)$$

Similarly to the linear case, the solution set of  $\xi$  plays the role of the axes of the balanced coordinates.

Please note that we do not require any state-space realization of any operator here. So Theorem 1 is applicable to very general nonlinear systems including both continuous and discontinuous both input-affine and input-non-affine dynamical system.

## 4 Balanced realization and model reduction

The result on the singular value analysis of Hankel operators given in Theorem 1 allows one to obtain the balanced realization of both continuous-time and discrete-time input-non-affine nonlinear systems. We can also apply the model reduction procedure via balanced truncation technique [11].

### 4.1 Input-normal/output-diagonal balancing

If the intermediate signal space  $X = \mathbb{R}^n$  which is the typical case for normal dynamical systems, the condition (18) reduces to (21) as shown in Example 5. In this case, we can apply the (input-normal) balancing procedure given in [3]. Here we need to employ the following assumption.

**Assumption A2** Suppose that  $X = \mathbb{R}^n$ , that  $(\partial^2 L_c(x)/\partial x^2)(0)$  and  $(\partial^2 L_o(x)/\partial x^2)(0)$  are positive definite and that the eigenvalues of  $((\partial^2 L_c/\partial x^2)(0))^{-1}((\partial^2 L_o/\partial x^2)(0))$  are distinct.

Under Assumption A2, we can prove the existence of  $n$  independent solutions  $\xi_i$ 's of (18) (or (21)).

**Theorem 2** Consider the Hankel operator  $\mathcal{H}$  in (1). Suppose that Assumptions A1 and A2 hold. Then there exists a neighborhood  $S_0 \subset \mathbb{R}$  of 0,  $n$  smooth functions  $\rho_i : S_0 \rightarrow \mathbb{R}_+$ ,  $i \in \{1, 2, \dots, n\}$  such that

$$\min\{\rho_i(s), \rho_i(-s)\} \geq \max\{\rho_{i+1}(s), \rho_{i+1}(-s)\}$$

holds for all  $s \in S_0$  and all  $i \in \{1, 2, \dots, n-1\}$  and that there exist  $n$  distinct smooth curves  $\xi_i : S_0 \rightarrow X$  satisfying  $\xi_i(0) = 0$  and

$$\begin{aligned}L_c(\xi_i(s)) &= \frac{s^2}{2}, \quad L_o(\xi_i(s)) = \frac{\rho_i^2(s) s^2}{2} \\ \frac{\partial L_o}{\partial x}(\xi_i(s)) &= \lambda_i(s) \frac{\partial L_c}{\partial x}(\xi_i(s))\end{aligned}$$

with

$$\lambda_i(s) := \rho_i^2(s) + \frac{s}{2} \frac{d\rho_i^2(s)}{ds}.$$

In particular, if  $S_0 = \mathbb{R}$ , then

$$\sup_{\substack{u \in U \\ u \neq 0}} \frac{\|\mathcal{H}(u)\|}{\|u\|} = \sup_{s \in S_0} \rho_1(s).$$

**Proof.** The proof is straightforwardly obtained from those of Theorems 6 and 7 in [2] which is based on Brouwer's fixed point theorem [8].  $\square$

Here the functions  $\rho_i$ 's are called *axis singular value functions* and they are the "singular values" of the corresponding Hankel operator  $\mathcal{H}$  indeed. Using Theorem 2 recursively, we can obtain the following input-normal/output-diagonal realization whose axes coincide with the solutions  $\xi_i$ 's in Theorem 2.

**Theorem 3** Consider the Hankel operator  $\mathcal{H}$  in (1). Suppose that Assumptions A1 and A2 hold. Then there exists a neighborhood  $X_0 \subset X$  of 0 and a coordinate transformation  $x = \Phi(z)$ ,  $\Phi(0) = 0$ , converting the system into an input-normal/output-diagonal form, i.e. there exist  $n$  smooth functions  $\tau_i : X_0 \rightarrow \mathbb{R}$ ,  $\rho_i : S_0 \rightarrow \mathbb{R}_+$  satisfying

$$\begin{aligned}L_c(\Phi(z)) &= \frac{1}{2} z^T z \\ L_o(\Phi(z)) &= \frac{1}{2} z^T \text{diag}(\tau_1(z), \dots, \tau_n(z)) z\end{aligned}$$

such that

$$z_i = 0 \Leftrightarrow \frac{\partial L_c(\Phi(z))}{\partial z_i} = 0 \Leftrightarrow \frac{\partial L_o(\Phi(z))}{\partial z_i} = 0 \quad (22)$$

