Nonlinear Input-Normal Realizations Based on the Differential Eigenstructure of Hankel Operators

Kenji Fujimoto, *Member, IEEE,* and Jacquelien M. A. Scherpen, *Senior Member, IEEE*

***Abstract*—This paper investigates the differential eigenstructure of Hankel operators for nonlinear systems. First, it is proven that the variational system and the Hamiltonian extension with extended input and output spaces can be interpreted as the Gâteaux differential and its adjoint of a dynamical input–output system, respectively. Second, the Gâteaux differential is utilized to clarify the main result the differential eigenstructure of the nonlinear Hankel operator which is closely related to the Hankel norm of the original system. Third, a new characterization of the nonlinear extension of Hankel singular values are given based on the differential eigenstructure. Finally, a balancing procedure to obtain a new input-normal/output-diagonal realization is derived. The results in this paper thus provide new insights to the realization and balancing theory for nonlinear systems.**

***Index Terms*—Balanced realization, model reduction, nonlinear control.**

I. Introduction

In the theory of continuous-time linear systems, the system Hankel operator plays an important role in a number of problems. For example, when viewed as mapping from past inputs to future outputs, it plays a direct role in the abstract definition of state [1]. It also plays a central role in minimality theory, in model reduction problems, in realization theory, and related to these, in linear identification methods. Specifically, the Hankel operator supplies a set of similarity invariants, the so called Hankel singular values, which can be used to quantify the importance of each state in the corresponding input–output system [2]. The Hankel operator can also be factored into the composition of an observability and controllability operators, from which Gramian matrices can be defined and the notion of a balanced realization follows, first introduced in [3], and further studied by many authors, e.g., [2] and [4]. The Hankel singular values are most easily computed in a state–space setting using the product of the Gramian matrices, though intrinsically they depend only on the given input–output mapping. The linear Hankel theory is rather complete and the relations between and interpretations in the state–space and input–output settings are fully understood.

Manuscript received November 1, 2000; revised January 20, 2002, March 17, 2003, and August 2, 2004. Recommended by Associate Editor Hua Wang. The work of K. Fujimoto was supported in part by a research fellowship of Delft University of Technology, The Netherlands, and by a Visitor Travel Grant from the Dutch NWO.

K. Fujimoto is with Department of Mechanical Science and Engineering, Nagoya University, Nagoya 464–8603, Japan (e-mail: k.fujimoto@ieee.org).

J. M. A. Scherpen is with Delft Center for Systems and Control, Delft University of Technology, 2628 CD Delft, The Netherlands (e-mail: j.m.a.scherpen@dcsc.tudelft.nl).

Digital Object Identifier 10.1109/TAC.2004.840476

The nonlinear extension of the state–space concept of balanced realizations has been introduced in [5], 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., [6]–[11]. Recently, the relation of the state–space notion of balancing for nonlinear systems with the nonlinear input–output Hankel operator has been considered; see, e.g., [6], [12], and [13]. In particular, singular value functions which are nonlinear state–space extension of the Hankel singular values for linear systems play an important role in the nonlinear Hankel theory. It has been shown that singular value functions are related to Hankel operators [12]–[14], [16]. However, there are some major differences with the linear theory, i.e., studying similarity invariance of singular value functions in relation to the nonlinear Hankel operator can be done via several interpretations of the concept of similarity invariance and may result in different conclusions. In this paper, we use the input–output interpretation to study the differential eigenstructure of the nonlinear Hankel operator, and show that such interpretation results in a new characterization of Hankel singular value functions for nonlinear systems. The relation with the state–space characterization of the singular value functions is also considered.

In order to study the singular value structure of nonlinear operators, we need to consider the concept of adjoint operators. Nonlinear adjoint operators can be found in the mathematics literature, e.g., [15], and they are expected to play a similar role in the nonlinear control systems theory. So called nonlinear Hilbert adjoint operators are introduced in [12], [13], and [16] as a special class of nonlinear adjoint operators. The existence of such operators in input–output sense has been shown in [12] and adjoint state–space realizations are only recently available in [14], [17], and [18], where the emphasis has been on the use of port-controlled Hamiltonian system methods.

However, these port-controlled Hamiltonian systems representing adjoint state–space realizations are not having clear relations with the past input and the future output energy functions of the original system, whereas the variational nonlinear adjoint systems given by Hamiltonian extensions, e.g., [19], do have a clear and strong relation with these energy functions. In order to fully exploit the latter relation, we study adjoint operators from a variational point of view, and provide a formal justification for the use of Hamiltonian extensions with extended input and output spaces using Gâteaux differentiation in a way similar to [19] and [20]. Then we apply these results to study the eigenstructure of the Gâteaux differential of the square norm of that operator. It is shown that the eigenstructure of these operators are closely related. This eigenstructure derives an alternative definition of the singular value functions, which have a stronger relationship with the Hankel norm of nonlinear systems, other than the singular value functions given in [5] have. Furthermore a new input-normal/output-diagonalization procedure for nonlinear systems is derived based on the differential eigenstructure.

In Section II, we present the linear system case as a paradigm, in order to present the line of thinking for the nonlinear case. In Section III, we provide the formal justification of the use of Hamiltonian extensions for nonlinear adjoint systems using Gâteaux differentiation. In Section IV, we concentrate on the Hankel operator, and correspondingly on the controllability and observability operators for nonlinear systems. In Section V, we clarify the eigenstructure of the Gâteaux differential of the square norm of the Hankel operator for nonlinear systems. In Section VI, a new procedure is derived for bringing the system in input-normal/output-diagonal form by using the differential eigenstructure clarified in Section V, repetitively. In Section VII, the proposed method is applied to a double pendulum system with the approximation technique based on Taylor series expansion. Finally, we end with some conclusions.

Notation: The mathematical notation used throughout is fairly standard. Vector norms are represented by $\|x\| = (x^{\mathrm{T}}x)^{1/2}$ for $x \in \mathbb{R}^n$. $L_2[a,b]$ represents the set of Lebesgue measurable functions, possibly vector-valued, with finite L_2 norm $\|x\|_2 = (\int_a^b \|x(t)\|^2 \mathrm{d}t)^{1/2}$. A condition about 0 means that this conditions holds for a neighborhood of 0. Finally, $x(\pm\infty)$ is an abbreviation for $\lim_{t\to\pm\infty} x(t)$. Throughout this paper, by smooth we generally mean C^∞, unless stated otherwise.

II. Linear Systems as a Paradigm

This section gives some examples of linear adjoint operators which play an important role in the linear systems theory; see, e.g., [21]. We present them here in a way that clarifies the line of thinking in the nonlinear case. Consider a causal linear input–output system $\Sigma : L_2^m(0,\infty) \to L_2^r(0,\infty)$ with a state–space realization

$$u \mapsto y = \Sigma(u) : \begin{cases} \dot{x} = Ax + Bu \\ y = Cx \end{cases} \tag{1}$$

where $x(0) = 0$ and A is Hurwitz. The Laplace transformation gives its transfer function matrix

$$\Sigma(s) := C(sI - A)^{-1}B.$$

Its adjoint operator $\Sigma^* : L_2^r(0,\infty) \to L_2^m(0,\infty)$ is given by

$$\Sigma(s)^* := \Sigma(-s)^{\mathrm{T}} = B^{\mathrm{T}}(-sI - A^{\mathrm{T}})^{-1}C^{\mathrm{T}}$$

with a state–space realization

$$u_a \mapsto y_a = \Sigma^*(u_a) : \begin{cases} \dot{x} = -A^{\mathrm{T}}x - C^{\mathrm{T}}u_a \\ y_a = B^{\mathrm{T}}x \end{cases}$$

where $x(\infty) = 0$. Here u_a and y_a have the same dimensions as y and u, respectively. It satisfies the definition for Hilbert adjoint operators, namely

$$\langle \Sigma(u), u_a \rangle_{L_2^r} = \langle u, \Sigma^*(u_a) \rangle_{L_2^m}.$$

Since u_a has the same dimension as y we can calculate the magnitude of operators as

$$\|\Sigma(u)\|_{L_2^r}^2 = \langle \Sigma(u), \Sigma(u) \rangle_{L_2^r} = \langle u, \Sigma^* \circ \Sigma(u) \rangle_{L_2^m}$$

by substituting $u_a = \Sigma(u)$. This relation can be utilized to derive the singular values of the corresponding input–output map.

For the finite-dimensional system Σ (1), the Hankel operator $\mathcal{H}$ is given by the composition of the controllability and observability operators $\mathcal{H} = \mathcal{O} \circ \mathcal{C}$, where the observability and controllability operators, $\mathcal{O} : \mathbb{R}^n \to L_2^r(0,\infty)$ and $\mathcal{C} : L_2^m(0,\infty) \to \mathbb{R}^n$, respectively, are given by

$$x^0 \mapsto y = \mathcal{O}(x^0) := C\mathrm{e}^{At}x^0 \tag{2}$$

$$u \mapsto x^0 = \mathcal{C}(u) := \int_0^\infty \mathrm{e}^{A\tau}Bu(\tau)\mathrm{d}\tau. \tag{3}$$

Note that these operators $\mathcal{O}$ and $\mathcal{C}$ are also operators on Hilbert spaces, hence, their adjoint operators are given by $\mathcal{O}^* : L_2^m(0,\infty) \to \mathbb{R}^n$ and $\mathcal{C}^* : \mathbb{R}^n \to L_2^r(0,\infty)$

$$u_a \mapsto x^0 = \mathcal{O}^*(u_a) := \int_0^\infty \mathrm{e}^{A^{\mathrm{T}}\tau}C^{\mathrm{T}}u_a(\tau)\mathrm{d}\tau$$

$$x^0 \mapsto y_a = \mathcal{C}^*(x^0) := B^{\mathrm{T}}\mathrm{e}^{A^{\mathrm{T}}t}x^0.$$

It can be easily checked that they satisfy

$$\left\langle \mathcal{O}(x^0), u_a \right\rangle_{L_2^r} = \left\langle x^0, \mathcal{O}^*(u_a) \right\rangle_{\mathbb{R}^n}$$

$$\left\langle \mathcal{C}(u), x^0 \right\rangle_{\mathbb{R}^n} = \left\langle u, \mathcal{C}^*(x^0) \right\rangle_{L_2^m}.$$

These adjoint operators can be used to calculate the observability and controllability Gramians, respectively

$$\begin{aligned} \left\|\mathcal{O}(x^0)\right\|_{L_2^r}^2 &= \left\langle x^0, \mathcal{O}^* \circ \mathcal{O}(x^0) \right\rangle_{\mathbb{R}^n} \\ &= \left\langle x^0, \int_0^\infty \mathrm{e}^{A^{\mathrm{T}}\tau}C^{\mathrm{T}}C\mathrm{e}^{A\tau}\mathrm{d}\tau x^0 \right\rangle_{\mathbb{R}^n} \\ &= \langle x^0, Qx^0 \rangle_{\mathbb{R}^n} \end{aligned} \tag{4}$$

$$\begin{aligned} \left\|\mathcal{C}^*(x^0)\right\|_{L_2^m}^2 &= \left\langle x^0, \mathcal{C}^{**} \circ \mathcal{C}^*(x^0) \right\rangle_{\mathbb{R}^n} \\ &= \left\langle x^0, \int_0^\infty \mathrm{e}^{A\tau}BB^{\mathrm{T}}\mathrm{e}^{A^{\mathrm{T}}\tau}\mathrm{d}\tau x^0 \right\rangle_{\mathbb{R}^n} \\ &= \langle x^0, Px^0 \rangle_{\mathbb{R}^n}. \end{aligned} \tag{5}$$

These imply $Q = \mathcal{O}^* \circ \mathcal{O}$ and $P = \mathcal{C}^{**} \circ \mathcal{C}^* = \mathcal{C} \circ \mathcal{C}^*$. Also, they fulfill Lyapunov equations

$$A^{\mathrm{T}}Q + QA + C^{\mathrm{T}}C = 0 \tag{6}$$

$$AP + PA^{\mathrm{T}} + BB^{\mathrm{T}} = 0. \tag{7}$$

Furthermore, from [21, Th. 8.1], we know the following fact.

Theorem 1: [21] The operator $\mathcal{H}^* \circ \mathcal{H}$ and the matrix QP have the same nonzero eigenvalues.

The square roots of the eigenvalues of QP are called the *Hankel singular values* of (1) and are denoted by σ_i's where $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_n$. In fact, the largest singular value characterizes the Hankel norm $\|\Sigma\|_H$ of the system Σ

$$\|\Sigma\|_H := \sup_{\substack{u \in L_2(0,\infty) \\ u \neq 0}} \frac{\|\mathcal{H}(u)\|_{L_2}}{\|u\|_{L_2}} = \sigma_1. \tag{8}$$

Further, using a similarity transformation (linear coordinate transformation), we can diagonalize both P and Q and furthermore let them coincide with each other, i.e.,

$$P = Q = \mathrm{diag}\{\sigma_1, \sigma_2, \ldots, \sigma_n\}.$$

The state–space realization satisfying the above equation is called *balanced realization* of Σ.

III. Variational and Adjoint Systems

This section is devoted to the state–space characterization of variational and adjoint systems based on [19], which were intensively utilized in [5] and [20] as a preparation for the main results in the following sections. In order to handle variational and adjoint systems of Hankel operators in this paper, we generalize the results in [19] using [18] and [22] so that they can treat initial and final states as input and output of the system. Since the proofs of the theorems are rather straightforwardly generalized versions of the results in [19], we include them in Appendix.

Consider an operator $\Sigma : \mathcal{U} \to L_2^r(t^0, t^1)$ defined on a (possibly infinite) time interval $(t^0, t^1) \subset \mathbb{R}$ described by the state–space realization

$$u \mapsto y = \Sigma(u) : \begin{cases} \dot{x} = f(x,u,t) & x(t^0) = x^0 \\ y = h(x,u,t) \end{cases} \tag{9}$$

with $x(t) \in \mathcal{X}^0 \subset \mathbb{R}^n$, $u(t) \in \mathbb{R}^m$, $y(t) \in \mathbb{R}^r$, and $\mathcal{U} \subset L_2^m(t^0, t^1)$ an open neighborhood of 0. Here, we assume $(x, u) = (0, 0)$ is an equilibrium, i.e. $f(0,0,t) = 0$ and $h(0,0,t) = 0$ hold for $\forall t \in \mathbb{R}$. and that all signals and functions are sufficiently smooth. This dynamical system can also be regarded as a mapping $\mathbb{R}^n \times L_2^m \to \mathbb{R}^n \times L_2^r$ defined by

$$(x^0, u) \mapsto (x^1, y) = \hat{\Sigma}(x^0, u) : \begin{cases} \dot{x} = f(x,u,t) & x(t^0) = x^0 \\ y = h(x,u,t) \\ x^1 = x(t^1) \end{cases}. \tag{10}$$

The variational system of $\hat{\Sigma}$ is given by

$$\left(x^0, u, x_v^0, u_v\right) \mapsto \left(x_v^1, y_v\right) = \hat{\Sigma}_v\left(x^0, u, x_v^0, u_v\right) : \begin{cases} \dot{x} = f(x,u,t) & x(t^0) = x^0 \\ \dot{x}_v = \frac{\partial f}{\partial x} x_v + \frac{\partial f}{\partial u} u_v & x_v(t^0) = x_v^0 \\ y_v = \frac{\partial h}{\partial x} x_v + \frac{\partial h}{\partial u} u_v \\ x_v^1 = x_v(t^1) \end{cases}. \tag{11}$$

The input-state-output set (u_v, x_v, y_v) are called variational input, state, and output, respectively.

