University of Groningen

Model Reduction by Differential Balancing Based on Nonlinear Hankel Operators

Kawano, Yu; Scherpen, Jacquélien M.A.

Published in:
IEEE Transactions on Automatic Control

DOI:
[10.1109/TAC.2016.2628201](https://doi.org/10.1109/TAC.2016.2628201)

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2017

[Link to publication in University of Groningen/UMCG research database](#)

Citation for published version (APA):
Kawano, Y., & Scherpen, J. M. A. (2017). Model Reduction by Differential Balancing Based on Nonlinear Hankel Operators. *IEEE Transactions on Automatic Control*, 62(7).
<https://doi.org/10.1109/TAC.2016.2628201>

Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

Model Reduction by Differential Balancing Based on Nonlinear Hankel Operators

Yu Kawano, *Member, IEEE*, and Jacquelin M. A. Scherpen, *Senior Member, IEEE*

Abstract—In this paper, we construct balancing theory for nonlinear systems in the contraction framework. First, we define two novel controllability and observability functions via prolonged systems. We analyze their properties in relation to controllability and observability, and use them for so-called differential balancing, and its application to model order reduction. One of the main contribution of this paper is showing that differential balancing has close relationships with the Fréchet derivative of the nonlinear Hankel operator. Inspired by [3], we provide a generalization in order to have a computationally more feasible method. Moreover, error bounds for model reduction by generalized balancing are provided.

Index Terms—Balanced truncation, contraction, Hankel operator, nonlinear systems.

I. INTRODUCTION

MODEL order reduction problems have been widely studied because reduced order models are useful for analysis, design and simulation [1], [2]. In both linear and nonlinear control theory, a balanced realization is a useful state-space representation when studying model reduction problems based on the Hankel norm [4]–[8]. Besides balancing, also moment matching [4] is a useful tool for model reduction for control, and for nonlinear systems, this method has only been recently developed, see [9], [10]. Balancing for nonlinear systems has a longer history [6], but there are still many recent developments, i.e., there are various other types of nonlinear balancing such as flow balancing [11], [12], incremental balancing [3] and dynamic balancing [13]. These methods are developed to take into account different characteristics of importance, such as incremental stability for example [3]. In general, it depends on the system analysis and the control goal which method is best. For instance, incremental balancing focuses on estimating an error bound for model reduction and preserving stability under model reduction by restricting a class of systems with odd functions. A common problem of flow, incremental and dynamic balancing is that the relationships between model reduction and the Hankel operator is unclear while the Hankel operator plays a central role in the linear balancing theory. Here we present yet another balancing theory which has close relationships with the Hankel operator, based upon contraction theory.

Manuscript received August 9, 2016; revised November 2, 2016; accepted November 8, 2016. Date of publication November 11, 2016; date of current version June 26, 2017. This work supported in part by JST CREST and JSPS KAKENHI Grant Numbers 15K18087. Recommended by Associate Editor M. R. Opmeer.

The authors are with the Jan C. Willems Center for Systems and Control, Engineering and Technology institute Groningen, Faculty of Science and Engineering, University of Groningen, Nijenborgh, 4, 9747 AG Groningen, The Netherlands (e-mail: y.kawano@rug.nl; j.m.a.scherpen@rug.nl).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TAC.2016.2628201

Contraction theory has been studied in recent decades, and deals with trajectories of nonlinear systems with respect to one another. One of the interesting ideas of contraction theory is considering the infinitesimal metric instead of a feasible distance function. That is, the nonlinear system and its variational system combined in the so called prolonged system are used in analysis and design. In this setting, for instance, stability [14]–[16], optimal and H^∞ control [17], [18], and dissipativity [19], [20] have been studied. However, if the system order becomes large, the analysis and control become difficult, which is one of the motivations for studying balancing in the contraction framework. We call it differential balancing theory.

In this paper, we construct differential balancing theory from two perspectives. Differential balancing itself is oriented towards analysis, and generalized differential balancing is oriented towards easier computations. Differential balancing is based on two energy functions defined by using inputs and outputs of prolonged systems, which we call differential controllability and observability functions. Differential balancing reduces to well known linear balancing in the linear case and has similar aspects as the flow and incremental balancing methods [3], [11], [12], which are based on studies around trajectories and feasible distance functions. However, for flow balancing, there is no concept of an energy function, except in a small sliding time window, and the relationship between balancing and system properties such as controllability, observability, and stability is unclear. Incremental balancing can be done when certain sufficient conditions on the structure of the system are fulfilled, such as being able to decouple the energy functions, or for the relation with asymptotic stability specific drift and input vector fields whose elements are odd functions are required. It is worth emphasizing that for both flow and incremental balancing, the existence of balanced realizations is not guaranteed in general even if there exist kinds of positive definite controllability and observability functions. For traditional nonlinear balancing, [6], the method is only valid in a neighborhood of a point, in general an equilibrium point, while the existence of balanced realization is guaranteed. In contrast, the differential balanced realization can be computed by taking a line integral of an infinitesimal metric, and the (global) balanced realization exists if the prolonged system has corresponding positive definite controllability and observability functions. Moreover, differential balancing does not require specific structures for the system in contrast to incremental balancing. Another advantage of differential balancing over traditional balancing is that model reduction by the differential balancing can be achieved for time-varying systems as well.

Furthermore, as mentioned before, flow, incremental and dynamic balancing methods [3], [11]–[13] have not been studied in relation to the nonlinear Hankel operator $\mathcal{H}(u)$ [7], [8], [21]. The Hankel operator is composed of the controllability and

observability operators which respectively have close relationships with controllability and observability Gramians. In the linear case, it is well known that the nonzero singular values of the Hankel operator λ defined by $\mathcal{H}^*(\mathcal{H}(u)) = \lambda u$, where $\mathcal{H}^*$ is the adjoint operator of $\mathcal{H}$, are equivalent to eigenvalues of the product of controllability and observability Gramians [5], which implies that a singular value is the gain between the input effort to reach a state and the output energy of that state in the balanced realization. In [7], [8], for the traditional nonlinear balancing, these relationships have been generalized by considering $(d\mathcal{H}(u))^* \circ \mathcal{H}(u) = \lambda u$, where $(d\mathcal{H}(u))^*(\cdot)$ is the adjoint operator of the Fréchet derivative of $\mathcal{H}(u)$. In comparison to the linear result, λ defined by $\mathcal{H}^*(\mathcal{H}(u), u) = \lambda u$, seems to be the most natural extension of the notion of a singular value; however, [7] points out this λ is not found so far. To overcome this, the Fréchet derivative has been employed. Here, we study the relationship between the nonlinear Hankel operator and differential balancing. Natural extensions of relationships of the linear case are studied, in particular between differential balancing and the differential singular Hankel value λ defined by $(d\mathcal{H}(u))^* \circ d\mathcal{H}(u)(v) = \lambda v$. Therefore, model reduction by differential balancing makes sense from the viewpoint of the nonlinear Hankel operator.