holds for all  $i \in \{1, 2, \dots, n\}$  on  $X_0$ . Furthermore

$$\begin{aligned} \tau_i(0, \dots, 0, \overset{i}{z_i}, 0, \dots, 0) &= \rho_i^2(z_i) \\ \frac{\partial \tau_i}{\partial z}(0, \dots, 0, \overset{i}{z_i}, 0, \dots, 0) &= (0, \dots, 0, \frac{d\rho_i^2(z_i)}{dz_i}, 0, \dots, 0) \end{aligned}$$

holds for all  $i \in \{1, 2, \dots, n\}$ . Here the notation  $(\overset{i}{\cdot})$  denotes the  $i$ -th element of a given vector. In particular, if  $X_0 = \mathbb{R}^n$ , then

$$\sup_{\substack{u \in U \\ u \neq 0}} \frac{\|\mathcal{H}(u)\|}{\|u\|} = \sup_{z_1 \in \mathbb{R}} \sqrt{\tau_1(z_1, 0, \dots, 0)}.$$

**Proof.** The proof is straightforwardly obtained from that of Theorem 8 in [2].  $\square$

In this section, we have assumed that the system has a finite-dimensional state-space (which is not assumed in Section 3). This means the balancing procedure given in Theorems 2 and 3 are applicable to any finite-dimensional state-space systems as given in Examples 1, 2 and 3. This result does not even require the ordinary dynamics of the system indeed.

## 4.2 Model reduction for dynamical systems

Here we will discuss model reduction procedure based on the balanced realization given in Theorem 3 for input-non-affine systems. To this end we need to employ dynamics as given in Example 2.

Let us now suppose that Assumptions A1 and A2 hold and that we already have the coordinate transformation  $z = \Phi(x)$  for the balanced realization in Theorem 3. Moreover, we employ the following setting.

**Assumption A3** Suppose that the Hankel operator  $\mathcal{H}$  is constructed via the procedure in Example 2 from a time-invariant continuous-time nonlinear system

$$\begin{cases} \dot{x} = f(x, u) \\ y = h(x, u) \end{cases}.$$

Let us consider the case

$$\min\{\rho_k(z_k), \rho_k(-z_k)\} > \max\{\rho_{k+1}(z_{k+1}), \rho_{k+1}(-z_{k+1})\} \quad (23)$$

holds for all  $z \in \Phi^{-1}(X_0)$ . Then the state variables  $z_1, \dots, z_k$  are more important in terms of the Hankel operator than those  $z_{k+1}, \dots, z_n$  due to the ordering of the axis singular value functions  $\rho_i$ 's.

Divide the coordinates into two parts corresponding to the division (23) as

$$\begin{aligned} z &= (z^a, z^b) \in \mathbb{R}^n \\ z^a &:= (z_1, \dots, z_k) \in \mathbb{R}^k \\ z^b &:= (z_{k+1}, \dots, z_n) \in \mathbb{R}^{n-k} \\ f^z(z, u) &:= \begin{pmatrix} f^a(z, u) \\ f^b(z, u) \end{pmatrix} := \frac{\partial \Phi^{-1}(x)}{\partial x} \Big|_{x=\Phi(z)} f(\Phi(z), u) \\ h^z(z, u) &:= h(\Phi(z), u). \end{aligned}$$

Moreover, divide the system  $\Sigma$  into two subsystems accordingly as follows

$$\Sigma^a : \begin{cases} \dot{z}^a = f^a(z^a, 0, u^a) \\ y^a = h^z(z^a, 0, u^a) \end{cases} \quad (24)$$

$$\Sigma^b : \begin{cases} \dot{z}^b = f^b(0, z^b, u^b) \\ y^b = h^z(0, z^b, u^b) \end{cases}. \quad (25)$$

For simplicity, let  $\mathcal{H}^a$  and  $\mathcal{H}^b$  denote the Hankel operators related to the divided state-space systems (24) and (25), and let  $L_c^z, L_o^z, L_c^a, L_o^a, L_c^b, L_o^b$  denote the controllability and observability functions of  $\mathcal{H}$  in the coordinate  $z$ , those of  $\mathcal{H}^a$ , and those of  $\mathcal{H}^b$ , respectively. That is, we have  $L_c^z(z) := L_c(\Phi(z))$  and  $L_o^z(z) := L_o(\Phi(z))$ . Then we obtain the following model reduction properties which are a nice and natural generalization of the linear case results in [10, 4].