The Hamiltonian extension $\hat{\Sigma}_a$ of $\hat{\Sigma}$ is given by a Hamiltonian control system of the following form:

$$(x^0, u, p^1, u_a) \mapsto (p^0, y_a) = \hat{\Sigma}_a(x^0, u, p^1, u_a) : \begin{cases} \dot{x} = \frac{\partial H}{\partial p}^{\mathrm{T}} = f(x,u,t) & x(t^0) = x^0 \\ \dot{p} = -\frac{\partial H}{\partial x}^{\mathrm{T}} = -\frac{\partial f}{\partial x}^{\mathrm{T}} p - \frac{\partial h}{\partial x}^{\mathrm{T}} u_a & p(t^1) = p^1 \\ y_a = \frac{\partial H}{\partial u}^{\mathrm{T}} = \frac{\partial f}{\partial u}^{\mathrm{T}} p + \frac{\partial h}{\partial u}^{\mathrm{T}} u_a \\ p^0 = p(t^0) \end{cases} \tag{12}$$

with the Hamiltonian

$$H(x, p, u, u_a) := p^{\mathrm{T}} f(x,u,t) + u_a^{\mathrm{T}} h(x,u,t).$$

The structure already reveals a form that corresponds to the linear adjoint notion. In the sequel, this issue is studied in more detail.

Here, the concept of Gâteaux differentiation for dynamical systems with initial and final states from an input–output point of view is considered. It is of importance for understanding the meaning of the Hamiltonian extensions and their relation with adjoint systems. Also, Gâteaux differentiation of Hankel operators plays an important role in the analysis of the properties of Hankel operators, which is the topic of Sections IV and V. To this end, we state the definition of Gâteaux differentiation.

Definition 1: (Gâteaux Differential) Suppose X and Y are Banach spaces, $U \subset X$ is open, and $\Sigma : U \to Y$. Then Σ is said to be Gâteaux differentiable at $x \in U$ if, for all $\zeta \in X$ the following limit exists:

$$\mathrm{d}\Sigma(x)(\zeta) = \lim_{\varepsilon \to 0} \frac{\Sigma(x + \varepsilon\zeta) - \Sigma(x)}{\varepsilon} = \frac{\mathrm{d}}{\mathrm{d}\varepsilon}\Sigma(x + \varepsilon\zeta)|_{\varepsilon=0}.$$

We write $\mathrm{d}\Sigma(x)(\zeta)$ for the Gâteaux differential of Σ at x in the "direction" ζ.

Then, we can prove the following property for the variational system (11) which is a generalized version of the results stated in [19].

Theorem 2: Consider an operator $\hat{\Sigma}$ with the state–space realization (10). Suppose that the trajectory of the state $x_v(t)$ of $\hat{\Sigma}_v(u, u_v)$ in (11) is uniquely determined for $\forall x^0 \in \mathcal{X}^0 \subset \mathbb{R}^n$, $\forall u \in \mathcal{U} \subset L_2^m(t^0, t^1)$, $\forall x_v^0 \in \mathbb{R}^n$ and $\forall u_v \in L_2^m(t^0, t^1)$ where $\mathcal{X}^0$ and $\mathcal{U}$ are open neighborhoods of 0 in $\mathbb{R}^n$ and $L_2^m(t^0, t^1)$, respectively. Then

$\hat{\Sigma} : \mathcal{X}^0 \times \mathcal{U} \to \mathbb{R}^n \times L_2^r(t^0, t^1)$ is Gâteaux differentiable $\Leftrightarrow$ $\hat{\Sigma}_v$ is a mapping[1] of $\mathcal{X}^0 \times \mathcal{U} \times \mathbb{R}^n \times L_2^m(t^0, t^1) \to \mathbb{R}^n \times L_2^r(t^0, t^1)$.

Furthermore, the Gâteaux differential of $\hat{\Sigma}$ is given by

$$\mathrm{d}\hat{\Sigma}(x^0, u)\left(x_v^0, u_v\right) = \hat{\Sigma}_v\left(x^0, u, x_v^0, u_v\right).$$

Perhaps more well-known than the Gâteaux differential is the Fréchet derivative, which is especially useful for analysis of nonlinear static functions. Fréchet derivative is a special class of Gâteaux differentiation and it coincides with Gâteaux differential if it is continuous and linear in the second argument "direction;" see, e.g., [22] and [23].

[1]It is also assumed that the trajectory of the corresponding state is well-defined.

Next, we give the formal justification of calling the Hamiltonian extension $\hat{\Sigma}_a$ the adjoint form of the variational system $\hat{\Sigma}_v$, as is done in [19]. The most general form of the Hamiltonian extension, i.e., including arbitrary initial conditions, can be seen as the differential version of [18, Prop. 2].

Theorem 3: Consider an operator $\hat{\Sigma}$ with the state–space realization (10). Suppose that the assumptions in Theorem 2 hold and that $\hat{\Sigma}_a$ is a mapping of $\mathcal{X}^0 \times \mathcal{U} \times \mathbb{R}^n \times L_2^r(t^0, t^1) \to \mathbb{R}^n \times L_2^m(t^0, t^1)$. Then, there holds

$$\left(\mathrm{d}\hat{\Sigma}(x^0, u)\right)^* (p^1, u_a) = \hat{\Sigma}_a(x^0, u, p^1, u_a) \tag{13}$$

with the inner product on $\mathbb{R}^n \times L_2(t^0, t^1)$.

Remark 1: In addition to the adjoint as given in Theorem 3, for nonlinear operators *the nonlinear Hilbert adjoint*, e.g., [16] and [18], is also a useful tool. A nonlinear Hilbert adjoint of a nonlinear operator $\Sigma : U \to Y$ with Hilbert spaces U and Y is an operator $\Sigma^* : Y \times U \to U$ satisfying

$$\langle \Sigma(u), y \rangle_Y = \langle u, \Sigma^*(y, u) \rangle_U$$

for all $u \in U$, $y \in Y$, where $\Sigma^*(y, u)$ is linear in y. Relations between the Hamiltonian extension concept and the nonlinear Hilbert adjoint can be established. Furthermore, we have established relations between nonlinear Hilbert adjoint operators and port-controlled Hamiltonian systems; see [18] for the details.

Summarizing we may conclude from this section that the Hamiltonian extension with the extended input and output spaces $\hat{\Sigma}_a$ is a control system that is a realization of the Hilbert adjoint of the Gâteaux differential of the original operator with the extended input and output $\hat{\Sigma}$. This interpretation results from taking the Gâteaux differential from the squared L_2 and $\mathbb{R}^n$ norm of the nonlinear operator.

IV. Hankel Operator and Its Differential

This section studies the state–space realizations for the adjoints of some energy functions and operators, and relates them to singular value analysis of nonlinear dynamical operators. We only consider special cases of system (9), namely, time invariant, input-affine, sufficiently smooth nonlinear systems without direct feed-through in the form of

$$\Sigma : \begin{cases} \dot{x} = f(x) + g(x)u \\ y = h(x) \end{cases} \tag{14}$$

defined on the time interval $t \in (-\infty, \infty)$. The system is supposed to be asymptotically stable on a neighborhood of 0, and L_2-stable in the sense that $u \in L_2^m(-\infty, 0)$ implies that $\Sigma(u)$ restricted to $(0, \infty)$ is in $L_2^r(0, \infty)$.

State–space characterization of nonlinear observability and controllability operators can be given as intuitively clear extensions from the linear case. The observability and controllability operators are mappings of $\mathbb{R}^n \to L_2^r(0, \infty)$ and $L_2^m(0, \infty) \to \mathbb{R}^n$, and their state–space realizations are given by

$$x^0 \mapsto y = \mathcal{O}(x^0) : \quad \begin{cases} \dot{x} = f(x) & x(0) = x^0 \\ y = h(x) \end{cases} \tag{15}$$

$$u \mapsto x^1 = \mathcal{C}(u) : \quad \begin{cases} \dot{x} = f(x) + g(x)\mathcal{F}_-(u) & x(-\infty) = 0 \\ x^1 = x(0) \end{cases}. \tag{16}$$

Here $\mathcal{F}_- : L_2^m(0, \infty) \to L_2^m(-\infty, \infty)$ is the *time flipping* operator defined by

$$\mathcal{F}_-(u) := \begin{cases} u(-t) : & t < 0 \\ 0 \quad : & t \geq 0]. \end{cases} \tag{17}$$

Furthermore the Hankel operator $\mathcal{H} : L_2^m(0, \infty) \to L_2^r(0, \infty)$ of Σ is given by

$$\mathcal{H} := \Sigma \circ \mathcal{F}_-.$$

The original definition of these operators was given in [12]. Clearly, there holds $\mathcal{H} = \mathcal{O} \circ \mathcal{C}$ [12].

The state–space realizations of the differentiations $\mathrm{d}\mathcal{O}$, $\mathrm{d}\mathcal{C}$ and $\mathrm{d}\mathcal{H}$ are given by the following lemma which will be utilized in the succeeding sections.

Lemma 1: Consider the operator Σ with the state–space realization (14). Suppose that $\hat{\Sigma}_v$ is a mapping of $\mathcal{X}^0 \times \mathcal{U} \times \mathbb{R}^n \times L_2^m(t^0, t^1) \to \mathbb{R}^n \times L_2^r(t^0, t^1)$. Then, the state–space realizations of $\mathrm{d}\mathcal{O} : \mathbb{R}^n \times \mathbb{R}^n \to L_2^r(0, \infty)$, $\mathrm{d}\mathcal{C} : L_2^m(0, \infty) \times L_2^m(0, \infty) \to \mathbb{R}^n$ and $\mathrm{d}\mathcal{H} : L_2^m(0, \infty) \times L_2^m(0, \infty) \to L_2^r(0, \infty)$ are given by

$$y_v = \mathrm{d}\mathcal{O}(x^0)\left(x_v^0\right) : \quad \begin{cases} \dot{x} = f(x) & x(0) = x^0 \\ \dot{x}_v = \frac{\partial f}{\partial x} x_v & x_v(0) = x_v^0 \\ y_v = \frac{\partial h}{\partial x} x_v \end{cases} \tag{18}$$

$$x_v^0 d = \mathrm{d}\mathcal{C}(u)(u_v) : \quad \begin{cases} \dot{x} = f(x) + g(x)\mathcal{F}_-(u), & x(-\infty) = 0 \\ \dot{x}_v = \frac{\partial (f + g\mathcal{F}_-(u))}{\partial x} x_v \\ \qquad + g(x)\mathcal{F}_-(u_v), & x_v(-\infty) = 0 \\ x_v^0 = x_v(0) \end{cases} \tag{19}$$

$$y_v = \mathrm{d}\mathcal{H}(u)(u_v) : \quad \begin{cases} \dot{x} = f(x) + g(x)\mathcal{F}_-(u), & x(-\infty) = 0 \\ \dot{x}_v = \frac{\partial (f + g\mathcal{F}_-(u))}{\partial x} x_v \\ \qquad + g(x)\mathcal{F}_-(u_v), & x_v(-\infty) = 0 \\ y_v = \frac{\partial h}{\partial x} x_v \end{cases}. \tag{20}$$

Proof: The proof directly follows from Theorem 2 and the definitions of $\mathcal{O}, \mathcal{C}, \mathcal{H}$. ■

Corollary 1: Consider the operator Σ with the state–space realization (14). Suppose that the assumptions in Lemma 1 hold. Then, there hold

$$\mathrm{d}\mathcal{O} = \mathcal{O}_{\mathrm{d}\Sigma} \quad (\mathcal{C}, \mathrm{d}\mathcal{C}) = \mathcal{C}_{\mathrm{d}\Sigma} \quad \mathrm{d}\mathcal{H} = \mathcal{H}_{\mathrm{d}\Sigma}.$$

Here, $\mathcal{H}_{\mathrm{d}\Sigma}$, $\mathcal{C}_{\mathrm{d}\Sigma}$, and $\mathcal{O}_{\mathrm{d}\Sigma}$ denote the Hankel operator, the controllability operator and the observability operator of $\mathrm{d}\Sigma$, respectively.

The adjoints of the operators of Lemma 1 can be obtained by applying Theorem 3, which results in the following lemma.

Lemma 2: Consider the operator Σ with the state–space realization (14). Suppose that the assumptions in Theorem 3 and Lemma 1 hold. Then state–space realizations of $(\mathrm{d}\mathcal{O}(x^0))^* : L_2^r(0, \infty)(\times \mathbb{R}^n) \to \mathbb{R}^n$, $(\mathrm{d}\mathcal{C}(u))^* : \mathbb{R}^n(\times L_2^m(0, \infty)) \to$

$L_2^m(0,\infty)$, and $(\mathrm{d}\mathcal{H}(u))^* : L_2^r(0,\infty)(\times L_2^m(0,\infty)) \to L_2^m(0,\infty)$ are given by

$$p^0 = \left(\mathrm{d}\mathcal{O}(x^0)\right)^*(u_a): \quad \begin{cases} \dot{x} = f(x) & x(0) = x^0 \\ \dot{p} = -\frac{\partial f}{\partial x}^{\mathrm{T}}(x)p - \frac{\partial h}{\partial x}^{\mathrm{T}}(x)u_a & p(\infty) = 0 \\ p^0 = p(0) \end{cases} \tag{21}$$

$$y_a = (\mathrm{d}\mathcal{C}(u))^*(p^1): \quad \begin{cases} \dot{x} = f(x) + g(x)\mathcal{F}_-(u) & x(-\infty) = 0 \\ \dot{p} = -\frac{\partial(f+g\mathcal{F}_-(u))}{\partial x}^{\mathrm{T}}(x)p & p(0) = p^1 \\ y_a = \mathcal{F}_+\left(g(x)^{\mathrm{T}}p\right) \end{cases} \tag{22}$$

$$y_a = (\mathrm{d}\mathcal{H}(u))^*(u_a): \quad \begin{cases} \dot{x} = f(x) + g(x)\mathcal{F}_-(u) & x(-\infty) = 0 \\ \dot{p} = -\frac{\partial(f+g\mathcal{F}_-(u))}{\partial x}^{\mathrm{T}}p - \frac{\partial h}{\partial x}^{\mathrm{T}}u_a & p(\infty) = 0 \\ y_a = \mathcal{F}_+\left(g(x)^{\mathrm{T}}p\right). \end{cases} \tag{23}$$

Proof: The proof is obtained by applying the adjoint Hamiltonian extensions of Section III, and using techniques from [18]. To begin with, substituting $t^0 = 0$, $t^1 = \infty$, $p^1 = p(\infty) = 0$, $u_v = 0$ for the (71) in the proof of Theorem 3 in Appendix yields

$$\begin{aligned}\langle y_v, u_a\rangle_{L_2^r} &= \left\langle \left(x_v^1, y_v\right), (0, u_a)\right\rangle_{\mathbb{R}^n\times L_2^r} \\ &= \left\langle \left(x_v^0, 0\right), \left(p^0, y_a\right)\right\rangle_{\mathbb{R}^n\times L_2^m} = \left\langle x_v^0, p^0\right\rangle_{\mathbb{R}^n}.\end{aligned}$$

Substituting moreover $y_v = \mathrm{d}\mathcal{O}(x^0)(x_v^0)$ and $p^0 = (\mathrm{d}\mathcal{O}(x^0))^*(u_a)$ as in (21) yields

$$\left\langle \mathrm{d}\mathcal{O}(x^0)\left(x_v^0\right), u_a\right\rangle_{L_2^r} = \left\langle x_v^0, \left(\mathrm{d}\mathcal{O}(x^0)\right)^*(u_a)\right\rangle_{\mathbb{R}^n}.$$

This proves the first part.