Generalized differential balancing is an extension of differential balancing in order to ease the computational efforts. We consider nonlinear systems with constant input vector fields and linear output functions. Generalized balancing relies on so-called generalized differential energy functions, which are bounds on differential energy functions, and consequently we have bounds on the differential singular Hankel values. Generalized differential balancing has several advantages over other computationally feasible methods such as in [3], [13]. First, generalized differential balancing does not require specific structures of the vector field of the system in contrast to the generalized incremental balancing [3]. Second, an error bound for model reduction is estimated differently from dynamic balancing [13]. Moreover, generalized differential balancing is applicable to time-varying systems.

The remainder of this paper is organized as follows. In Section II, we define two differential energy functions, and we provide characterizations for them. Furthermore, we investigate relationships between these two differential energy functions and a minimal realization of the system in the time-invariant case. In Section III, we define differential balanced realizations and study model reduction for such realizations for the time-varying nonlinear systems case. Moreover, we characterize differential balanced realizations from the viewpoint from the nonlinear Hankel operator in the time-invariant case. An example illustrates our method. In Section IV, we present generalized differential balancing, and model reduction based on that. It is illustrated by a system composed of 50 mass-spring-damper systems with nonlinear springs. Finally in Section V we conclude the paper.

Notation: Let $\mathbb{R}$ be the field of real numbers. Denote $\mathbb{R}_{\geq 0} := [0, \infty) \subset \mathbb{R}$. It is said that $u : [a, b] \rightarrow \mathbb{R}^m$ is in $L_2^m[a, b]$ if $\|u(t)\|_{L_2^m[a, b]} := \sqrt{\int_a^b \|u(t)\|^2 dt} < \infty$, where $\|u(t)\| := \sqrt{u^T(t)u(t)}$. A path γ on $\mathbb{R}^n$ is a class C^2 mapping $\gamma : \mathbb{R} \supset [0, 1] \rightarrow \mathbb{R}^n$. A locally Lipschitz function $\alpha : \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}_{\geq 0}$ is said to belong to class $\mathcal{K}_\infty$ if it is strictly increasing, $\alpha(0) = 0$, and $\lim_{s \rightarrow \infty} \alpha(s) \rightarrow \infty$. For matrix $A(x, t) = (a_{ij})$, denote $\delta_f(A) := (\partial a_{ij} / \partial t + (\partial a_{ij} / \partial x)f)$ with a vector field $f : \mathbb{R} \times$

$\mathbb{R}^n \rightarrow \mathbb{R}^n$. If A is invertible, we use the notation A^{-T} to denote $(A^{-1})^T$. For vector valued function $F : \mathbb{R}^n \times \mathbb{R} \rightarrow \mathbb{R}^m$, denote $\partial F(x, t) / \partial x := [\partial F(x, t) / \partial x_1, \dots, \partial F(x, t) / \partial x_n]$, and $\partial^T F(x, t) / \partial x := (\partial F(x, t) / \partial x)^T$.

II. DIFFERENTIAL ENERGY FUNCTIONS

A. Preliminaries

In contraction theory the so-called nonlinear prolonged system is studied, which consists of the nonlinear time-varying system and its associated system of differential dynamics (also called variational system). I.e., the system Σ_{gh} , and its corresponding variational system $d\Sigma_{gh}$ together form the nonlinear prolonged system and are given as follows

$$\Sigma_{gh} : \begin{cases} \dot{x}(t) := dx(t)/dt \\ \quad = f(x(t), t) + g(x(t), t)u(t), \quad x(t_0) := x_0, \\ y(t) = h(x(t), t), \end{cases}$$

$$d\Sigma_{gh} : \begin{cases} \delta \dot{x}(t) := \frac{d\delta x(t)}{dt} = \frac{\partial(f(x(t), t) + g(x(t), t)u(t))}{\partial x} \delta x(t) \\ \quad + g(x(t), t)\delta u(t), \quad \delta x(t_0) := \delta x_0, \\ \delta y(t) = \frac{\partial h(x(t), t)}{\partial x} \delta x(t), \end{cases}$$

where $x(t) \in \mathbb{R}^n$, $u(t) \in \mathbb{R}^m$ and $y(t) \in \mathbb{R}^p$ are respectively the state, input and output of Σ_{gh} ; $\delta x(t) \in \mathbb{R}^n$, $\delta u(t) \in \mathbb{R}^m$ and $\delta y(t) \in \mathbb{R}^p$ are respectively the state, input and output of $d\Sigma_{gh}$; $f : \mathbb{R}^n \times \mathbb{R} \rightarrow \mathbb{R}^n$, $g : \mathbb{R}^n \times \mathbb{R} \rightarrow \mathbb{R}^{n \times m}$ and $h : \mathbb{R}^n \times \mathbb{R} \rightarrow \mathbb{R}^p$ are class C^2 .

In the special case when $u(t) \equiv 0$ and $\delta u(t) \equiv 0$, we denote Σ_{gh} and $d\Sigma_{gh}$ by Σ_h and $d\Sigma_h$, respectively.

Note that the state trajectory of $d\Sigma_{gh}$ depends on that of Σ_{gh} . Let $\psi(\cdot, t_0, x_0, u)$ denote the state trajectory of Σ_{gh} starting from $x_0 \in \mathbb{R}^n$ for each choice of $u \in L_2^m[t_0, \infty)$. We also use ψ with only 3 arguments (a slight abuse of notation), i.e., $\psi(\cdot, t_0, x_0)$, if $u(t) \equiv 0$, to denote the state trajectory of Σ_h starting from $x_0 \in \mathbb{R}^n$. By using the state trajectory ψ , the variational part $d\Sigma_{gh}$ can be represented as

$$\delta \dot{x} = \frac{\partial(f(\psi, t) + g(\psi, t)u)}{\partial x} \delta x + g(\psi, t)\delta u.$$

Thus, the state trajectory of $d\Sigma_{gh}$ depends on state and input trajectories in addition to $(t, t_0, \delta x_0, \delta u)$. Since ψ depends on (t, t_0, x_0, u) , the state trajectory of $d\Sigma_{gh}$ depends on $(t, t_0, x_0, u, \delta x_0, \delta u)$.