**Theorem 4** Consider the Hankel operator  $\mathcal{H}$  in (1) and suppose that Assumptions A1, A2 and A3 hold. Then the related controllability and observability functions satisfy

$$L_c^a(z^a) = L_c^z(z^a, 0), \quad L_o^a(z^a) = L_o^z(z^a, 0) \quad (26)$$

$$L_c^b(z^b) = L_c^z(0, z^b), \quad L_o^b(z^b) = L_o^z(0, z^b). \quad (27)$$

Furthermore, the state-space systems  $\Sigma^a$  and  $\Sigma^b$  are in an input-normal/output-diagonal form in the sense of Theorem 3, and

$$\rho_i^a(z_i^a) = \rho_i(z_i^a) \quad i \in \{1, 2, \dots, k\} \quad (28)$$

$$\rho_i^b(z_i^b) = \rho_{i+k}(z_i^b) \quad i \in \{1, 2, \dots, n-k\} \quad (29)$$

hold with  $\rho_i^a$ 's and  $\rho_i^b$ 's the singular values of the Hankel operators  $\mathcal{H}^a$  and  $\mathcal{H}^b$ , respectively. In particular, if  $X_0 = \mathbb{R}^n$ , then the Hankel norm is also preserved as

$$\sup_{\substack{u^a \in U \\ u^a \neq 0}} \frac{\|\mathcal{H}^a(u^a)\|}{\|u^a\|} = \sup_{\substack{u \in U \\ u \neq 0}} \frac{\|\mathcal{H}(u)\|}{\|u\|}. \quad (30)$$

**Proof.** Following the argument in [11], the observability function  $L_o^z$  of the system  $\Sigma$  in the coordinate  $z$  is given by a solution of a Hamilton-Jacobi equation

$$\frac{\partial L_o^z(z)}{\partial z} f^z(z, 0) + \frac{1}{2} h^z(z, 0)^T h^z(z, 0) = 0.$$

Substituting  $z = (z^a, 0)$  for this equation, we obtain

$$\begin{aligned}
0 &= \left( \frac{\partial L_o^z(z^a, z^b)}{\partial z^a}, \frac{\partial L_o^z(z^a, z^b)}{\partial z^b} \right) \Big|_{z^b=0} \begin{pmatrix} f^a((z^a, 0), u) \\ f^b((z^a, 0), u) \end{pmatrix} \\
&\quad + \frac{1}{2} h^z((z^a, 0), 0)^T h^z((z^a, 0), 0) \\
&= \frac{\partial L_o^z(z^a, 0)}{\partial z^a} f^a((z^a, 0), u) + \frac{1}{2} h^z((z^a, 0), 0)^T h^z((z^a, 0), 0)
\end{aligned}$$

because of (22). Clearly, this equation coincides with the Hamilton-Jacobi equation for the observability function  $L_o^z$  of  $\mathcal{H}^a$ . That is, we have proven the relation (26). The relation (27) can be obtained in the same way.

Next we consider the controllability function  $L_c^z$ . By the definition of  $\mathcal{C}$  and  $L_c$  in (2) and (4), it can be observed that  $L_c^z(z)$  can be obtained by solving a Hamilton-Jacobi equation

$$\frac{\partial L_c^z(z)}{\partial z} f^z(z, u_*(z)) + u_*(z)^T u_*(z) = 0 \quad (31)$$

which is related to the optimal control problem in (4). Here  $u_*(z)$  is the solution of

$$u = -\frac{1}{2} \frac{\partial f^z(z, u)}{\partial u} \frac{\partial L_c^z(z)}{\partial z} \quad (32)$$

The existence (and smoothness) of  $\mathcal{C}^\dagger$  in Assumption A1 implies the existence of the solution  $u = u_*(z)$  here. Substituting  $z = (z^a, 0)$  for the equation (32) yields

$$u = -\frac{1}{2} \frac{\partial f^a((z^a, 0), u)}{\partial u} \frac{\partial L_c^z(z^a, 0)}{\partial z^a} \quad (33)$$

which is equivalent to the constraint equation for the controllability function  $L_c^a$ . Obviously,  $u = u_*(z^a, 0)$  is also the solution of this equation. Further, substituting  $z = (z^a, 0)$  for the equation (31), we obtain