The second part can be proven in a similar way as in the first part. Substituting $t^0 = -\infty$, $t^1 = 0$, $x_v^0 = x_v(-\infty) = 0$, $u_a = 0$ for the (71) yields

$$\begin{aligned}\left\langle x_v^1, p^1\right\rangle_{\mathbb{R}^n} &= \left\langle \left(x_v^1, y_v\right), \left(p^1, 0\right)\right\rangle_{\mathbb{R}^n\times L_2^r} \\ &= \left\langle \left(0, \mathcal{F}_-(u_v)\right), \left(p^0, y_a\right)\right\rangle_{\mathbb{R}^n\times L_2^m} \\ &= \langle \mathcal{F}_-(u_v), y_a\rangle_{L_2^m} = \langle u_v, \mathcal{F}_+(y_a)\rangle_{L_2^m}.\end{aligned}$$

Substituting moreover $x_v^1 = \mathrm{d}\mathcal{C}(u)(u_v)$ and $y_a = (\mathrm{d}\mathcal{C}(u))^*(p^1)$ as in (22) yields

$$\left\langle \mathrm{d}\mathcal{C}(u)(u_v), p^1\right\rangle_{\mathbb{R}^n} = \left\langle u_v, (\mathrm{d}\mathcal{C}(u))^*\left(p^1\right)\right\rangle_{L_2^m}.$$

This proves the second part.

In order to prove the last part, we use the relation (20) for arbitrary signals $u_v \in \mathcal{U}_v \subset L_2^m(0,\infty)$ and $u_a \in L_2^r(0,\infty)$. Then

$$\begin{aligned}\langle u_a, \mathrm{d}\mathcal{H}(u)(u_v)\rangle_{L_2^r} &= \langle u_a, \mathrm{d}\mathcal{O}\left(\mathcal{C}(u)\right)\left(\mathrm{d}\mathcal{C}(u)(u_v)\right)\rangle_{L_2^r} \\ &= \left\langle (\mathrm{d}\mathcal{C}(u))^* \circ \left(\mathrm{d}\mathcal{O}\left(\mathcal{C}(u)\right)\right)^*(u_a), u_v\right\rangle_{L_2^r} \\ &\equiv \left\langle (\mathrm{d}\mathcal{H}(u))^*(u_a), u_v\right\rangle_{L_2^r}.\end{aligned}$$

Therefore, we obtain

$$(d\mathcal{H}(u))^*(u_a) = (d\mathcal{C}(u))^* \circ \left(\mathrm{d}\mathcal{O}\left(\mathcal{C}(u)\right)\right)^*(u_a). \tag{24}$$

We can check the state–space realization of the right-hand side of the aforementioned equation using (21) and (22) coincide with the left-hand side given by (23). This completes the proof. ■

Lemma 2 can be seen as the differential version of [18, Prop. 3]. It is readily checked that for linear systems the above characterizations yield the well-known state–space characterizations of these operators as shown in Section II.

V. Differential Eigenstructure of Hankel Operators

This section clarifies the eigenstructure of the Gâteaux differential of the square norm of the Hankel operator based on the results developed in the previous section. Furthermore, we introduce a new definition of singular value functions which are closely related to the differential eigenstructure and the Hankel norm of nonlinear dynamical systems.

A. Preliminary Results on Singular Value Functions

In order to proceed, we need to define the energy functions.

Definition 2: The observability function $L_o(x)$ and the controllability function $L_c(x)$ of Σ as in (14) are defined by

$$L_o(x^0) := \frac{1}{2}\int_0^\infty \|y(t)\|^2\, dt \quad x(0) = x^0, u(t) \equiv 0$$

$$L_c(x^1) := \inf_{\substack{u\in L_2^m(-\infty,0)\\ x(-\infty)=0,\ x(0)=x^1}} \frac{1}{2}\int_{-\infty}^0 \|u(t)\|^2\, dt.$$

It is assumed throughout that there exist well-defined controllability operators and Gramians in the linear case. These functions are closely related to observability and controllability functions. The relation between the observability function, operator and Gramian is given by, e.g., [12] and [13],

$$L_o(x^0) = \frac{1}{2}\left\|\mathcal{O}(x^0)\right\|_{L_2^r}^2 = {x^0}^{\mathrm{T}} Q(x^0) x^0 \tag{25}$$

with a square symmetric matrix $Q(x^0)$. In the linear case $Q(x^0)$ is constant and equals the observability Gramian; see (4).

For the controllability function we have a different relation, since we have to deal with the minimum control energy, i.e.,

$$L_c(x^1) = \frac{1}{2}\left\|\mathcal{C}^\dagger(x^1)\right\|_{L_2^m}^2 = {x^1}^T \tilde{P}(x^1) x^1 \tag{26}$$

with a square symmetric matrix $\tilde{P}(x^1)$, where $\mathcal{C}^\dagger : \mathbb{R}^n \to L_2^m(0,\infty)$, which is the pseudoinverse of $\mathcal{C}$ defined by

$$\mathcal{C}^\dagger(x^1) := \arg\min_{\mathcal{C}(u)=x^1} \|u\|_{L_2^m}.$$

In the linear case $\tilde{P}(x^1)$ is constant and equals the inverse of the controllability Gramian; see (5).

In a former result [5], the energy functions have been used for the definition of balanced realizations and singular value functions of nonlinear systems. Also, they fulfill corresponding Hamilton–Jacobi equations

$$\frac{\partial L_o(x)}{\partial x} f(x) + \frac{1}{2} h(x)^{\mathrm{T}} h(x) = 0 \tag{27}$$

$$\frac{\partial L_c(x)}{\partial x} f(x) + \frac{1}{2}\frac{\partial L_c(x)}{\partial x} g(x) g(x)^{\mathrm{T}} \frac{\partial L_c(x)}{\partial x}^{\mathrm{T}} = 0 \tag{28}$$

(with $\dot{x} = -f - g\, g^{\mathrm{T}}(\partial L_c/\partial x)^{\mathrm{T}}$ asymptotically stable) in a similar way to the observability Gramian and the inverse of the controllability Gramian are the solutions of the Lyapunov equations; see (6) and (7).

In the following theorem, we review what we mean by *input-normal/output-diagonal form*.

Theorem 4: [5] Consider an operator Σ with an asymptotically stable state–space realization (14). Assume that there exists a neighborhood W of the origin such that $Q(x^0)$ as defined above has a constant number of different eigenvalues. Then there exists a smooth coordinate transformation $x = \Phi(z)$, $\Phi(0) = 0$, on W, which converts Σ into an input-normal/output-diagonal form, where

$$L_c(\Phi(z)) = \frac{1}{2} z^{\mathrm{T}} z \tag{29}$$

$$L_o(\Phi(z)) = z^{\mathrm{T}} \begin{pmatrix} \tau_1(z) & & 0 \\ & \ddots & \\ 0 & & \tau_n(z) \end{pmatrix} z \tag{30}$$

with $\tau_1(z) \geq \ldots \geq \tau_n(z)$ being the so-called smooth singular value functions on W.

The following example exhibits how Theorem 4 works.

Example 1: Let us take the system (14) with $x = (x_1, x_2) \in \mathbb{R}^2$, $u = (u_1, u_2) \in \mathbb{R}^2$ and $y = (y_1, y_2) \in \mathbb{R}^2$ and f, g and h as follows which fulfills the assumptions in Theorem 4; see the equation at the bottom of the page. This system is zero-state observable and asymptotically stable about 0. Solving the Hamilton–Jacobi equations for L_o and L_c in (27) and (28) we obtain the equation shown at the bottom of the next page, on $\mathcal{X}^0 = \mathbb{R}^n$. We see that the controllability function is already in input-normal form and that the observability function is in output-diagonal form. It should be noted that the choice of the singular value functions $\tau_i(z)$'s are not unique and this property was investigated in [6]. The neighborhood W of 0, where the number of distinct singular value functions is constant, is

$$W = \left\{ x \middle| -x_1^4 + 3x_1^2 x_2^2 + 4x_2^4 < 27 \right\}$$

i.e., $\tau_1(x) > \tau_2(x)$ for $\forall x \in W$.

B. Differential Eigenstructure

This section discusses the nonlinear extension of Theorem 1. It concerns the eigenstructure of the operator $\mathcal{H}^* \circ \mathcal{H}$, that is, it is on the solution $\lambda \in \mathbb{R}$ and $v \in L_2^m(0,\infty) \setminus \{0\}$ of a linear equation

$$\mathcal{H}^* \circ \mathcal{H}(v) = \lambda\, v. \tag{31}$$

It is claimed in Theorem 1 that all nonzero solutions of λ are given by $\lambda = \sigma_i^2$, $i \in \{1, 2, \ldots, n\}$ where σ_i's are Hankel singular values.

Recently, in [12], its nonlinear extension based on *nonlinear Hilbert adjoint* as given in Remark 1 was discussed. A simplest nonlinear generalization of (31) with a nonlinear Hilbert adjoint $\mathcal{H}^*(\cdot, \cdot)$ may be

$$\mathcal{H}^*\left(\mathcal{H}(v), v\right) = \lambda\, v.$$

Unfortunately the solution of the previous equation is not found so far. A weaker version of this equation was investigated in [12]. However, this is also insufficient in the sense that the result is coordinate dependent though the above equation does not employ any local coordinates.

To overcome the problem explained above, we consider the eigenstructure of another operator $u \mapsto (\mathrm{d}\mathcal{H}_\Sigma(u))^* \circ \mathcal{H}_\Sigma(u)$ characterized by

$$(\mathrm{d}\mathcal{H}(v))^* \circ \mathcal{H}(v) = \lambda\, v \tag{32}$$

where $\lambda \in \mathbb{R}$ is an eigenvalue and $v \in L_2^m(0,\infty) \setminus \{0\}$ the corresponding eigenvector, which employs a similar idea to [5], [19], and [20]. Note that the operator $u \mapsto (\mathrm{d}\mathcal{H}_\Sigma(u))^* \circ \mathcal{H}_\Sigma(u)$ is *nonlinear* so in general the eigenstructure is different from the linear case. First of all, we prove the fact that this eigenstructure has a close relationship with the Hankel norm of Σ defined by

$$\|\Sigma\|_H := \sup_{\substack{u \in L_2^m(0,\infty) \\ u \neq 0}} \frac{\|\mathcal{H}(u)\|_2}{\|u\|_2}. \tag{33}$$

Theorem 5: Consider an operator Σ with its Hankel operator $\mathcal{H}$. Assume that the Hankel operator $\mathcal{H}$ is continuously differentiable. Let $v \in L_2^m(0,\infty) \setminus \{0\}$ denote the input which achieves the maximization in the definition of Hankel norm in (33), namely

$$\|\Sigma\|_H = \frac{\|\mathcal{H}(v)\|_2}{\|v\|_2}.$$

Then, v satisfies (32) with the eigenvalue $\lambda = \|\Sigma\|_H^2$.

$$f(x) = \begin{pmatrix} -9x_1 + 6x_1^2 x_2 + 6x_2^3 - x_1^5 - 2x_1^3 x_2^2 - x_1 x_2^4 \\ -9x_2 - 6x_1^3 - 6x_1 x_2^2 - x_1^4 x_2 - 2x_1^2 x_2^3 - x_2^5 \end{pmatrix} \quad h(x) = \begin{pmatrix} \frac{2\sqrt{2}\left(3x_1 + x_1^2 x_2 + x_2^3\right)\left(3 - x_1^4 - 2x_1^2 x_2^2 - x_2^4\right)}{1 + x_1^4 + 2x_1^2 x_2^2 + x_2^4} \\ \frac{\sqrt{2}\left(3x_2 - x_1^3 - x_1 x_2^2\right)\left(3 - x_1^4 - 2x_1^2 x_2^2 - x_2^4\right)}{1 + x_1^4 + 2x_1^2 x_2^2 + x_2^4} \end{pmatrix}$$

$$g(x) = \begin{pmatrix} \frac{3\sqrt{2}\left(9 - 6x_1 x_2 + x_1^4 - x_2^4\right)}{9 + x_1^4 + 2x_1^2 x_2^2 + x_2^4} & \frac{\sqrt{2}\left(-9x_1^2 - 27x_2^2 + 6x_1^3 x_2 + 6x_1 x_2^3 - \left(x_1^2 + x_2^2\right)^3\right)}{9 + x_1^4 + 2x_1^2 x_2^2 + x_2^4} \\ \frac{\sqrt{2}\left(27x_1^2 + 9x_2^2 + 6x_1^3 x_2 + 6x_1 x_2^3 + \left(x_1^2 + x_2^2\right)^3\right)}{9 + x_1^4 + 2x_1^2 x_2^2 + x_2^4} & \frac{3\sqrt{2}\left(9 + 6x_1 x_2 - x_1^4 + x_2^4\right)}{9 + x_1^4 + 2x_1^2 x_2^2 + x_2^4} \end{pmatrix}.$$

Proof: The differential of $\|\mathcal{H}(u)\|_2/\|u\|_2$ (in the direction δu) satisfies

$$\begin{aligned}
&d\left(\frac{\|\mathcal{H}(u)\|_2}{\|u\|_2}\right)(\delta u)\\
&=\frac{\|u\|_2 d\left(\|\mathcal{H}(u)\|_2\right)(\delta u)-\|\mathcal{H}(u)\|_2\, d\left(\|u\|_2\right)(\delta u)}{\|u\|_2^2}\\
&=\frac{\frac{\|u\|_2}{\|\mathcal{H}(u)\|_2\langle(\mathrm{d}\mathcal{H}(u))^*\circ\mathcal{H}(u),\delta u\rangle}-\frac{\|\mathcal{H}(u)\|_2}{\|u\|_2\langle u,\delta u\rangle}}{\|u\|_2^2}\\
&=\frac{\left\langle(\mathrm{d}\mathcal{H}(u))^*\circ\mathcal{H}(u)-\left(\frac{\|\mathcal{H}(u)\|_2}{\|u\|_2}\right)^2 u,\delta u\right\rangle}{\|u\|_2\,\|\mathcal{H}(u)\|_2}\\
&\equiv 0
\end{aligned}$$

for all variations δu at $u=v$ because it is a critical point. This reduces to

$$(d\mathcal{H}(v))^*\circ\mathcal{H}(v)\equiv\left(\frac{\|\mathcal{H}(v)\|_2}{\|v\|_2}\right)^2 v=\|\Sigma\|_H^2 v$$

which proves the theorem. ■

Theorem 5 uses the necessary condition for maximization that the differential should be zero at the maximum. For the maximization in the definition of the Hankel norm (33) it is necessary that the input v satisfies (32). Therefore, the eigenstructure (32) is worth investigating and the solutions of (32) derive fruitful results in what follows indeed. Now, we continue to study more precise properties of the differential eigenstructure (32) by assuming the following assumption.

- *Assumption A1:* Suppose that the system Σ in (14) is asymptotically stable about 0, that there exist open neighborhoods $\mathcal{X}^0\subset\mathbb{R}^n$ of 0 and $\mathcal{U}\subset L_2^m(0,\infty)$ of 0 such that the operators $\mathcal{O}:\mathcal{X}^0\to L_2^r(0,\infty)$, $\mathcal{C}:\mathcal{U}\to\mathcal{X}^0$ and $\mathcal{C}^\dagger:\mathcal{X}^0\to\mathcal{U}$ exist and are continuously differentiable.

First, the following lemma gives the complete characterization of the eigenvectors corresponding to nonzero eigenvalues.