Remark 2.1: For each $s \in [0, 1]$, let a class C^2 vector $\gamma(s)$ be an initial condition for Σ_{gh} and $\nu(\cdot, s)$ be an input signal. Note that $\gamma(s)$ can be seen as a curve connecting two states x_1 and x_2 , i.e. $\gamma(0) = x_1$ and $\gamma(1) = x_2$, and $\nu(t, s)$ can be seen as a curve connecting two inputs $u_1(t)$ and $u_2(t)$ at each t , i.e. $\nu(t, 0) = u_1(t)$ and $\nu(t, 1) = u_2(t)$. If $\nu(\cdot, \cdot)$ is class C^2 , then $\psi(\cdot, t_0, \gamma(s), \nu(t, s))$ is a solution to the system Σ_{gh} . Define $\delta\psi(t, t_0, \gamma(s), \nu(t, s)) := \partial\psi(t, t_0, \gamma(s), \nu(t, s)) / \partial s$ and $\delta\nu(t, s) := \partial\nu(t, s) / \partial s$. Then, $\delta\psi(\cdot, t_0, \gamma(s), \nu(t, s))$ is a solution to $d\Sigma_{gh}$ from the initial condition $\partial\gamma(s) / \partial s$ with the input $\partial\nu(t, s) / \partial s$. Also, the output signal of $d\Sigma_{gh}$ is given by $\delta y(t, t_0, s)$. Further on in this paper it will become clear from the distance definition in Definition 2.9 that $\delta\psi(\cdot, t_0, \gamma(s), \nu(s))$, $\delta\nu(\cdot, s)$, and $\delta y(\cdot, t_0, s)$ characterize distances of any pair of the state, input, and output, respectively. That is, the prolonged

system characterizes distances of any pair of the state, input, and output trajectories. $\triangleleft$

B. Differential Controllability and Observability Functions

In order to develop differential balancing theory, we first need to define the corresponding energy functions representing the controllability and observability functions in a similar way as is done in other balancing approaches, e.g. [3], [8]. We define the following two energy functions.

Definition 2.2: The differential controllability function of the prolonged system consisting of Σ_{gh} and $d\Sigma_{gh}$ is defined as

$$L_C(x_0, u, \delta x_0) := \inf_{u \in L_2^m(-\infty, t_0]} \frac{1}{2} \int_{-\infty}^{t_0} \|\delta u(t)\|^2 dt,$$

where $x(t_0) = x_0 \in \mathbb{R}^n$, $u \in L_2^m(-\infty, t_0]$, $\delta x(t_0) = \delta x_0 \in \mathbb{R}^n$ and $\delta x(-\infty) = 0$.

Definition 2.3: The differential observability function of the prolonged system consisting of Σ_h and $d\Sigma_h$ is defined as

$$L_O(x_0, \delta x_0) := \frac{1}{2} \int_{t_0}^{\infty} \|\delta y(t)\|^2 dt,$$

where $x(t_0) = x_0 \in \mathbb{R}^n$, $\delta x(t_0) = \delta x_0 \in \mathbb{R}^n$, $\delta x(\infty) = 0$.

Remark 2.4: The differential controllability function has a close relationship with the inverse time state space trajectories of Σ_{gh} and $d\Sigma_{gh}$. These inverse time state space trajectories are characterized as follows by Σ_g^- and $d\Sigma_g^-$

$$\begin{aligned} \Sigma_g^- : \dot{x}^-(t) &= -f(x^-(t), -t) \\ &\quad - g(x^-(t), -t)u^-(t), \quad x^-(t_0) := x_0, \\ d\Sigma_g^- : \dot{\delta x}^-(t) &= -\frac{\partial(f(x^-(t), -t) + g(x^-(t), -t)u^-(t))}{\partial x^-} \delta x^-(t) \\ &\quad - g(x^-(t), -t)\delta u^-(t), \quad \delta x^-(t_0) := \delta x_0. \end{aligned}$$

Now if we define the following energy function

$$L_C^-(x_0^-, u^-, \delta x_0^-) := \inf_{u^- \in L_2^m[-t_0, \infty)} \frac{1}{2} \int_{-t_0}^{\infty} \|\delta u^-(t)\|^2 dt,$$

where $x(-t_0) = x_0^- \in \mathbb{R}^n$, $u^- \in L_2^m[-t_0, \infty)$, $\delta x^-(-t_0) = \delta x_0^- \in \mathbb{R}^n$ and $\delta x^-(\infty) = 0$. From the definitions, $L_C^-(x_0^-, u^-, \delta x_0^-) = L_C(x_0, u, \delta x_0)$ if $x^-(t) \equiv x(-t)$, $u^-(t) \equiv u(-t)$, and $\delta x_0^- = \delta x_0$. For linear systems, the inverse time system definition of the controllability function L_C^- corresponds to the controllability Gramian, whereas L_C corresponds to the inverse of the controllability Gramian. The analysis is similar to the analysis in [6]. We will use the time reverse flow further on in the paper, and thus hereafter, we denote $\psi^-(\cdot, t_0, x_0^-, u^-)$ as the state trajectory of Σ_g^- starting from $x_0^- \in \mathbb{R}^n$ for $u^- \in L_2^m[t_0, \infty)$. $\triangleleft$

Note that the differential controllability or observability function can be viewed as the controllability or observability function for $d\Sigma_{gh}$ or $d\Sigma_h$ along the state and input trajectories of Σ_{gh} or state trajectories of Σ_h , respectively.

As mentioned for the controllability case in Remark 2.4, in the linear case, these two functions correspond to the controllability and observability Gramians, respectively. Similar to the linear case, the controllability and observability functions are related to Lyapunov type of equations. The differential controllability function is characterized by the following Lyapunov type of

equations (note that for ease of notation we leave out arguments when clear from the context).

Theorem 2.5: Let $P(x, t) : \mathbb{R}^n \times \mathbb{R} \rightarrow \mathbb{R}^{n \times n}$ be a solution to

$$\begin{aligned} -\delta_f(P(x, t)) + \frac{\partial f(x, t)}{\partial x} P(x, t) + P(x, t) \frac{\partial^T f(x, t)}{\partial x} \\ = -g(x, t)g^T(x, t), \end{aligned} \quad (1)$$

$$\begin{aligned} -\delta_{g_i}(P(x, t)) + \frac{\partial g_i(x, t)}{\partial x} P(x, t) + P(x, t) \frac{\partial^T g_i(x, t)}{\partial x} \\ = 0, \quad i = 1, 2, \dots, m, \end{aligned} \quad (2)$$

and let P be nonsingular, real symmetric class C^1 , and bounded for all $(x, t) \in \mathbb{R}^n \times \mathbb{R}$. Suppose that any trajectory $\delta x^-(t)$ of the following system

$$\begin{aligned} \dot{\delta x}^-(t) &= -\frac{\partial f(\psi^-, -t) + g(\psi^-, -t)u^-}{\partial x^-} \delta x^-(t) \\ &\quad - g(\psi^-, -t)g^T(\psi^-, -t)P^{-1}(\psi^-, -t)\delta x^-(t). \end{aligned} \quad (3)$$

is bounded and continuous for all $t \geq t_0$ and $\lim_{t \rightarrow \infty} \delta x^-(t) = 0$ along any bounded and continuous trajectory $(\psi^-(\cdot, t_0, x_0, u^-), u^-)$ of the inverse-time system Σ_g^- . Then, $L_C = (1/2)\delta x_0^T P^{-1}(x_0, t_0)\delta x_0$.