$$\begin{aligned}
0 &= \left( \frac{\partial L_c^z(z^a, z^b)}{\partial z^a}, \frac{\partial L_c^z(z^a, z^b)}{\partial z^b} \right) \Big|_{z^b=0} \begin{pmatrix} f^a((z^a, 0), u_*(z^a, 0)) \\ f^b((z^a, 0), u_*(z^a, 0)) \end{pmatrix} \\
&\quad + u_*(z^a, 0)^T u_*(z^a, 0) \\
&= \frac{\partial L_c^z(z^a, 0)}{\partial z^a} f^a((z^a, 0), u_*(z^a, 0)) + u_*(z^a, 0)^T u_*(z^a, 0)
\end{aligned}$$

which coincides with the Hamilton-Jacobi equation for the controllability function  $L_c^a(z^a)$  for  $\mathcal{H}^a$ . That is, we have the relation (26). The relation (27) can be obtained in the same manner.

Since the balanced realization given in Theorem 3 is characterized only by the controllability and observability functions, the systems  $\Sigma^a$  and  $\Sigma^b$  are also balanced. Then Equations (28)–(30) follow immediately from Theorem 3. This completes the proof.  $\square$

This theorem is a generalized version of Theorem 5 in [3]. The proposed result can handle continuous time input-non-affine nonlinear systems whereas the former result is only valid for input-affine systems. In this model reduction procedure, the reduced systems  $\Sigma^a$  and  $\Sigma^b$  are uniquely determined (coordinate free), although the input-normal coordinate  $z = \Phi^{-1}(x)$  itself is not unique.

Unfortunately, however, we did not obtain similar invariance properties in the discrete-time case yet, although the model reduction procedure itself works in a similar way. Model reduction with preserving several gain properties for nonlinear discrete time systems is still an open problem.

## References

- [1] K. Fujimoto and J. M. A. Scherpen. Eigenstructure of nonlinear Hankel operators. In A. Isidori, F. Lamnabhi-Lagarrigue, and W. Respondek, editors, *Nonlinear Control in the Year 2000*, volume 258 of *Lecture Notes on Control and Information Science*, pages 385–398. Springer-Verlag, Paris, 2000.
- [2] K. Fujimoto and J. M. A. Scherpen. Nonlinear input-normal realizations based on the differential eigenstructure of Hankel operators. Submitted, 2000.
- [3] K. Fujimoto and J. M. A. Scherpen. Balancing and model reduction for nonlinear systems based on the differential eigenstructure of Hankel operators. In *Proc. 40th IEEE Conf. on Decision and Control*, pages 3252–3257, 2001.
- [4] K. Glover. All optimal Hankel-norm approximations of linear multivariable systems and their  $l^\infty$ -error bounds. *Int. J. Control*, 39:1115–1193, 1984.
- [5] W. S. Gray and J. M. A. Scherpen. On the nonuniqueness of balanced nonlinear realizations. In *Proc. 38th IEEE Conf. on Decision and Control*, pages 729–734, 1999.
- [6] J. Hahn and T. F. Edgar. Reduction of nonlinear models using balancing of empirical gramians and Galerkin projections. In *Proc. American Control Conference*, pages 2864–2868, 2000.
- [7] R. Lopezlena and J. M. A. Scherpen. Energy functions and balancing for nonlinear discrete-time systems. In *Proc. of IFAC World Congress*, pages 659–664, 2002.
- [8] J. W. Milnor. *Topology from the Differential Viewpoint*. Princeton University Press, New Jersey, 1965.
- [9] A. J. Newman and P. S. Krishnaprasad. Computation for nonlinear balancing. In *Proc. 37th IEEE Conf. on Decision and Control*, pages 4103–4104, 1998.
- [10] L. Pernebo and L. M. Silverman. Model reduction via balanced state space representations. *IEEE Trans. Autom. Contr.*, AC-27:382–387, 1982.
- [11] J. M. A. Scherpen. Balancing for nonlinear systems. *Systems & Control Letters*, 21:143–153, 1993.
- [12] J. M. A. Scherpen and W. S. Gray. Minimality and local state decompositions of a nonlinear state space realization using energy functions. *IEEE Trans. Autom. Contr.*, AC-45(11):2079–2086, 2000.
- [13] J. M. A. Scherpen and W. S. Gray. Nonlinear Hilbert adjoints: Properties and applications to Hankel singular value analysis. *Nonlinear Analysis: Theory, Methods and Applications*, 51(5):883–901, 2002.
- [14] E. I. Verriest and W. S. Gray. Discrete-time nonlinear balancing. In *Proc. 5th IFAC Symp. Nonlinear Control Systems*, pages 515–520, 2001.