Lemma 3: Consider the system Σ in (14). Suppose that Assumption A1 holds. Then a pair $\lambda\in\mathbb{R}\setminus\{0\}$ and $v\in\mathcal{U}\setminus\{0\}$ is a pair of eigenvalues and eigenvectors of the mapping $u\mapsto(\mathrm{d}\mathcal{H}(u))^*\circ\mathcal{H}(u)$ if and only if there exists $x^0\in\mathcal{X}^0\setminus\{0\}$ such that λ and v satisfy

$$\begin{cases}\dot{x}=\frac{\partial H(x,p)}{\partial p}^{\mathrm{T}} & x(0)=x^0,\ x(\infty)=0\\ \dot{p}=-\frac{\partial H(x,p)}{\partial x}^{\mathrm{T}} & p(0)=\frac{1}{\lambda}\frac{\partial L_o}{\partial x}^{\mathrm{T}}(x^0)\\ v=g(x)^{\mathrm{T}}p\end{cases}\tag{34}$$

with the Hamiltonian

$$H(x,p)=-p^{\mathrm{T}}\left(f(x)+\frac{1}{2}g(x)g(x)^{\mathrm{T}}p\right).$$

Proof: Necessity is proven first. Instead of considering the state–space realization of the operator $y_a=(\mathrm{d}\mathcal{H}(u))^*\circ\mathcal{H}(u)$ directly, we use the (24) as in the proof of Lemma 2. Note that both $\mathcal{O}$ and $\mathcal{C}$ are differentiable because of Assumption A1. We can observe

$$y_a=(\mathrm{d}\mathcal{H}(u))^*\circ\mathcal{H}(u)=(\mathrm{d}\mathcal{C}(u))^*\circ(\mathrm{d}\mathcal{O}\left(\mathcal{C}(u)\right))^*\circ\mathcal{O}\circ\mathcal{C}(u).\tag{35}$$

Let $x^0:=\mathcal{C}(u)$, then (35) reduces to

$$y_a=(\mathrm{d}\mathcal{C}(u))^*\circ\left(\mathrm{d}\mathcal{O}(x^0)\right)^*\circ\mathcal{O}(x^0).\tag{36}$$

Next, we consider the Gâteaux differential of $L_o(x^0)$ in the direction ζ

$$\begin{aligned}\frac{\partial L_o(x^0)}{\partial x^0}\zeta&=\mathrm{d}L_o(x^0)(\zeta)=\frac{1}{2}\mathrm{d}\left\|\mathcal{O}(x^0)\right\|_2^2(\zeta)\\&=\left\langle\mathcal{O}(x^0),\mathrm{d}\mathcal{O}(x^0)(\zeta)\right\rangle_{L_2^r}\\&=\left\langle\left(\mathrm{d}\mathcal{O}(x^0)\right)^*\circ\mathcal{O}(x^0),\zeta\right\rangle_{\mathbb{R}^n}.\end{aligned}$$

Note that Assumption A1 implies that $x^0\in\mathcal{X}^0$ and and that $L_o(x)$ is differentiable on $\mathcal{X}^0$. This means that

$$\mathrm{d}\mathcal{O}(x^0))^*\circ\mathcal{O}(x^0)=\frac{\partial L_o(x^0)}{\partial x^0}^{\mathrm{T}}.$$

Hence, from (36) it follows that

$$y_a=(\mathrm{d}\mathcal{C}(u))^*\left(\frac{\partial L_o(x^0)}{\partial x^0}^{\mathrm{T}}\right).$$

It follows from Lemma 2 that the state–space realization of this operator is given by

$$\begin{cases}\dot{x}=f(x)+g(x)\mathcal{F}_-(u) & x(-\infty)=0\\ \dot{p}=-\frac{\partial(f+g\mathcal{F}_-(u))}{\partial x}^{\mathrm{T}}(x)p & p(0)=\frac{\partial L_o}{\partial x}^{\mathrm{T}}(x^0)\\ y_a=\mathcal{F}_+\left(g(x)^{\mathrm{T}}p\right)\end{cases}.$$

If we consider the reverse-time expression of this system (with x and p now representing the reverse time state variables) given by

$$\begin{cases}\dot{x}=-f(x)-g(x)u & x(0)=x^0\,(x(\infty)=0)\\ \dot{p}=\frac{\partial(f+gu)}{\partial x}^{\mathrm{T}}(x)p & p(0)=\frac{\partial L_o}{\partial x}^{\mathrm{T}}(x^0)\\ y_a=g(x)^{\mathrm{T}}p\end{cases}.$$

$$\begin{aligned}L_c(x)&=\frac{1}{2}x^Tx\\ L_o(x)&=\frac{1}{2}\frac{36x_1^2+9x_2^2+18x_1^3x_2+18x_1x_2^3+x_1^6+6x_1^4x_2^2+9x_1^2x_2^4+4x_2^6}{1+x_1^4+2x_1^2x_2^2+x_2^4}\\ &=\frac{1}{2}x^{\mathrm{T}}\begin{pmatrix}\frac{36+18x_1x_2+x_1^4+6x_1^2x_2^2}{1+x_1^4+2x_1^2x_2^2+x_2^4} & 0\\ 0 & \frac{9+18x_1x_2+9x_1^2x_2^2+4x_2^4}{1+x_1^4+2x_1^2x_2^2+x_2^4}\end{pmatrix}x=:\frac{1}{2}x^{\mathrm{T}}\begin{pmatrix}\tau_1(x) & 0\\ 0 & \tau_2(x)\end{pmatrix}x\end{aligned}$$

Now, we have the causal state–space expression of the operator $y_a = (\mathrm{d}\mathcal{H}(u))^* \circ \mathcal{H}(u)$ using x^0. Suppose $u = v$ and $y_a = \lambda v$ hold, i.e., the pair λ and v is the pair of eigenvalue and eigenvector. Then, we have

$$v = \frac{1}{\lambda} y_a = \frac{1}{\lambda} g(x)^{\mathrm{T}} p.$$

Let $\bar{p} := (1/\lambda)p$, then

$$\begin{cases} \dot{x} = -f(x) - g(x)g(x)^{\mathrm{T}}\bar{p} & x(0) = x^0 \\ \dot{\bar{p}} = \frac{1}{\lambda}\dot{p} & \\ \quad = \frac{1}{\lambda}\frac{\partial\left(f + \left(\frac{1}{2}\right) gg^{\mathrm{T}}\bar{p}\right)}{\partial x}^{\mathrm{T}} p & \\ \quad = \frac{\partial\left(f + \left(\frac{1}{2}\right) gg^{\mathrm{T}}\bar{p}\right)}{\partial x}^{\mathrm{T}} \bar{p} & \bar{p}(0) = \frac{1}{\lambda}\frac{\partial L_o}{\partial x}^{\mathrm{T}}(x^0) \\ v = g(x)^{\mathrm{T}}\bar{p} & \end{cases}.$$

Sufficiency follows straightforwardly from the converse arguments. This completes the proof. ■

Lemma 3 gives the characterization of all pairs of eigenvalues and eigenvectors. Next, we concentrate on a special class of these pairs. They are closely related to the energy functions $L_o(x)$ and $L_c(x)$.

Lemma 4: Consider the system Σ in (14). Suppose that Assumption A1 holds and that there exist $\lambda \in \mathbb{R} \setminus \{0\}$ and $x^0 \in \mathcal{X}^0 \setminus \{0\}$ satisfying

$$\frac{\partial L_o}{\partial x}(x^0) = \lambda \frac{\partial L_c}{\partial x}(x^0). \tag{37}$$

Then, λ is an eigenvalue of the mapping $u \mapsto (\mathrm{d}\mathcal{H}(u))^* \circ \mathcal{H}(u)$ corresponding to the eigenvector

$$v = \mathcal{C}^\dagger(x^0).$$

Proof: For Assumption A1, the operator $\mathcal{C}^\dagger$ exists. Its state–space realization of is given in [18] as

$$\mathcal{C}^\dagger : x^0 \mapsto y_a \begin{cases} \dot{x} = -f(x) - g(x)g(x)^{\mathrm{T}}\frac{\partial L_c}{\partial x}^{\mathrm{T}} & x(0) = x^0 \\ y_a = g(x)^{\mathrm{T}}\frac{\partial L_c}{\partial x}^{\mathrm{T}} & \end{cases}.$$

If (37) holds, then

$$p(0) = \frac{1}{\lambda}\frac{\partial L_o}{\partial x}^{\mathrm{T}}(x^0) = \frac{\partial L_c}{\partial x}^{\mathrm{T}}(x^0).$$

Combined with the dynamics for p in (34), with the Hamilton–Jacobi–Bellman equation for L_c in (28), and with Lemma 2, this implies that

$$p(t) \equiv \frac{\partial L_c}{\partial x}^{\mathrm{T}}(x(t))$$

holds along the trajectory of $(\mathrm{d}\mathcal{H}(u))^* \circ \mathcal{H}(u)$. This concludes the proof. ■

Although Lemma 4 gives a sufficient condition for pairs of eigenvalues and eigenvectors in terms of the energy functions, it does not say anything about the existence of pairs x_0 and λ that fulfill (37). We now continue to investigate the necessity of the condition. In order to proceed, we define two scalar functions

$$\rho_{\max}(c) := \sup_{\substack{u \in \mathcal{C}^\dagger(\mathcal{X}^0) \\ \|u\|_2 = c}} \frac{\|\mathcal{H}(u)\|_2}{\|u\|_2} = \sup_{\substack{u \in \mathcal{U} \\ \|u\|_2 = c}} \frac{\|\mathcal{H}(u)\|_2}{\|u\|_2} \tag{38}$$

$$\rho_{\min}(c) := \inf_{\substack{u \in \mathcal{C}^\dagger(\mathcal{X}^0) \\ \|u\|_2 = c}} \frac{\|\mathcal{H}(u)\|_2}{\|u\|_2} \tag{39}$$

$\rho_{\max}$ is closely related to the Hankel norm because if $\mathcal{U} = L_2^m(0,\infty)$, then

$$\|\Sigma\|_H = \sup_{\substack{u \in L_2^m(0,\infty) \\ u \neq 0}} \frac{\|\mathcal{H}(u)\|_2}{\|u\|_2} = \sup_{c > 0} \rho_{\max}(c)$$

i.e., $\rho_{\max}(c)$ represents the Hankel norm under the fixed input magnitude $\|u\|_2 = c$. In fact, this is a natural nonlinear generalization of the property (8) in the linear case. For $\rho_{\min}(c)$, we do not have such an interpretation. Furthermore, $\rho_{\max}$ and $\rho_{\min}$ equal the maximum and minimum Hankel singular value, respectively, in the linear case. Therefore, these functions can be seen as an alternative nonlinear extension of Hankel singular values other than the singular value functions $\tau_i(z)$, $i = 1, \ldots, n$ in Theorem 4. Now, the results on the necessity is stated.

Theorem 6: Consider the system Σ in (14). Suppose that Assumption A1 holds. Let $v_{\max}(c)$ and $v_{\min}(c)$ denote the inputs which achieve the maximization and minimization under an arbitrary input magnitude $\|u\|_2 = c$, in the definition of $\rho_{\max}$ and $\rho_{\min}$ in (38) and (39), respectively, i.e., they satisfy

$$c = \|v_i(c)\|_2 \tag{40}$$

$$\rho_i(c) = \frac{\|\mathcal{H}(v_i(c))\|_2}{\|v_i(c)\|_2} \tag{41}$$

for $i \in \{\max, \min\}$. Then, $v_{\max}(c)$ and $v_{\min}(c)$ are the eigenvectors of $u \mapsto (\mathrm{d}\mathcal{H}(u))^* \circ \mathcal{H}(u)$ with respect to the following eigenvalues $\lambda_{\max}(c)$ and $\lambda_{\min}(c)$, respectively

$$\lambda_i(c) = \rho_i^2(c) + \frac{c}{2}\frac{\mathrm{d}\rho_i^2(c)}{\mathrm{d}c}, \qquad i = \{\max, \min\}. \tag{42}$$

Furthermore, both pairs $(\lambda, x^0) = (\lambda_{\max}(c), \mathcal{C}(v_{\max}(c)))$ and $(\lambda_{\min}(c), \mathcal{C}(v_{\min}(c)))$ satisfy (37) in Lemma 4.

Proof: Let i denote the index such that $i \in \{\max, \min\}$. Firstly we define $\xi_i(c) := \mathcal{C}(v_i(c))$ and show the existence of $\lambda_i(c)$ such that

$$\frac{\partial L_o}{\partial x}(\xi_i(c)) = \lambda_i(c)\frac{\partial L_c}{\partial x}(\xi_i(c)). \tag{43}$$

To this effect, let the level set of $L_c(x)$ be given by

$$\mathcal{X}_{L_c = k} := \{x | L_c(x) = k\}.$$

Then, $\xi_i(c) \in \mathcal{X}_{L_c = (c^2/2)}$ follows from the fact that $v_i(c)$ is the input which minimizes the input energy. Indeed $\xi_i(c) \in \mathcal{X}_{L_c = (c^2/2)}$ denotes the set of the states derived by the input $\|u\|_2 = c$. Consider a curve $\eta(s) \in \mathcal{X}_{L_c = (c^2/2)}$ parameterized by a scalar variable s such that $\eta(0) = \xi_i(c)$ holds. Since $\eta(s)$ is contained in the level set $\mathcal{X}_{L_c=(c^2/2)}$

$$\frac{\mathrm{d}L_c(\eta(s))}{\mathrm{d}s} = \frac{\partial L_c(\eta)}{\partial \eta}\frac{\mathrm{d}\eta(s)}{\mathrm{d}s} = 0 \tag{44}$$

holds along $\eta(s)$. Next, from the definition, we can observe the following relations:

$$\sup_{\substack{u\in\mathcal{C}^\dagger(\mathbf{R}^n)\\ \|u\|_2=c}} \frac{\|\mathcal{H}(u)\|_2}{\|u\|_2} = \sup_{x\in\mathcal{X}_{L_c=\frac{c^2}{2}}} \sqrt{\frac{L_o(x)}{L_c(x)}}$$

$$\inf_{\substack{u\in\mathcal{C}^\dagger(\mathbf{R}^n)\\ \|u\|_2=c}} \frac{\|\mathcal{H}(u)\|_2}{\|u\|_2} = \inf_{x\in\mathcal{X}_{L_c=\frac{c^2}{2}}} \sqrt{\frac{L_o(x)}{L_c(x)}}.$$

This implies that $\eta(0) = \xi_i(c)$ maximizes (minimizes) the value (L_o/L_c) in the level set $\mathcal{X}_{L_c=(c^2/2)}$. Therefore, we obtain

$$\begin{aligned} \left.\frac{\mathrm{d}\frac{L_o(\eta(s))}{L_c(\eta(s))}}{\mathrm{d}s}\right|_{s=0} &= \frac{2}{c^2}\left.\frac{\mathrm{d}L_o(\eta(s))}{\mathrm{d}s}\right|_{s=0} \\ &= \frac{2}{c^2}\left.\frac{\partial L_o(\eta)}{\partial \eta}\frac{\mathrm{d}\eta(s)}{\mathrm{d}s}\right|_{s=0} \\ &= 0. \end{aligned} \tag{45}$$

Equations (44) and (45) have to hold for all curves $\eta(s) \in \mathcal{X}_{L_c=(c^2/2)}$. Namely both $(\partial L_o/\partial x)$ and $(\partial L_c/\partial x)$ are orthogonal to the tangent space of $\mathcal{X}_{L_c=(c^2/2)}$ at $x = \xi_i(c)$. Because this tangent space is $(n-1)$-dimensional, we can conclude $(\partial L_o/\partial x)$ and $(\partial L_c/\partial x)$ are linearly dependent at $x = \xi_i(c)$. Therefore, there exists a scalar constant $\lambda_i(c)$ such that (43) holds. Remember that $v_i(c)$ can be described by $v_i(c) = \mathcal{C}^\dagger(\xi_i(c))$. Then it follows directly from Lemma 4 that $v_i(c) = \mathcal{C}^\dagger(\xi_i(c))$ is the eigenvector of $u \mapsto (\mathrm{d}\mathcal{H}(u))^* \circ \mathcal{H}(u)$ with respect to $\lambda_i(c)$.