Proof: In this proof, we use

$$\frac{dP^{-1}}{dt} = -P^{-1} \left(\delta_f(P) + \sum_{i=1}^n \delta_{g_i}(P)u_i \right) P^{-1}, \quad (4)$$

which is obtained by expanding $d(P P^{-1})/dt = 0$.

Define

$$\bar{L}_C(x(t), \delta x(t), t) := \frac{1}{2} \delta x^T(t) P^{-1}(x(t), t) \delta x(t). \quad (5)$$

By differentiating $\bar{L}_C$ with respect to t along the state trajectories of the systems Σ_{gh} and $d\Sigma_{gh}$, we have, from (1), (2), and (4),

$$\begin{aligned} 2 \frac{d\bar{L}_C}{dt} &= \delta x^T P^{-1} \left(-\delta_f(P) + P \frac{\partial^T f}{\partial x} + \frac{\partial f}{\partial x} P \right) P^{-1} \delta x \\ &\quad + \sum_{i=1}^m u_i \delta x^T P^{-1} \left(-\delta_{g_i}(P) + P \frac{\partial^T g_i}{\partial x} + \frac{\partial g_i}{\partial x} P \right) P^{-1} \delta x \\ &\quad + \delta u^T g^T P^{-1} \delta x + \delta x^T P^{-1} g \delta u \\ &= -\|g^T P^{-1} \delta x\|^2 + \delta u^T g^T P^{-1} \delta x + \delta x^T P^{-1} g \delta u \\ &= \|\delta u\|^2 - \|\delta u - g^T P^{-1} \delta x\|^2, \end{aligned}$$

and thus

$$\begin{aligned} \int_{-\infty}^{t_0} \frac{d\bar{L}_C(x(t), \delta x(t), t)}{dt} dt \\ = \frac{1}{2} \int_{-\infty}^{t_0} \|\delta u\|^2 - \|\delta u - g^T P^{-1} \delta x\|^2 dt \leq \frac{1}{2} \int_{-\infty}^{t_0} \|\delta u(t)\|^2 dt. \end{aligned}$$

Hence, $\int_{-\infty}^{t_0} (d\bar{L}_C/dt) dt$ is a lower bound of L_C . It is clear that L_C is equal to this lower bound when $\delta u(t) = g^T(x(t), t) P^{-1}(x(t), t) \delta x(t)$. Thus, we consider $d\Sigma_{gh}$ controlled by this input. From the stability assumption for (3), we have $\delta x^-(\infty) = \delta x(-\infty) = 0$, which implies $\bar{L}_C(x(-\infty),$

$\delta x(-\infty), -\infty) = 0$. In summary, we have

$$\begin{aligned} L_C &= \inf_{\delta u \in L_2(-\infty, t_0]} \frac{1}{2} \int_{-\infty}^{t_0} \|\delta u(t)\|^2 dt \\ &= \int_{-\infty}^{t_0} \frac{d\bar{L}_C(x(t), \delta x(t), t)}{dt} dt = \bar{L}_C(x_0, \delta x_0, t_0). \end{aligned}$$

That completes the proof. $\blacksquare$

In this theorem, we assumed a kind of stabilizability of the inverse-time system $d\Sigma_g^-$. This is a standard assumption for analysis of the input energy function, and a similar assumption is found in e.g., [6]. It corresponds to the controllability assumption for linear systems, and can be seen as an asymptotic reachability assumption. For further elaboration on this for classical nonlinear balancing, and the relation with strong accessibility, we refer to [22]. In the differential balancing case asymptotic reachability and strong accessibility are equivalent. We use asymptotic reachability for accessibility analysis of system Σ_{gh} in the proof of Lemma A.2.

Equation (1) is a generalization of the Lyapunov equation for the controllability Gramian of a linear system, and in the linear case, (2) vanishes. Equation (2) guarantees that u does not appear in the time derivative of $\delta x^T P^{-1}(x, t) \delta x$. That is, the differential controllability function is independent from u .

According to its definition, the differential controllability function depends on the input trajectory u . The input trajectory is a function of the time. If u is determined by a state feedback controller, this also depends on the states. When we study the differential controllability function depending on $u(x, t)$, the following corollary of Theorem 2.5 can be used.

Corollary 2.6: Suppose that any trajectory $\delta x^-(t)$ of (3) is bounded and continuous for all $t \geq t_0$ and $\lim_{t \rightarrow \infty} \delta x^-(t) = 0$ along a bounded and continuous trajectory $(\psi^-(\cdot, t_0, x_0, u^-), u^-)$ of the inverse-time systems Σ_g^- , where $u^-(x^-, t) := u(x^-, -t)$ for given $u(x, t)$.

Let $P(x, t) : \mathbb{R}^n \times \mathbb{R} \rightarrow \mathbb{R}^{n \times n}$ be a solution¹ to

$$\begin{aligned} & -\delta_f(P(x, t)) - \sum_{i=1}^m u_i(x, t) \delta_{g_i}(P(x, t)) \\ & + \frac{\partial f(x, t)}{\partial x} P(x, t) + P(x, t) \frac{\partial^T f(x, t)}{\partial x} \\ & + \sum_{i=1}^m u_i(x, t) \left(\frac{\partial g_i(x, t)}{\partial x} P(x, t) + P(x, t) \frac{\partial^T g_i(x, t)}{\partial x} \right) \\ & = -g(x, t) g^T(x, t), \end{aligned} \quad (6)$$

and nonsingular, real symmetric class C^1 , and bounded for all $x \in \mathbb{R}^n$ and for all $t \in \mathbb{R}$. Then, $L_C = (1/2) \delta x_0^T P^{-1}(x_0, t_0) \delta x_0$. $\triangleleft$

Note that if $u \equiv 0$, (6) is nothing but (1).

In a similar manner, the differential observability function is characterized by a Lyapunov type of equation as follows. Naturally (see e.g., [6]–[8]), the results follow somewhat more straightforwardly, since we do not have to consider the time reverse system.

Theorem 2.7: Suppose that any trajectory $\delta x(t)$ of $d\Sigma_h$ is bounded and continuous for all $t \geq t_0$ and $\lim_{t \rightarrow \infty} \delta x^-(t) = 0$

along a bounded and continuous trajectory $\psi(\cdot, t_0, x_0)$ of Σ_h . Let $Q(x, t) : \mathbb{R}^n \times \mathbb{R} \rightarrow \mathbb{R}^{n \times n}$ be a solution to

$$\begin{aligned} & \delta_f(Q(x, t)) + \frac{\partial^T f(x, t)}{\partial x} Q(x, t) + Q(x, t) \frac{\partial f(x, t)}{\partial x} \\ & = -\frac{\partial^T h(x, t)}{\partial x} \frac{\partial h(x, t)}{\partial x}, \end{aligned} \quad (7)$$

and real symmetric, class C^1 , and bounded for all $(x, t) \in \mathbb{R}^n \times \mathbb{R}$. Then $L_O = (1/2) \delta x_0^T Q(x_0, t_0) \delta x_0$.

Proof: Let $\bar{L}_O(x_0, \delta x_0, t_0) := (1/2) \delta x_0^T Q(x_0, t_0) \delta x_0$. By differentiating $\bar{L}_O(x(t), \delta x(t), t)$ with respect to t along the state trajectories of the systems Σ_h and $d\Sigma_h$, we have, from (7),

$$\begin{aligned} 2 \frac{d\bar{L}_O}{dt} &= \delta x^T \left(\delta_f(Q) + \frac{\partial^T f}{\partial x} Q + Q \frac{\partial f}{\partial x} \right) \delta x \\ &= -\delta x^T \frac{\partial^T h}{\partial x} \frac{\partial h}{\partial x} \delta x = -\|\delta y\|^2, \end{aligned}$$

which implies that

$$\begin{aligned} L_O &= \frac{1}{2} \int_{t_0}^{\infty} \|\delta y(t)\|^2 dt \\ &= - \int_{t_0}^{\infty} \frac{d\bar{L}_O(x(t), \delta x(t), t)}{dt} dt \\ &= \bar{L}_O(x_0, \delta x_0, t_0) - \bar{L}_O(x(\infty), \delta x(\infty), \infty). \end{aligned}$$

From the assumption for Σ_h and $d\Sigma_h$, we have $\delta x(\infty) = 0$, and consequently $\bar{L}_O(x(\infty), \delta x(\infty), \infty) = 0$. Therefore, $L_O = \bar{L}_O(x_0, \delta x_0, t_0)$. $\blacksquare$

In Theorems 2.5 and 2.7, we assume that $P(x, t)$ and $Q(x, t)$ are non-singular for all $x \in \mathbb{R}^n$ and $t \in \mathbb{R}$. From Definitions 2.2 and 2.3, it is clear that the differential controllability and observability functions are non-negative. Thus, $P(x, t)$ and $Q(x, t)$ satisfying the conditions in the theorems are positive definite for all $x \in \mathbb{R}^n$ and $t \in \mathbb{R}$.

Remark 2.8: If we consider linear systems, the variational system is the same as the original linear system. Thus, the characterizing Lyapunov type of equations of Theorem 2.5 and 2.7 reduce to the standard Lyapunov equations for the controllability and observability Gramians of linear systems. $\triangleleft$

C. Controllability Analysis for Time-Invariant Systems

With the Lyapunov type of equations characterizing the differential controllability and observability functions, we are now ready to relate the existence of these functions to controllability (strong accessibility), observability and incremental stability for *time invariant systems*. In order to make this paper self contained, we provide the definitions of distance and incremental stability here.

Definition 2.9: Let $\Gamma(x_1, x_2)$ be the set of all smooth paths $\gamma(s) : \mathbb{R} \supset [0, 1] \rightarrow \mathbb{R}^n$ connecting x_1 and x_2 . The distance between two points is defined as

$$d(x_1, x_2) = \inf_{\Gamma(x_1, x_2)} \int_0^1 D \left(\gamma(s), \frac{d\gamma(s)}{ds} \right) ds,$$

where D is a Finsler function [23].

In this paper, we consider Finsler functions in the following form, i.e., Riemannian metric $D := \sqrt{\delta x^T X \delta x}$ for a positive definite $X(x)$. Thanks to the Hopf-Rinow theorem [23], there

¹A solution to (6) depends on u but u is a function of x and t . Thus, the solution is a function of x and t .

always exists a minimizing path $\gamma(s)$ with respect to $X(x)$. This minimizing path is not necessarily smooth but piecewise smooth. Throughout the paper, we assume that the minimizing path is smooth.

Definition 2.10: [16] Let $\psi(\cdot, t_0, x_0)$ be a solution to the autonomous system Σ_h starting from the initial condition $x(0) = x_0 \in \mathbb{R}^n$. Let $d: \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}_{\geq 0}$ be a continuous distance. The autonomous system Σ_h is said to be incrementally globally stable if there exists a class $\mathcal{K}_\infty$ function α such that $\forall x_1, x_2 \in \mathbb{R}^n, \forall t \geq 0, d(\psi(t, t_0, x_1), \psi(t, t_0, x_2)) \leq \alpha(d(x_1, x_2))$. Moreover, an incrementally globally stable system Σ_h is said to be incrementally globally asymptotically stable if $\lim_{t \rightarrow \infty} d(\psi(t, t_0, x_1), \psi(t, t_0, x_2)) = 0$.

Note that the system Σ_h is incrementally globally asymptotically stable if and only if $\delta x(t)$ of $d\Sigma_h$ is bounded and continuous and $\lim_{t \rightarrow \infty} \delta x(t) = 0$ for all initial states $\delta x(0) = \delta x_0 \in \mathbb{R}^n$.

As mentioned before, the controllability function is related to asymptotic reachability and strong accessibility for nonlinear systems. For this reason, we use the concept of strong accessibility, which is a well known concept for nonlinear systems, e.g., [24].

Hereafter, to study the relationship between local strong accessibility and the time-invariant solution to (1) and (2), we assume the existence of a solution to (1) and (2), i.e.,

Assumption 2.11: Equations (1) and (2) have a real symmetric and class C^1 solution $P(x)$.

In order to be able to use the notion of strong accessibility, we also make the rather standard assumption that the strong accessibility distribution [24] has a constant dimension.

Assumption 2.12: For the system Σ_{gh} , f and g are analytic, and its strong accessibility distribution has a constant dimension for all $x \in \mathbb{R}^n$.

Under these assumptions, we have the following theorem, which is the nonlinear differential extension of the relation between controllability, and existence and positivity of the controllability Gramian for linear systems, e.g., [4].