Second, we prove (42). By the previous discussion, for any vector $\zeta \in \mathbb{R}^n$ which is not orthogonal to the tangent space of $\mathcal{X}_{L_c=(c^2/2)}$ at $x = \xi_i(c)$, $\lambda_i(c)$ can be expressed as

$$\lambda_i(c) = \frac{\frac{\partial L_o}{\partial x}\zeta}{\frac{\partial L_c}{\partial x}\zeta}.$$

Let ζ be the directional differential $(\mathrm{d}\varsigma(s)/\mathrm{d}s)$ of another curve $\varsigma(s) = x_i^s$ passing through the maximizing (minimizing) state. Namely, it goes across the level set $\mathcal{X}_{L_c=(c^2/2)}$ through $x = \xi_i(c)$ and $\varsigma(c) = \xi_i(c)$ holds. Then, we can obtain

$$\begin{aligned} \lambda_i(c) &= \left.\frac{\frac{\partial L_o(\varsigma)}{\partial\varsigma}\frac{\mathrm{d}\varsigma(s)}{\mathrm{d}s}}{\frac{\partial L_c(\varsigma)}{\partial\varsigma}\frac{\mathrm{d}\varsigma(s)}{\mathrm{d}s}}\right|_{s=c} = \left.\frac{\frac{\mathrm{d}L_o(\varsigma(s))}{\mathrm{d}s}}{\frac{\mathrm{d}L_c(\varsigma(s))}{\mathrm{d}s}}\right|_{s=c} = \frac{\frac{\mathrm{d}\left(\rho_i^2(c)\frac{c^2}{2}\right)}{\mathrm{d}c}}{\frac{\mathrm{d}\left(\frac{c^2}{2}\right)}{\mathrm{d}c}} \\ &= \frac{c\rho_i^2(c) + \frac{c^2}{2}\frac{\mathrm{d}\rho_i^2(c)}{\mathrm{d}c}}{c} = \rho_i^2(c) + \frac{c}{2}\frac{\mathrm{d}\rho_i^2(c)}{\mathrm{d}c} \end{aligned}$$

because of the definition of $\rho_i(c)$ (38) and (39). This proves the theorem. ∎

The above result only focuses on the maximum and minimum values on the level sets. For linear systems this results in the maximum and minimum Hankel singular values. In that case, by "ruling" out these directions, and further using similar arguments, the other Hankel singular values can also be obtained. By using a similar method we can extend this result for *axis singular value functions* $\rho_i(s)$'s, $i \in \{1, 2, \ldots, n\}$ which are mappings from $\mathbb{R} \to \mathbb{R}$, under the following assumption.

- *Assumption A2:* Suppose that the Hankel singular values of the Jacobian linearization of the system Σ are nonzero and distinct.

Theorem 7: Consider the system Σ in (14). Suppose that Assumptions A1 and A2 hold. Then there exists a neighborhood $U \subset \mathbb{R}$ of 0, n smooth functions $\rho_i : U \to (0, \infty)$'s, $i \in \{1, 2, \ldots, n\}$ such that

$$\min\{\rho_i(s), \rho_i(-s)\} \geq \max\{\rho_{i+1}(s), \rho_{i+1}(-s)\} \tag{46}$$

holds for all $s \in U$ and $\forall i \in \{1, 2, \ldots, n-1\}$ and that there exist n distinct smooth curves $\xi_i : U \to \mathbb{R}^n$ satisfying $\xi_i(0) = 0$ and

$$L_c(\xi_i(s)) = \frac{s^2}{2} \quad L_o(\xi_i(s)) = \frac{\rho_i^2(s)s^2}{2} \tag{47}$$

$$\frac{\partial L_o}{\partial x}(\xi_i(s)) = \lambda_i(s)\frac{\partial L_c}{\partial x}(\xi_i(s)) \tag{48}$$

with

$$\lambda_i(s) := \rho_i^2(s) + \frac{s}{2}\frac{\mathrm{d}\rho_i^2(s)}{\mathrm{d}s}. \tag{49}$$

Furthermore, $\max\{\rho_1(s), \rho_1(-s)\}$ and $\min\{\rho_n(s), \rho_n(-s)\}$ coincide with $\rho_{\max}(s)$ and $\rho_{\min}(s)$ respectively for all $s \geq 0$. In particular, if $U = \mathbb{R}$, then

$$\|\Sigma\|_H = \sup_{s\in\mathbb{R}} \rho_1(s). \tag{50}$$

Proof: Suppose the state–space realization is in input-normal form. Consider (37). For the smoothness of $\partial L_o/\partial x$, there exists a smooth matrix valued function $Q : \mathbb{R}^n \to \mathbb{R}^{n\times n}$ such that

$$\frac{\partial L_o}{\partial x} = x^{\mathrm{T}}Q(x)^{\mathrm{T}}.$$

Assumption A2 that the Jacobian linearization of the system has n nonzero distinct Hankel singular values implies that there exists a neighborhood of 0 on which $Q(x)$ is decomposed as

$$W(x)^{-1}Q(x)W(x) = \mathrm{diag}\,(q_1(x), q_2(x), \ldots, q_n(x))$$

where $q_1 \geq q_2 \geq \ldots \geq q_n$ and where $W : \mathbb{R}^n \to \mathbb{R}^{n\times n}$ is a nonsingular smooth matrix valued function with $w_i : \mathbb{R}^n \to \mathbb{R}^n$'s

$$W(x) = (w_1(x), w_2(x), \ldots, w_n(x)).$$

Hence, (37) reduces to

$$W(x)\mathrm{diag}\,(q_1(x), q_2(x), \ldots, q_n(x))\,W^{-1}(x)x = \lambda x. \tag{51}$$

Consider the generalized sphere $S_c^{n-1} := \{\, x \mid \|x\| = c\}$ and mappings $\tilde{w}_i : S_c^{n-1} \to S_c^{n-1}$'s defined by

$$\tilde{w}_i : x \mapsto \frac{c}{\|w_i(x)\|}w_i(x). \tag{52}$$

Then, there exists a neighborhood $U \subset \mathcal{X}^0$ of 0 on which, for sufficiently small $c > 0$, we can choose $2n$ closed sets[2] $X_i^j \subset S_r^{n-1}$'s ($i \in \{1, 2, \ldots, n\}$, $j \in \{+, -\}$) which are homeomorphic to the $n-1$-dimensional unit disc $D^{n-1} = \{x \in \mathbb{R}^{n-1} \mid \|x\| \leq 1\}$ satisfying

$$x \in X_i^j \Rightarrow \tilde{w}_i(x) \in X_i^j.$$

This results from the fact that $w_i(0)$ equals the eigenvector of the observability Gramian of the Jacobian linearization of the original system and $w_i(x)$ is smooth in a neighborhood of the origin. (The image $\tilde{w}_i(S_c^{n-1})$ can be chosen small enough by choosing a sufficiently small $c > 0$.) Then, it follows from Brouwer's fixed point theorem (see, e.g., [24]) that the mapping $\tilde{w}_i$ in (52) has a fixed point, i.e. $\exists \xi_i^j \in X_i^j$ s.t.

$$\xi_i^j = \tilde{w}_i\left(\xi_i^j\right) = \frac{c}{\left\|w_i\left(\xi_i^j\right)\right\|} w_i\left(\xi_i^j\right).$$

Then, it can be easily checked that (51) holds with the eigenvectors $x = \xi_i^j$'s and the eigenvalues $\lambda = q_i(\xi_i^j)$'s. Finally, define

$$\xi_i(s) := \begin{cases} \xi_i^+ \in S_s^{n-1}, & (s \geq 0) \\ \xi_i^- \in S_{-s}^{n-1}, & (s < 0) \end{cases}.$$

Then, (48) holds. Property (47) can be proven in a way similar to the proof of Theorem 6. Furthermore, the ordering $\min\{\rho_i(s), \rho_i(-s)\} \geq \max\{\rho_{i+1}(s), \rho_{i+1}(-s)\}$ follows from the fact that $\rho_i(0) = \sigma_i$ holds with the Hankel singular value σ_i of the Jacobian linearization of the system. This completes the proof. ■

Notice that the scalar variable s of $\rho_i(s)$'s, $i \in \{1, 2, \ldots, n\}$ can be negative, whereas the variable c of $\rho_j(c)$'s, $j \in \{\max, \min\}$ is nonnegative, since it represents the input energy level. Both $\rho_i^2(c)$ and $\rho_i^2(-c)$ denote the ratio L_o/L_c with respect to the prescribed input energy $L_c = (1/2)c^2$. The eigenstructure given in Theorems 6 and 7 is particularly important because it is closely related to the Hankel norm and the corresponding axis singular value functions. Indeed the axis singular value functions represent the gain of the Hankel operator at the eigenvector $u = \mathcal{C}^\dagger(\xi_i(s))$ as in (41), i.e.,

$$\rho_i(s) = \left.\frac{\|\mathcal{H}(u)\|}{\|u\|}\right|_{u=\mathcal{C}^\dagger(\xi_i(s))} \tag{53}$$

holds. By its definition, the eigenvector $u = \mathcal{C}^\dagger(\xi_i(s))$ represents the stationary point of this gain which follows from the same arguments to the proof of Theorem 5. Furthermore, it should be noted that Theorems 6 and 7 give an input–output characterization of the Hankel operator without using any local coordinates, that is, they are coordinate free.

For linear systems, we have that $\lambda_i(s) = \rho_i^2(s) = \sigma_i^2$, $i = 1, \ldots, n$ where σ_i's are the Hankel singular values. Indeed, (37) reduces to

$$x^{0^{\mathrm{T}}} QP = \lambda x^{0^{\mathrm{T}}}$$

[2] $2n$ is equivalent to the number of the axis intersecting the generalized sphere S_c^{n-1}.

with observability and controllability Gramians Q and P, respectively. This equation implies that the result obtained here is a natural nonlinear extension of the linear case result in Theorem 1. The effectiveness of Theorem 7 is demonstrated in the following example.

Example 2: Consider the state–space system in the form of (14) which is in an input-normal/output-diagonal form, as given in Example 1. In order to obtain the $\xi_i(s)$'s we have to compute the solution of (47) and (48), which reduces down to

$$s^2 = 2L_c(x) = x_1^2 + x_2^2 \tag{54}$$

$$\begin{aligned} 0 &= \det\begin{pmatrix} \frac{\partial L_c}{\partial x} \\ \frac{\partial L_o}{\partial x} \end{pmatrix} \\ &= \frac{-27x_1x_2 + 9x_1^4 - 9x_2^4 + 3x_1^5x_2 + 6x_1^3x_2^3 + 3x_1x_2^5}{\left(1 + x_1^4 + 2x_1^2x_2^2 + x_2^4\right)^2}. \end{aligned} \tag{55}$$

Here, the second equation follows from the fact that $\partial L_c/\partial x$ is parallel to $\partial L_o/\partial x$ and that x is two-dimensional. These equations have the following two solutions which can be obtained by a standard CAD such as Maple:

$$\xi_1(s) = \begin{pmatrix} \frac{3s}{\sqrt{9+s^4}} \\ \frac{s^3}{\sqrt{9+s^4}} \end{pmatrix} \quad \xi_2(s) = \begin{pmatrix} -\frac{s^3}{\sqrt{9+s^4}} \\ \frac{3s}{\sqrt{9+s^4}} \end{pmatrix}. \tag{56}$$

For (47), the axis singular value functions $\rho_i(s)$'s can be obtained by a direct calculation

$$\rho_1(s) = \sqrt{\frac{L_o(\xi_1(s))}{L_c(\xi_1(s))}} = 2\sqrt{\frac{9 + s^4}{1 + s^4}} \tag{57}$$

$$\rho_2(s) = \sqrt{\frac{L_o(\xi_2(s))}{L_c(\xi_2(s))}} = \sqrt{\frac{9 + s^4}{1 + s^4}}. \tag{58}$$

Notice that both functions $\xi_i(s)$'s and $\rho_i(s)$'s are defined for $\forall s \in \mathbb{R}$. We can easily check that the $\lambda_i(s)$'s given by (49) satisfy (48). Furthermore, it can be observed that

$$\begin{aligned} \min\{\rho_1(s), \rho_1(-s)\} &= 2\sqrt{\frac{9 + s^4}{1 + s^4}} \\ &> \sqrt{\frac{9 + s^4}{1 + s^4}} = \max\{\rho_2(s), \rho_2(-s)\} \end{aligned}$$

holds for $\forall s \in \mathbb{R}$. This implies that the relation (46) holds on $U = \mathbb{R}$. Therefore, it follows from (50) that

$$\|\Sigma\|_H = \sup_{s \in \mathbb{R}} \rho_1(s) = 6.$$

Thus, Theorem 7 is a powerful tool in investigating the gain structure of the Hankel operator and also it is a natural nonlinear extension of the linear case result Theorem 1.

VI. Input-Normal/Output-Diagonal Realizations

The previous section introduced *axis singular value functions* $\rho_i(s)$'s which are alternative nonlinear generalizations of Hankel singular values in the linear case. Furthermore, it was shown in Theorem 7 that those functions are closely related to the curves on the state–space $\xi_i(s)$'s. This section utilizes Theorem 7 to derive a new characterization of the input-normal/output-diagonal realization. More precisely, it will be proven that there exists an input-normal/output-diagonal realization whose (conventional) singular value functions $\tau_i(z)$'s coincide with the axis singular value functions $\rho_i(s)$'s on the coordinate axes.

To this end, at first let us consider a coordinate transformation which converts the curves $x = \xi_i(s)$'s into the coordinate axes z_i's.

Lemma 5: Suppose that Assumptions A1 and A2 hold. Then there exists a neighborhood of the origin U and a coordinate transformation $x = \Phi(z)$ on U satisfying

$$L_c(\Phi(z)) = \frac{1}{2} z^{\mathrm{T}} z \tag{59}$$

$$\Phi(0, \ldots, 0, z_i, 0, \ldots, 0) = \xi_i(z_i) \tag{60}$$

i.e., the system is described by an input-normal form on the coordinate z and the condition (37) holds on the coordinate axes z_i's.

Proof: It is noted that the case $n = 1$ is trivial from Theorem 4. Hence, $n \geq 2$ is assumed in what follows. It is also assumed that without loss of generality that the system is already in an input-normal form. Let us consider the generalized polar coordinates

$$\begin{pmatrix} r(x) \\ \theta_1(x) \\ \theta_2(x) \\ \vdots \\ \theta_{n-1}(x) \end{pmatrix} = \begin{pmatrix} \left(x_1^2 + x_2^2 + \ldots + x_n^2\right)^{\frac{1}{2}} \\ \operatorname{atan2}(x_2, x_1) \\ \operatorname{atan2}\left(x_3, \left(x_1^2 + x_2^2\right)^{\frac{1}{2}}\right) \\ \vdots \\ \operatorname{atan2}\left(x_n, \left(x_1^2 + x_2^2 + \ldots + x_{n-1}^2\right)^{\frac{1}{2}}\right) \end{pmatrix}.$$

Here, $\operatorname{atan2} : \mathbb{R}^2 \to (-\pi, \pi] \subset \mathbb{R}$ denotes the function defined by $\operatorname{atan2}(y, x) := \arg(x + y\Im)$ with $\Im$ the imaginary unit. It satisfies $\operatorname{atan2}(y, x) = \arctan(y/x)$ for all $x > 0$. On these coordinates, consider a rotational matrix $R(\varphi, \theta) \in \mathbb{R}^{n \times n}$ with $\theta := (\theta_1, \theta_2, \ldots, \theta_{n-1})$, $\varphi := (\varphi_1, \varphi_2, \ldots, \varphi_{n-1}) \in \mathbb{R}^{n-1}$ changing the generalized polar coordinate θ into φ defined by

$$\begin{aligned} R(\varphi, \theta) &:= R_{n-1}(\varphi_{n-1}) R_{n-2}(\varphi_{n-2}) \cdots R_1(\varphi_1) \\ &\quad \times R_1(-\theta_1) \cdots R_{n-2}(-\theta_{n-2}) R_{n-1}(-\theta_{n-1}) \in \mathbb{R}^{n \times n} \end{aligned} \tag{61}$$

with $R_i(\theta_i)$'s the rotation matrices for the component angles θ_i, $i = 1, \ldots, n$ defined by

$$R_1(\theta_1) := \begin{pmatrix} 1 & 0 & 0 & & & 0 \\ 0 & \cos\theta_1 & -\sin\theta_1 & & & \\ 0 & \sin\theta_1 & \cos\theta_1 & & & \\ & & & 1 & & 0 \\ & & & & \ddots & \\ 0 & & & 0 & & 1 \end{pmatrix} \in \mathbb{R}^{n \times n}$$

$$R_2(\theta_2) := \begin{pmatrix} 1 & 0 & 0 & 0 & & & 0 \\ 0 & \cos\theta_2 & 0 & -\sin\theta_2 & & & \\ 0 & 0 & 1 & 0 & & & \\ 0 & \sin\theta_2 & 0 & \cos\theta_2 & & & \\ & & & & 1 & & 0 \\ & & & & & \ddots & \\ 0 & & & & 0 & & 1 \end{pmatrix} \in \mathbb{R}^{n \times n}$$

and so on. Here, let us define a coordinate transformation

$$z = R(\varphi(x), \theta(x))\, x =: \Phi^{-1}(x) \tag{62}$$

where $\varphi(x) \in \mathbb{R}^{n-1}$ will be defined later on. Note that this coordinate transformation converts (r, θ) into (r, φ). It can be readily observed that it satisfies the isometric property (59) because the rotation matrix R is unitary.