Theorem 2.13: Suppose that Assumptions 2.11 and 2.12 hold, the system Σ_h is incrementally globally asymptotically stable, and the system Σ_h has a nonempty α limit set L^- . Then, the system Σ_{gh} is locally strongly accessible if and only if $0 < P(x) < \infty$ at each $x \in \mathbb{R}^n$.

Proof: Sufficiency immediately follows from Lemmas A.2 and A.4. We show the necessity. If Σ_{gh} is locally strongly accessible, then the strong accessibility rank condition holds for all $x \in \mathbb{R}^n$ under Assumption 2.12 [24]. Therefore, $d\Sigma_{gh}$ satisfies the Kalman-like controllability rank condition for linear time-varying systems [25] along any feasible solution $x(t)$ of Σ_{gh} with $u(t) \equiv 0$, i.e., $d\Sigma_{gh}$ can be considered as a linear time-varying system and is controllable, [25]. Therefore, controllability Gramian $C(t_0, t)$ for linear time-varying system $d\Sigma_{gh}$ satisfies $C(t_0, t) > 0$ along any feasible solution with $u(t) \equiv 0$ [26]. The stability assumption of Σ_h implies that any trajectory Σ_{gh} is feasible and $\lim_{t \rightarrow \infty} C(t_0, t)$ exists if $u(t) \equiv 0$. From Theorem 2.5, $\lim_{t \rightarrow \infty} C(t_0, t) = P(x(0))$ for any feasible solution $(x(t), u(t))$. Therefore, $0 < P(x) < \infty$ at each $x \in \mathbb{R}^n$. ■

See [22] for the sufficiency result of the non-differential case. In contrast to the differential case, for the non-differential case a necessary condition for local strong accessibility in terms of the controllability function has never been derived.

Note that the incremental stability assumption is only made on the autonomous system, as is also standard for the non-differential case, e.g., [6]–[8]. If the system Σ_h has an equilibrium point, this is contained in the α limit set L^- .

The above theorem implies that local strong accessibility of Σ_{gh} can be verified by using the differential controllability function. According to the proof, the strong accessibility distribution of Σ_{gh} corresponds to the Kalman-like controllability rank condition of $d\Sigma_{gh}$ as a linear time-varying system along the trajectory of Σ_{gh} , which implies that the control effort of the nonlinear system Σ_{gh} can be evaluated by using that of its variational system $d\Sigma_{gh}$.

According to the proof of Lemma A.2, if $0 < P(x) < \infty$ then the stability assumption for (3) in Theorem 2.5 holds under mild assumptions. Moreover, that stability assumption implies incremental global asymptotic stabilizability of the time-reverse system Σ_g^- , and consequently asymptotic stability of (3). Thus, now asymptotic stability of (3) is connected to positive definiteness of $P(x)$.

Next, the connection between positive definiteness of $P(x)$ and incremental asymptotic stability of Σ_h is established. The proof is given in Appendix A.

Theorem 2.14: Suppose that Assumptions 2.11 and 2.12 hold, the system Σ_h has a nonempty α limit set L^- , and the inverse-time system Σ_g^- is incrementally globally asymptotically stabilizable with input (31) in Appendix A. Then, the system Σ_h is incrementally globally asymptotically stable if and only if $0 < P(x) < \infty$ at each $x \in \mathbb{R}^n$.

D. Observability Analysis in the Time-Invariant Case

Similar to controllability above, we now address the relation between the existence of the differential observability function, and well known local observability [24] and incremental stability for the differential nonlinear time-invariant case.

Hereafter, to study the relationship between local observability, a solution to (7), and the differential observability function, we assume the existence of such solution.

Assumption 2.15: Equation (7) has a real symmetric and class C^1 solution $Q(x)$.

In addition, in order to use the notion of nonlinear observability, we make the rather standard assumption that the observability codistribution [24] has a constant dimension.

Assumption 2.16: For the system Σ_{gh} , f, h, g are analytic, and its observability codistribution [24] has constant dimension for all $x \in \mathbb{R}^n$.

Under these assumptions, we are now ready to state the following theorems, which use Lemmas B.1, B.2 and B.3 from the Appendix. The theorems are the nonlinear differential extension of the relation between observability, stability, and existence and positivity of the observability Gramian for linear systems, e.g. [4].

Theorem 2.17: Suppose that Assumptions 2.15 and 2.16 hold and that Σ_h is locally observable. Then, system Σ_h is incrementally globally asymptotically stable if and only if $0 < Q(x) < \infty$ at each $x \in \mathbb{R}^n$. ■

Theorem 2.18: Suppose that Assumptions 2.15 and 2.16 hold, and the system Σ_h is incrementally globally asymptotically stable. Then, the system Σ_h is locally observable if and only if $0 < Q(x) < \infty$ at each $x \in \mathbb{R}^n$. ■

See [22] for the non-differential version of these results.

In summary, positivity of the differential controllability and observability functions characterize when a realization of a system is minimal, i.e.,

Theorem 2.19: Suppose that Assumptions 2.11, 2.12, 2.15 and 2.16 hold, the system Σ_h is incrementally globally asymptotically stable, and the system Σ_h has a nonempty α limit set L^- . The systems Σ_{gh} and Σ_h are respectively locally strongly accessible and locally observable if and only if $0 < P(x) < \infty$ and $0 < Q(x) < \infty$ hold at each $x \in \mathbb{R}^n$. ■

Finally, we can consider the system with input Σ_{gh} . It is immediate that local observability of Σ_h corresponds to local zero-state observability of Σ_{gh} , and since it is well-known that local zero-state observability implies local observability (e.g., [6], [22]), we obtain the following corollary.

Corollary 2.20: Suppose that Assumptions 2.11, 2.12, 2.15 and 2.16 hold, the system Σ_h is incrementally globally asymptotically stable, and the system Σ_h has a nonempty α limit set L^- . Then the system Σ_{gh} is locally strongly accessible and locally (zero-state) observable if $0 < P(x) < \infty$ and $0 < Q(x) < \infty$ at each $x \in \mathbb{R}^n$. ■

In fact, we can remove Assumption 2.16 from Corollary 2.20 because it has been shown in e.g., [24] that the observability codistribution of an analytic and locally accessible system Σ_{gh} has a constant dimension for all $x \in \mathbb{R}^n$.

III. DIFFERENTIALLY BALANCED REALIZATION

A. Balanced Realization and Model Reduction

The previous analysis was done for time-invariant systems, because this allows us to study some properties in relation of our differential controllability and observability functions to strong accessibility, observability and incremental stability. Next, we will only use the differential controllability and observability functions to define a differentially balanced realization. Therefore, we allow the system to be *time-variant* again until further notice. Similar to other linear and nonlinear balancing approaches, we define a differentially balanced realization for the system Σ_{gh} in terms of the differential controllability and observability functions.