Next, we need to prove the second property (60). To this end, a generalized sphere

$$S_r^{n-1} := \{x \mid \|x\| = r, x \in \mathbb{R}^n\}$$

is considered. The $(n-1)$-dimensional vector $\theta = (\theta_1, \theta_2, \ldots, \theta_n)$ can be regarded as the coordinates of S_r^{n-1}. What we have to do is to find a coordinate transformation $\varphi = \Psi(\theta)$ with $\Psi : S_r^{n-1} \twoheadrightarrow S_r^{n-1}$ satisfying

$$\begin{aligned} \Psi\left(\theta_{r,1}^{+}\right) &= \varphi_{r,1}^{+} := (0, 0, \ldots, 0) \\ \Psi\left(\theta_{r,1}^{-}\right) &= \varphi_{r,1}^{-} := (\pi, 0, \ldots, 0) \\ \Psi\left(\theta_{r,i}^{+}\right) &= \varphi_{r,i}^{+} := \left(0, \ldots, 0, \underbrace{\left(\frac{1}{2}\right)\pi}_{(i-1)-\mathrm{th}}, 0, \ldots, 0\right) \\ \Psi\left(\theta_{r,i}^{-}\right) &= \varphi_{r,i}^{-} := \left(0, \ldots, 0, \underbrace{-\left(\frac{1}{2}\right)\pi}_{(i-1)-\mathrm{th}}, 0, \ldots, 0\right) \end{aligned} \tag{63}$$

with $\theta_{r,i}^{+} := \theta(\xi_i(+r))$ and $\theta_{r,i}^{-} := \theta(\xi_i(-r))$. Since Assumption A2 guarantees that each curve ξ_i coincides with the coordinate axis z_i at the origin, for a sufficiently small $r > 0$, there exists $2(n-1)$ open sets $U_{r,i}^{j} \subset S_r^{n-1}$, $i \in \{1, 2, \ldots, n-1\}$, $j \in \{+, -\}$ containing both $\theta_{r,i}^{j}$ and $\varphi_{r,i}^{j}$ in such a way that $U_{r,i}^{j}$'s are separated from each other. We can find $2(n-1)$ subsets $N_{r,i}^{j} \supset U_{r,i}^{j}$ separated from each other, since $U_{r,i}^{j}$'s are open. Then, Lemma 7 in the Appendix implies that there exist diffeomorphism $\Psi_{r,i}^{j}$'s on $N_{r,i}^{j}$'s which coincide with the identity on $N_{r,i}^{j} \setminus U_{r,i}^{j}$. Therefore, the diffeomorphism Ψ defined by

$$\varphi = \Psi(\theta) := \begin{cases} \Psi_{r,i}^{j}(\theta) & \left(\theta \in U_{r,i}^{j}\right) \\ \theta & \left(\theta \notin U_{r,i}^{j}\right) \end{cases} \tag{64}$$

satisfies the constraints (63). Finally, the coordinate transformation in (62) is given by

$$z = R\left(\Psi\left(\theta(x)\right), \theta(x)\right) x =: \Phi^{-1}(x)$$

with Ψ defined in (64). Obviously, this coordinate transformation satisfies (60) for the property (63) which completes the proof. ∎

Lemma 5 provides *balanced* coordinates in the sense that the coordinate axis z_1 plays the biggest role in the input-output behavior from the view point of Hankel norm. Unfortunately, this state–space realization is not input-normal/output-diagonal and has no relationship to the former result given in Theorem

4. However, if we look carefully at the coordinate transformation in Lemma 5 and apply it recursively, then we can obtain a unified realization which has both of the properties in Theorem 4 and Lemma 5 as follows. That is, there exists an input-normal/output-diagonal realization whose (conventional) singular value functions $\tau_i(z)$'s coincide with the *axis singular value functions* $\rho_i(z_i)$'s.

Theorem 8: Consider the system Σ in (14). Suppose that Assumptions A1 and A2 hold. Then there exist a neighborhood U of 0 and a coordinate transformation $x = \Phi(z)$ on U converting the system into an input-normal form (29) and (30) satisfying the following properties:

$$z_i = 0 \Leftrightarrow \frac{\partial L_c(\Phi(z))}{\partial z_i} = 0 \Leftrightarrow \frac{\partial L_o(\Phi(z))}{\partial z_i} = 0 \tag{65}$$

$$\tau_i(0,\ldots,0,\underbrace{z_i}_{i-\text{th}},0,\ldots,0) = \rho_i^2(z_i) \tag{66}$$

$$\left.\frac{\partial \tau_i}{\partial z}\right|_{z=(0,\ldots,0,\underbrace{z_i}_{i-\text{th}},0,\ldots,0)} = \left(0,\ldots,0,\underbrace{\frac{\mathrm{d}\rho_i^2(z_i)}{\mathrm{d}z_i}}_{i-\text{th}},0,\ldots,0\right) \tag{67}$$

hold for all $i \in \{1,2,\ldots,n\}$. In particular, if $U = \mathbb{R}^n$, then

$$\|\Sigma\|_H = \sup_{z_1\in\mathbb{R}} \sqrt{\tau_1(z_1,0,\ldots,0)}. \tag{68}$$

The proof of this theorem is very long and tedious, and breaks down into several steps. Equation (65) is proven by induction with respect to the dimension n. Then, the output-diagonal form in the z coordinates is proven. The proof can be found in Appendix. It should be noted that the proof is constructive and it actually gives a procedure to obtain the new input-normal/output-diagonal realization.

In Theorem 8, the existence of an input-normal/output-diagonal form is proven so that (65) and the properties in Theorem 7 hold along each coordinate axis. The relationship (65) is a stronger version of the (37) which can be achieved by applying Theorem 7 repetitively. This property is quite important because the input-normal/output-diagonal structure is preserved under the projection to lower dimensional subspaces spanned by each coordinate axis $\{z_1, z_2, \ldots\}$ and this will play an important role in model reduction of nonlinear systems. We illustrate our final result in the following example.

Example 3: Consider the state–space system in the form of (14) as given in Examples 1 and 2 again. Equations in (56) imply that the coordinate transformation $x = \Phi(z)$ is given by

$$x = \Phi(z) = \begin{pmatrix} \frac{3}{\sqrt{9+\left(z_1^2+z_2^2\right)^2}} & \frac{-\left(z_1^2+z_2^2\right)}{\sqrt{9+\left(z_1^2+z_2^2\right)^2}} \\ \frac{\left(z_1^2+z_2^2\right)}{\sqrt{9+\left(z_1^2+z_2^2\right)^2}} & \frac{3}{\sqrt{9+\left(z_1^2+z_2^2\right)^2}} \end{pmatrix} \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} \tag{69}$$

in the form of the rotation coordinate transformation (61), which maps the z_i-axis into ξ_i, i.e.,

$$\Phi(s,0) = \xi_1(s)$$
$$\Phi(0,s) = \xi_2(s).$$

See the proof of Theorem 8 in the Appendix for the detailed procedure to obtain (69). The coordinate transformation (69) converts the system vector fields and the output mapping into

$$f(z) = \begin{pmatrix} -9z_1 - z_1^5 - 2z_1^3z_2^2 - z_1z_2^4 \\ -9z_2 - z_1^4z_2 - 2z_1^2z_2^3 - z_2^5 \end{pmatrix}$$
$$g(z) = \begin{pmatrix} \sqrt{18+2z_1^4+4z_1^2z_2^2+2z_2^4} & 0 \\ 0 & \sqrt{18+2z_1^4+4z_1^2z_2^2+2z_2^4} \end{pmatrix}$$
$$h(z) = \begin{pmatrix} \frac{\left(6z_1-2z_1^5-4z_1^3z_2^2-2z_1z_2^4\right)\sqrt{18+2z_1^4+4z_1^2z_2^2+2z_2^4}}{1+z_1^4+2z_1^2z_2^2+z_2^4} \\ \frac{\left(3z_2-z_1^4z_2-2z_1^2z_2^3-z_2^5\right)\sqrt{18+2z_1^4+4z_1^2z_2^2+2z_2^4}}{1+z_1^4+2z_1^2z_2^2+z_2^4} \end{pmatrix}.$$

The observability and controllability functions in the new coordinates are given as follows:

$$L_c(\Phi(z)) = \frac{1}{2}z^{\mathrm{T}}z$$
$$L_o(\Phi(z)) = \frac{1}{2}z^{\mathrm{T}} \begin{pmatrix} \frac{4\left(9+z_1^4+2z_1^2z_2^2+z_2^4\right)}{1+z_1^4+2z_1^2z_2^2+z_2^4} & 0 \\ 0 & \frac{9+z_1^4+2z_1^2z_2^2+z_2^4}{1+z_1^4+2z_1^2z_2^2+z_2^4} \end{pmatrix} z$$
$$=: \frac{1}{2}z^{\mathrm{T}} \begin{pmatrix} \tilde{\tau}_1(z) & 0 \\ 0 & \tilde{\tau}_2(z) \end{pmatrix} z$$

which of course satisfy the HJB equations (28) and (27). It can be readily checked that the aforementioned energy functions L_o and L_c satisfy the balanced properties (65), (66) and (67) on the valid region $U = \mathbb{R}^n$. Hence, (68) implies

$$\|\Sigma\|_H = \sup_{z_1\in\mathbb{R}} \sqrt{\tilde{\tau}_1(z_1,0)} = 6.$$

which indeed equals the outcome of Example 2.

It can be observed from the previous example that the singular value functions $\tilde{\tau}_i$'s have a close relationship with the Hankel norm of the system in the new input-normal/output-diagonal realization. The gain structure of the Hankel operator is clearly exhibited in this coordinate.

Remark 2: The balancing method in the linear case [3], [21] provides the balancing between input and output in the sense $Q = P$ with the observability and controllability Gramians Q and P as well as the balancing between each coordinate axis as in the input-normal/output-diagonal realization given in Theorem 8. Future research should cope with these two balancing properties simultaneously. So far, a procedure to convert an input-normal/output-diagonal realization into a balanced form in the sense of input and output on each coordinate axis z_i was given in [5].

VII. EXAMPLE

This section demonstrates how the input-normal/output-diagonal balanced realization procedure achieved in Theorem 8 works with a physical system, where Taylor series approximations are utilized. Consider a system given by

$$\Sigma : \begin{cases} \dot{x} = -M(x)^{-1}\frac{\partial V(x)}{\partial x}^{\mathrm{T}} + M(x)^{-1}\begin{pmatrix} 1 \\ 0 \end{pmatrix} u \\ y = (1,0)x \end{cases} \tag{70}$$

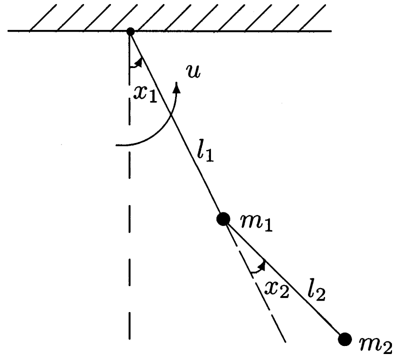


Fig. 1. Double pendulum.

with

$$V(x) = -m_1gl_1\cos x_1 - m_2gl_1\cos x_1 - m_2gl_2\cos(x_1+x_2)$$
$$M(x) = \begin{pmatrix} m_1l_1^2+m_2l_1^2+m_2l_2^2 \\ +2m_2l_1l_2\cos x_2 & m_2l_2^2Lm_2l_1\cos x_2 \\ m_2l_2^2Lm_2l_1\cos x_2 & m_2l_2^2 \end{pmatrix}.$$

The system Σ of (70) is the gradient system of the simple Hamiltonian system that describes the equations of motion of the frictionless double pendulum depicted in Fig. 1. The gradient system is of lower order than the Hamiltonian system, and therefore computationally easier to handle, but still captures the physical properties of the system. Furthermore, the frictionless system is only Lyapunov stable, but not asymptotically stable, while the associated gradient system is asymptotically stable, and thus fulfills the requirements of this paper. See [25] for more details. $V(x)$ and $M(x)$ denote the potential energy and the inertia matrix of the double pendulum, respectively. The constant parameters are given by $m_1 = 1.0$, $m_2 = 10.0$, $l_1 = 10.0$, $l_2 = 1.0$ and the gravity constant is given by $g = 9.81$. See [25] for the details.

Solving the Hamilton–Jacobi equations (27) and (28) based on Taylor series approximation up to fourth order gives the following controllability and observability functions.

$$\begin{aligned} L_c(x) =& 1.5918\times 10^4x_1^2+3.1100\times 10^5x_1\times x_2 \\ &+ 2.4927\times 10^4x_2^2-1.8475\times 10^1x_1^4 \\ &- 7.1487\times 10^4x_1^3x_2-1.5405\times 10^4x_1^2x_2^2 \\ &+ 1.1767\times 10^5x_1x_2^3+6.3707\times 10^4x_2^4+o\left(\|x\|^4\right) \\ L_o(x) =& -0.8207\times 10^{-3}x_1x_2^3+4.5952\times 10^{-2}x_1x_2 \\ &+ 2.7782\times 10^{-1}x_1^2-6.0976\times 10^{-4}x_2^4 \\ &+ 8.1656\times 10^{-4}x_1^2x_2^2+7.4077\times 10^{-3}x_1^3x_2 \\ &+ 1.9147\times 10^{-3}x_2^2+2.3089\times 10^{-2}x_1^4+o\left(\|x\|^4\right). \end{aligned}$$

Using a coordinate transformation $x = \bar{\Phi}(\bar{x}) = (\bar{\phi}_1(\bar{x}), \bar{\phi}_2(\bar{x}))$ with

$$\begin{aligned} \bar{\phi}_1(\bar{x}) :=& 6.8308\times 10^{-4}\bar{x}_1 + 5.8100\times 10^{-4}\bar{x}_2 \\ &- 3.7538\times 10^{-11}\bar{x}_1^3 - 4.1356\times 10^{-9}\bar{x}_1^2\bar{x}_2 \\ &- 3.8164\times 10^{-8}\bar{x}_1\bar{x}_2^2 - 9.3423\times 10^{-8}\bar{x}_2^3 \\ \bar{\phi}_2(\bar{x}) :=& -1.3594\times 10^{-3}\bar{x}_1 - 7.0359\times 10^{-3}\bar{x}_2 \\ &+ 5.0227\times 10^{-10}\bar{x}_1^3 + 4.2708\times 10^{-8}\bar{x}_1^2\bar{x}_2 \\ &+ 3.9365\times 10^{-7}\bar{x}_1\bar{x}_2^2 + 9.6278\times 10^{-7}\bar{x}_2^3 \end{aligned}$$

we can obtain an input-normal/output-diagonal form as follows which is the outcome of Theorem 4:

$$\begin{aligned} L_c\left(\bar{\Phi}(\bar{x})\right) =& \frac{1}{2}\bar{x}^{\mathrm{T}}\bar{x} + o\left(\|\bar{x}^4\|\right) \\ L_o\left(\bar{\Phi}(\bar{x})\right) =& 9.0497\times 10^{-8}\bar{x}_1^2 + 7.1876\times 10^{-10}\bar{x}_2^2 \\ &+ 1.5771\times 10^{-14}\bar{x}_1^4 - 4.8377\times 10^{-15}\bar{x}_1^3\bar{x}_2 \\ &- 8.1507\times 10^{-13}\bar{x}_1^2\bar{x}_2^2 - 2.7193\times 10^{-12}\bar{x}_1\bar{x}_2^3 \\ &- 1.7334\times 10^{-14}\bar{x}_2^4 + o\left(\|\bar{x}^4\|\right) \\ =& \frac{1}{2}\bar{x}^{\mathrm{T}}\begin{pmatrix} \bar{\tau}_1(\bar{x}) & 0 \\ 0 & \bar{\tau}_2(\bar{x}) \end{pmatrix}\bar{x} + o\left(\|\bar{x}^4\|\right) \\ \bar{\tau}_1(\bar{x}) =& 1.8099\times 10^{-7} - 9.6754\times 10^{-15}\bar{x}_1\bar{x}_2 \\ &+ 3.1542\times 10^{-14}\bar{x}_1^2 \\ \bar{\tau}_2(\bar{x}) =& 1.4375\times 10^{-9} - 3.4668\times 10^{-14}\bar{x}_2^2 \\ &- 5.4386\times 10^{-12}\bar{x}_1\bar{x}_2 - 1.6301\times 10^{-12}\bar{x}_1^2. \end{aligned}$$

Using (54) and (55), the solution pair of the differential eigenstructure (37) can be obtained as

$$\xi_1(s) = \begin{pmatrix} s - 2.9874\times 10^{-17}s^3 \\ -0.1109\times 10^{-9}s - 2.6943\times 10^{-8}s^3 \end{pmatrix}$$
$$\xi_2(s) = \begin{pmatrix} 8.8065\times 10^{-12}s + 1.5144\times 10^{-5}s^3 \\ s - 1.3339\times 10^{-16}s^3 \end{pmatrix}$$

parametrized by a scalar $s \in \mathbb{R}$. The property in (57) and (58) gives the axis singular value functions as

$$\begin{aligned} \rho_1(s) =& 4.2543\times 10^{-4} + 3.7070\times 10^{-11}s^2 \\ &- 1.4619\times 10^{-17}s^4 + o(s^4) \\ \rho_2(s) =& 3.7915\times 10^{-5} - 4.5718\times 10^{-10}s^2 \\ &- 5.4584\times 10^{-13}s^4 + o(s^4). \end{aligned}$$

Furthermore, constructing a rotational coordinate transformation as in the proof of Lemma 5, one can find that a coordinate transformation $x = \Phi(z) = (\phi_1(z), \phi_2(z))$ with

$$\begin{aligned} \phi_1(z) :=& z_1 + 2.6943\times 10^{-8}z_1^zz_2 + 3.0343\times 10^{-5}z_1z_2^2 \\ &+ 1.5144\times 10^{-5}z_2^3 \\ \phi_2(z) :=& z_2 - 2.6943\times 10^{-8}z_1^3 - 3.0343\times 10^{-5}z_1^2z_2 \\ &- 1.5144\times 10^{-5}z_1z_2^2 \end{aligned}$$

converts the controllability and observability functions into

$$\begin{aligned} L_c\left(\bar{\Phi}\circ\Phi(z)\right) =& \frac{1}{2}z^{\mathrm{T}}z+o\left(\|z\|^4\right) \\ L_o\left(\bar{\Phi}\circ\Phi(z)\right) =& 9.0497\times 10^{-8}z_1^2+7.1876\times 10^{-10}z_2^2 \\ &+ 1.5771\times 10^{-14}z_1^4+4.6331\times 10^{-12}z_1^2z_2^2 \\ &- 1.7334\times 10^{-14}z_2^4+o\left(\|z\|^4\right) \\ =& \frac{1}{2}z^{\mathrm{T}}\begin{pmatrix} \tau_1(z) & 0 \\ 0 & \tau_2(z) \end{pmatrix}z+o\left(\|z^4\|\right) \\ \tau_1(z) =& 1.8099\times 10^{-7}z_1^2+3.1542\times 10^{-14}z_1^4 \\ &+ 9.2662\times 10^{-12}z_1^2z_2^2 \\ \tau_2(z) =& 1.4375\times 10^{-9}z_2^2-3.4668\times 10^{-14}z_2^4 \end{aligned}$$

which is the input-normal/output-diagonal balanced realization in the sense of Theorem 8. Finally, the Hankel norm of the system in a small neighborhood U of the origin, e.g, $U = \{ z \mid \|z\| \leq 10\}$, (based on fourth-order Taylor series approximation) is readily given by

$$\begin{aligned}\|\Sigma\|_H &\approx \sup_{s\in[-10,10]} \rho_1(s) = \sup_{s\in[-10,10]} \sqrt{\tau_1(s,0)} \\ &= 4.2543 \times 10^{-4}.\end{aligned}$$

Thus this realization clarifies the gain structure of the Hankel operator.

VIII. Conclusion

The differential eigenstructure of Hankel operators for nonlinear systems has been studied in this paper. First, it has been proven that a variational system and a Hamiltonian extension with extended input and output spaces can be interpreted as the Gâteaux differentiation and its adjoint of a dynamical input–output system respectively, as a preparation. Second, the Gâteaux differentiation has been utilized in order to clarify the differential eigenstructure of the Hankel operator for nonlinear systems, which is closely related to the Hankel norm of the original system. Third, a new characterization of the nonlinear version of the Hankel singular values was given. Based on this characterization, an input-normal/output-diagonalization procedure has been derived. Future directions will involve the extension of the presented methods to model reduction for nonlinear systems.

Appendix
Lemmas

This appendix recalls two lemmas from [24] and [26], which are necessary to prove Theorem 8 and Lemma 5.

Lemma 6: [26] Let L be a smooth function in a convex neighborhood U of 0 in $\mathbb{R}^n$, with $L(0) = 0$. Then

$$L(x_1, \ldots, x_n) = \sum_{i=1}^{n} x_i a_i(x_1, \ldots, x_n)$$

for some suitable smooth functions a_i's defined on U, with $a_i(0) = (\partial L/\partial x)(0)$.

The second one is the Brouwer's fixed point theorem.

Theorem 9: [24] Let D^n denote the unit disc on $\mathbb{R}^n$

$$D^n = \left\{ x \in \mathbb{R}^n | x_1^2 + \ldots + x_n^2 \leq 1 \right\}.$$

Then, any continuous function $G : D^n \to D^n$ has a fixed point.

The last one is an extended version of the *homogeneity lemma*.

Lemma 7: [24] Consider two interior points y and z of a smooth connected manifold N. Then, there exists a diffeomorphism $h : N \to N$ converting y into z which is smoothly isotopic to the identity. Furthermore, for any open set $U \subset N$ containing y and z, the function h can be chosen in such a way that $h|_{N\setminus U}$ coincides with the identity.

Proof of Theorem 2

This proof is based on the results in [19]. Let $(u(t,\varepsilon), x(t,\varepsilon), y(t,\varepsilon))$, $t \in (t^0, t^1)$ denote a family of input-state-output trajectories of $\hat{\Sigma}$ parameterized by ε. Then, we have

$$\begin{aligned}\frac{\partial y}{\partial \varepsilon}(t,0) &= \frac{\partial h}{\partial x}(x(t,0), u(t,0), t) \frac{\partial x}{\partial \varepsilon}(t,0) \\ &\quad + \frac{\partial h}{\partial u}(x(t,0), u(t,0), t) \frac{\partial u}{\partial \varepsilon}(t,0) \\ \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial x}{\partial \varepsilon}(t,0) &= \frac{\partial}{\partial \varepsilon}\frac{\mathrm{d}x(t,0)}{\mathrm{d}t} \\ &= \frac{\partial f}{\partial x}(x(t,0), u(t,0), t) \frac{\partial x}{\partial \varepsilon}(t,0) \\ &\quad + \frac{\partial f}{\partial u}(x(t,0), u(t,0), t) \frac{\partial u}{\partial \varepsilon}(t,0)\end{aligned}$$

and, moreover

$$\begin{aligned}\frac{\mathrm{d}x(t,0)}{\mathrm{d}t} &= f(x(t,0), u(t,0), t) \\ y(t,0) &= h(x(t,0), u(t,0), t).\end{aligned}$$

Therefore, the trajectories $(\partial u/\partial \varepsilon(t,0),\ \partial x/\partial \varepsilon(t,0),\ \partial y/\partial \varepsilon(t,0))$ coincide with the input-state-output trajectories of the variational system $\hat{\Sigma}_v$. Now, let $u(t,\varepsilon) = u(t) + \varepsilon v(t)$. Then, we obtain

$$\begin{aligned}&\mathrm{d}\hat{\Sigma}\left((x^0, u), (x_v^0, v)\right) \\ &\quad = \lim_{\varepsilon\to 0} \frac{\hat{\Sigma}\left(x^0 + \varepsilon x_v^0, u + \varepsilon v\right) - \hat{\Sigma}(x^0, u)}{\varepsilon} \\ &\quad = \lim_{\varepsilon\to 0} \frac{\left(x(t^1,\varepsilon) - x(t^1,0), y(t,\varepsilon) - y(t,0)\right)}{\varepsilon} \\ &\quad = \left(\frac{\partial x}{\partial \varepsilon}(t^1,0), \frac{\partial y}{\partial \varepsilon}(t,0)\right).\end{aligned}$$

Due to the assumption that the state trajectory $x_v(t)$ of the variational system $\hat{\Sigma}_v$ is uniquely determined, we can conclude that the existence of $\mathrm{d}\hat{\Sigma}$ is equivalent to that $\hat{\Sigma}_v$ is an operator on L_2 spaces. As a result, $\hat{\Sigma}_v$ coincides with the Gâteaux differential $\mathrm{d}\hat{\Sigma}$. This proves the theorem. □

Proof of Theorem 3

The proof follows similar arguments as [18, Prop. 2], i.e., it uses the port-controlled Hamiltonian systems structure. Let the Hamiltonian function be given by $H_v = p^{\mathrm{T}} x_v$, and denote $x_v^0 := x_v(t^0)$, $x_v^1 := x_v(t^1)$, $p^0 := p(t^0)$ and $p^1 := p(t^1)$ for simplicity. Then, we have

$$\begin{aligned}\frac{\mathrm{d}H_v}{\mathrm{d}t} &= p^{\mathrm{T}} \dot{x}_v + x_v^{\mathrm{T}} \dot{p} \\ &= p^{\mathrm{T}}\left(\frac{\partial f}{\partial x} x_v + \frac{\partial f}{\partial u} u_v\right) - x_v^{\mathrm{T}}\left(\frac{\partial f}{\partial x}^{\mathrm{T}} p + \frac{\partial h}{\partial x}^{\mathrm{T}} u_a\right) \\ &= p^{\mathrm{T}} \frac{\partial f}{\partial u} u_v - x_v^{\mathrm{T}} \frac{\partial h}{\partial x}^{\mathrm{T}} u_a \\ &= \left(p^{\mathrm{T}} \frac{\partial f}{\partial u} + u_a^{\mathrm{T}} \frac{\partial h}{\partial u}\right) u_v - \left(x_v^{\mathrm{T}} \frac{\partial h}{\partial x}^{\mathrm{T}} + u_v^{\mathrm{T}} \frac{\partial h}{\partial u}^{\mathrm{T}}\right) u_a \\ &= y_a^{\mathrm{T}} u_v - y_v^{\mathrm{T}} u_a.\end{aligned}$$

This reduces to

$$\begin{aligned}\langle y_a, u_v\rangle_{L_2^m} - \langle y_v, u_a\rangle_{L_2^r} &= \int_{t^0}^{t^1} \left(y_a^{\mathrm{T}} u_v - y_v^{\mathrm{T}} u_a\right) \mathrm{d}t \\ &= \int_{t^0}^{t^1} \frac{\mathrm{d}H_v}{\mathrm{d}t}\mathrm{d}t \\ &= H_v|_{t=t^1} - H_v|_{t=t^0} \\ &= \left\langle x_v^1, p^1\right\rangle_{\mathbb{R}^n} - \left\langle x_v^0, p^0\right\rangle_{\mathbb{R}^n}.\end{aligned}$$

Therefore

$$\left\langle \left(x_v^1, y_v\right), \left(p^1, u_a\right)\right\rangle_{\mathbb{R}^n \times L_2^r} = \left\langle \left(x_v^0, u_v\right), \left(p^0, y_a\right)\right\rangle_{\mathbb{R}^n \times L_2^m} \tag{71}$$

holds with the inner product on $\mathbb{R}^n \times L_2$. Substituting $(x_v^1, y_v) = \hat{\Sigma}_v((x^0, u), (x_v^0, u_v))$ and $(p^0, y_a) = \hat{\Sigma}_a((x^0, u), (p^1, u_a))$ implies (13) and completes the proof. □

Proof of Theorem 8

The theorem is proven by induction with respect to the dimension n of the state x.

a) **Case** $n = 1$: Trivial from Lemma 5.

b) **Case** $n = k(\geq 2)$: Suppose that the theorem holds in the case $n = k - 1$. It is assumed without loss of generality that the system is already applied the coordinate transformation in Lemma 5. The coordinate transformation $z = \Phi(x)$ with $\Phi : \mathbb{R}^k \to \mathbb{R}^k$ is constructed by

$$z = \Phi(x) = \Phi^k \circ \Phi^{k-1} \circ \cdots \circ \Phi^1(x)$$

where the intermediate coordinates are described as

$$z \overset{\mathrm{Id}}{\mapsto} z^k \overset{\Phi^k}{\mapsto} z^{k-1} \overset{\Phi^{k-1}}{\mapsto} \cdots \overset{\Phi^2}{\mapsto} z^1 \overset{\Phi^1}{\mapsto} x.$$

Each function Φ^j has a form

$$z^{j-1} = \begin{pmatrix} z_1^{j-1} \\ \vdots \\ z_{j-1}^{j-1} \\ z_j^{j-1} \\ z_{j+1}^{j-1} \\ \vdots \\ z_k^{j-1} \end{pmatrix} = \begin{pmatrix} \psi_1^j\left(z_1^j, \ldots, z_k^j\right) \\ \vdots \\ \psi_{j-1}^j\left(z_1^j, \ldots, z_k^j\right) \\ z_j^j \\ \psi_{j+1}^j\left(z_1^j, \ldots, z_k^j\right) \\ \vdots \\ \psi_k^j\left(z_1^j, \ldots, z_k^j\right) \end{pmatrix} = \Phi^j(z^j)$$

satisfying

$${z^{j-1}}^{\mathrm{T}} z^{j-1} = \Phi^j(z^j)^{\mathrm{T}}\Phi^j(z^j) = {z^j}^{\mathrm{T}} z^j.$$

Here, the functions ψ_i^j's are constructed by the the transformation in Theorem 8 in the case $n = k-$ by regarding the parameter z_j as a constant. That is, the function Φ^j satisfies

$$\begin{aligned} z_i^j = 0 &\Leftrightarrow \frac{\partial L_c\left(\Phi^1 \circ \cdots \circ \Phi^j(z^j)\right)}{\partial z_i^j} = 0 \\ &\Leftrightarrow \frac{\partial L_o\left(\Phi^1 \circ \cdots \circ \Phi^j(z^j)\right)}{\partial z_i^j} = 0 \qquad \forall i \neq j. \end{aligned}$$