Definition 3.1: A realization of the associated system $d\Sigma_{gh}$ is said to be a differentially balanced realization on an open subset $D \subset \mathbb{R}^n \times \mathbb{R}$ if there exists a diagonal matrix

$$\Lambda(x, t) = \text{diag}\{\sigma_1(x, t), \sigma_2(x, t), \dots, \sigma_n(x, t)\}, \quad (8)$$

where $\sigma_1(x, t) \geq \sigma_2(x, t) \geq \dots \geq \sigma_n(x, t) > 0$ holds on D , and $P(x, t) = \Lambda(x, t)$ and $Q(x, t) = \Lambda(x, t)$ respectively satisfy (1), (2) and (7).

In the linear case, the differentially balanced realization is nothing but the balanced realization. It is well known that a linear system having positive definite controllability and observability Gramians, can be transformed into a balanced realization. Similar to the linear case, we have the following theorem in the nonlinear differential case.

Theorem 3.2: Let $P(x, t)$ and $Q(x, t)$ be respectively real symmetric and class C^1 solutions to (1), (2) and (7), where $0 < P(x, t) < \infty$ and $0 < Q(x, t) < \infty$ for all $(x, t) \in \mathbb{R}^n \times \mathbb{R}$. The system $d\Sigma_{gh}$ can be transformed into a differentially balanced realization on an open subset $D \subset \mathbb{R}^n \times \mathbb{R}$ by a differential coordinate transformation $\omega = T(x, t)\delta x$, where $T(x, t)$ is class C^1 and nonsingular for all $(x, t) \in \mathbb{R}^n \times \mathbb{R}$. Moreover, $\sigma_i^2(x, t)$ ($i = 1, \dots, n$) in (8) are the eigenvalues of the product $P(x, t)Q(x, t)$.

Proof: In a similar manner as for the linear case, it can be shown that there exists a class C^1 and invertible matrix $T(x, t) : \mathbb{R}^n \times \mathbb{R} \rightarrow \mathbb{R}^{n \times n}$ which achieves

$$\begin{aligned} T(x, t)P(x, t)T^T(x, t) &= T^{-T}(x, t)Q(x, t)T^{-1}(x, t) \\ &= \Lambda(x, t), \end{aligned}$$

where $\Lambda(x, t) = \text{diag}\{\sigma_1(x, t), \dots, \sigma_n(x, t)\}$, and $\sigma_i(x, t) > 0$ ($i = 1, \dots, n$). Moreover, $T(x, t)$ can be chosen such that $\sigma_1(x, t) \geq \dots \geq \sigma_n(x, t)$ in a sufficiently small open subset $D \subset \mathbb{R}^n \times \mathbb{R}$. Next, $P(x, t)Q(x, t) = T^{-1}(x, t)\Lambda^2(x, t)T(x, t)$ implies that $\sigma_i^2(x, t)$ ($i = 1, \dots, n$) are eigenvalues of $P(x, t)Q(x, t)$. After a coordinate transformation $\omega = T(x, t)\delta x$, the variational system $d\Sigma_{gh}$ becomes

$$\begin{cases} \dot{\omega} = \left(\delta_{f+gu}(T) + T \frac{\partial(f+gu)}{\partial x} \right) T^{-1}\omega + Tg\delta u, \\ \delta y = \frac{\partial h}{\partial x} T^{-1}\omega, \end{cases} \quad (9)$$

We now have to show that $\Lambda(x, t)$ is a solution to (1), (2) and (7) for this system. Here, we only prove it for (1), as the other proofs follow similarly. By multiplying T from the left and T^T from the right to (1), we have

$$-T\delta_f(P)T^T + T \frac{\partial f}{\partial x} P T^T + T P \left(\frac{\partial f}{\partial x} \right)^T T^T = -Tg g^T T^T.$$

The left hand side can be computed as

$$\begin{aligned} & -\delta_f(TPT^T) + \delta_f(T)PT^T + TP\delta_f(T)^T \\ & + T \frac{\partial f}{\partial x} P T^T + T P \left(\frac{\partial f}{\partial x} \right)^T T^T \\ & = -\delta_f(\Lambda) + \left(T \frac{\partial f}{\partial x} + \delta_f(T) \right) T^{-1}\Lambda \\ & + \Lambda T^{-T} \left(T \frac{\partial f}{\partial x} + \delta_f(T) \right)^T. \end{aligned}$$

Thus, Λ satisfies (1) for (9). ■

In order to obtain the balanced realization in this theorem, we require $P(x, t)$ satisfies (1) and (2), i.e. the differential controllability function is independent from u . This $P(x, t)$ does not always exist, but in that case, we may consider to use a solution to (6) instead. However, in case we are interested in the relationship between the solution to (6) and local strong accessibility of the original system such as we study for time-invariant systems in Section II.B, it should be noted that at present we are unable to establish such relation.

We consider model reduction based on the differentially balanced realization. The differentially balanced realization is obtained by using the differential coordinate transformation $w = T(x, t)\delta x$ with $T(x, t)$ in Theorem 3.2, and the order of the balanced variational system can be reduced accordingly. To achieve model reduction for the original system Σ_{gh} , the corresponding coordinate transformation for Σ_{gh} is also needed. This coordinate transformation can be constructed by using a path integral.

Let us consider a pair of reference state and input trajectories $(x^*, u^*) := (\psi(\cdot, t_0, x_0^*, u^*), u^*)$ of Σ_{gh} . Suppose that this is continuous and bounded. Next, consider a path $\gamma(\cdot)$ connecting x_0^* and x_0 parametrized by $s \in [0, 1]$, i.e., $\gamma(0) = x_0^*$ and $\gamma(1) = x_0$, where x_0^* is fixed, and a path $\nu(t, \cdot)$ connecting $u^*(t)$ and

$u(t)$, where $u^*(t)$ is fixed, parametrized by $s \in [0, 1]$. Suppose that $\nu(\cdot, \cdot)$ is class C^2 . Then, the state trajectory of Σ_{gh} starting from $\gamma(s)$ controlled by $\nu(t, s)$ is $\psi(\cdot, t_0, \gamma(s), \nu(t, s))$. Based on $w = T(x, t)\delta x$, define

$$\begin{aligned} z(t) &= \varphi(x(t), t) \\ &:= \int_0^1 T(\psi, t) \frac{\partial \psi(\cdot, t_0, \gamma(s), \nu(t, s))}{\partial s} ds. \end{aligned} \quad (10)$$

Although it seems that $x(t)$ does not appear in the right hand side of (10), it does. This is because from $\gamma(s)$ connects x_0^* and x_0 , and $\nu(t, s)$ connects $u^*(t)$ and $u(t)$, it follows that $\psi(\cdot, t_0, \gamma(s), \nu(t, s))$ is a path connecting $x^*(t)$ and $x(t)$, where the reference trajectory $x^*(t)$ is given as a function of the time. Thus, the right hand side is a function of $x(t)$, i.e. φ is a function of $x(t)$.