On this coordinate, we have

$$\begin{aligned} &\frac{\partial L_o\left(\Phi^1 \circ \cdots \circ \Phi^j(z^j)\right)}{\partial z_j^j} \\ &\quad = \sum_{i=1}^{k} \frac{\partial L_o\left(\Phi^1 \circ \cdots \circ \Phi^{j-1}(z^{j-1})\right)}{\partial z_i^{j-1}} \frac{\partial z_i^{j-1}}{\partial z_j^j} \\ &\quad = \frac{\partial L_o\left(\Phi^1 \circ \cdots \circ \Phi^{j-1}(z^{j-1})\right)}{\partial z_j^{j-1}} \\ &\qquad + \sum_{\substack{i=1 \\ i\neq j}}^{k} \frac{\partial L_o\left(\Phi^1 \circ \cdots \circ \Phi^{j-1}(z^{j-1})\right)}{\partial z_i^{j-1}} \frac{\partial \psi_i^j(z^j)}{\partial z_j^j}. \end{aligned}$$

Note that Assumption 2 guarantees that

$$\Phi^1 = \cdots = \Phi^k = \mathrm{Id} \tag{72}$$

holds at the origin. This implies

$$\frac{\partial L_o\left(\Phi^1 \circ \cdots \circ \Phi^j(z^j)\right)}{\partial z_i^j} \approx \sigma_i z_i^j$$

holds near the origin, where σ_i is the Hankel singular value of the Jacobian linearization of the system. Hence, for any $\epsilon_1 > 0$, there exists a $\delta_1 > 0$ satisfying

$$\begin{aligned} &\|z^j\| < \delta_1 \Rightarrow \\ &\sigma_i z_i^j - \epsilon_1\|z^j\| < \frac{\partial L_o\left(\Phi^1 \circ \cdots \circ \Phi^j(z^j)\right)}{\partial z_i^j} < \sigma_i z_i^j + \epsilon_1\|z^j\|. \end{aligned}$$

On the other hand, (72) also implies

$$\left.\frac{\partial \psi i^j(z^j)}{\partial z_j^j}\right|_{z^j=0} = 0.$$

Hence, for any $\epsilon_2 > 0$, there exists $\delta_2 > 0$ satisfying

$$\|z^j\| < \delta_2 \Rightarrow \left\|\frac{\partial \psi_i^j(z^j)}{\partial z_j^j}\right\| < \epsilon_2.$$

Therefore, choosing small enough ϵ_1 and ϵ_2 satisfying

$$\epsilon_1 + \epsilon_1\epsilon_2(k-1) + \epsilon_2 \sum_{\substack{i=1 \\ i\neq j}}^{k} \sigma_i \leq \frac{\sigma_j}{\sqrt{k}}$$

we obtain a relation

$$\begin{aligned}
&\frac{\partial L_o\left(\Phi^1\circ\cdots\circ\Phi^j(z^j)\right)}{\partial z_j^j}\\
&\quad > \sigma_j z_j^j - \epsilon_1\|z^j\| + \sum_{\substack{i=1\\ i\neq j}}^{k}\left(\sigma_i z_i^j - \epsilon_1\|z^j\|\right)\epsilon_2\\
&\quad = \sigma_j z_j^j + \sum_{\substack{i=1\\ i\neq j}}^{k}\epsilon_2\sigma_i z_i^j - \left(\epsilon_1+\epsilon_1\epsilon_2(k-1)\right)\|z^j\|\\
&\quad \geq \sigma_j z_j^j - \|z^j\|\left(\epsilon_1+\epsilon_1\epsilon_2(k-1)+\epsilon_2\sum_{\substack{i=1\\ i\neq j}}^{k}\sigma_i\right)\\
&\quad \geq \sigma_j\left(z_j^j - \frac{\|z^j\|}{\sqrt{k}}\right).
\end{aligned}$$

In the same way, we can prove the upper bound

$$\frac{\partial L_o\left(\Phi^1\circ\cdots\circ\Phi^j(z^j)\right)}{\partial z_j^j} < \sigma_j\left(z_j^j + \frac{\|z^j\|}{\sqrt{k}}\right).$$

That is

$$\frac{\partial L_o\left(\Phi^1\circ\cdots\circ\Phi^j(z^j)\right)}{\partial z_j^j} \neq 0$$

holds for all $z^j \in Z_j^j$ where

$$Z_j^j := \left\{ z^j \middle| \left|z_j^j\right| \geq \frac{\|z^j\|}{\sqrt{k}} \quad \|z^j\| < \min\{\delta_1,\delta_2\}\right\}.$$

This proves that (65) holds on the region Z_j^j. Therefore, the coordinate transformation $\Phi^i, i > j$ maps $\forall z^j \in Z_j^j$ identically since the region Z_j^j is already balanced. In fact, the coordinate transformation given in Lemma 5, maps the balanced region identically by its construction. This fact implies that (65) holds on their union

$$\begin{aligned}
&Z := \cup_{j=1}^{k} Z_j\\
&Z_j := \left\{ z \middle| |z_j| \geq \frac{\|z\|}{\sqrt{k}} \quad \|z\| < \min\{\delta_1,\delta_2\}\right\}.
\end{aligned}$$

Furthermore, it can be easily checked that

$$Z = \{z \mid \|z\| < \min\{\delta_1,\delta_2\}\}$$

which implies that the theorem also holds in the case $n = k$ (with $U = \Phi(Z)$).

The cases (**a**) and (**b**) prove the (65) by induction.

Next, we show the fact that the system can be described by an output-diagonal form in the z coordinate. It follows from Theorem 7 with

$$\frac{\partial L_o\left(\Phi(z)\right)}{\partial z} = \lambda_i \frac{\partial L_c\left(\Phi(z)\right)}{\partial z}$$

which holds along the coordinate axes z_i's, that

$$L_o\left(\Phi(0,\ldots,0,z_i,0,\ldots,0)\right) = \frac{1}{2}\rho_i^2(z_i)z_i^2 \tag{73}$$

$$\begin{aligned}
&\left.\frac{\partial L_o\left(\Phi(z)\right)}{\partial z}\right|_{z=(0,\ldots,0,z_i,0,\ldots,0)}\\
&\quad = \lambda_i(z_i)\left.\frac{\partial L_c\left(\Phi(z)\right)}{\partial z}\right|_{z=(0,\ldots,0,z_i,0,\ldots,0)}\\
&\quad = \lambda_i(z_i)(0,\ldots,0,z_i,0,\ldots,0)\\
&\quad = \left(0,\ldots,0,\frac{\mathrm{d}}{\mathrm{d}z_i}\left(\frac{1}{2}\rho_i^2(z_i)z_i^2\right),0,\ldots,0\right).
\end{aligned} \tag{74}$$

Let $L_o(\Phi(z))$ be described by

$$L_o\left(\Phi(z)\right) = \frac{1}{2}\sum_{i=1}^{n}\rho_i^2(z_i)z_i^2 + e(z)$$

with a smooth function $e(z)$. Now, (65), (73), and (74) imply that

$$z_i = 0 \Rightarrow \frac{\partial e(z)}{\partial z_i} = 0 \tag{75}$$

$$z = (0,\ldots,0,z_i,0,\ldots,0) \Rightarrow e(z) = 0,\ \frac{\partial e(z)}{\partial z} = 0. \tag{76}$$

For (75) and Lemma 6, we have

$$\begin{aligned}
e(z) = \sum_{i=1}^{n} z_i^2 e_i(z_i) + \sum_{i=1}^{n}\sum_{\substack{j=1\\ i<j}}^{n} z_i^2 z_j^2 e_{i,j}(z_i,z_j)\\
+\cdots+ z_1^2 z_2^2\ldots z_n^2 e_{1,2,\ldots,n}(z)
\end{aligned}$$

with smooth functions e_i's. Furthermore, from (76) we obtain

$$e(z) = \sum_{i=1}^{n}\sum_{\substack{i=1\\ i<j}}^{n} z_i^2 z_j^2 e_{i,j}(z_i,z_j) + \cdots + z_1^2 z_2^2\ldots z_n^2 e_{1,2,\ldots,n}(z).$$

For example, define the scalar functions $\tau_i : \mathbb{R}^n \supset U \to \mathbb{R}$'s ($i \in \{1,2,\ldots,n\}$) by

$$\begin{aligned}
\tau_i(z) := \rho_i^2(z_i) + 2\sum_{\substack{j=1\\ i<j}}^{n} z_j^2 e_{i,j}(z_i,z_j)\\
+\cdots+ 2z_{i+1}^2\cdots z_n^2 e_{i,i+1,\ldots,n}(z_i,z_{i+1},\ldots,z_n).
\end{aligned}$$

Then, we can see that

$$L_o\left(\Phi(z)\right) = \frac{1}{2}\sum_{i=1}^{n}\rho_i^2(z_i)z_i^2 + e(z) = \frac{1}{2}\sum_{i=1}^{n}\tau_i(z)z_i^2 \tag{77}$$

and the observability function is in output-diagonal form (30). Equations (66) and (67) follow straightforwardly from (73), (74), and (77). This completes the proof. $\square$

REFERENCES

[1] T. Kailath, *Linear Systems*. Upper Saddle River, NJ: Prentice-Hall, 1980.

[2] E. A. Jonckheere and L. M. Silverman, "Singular value analysis of deformable systems," *Circuits, Syst. Signal Processing*, vol. 1, no. 3–4, pp. 447–470, 1982.

[3] B. C. Moore, "Principal component analysis in linear systems: controllability, observability and model reduction," *IEEE Trans. Autom. Control*, vol. AC-26, no. 1, pp. 17–32, Jan. 1981.

[4] K. Glover, "All optimal Hankel-norm approximations of linear multivariable systems and their l^∞-error bounds," *Int. J. Control*, vol. 39, pp. 1115–1193, 1984.

[5] J. M. A. Scherpen, "Balancing for nonlinear systems," *Syst. Control Lett.*, vol. 21, pp. 143–153, 1993.

[6] W. S. Gray and J. M. A. Scherpen, "State dependent matrices in quadratic forms," *Syst. Control Lett.*, vol. 44, no. 3, pp. 219–232, 2001.

[7] J. Hahn and T. F. Edgar, "Reduction of nonlinear models using balancing of empirical gramians and Galerkin projections," in *Proc. Amer. Control Conf.*, 2000, pp. 2864–2868.

[8] ——, "An improved method for nonlinear model reduction using balancing of empirical Gramians," *Comput. Chem. Eng.*, vol. 26, no. 10, pp. 1379–1397, 2002.

[9] A. J. Newman and P. S. Krishnaprasad, "Computation for nonlinear balancing," in *Proc. 37th IEEE Conf. Decision and Control*, 1998, pp. 4103–4104.

[10] A. J. Newman, "Modeling and reduction with applications to semiconductor processing," Ph.D. dissertation, Univ. Maryland, College Park, MD, 1999.

[11] S. Lall, J. E. Marsden, and G. Glavaski, "A subspace approach to balanced truncation for model reduction of nonlinear systems," *Int. J. Robust Nonlinear Control*, vol. 12, no. 6, pp. 519–535, 2002.

[12] W. S. Gray and J. M. A. Scherpen, "Hankel operators and Gramians for nonlinear systems," in *Proc. 37th IEEE Conf. Decision and Control*, 1998, pp. 3349–3353.

[13] J. M. A. Scherpen and W. S. Gray, "On singular value functions and Hankel operators for nonlinear systems," in *Proc. Amer. Control Conf.*, 1999, pp. 2360–2364.

[14] J. M. A. Scherpen, K. Fujimoto, and W. S. Gray, "On Hankel structure, Hilbert adjoints, and Hamiltonian adjoint realizations for nonlinear systems," in *Proc. 2000 Conf. Decision and Control*, Sydney, Australia, Dec. 2000, pp. 5102–5107.

[15] J. Batt, "Nonlinear compact mappings and their adjoints," *Math. Ann.*, vol. 189, pp. 5–25, 1970.

[16] J. M. A. Scherpen and W. S. Gray, "Nonlinear Hilbert adjoints: properties and applications to Hankel singular value analysis," *Nonlinear Anal.: Theory, Meth., Appl.*, vol. 51, no. 5, pp. 883–901, 2002.

[17] K. Fujimoto and J. M. A. Scherpen, "Eigenstructure of nonlinear Hankel operators," in *Nonlinear Control in the Year 2000*. ser. Lecture Notes on Control and Information Science, A. Isidori, F. Lamnabhi-Lagarrigue, and W. Respondek, Eds. Paris, France: Springer-Verlag, 2000, vol. 258, pp. 385–398.

[18] K. Fujimoto, J. M. A. Scherpen, and W. S. Gray, "Hamiltonian realizations of nonlinear adjoint operators," *Automatica*, vol. 38, no. 10, pp. 1769–1775, 2002.

[19] P. E. Crouch and A. J. van der Schaft, *Variational and Hamiltonian Control Systems*, ser. Lecture Notes on Control and Information Science. Berlin, Germany: Springer-Verlag, 1987, vol. 101.

[20] J. A. Ball and A. J. van der Schaft, "J-inner-outer factorization, J-spectral factorization, and robust control for nonlinear systems," *IEEE Trans. Autom. Control*, vol. 41, no. 3, pp. 379–392, Mar. 1996.

[21] K. Zhou, J. C. Doyle, and K. Glover, *Robust and Optimal Control*. Upper Saddle River, NJ: Prentice-Hall, Inc., 1996.

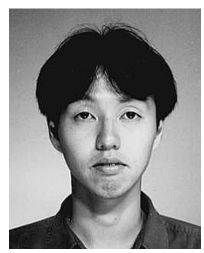
[22] K. Fujimoto, "Synthesis and analysis of nonlinear control systems based on transformations and factorizations," Ph.D. dissertation, Kyoto Univ., Kyoto, Japan, 2000.

[23] P. Blanchard and E. Bruning, *Variational Methods in Mathematical Physics: A Unified Approach*. Berlin, Germany: Springer-Verlag, 1992.

[24] J. W. Milnor, *Topology From the Differential Viewpoint*. Princeton, NJ: Princeton Univ. Press, 1965.

[25] 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. Control*, vol. 45, no. 11, pp. 2079–2086, Nov. 2000.

[26] J. W. Milnor, *Morse Theory*. Princeton, NJ: Princeton Univ. Press, 1963.



Kenji Fujimoto (S'97–A'97–M'02) received the B.Sc. and M.Sc. degrees in engineering and Ph.D. degree in informatics from Kyoto University, Kyoto, Japan, in 1994, 1996, and 2001, respectively.

He is currently an Associate Professor with the Department of Mechanical Science and Engineering, Nagoya University, Nagoya, Japan. From 1997 to 2004, he was a Research Associate with the Department of Systems Science, Kyoto University. From 1999 to 2000, he was a Research Fellow with the Department of Electrical Engineering, Delft University of Technology, The Netherlands. He has held visiting research positions at the Australian National University, Australia and Delft University of Technology, in 1999 and 2002, respectively. His research interests include nonlinear systems theory and control.

He is currently an Associate Editor of the IEEE Conference Editorial Board.

Jacquelien M. A. Scherpen (M'95–SM'04) received the M.Sc. and Ph.D. degrees in applied mathematics from the University of Twente, The Netherlands, in 1990 and 1994, respectively.

Currently, she is an Associate Professor with the Delft Center for Systems and Control, Delft University of Technology, The Netherlands. She has held visiting research positions at the Universite de Compiegne, France, SUPELEC, Gif-sur-Yvette, France, the University of Tokyo, Tokyo, Japan, and Old Dominion University, Norfolk, VA. Her research interests include nonlinear model reduction methods, realization theory, and nonlinear control methods, with, in particular, modeling and control of physical systems with applications to electrical circuits.

Dr. Scherpen has served as an Associate Editor of the IEEE TRANSACTIONS ON AUTOMATIC CONTROL, and is currently Associate Editor of the *International Journal of Robust and Nonlinear Control*.