Suppose that $\varphi(x(t), t)$ is invertible. Then, denote the system in the z -coordinate as

$$\Sigma_{\hat{g}\hat{h}} \begin{cases} \dot{z}(t) = \hat{f}(z(t), t) + \hat{g}(z(t), t)u(t), \\ y = \hat{h}(z(t), t). \end{cases}$$

Its differential dynamics is

$$d\Sigma_{\hat{g}\hat{h}} \begin{cases} \delta \dot{z}(t) = \frac{\partial \hat{f}(z(t), t) + \hat{g}(z(t), t)u(t)}{\partial z} \delta z(t) \\ \quad + \hat{g}(z(t), t)\delta u(t), \\ \delta y(t) = \frac{\partial \hat{h}(z(t), t)}{\partial z}. \end{cases}$$

We are now ready to state that the Σ_{gh} and $d\Sigma_{gh}$ in the $z, \delta z$ coordinates are differentially balanced, i.e.,

Theorem 3.3: In the z -coordinates, the system is differentially balanced along path $\psi(t, t_0, \gamma(\cdot), \nu(t, \cdot))$.

The proof is given in Appendix C.

In general, the z -coordinate depends on fixed reference $x^*(t)$ for all $x(t)$ and path γ . These are design parameters because for every x^* and γ there exists a balanced realization. The reference trajectory $x^*(t)$ is fixed for any flow of the system $x(t)$, and not $x^*(t)$ but $x(t)$ corresponds to trajectories used to compute the differential controllability and observability functions. The differential controllability and observability functions are functions of x_0 , and thus we compute these functions for all x_0 . In the computation of differential controllability and observability functions (and Hankel singular value analysis), the reference trajectory $x^*(t)$ does not appear. The reference trajectory $x^*(t)$ is used only for constructing the coordinate transformation to obtain the reduced model which gives a good approximation of the path connecting any $x(t)$ and a fixed $x^*(t)$. We may choose the reference trajectory as an element in the α limit set. In the linear case, $x^*(t) \equiv 0$ and the path is a straight line. In the nonlinear case, the α limit set is not necessarily the set of equilibrium points.

If there exists a $\varphi(x, t)$ such that

$$\frac{\partial \varphi(x, t)}{\partial x} = T(x, t), \quad (11)$$

then the coordinate transformation $z = \varphi(x, t)$ is independent from the path. It only depends on the reference trajectory x^* , which is a design parameter.

In (8), suppose that $\sigma_k(x, t) > \sigma_{k+1}(x, t)$ for $k < n$, which implies that z_k is more important than z_{k+1} in the sense of

differential energy, i.e., differential observability and controllability. Hence, z_1 until z_k are more important than z_{k+1} until z_n . We partition the system in the z -coordinates correspondingly as follows:

$$\begin{aligned} \hat{f}(z, t) &= \begin{bmatrix} \hat{f}_a(z_a, z_b, t) \\ \hat{f}_b(z_a, z_b, t) \end{bmatrix} \\ &:= T(\varphi^{-1}(z, t), t)f(\varphi^{-1}(z, t), t), \\ \hat{g}(z, t) &= \begin{bmatrix} \hat{g}_a(z_a, z_b, t) \\ \hat{g}_b(z_a, z_b, t) \end{bmatrix} \\ &:= T(\varphi^{-1}(z, t), t)g(\varphi^{-1}(z, t), t), \\ \hat{h}(z_a, z_b, t) &:= h(\varphi^{-1}(z, t), t), \end{aligned}$$

where $z_a := [z_1, \dots, z_k]^T$ and $z_b := [z_{k+1}, \dots, z_n]^T$.

Its reduced order system is

$$\Sigma_{\hat{g}\hat{h}}^r \begin{cases} \dot{\hat{z}}_a(t) = \hat{f}_a(\hat{z}_a(t), 0, t) + \hat{g}_a(\hat{z}_a(t), 0, t)u(t) \\ \hat{y}_a(t) = \hat{h}(\hat{z}_a(t), 0, t) \end{cases}.$$

In a similar manner as with traditional balancing as defined in [6], under a triangular structure assumption (i.e., the large singular values do not depend on the coordinates corresponding to the small singular values), we have the following theorem for the reduced order system.

Theorem 3.4: Assume that $\sigma_i(\varphi^{-1}(z_a, z_b, t), t) \equiv \sigma_i(\varphi^{-1}(z_a, 0, t), t)$, $i = 1, \dots, k$. Then, the reduced order system $\Sigma_{\hat{g}\hat{h}}^r$ is differentially balanced along the path $\psi(t, t_0, \gamma(\cdot), u(t, \cdot))$ with

$$\hat{\Lambda}_k(\hat{z}_a, t) := \text{diag}\{\sigma_1(\varphi^{-1}(\hat{z}_a, 0, t), t), \dots, \sigma_k(\varphi^{-1}(\hat{z}_a, 0, t), t)\},$$

where $\sigma_1(\varphi^{-1}(\hat{z}_a, 0, t), t) \geq \dots \geq \sigma_k(\varphi^{-1}(\hat{z}_a, 0, t), t)$.

The proof is given in Appendix C.

When differential balancing depends on the path, it is not clear if properties such as strong accessibility, observability and incremental stability of the original system are preserved in the reduced order system. However, in a special case we can say more, i.e., if we suppose the existence of $\varphi(x, t)$ such that (11) holds, then we have the following.

Theorem 3.5: Suppose that the original system Σ_{gh} satisfies the assumptions of Corollary 2.20. Suppose that there exists $\varphi(x, t)$ such that (11) holds, and the reduced order system satisfies $\hat{f}_a(z_a, z_b, t) \equiv \hat{f}_a(z_a, 0, t)$ and $\sigma_i(\varphi^{-1}(z_a, z_b, t), t) \equiv \sigma_i(\varphi^{-1}(z_a, 0, t), t)$, $i = 1, \dots, k$ in addition to Assumptions 2.12 and 2.16. Then, the reduced order model $\Sigma_{\hat{g}\hat{h}}^r$ is incrementally globally asymptotically stable, locally strongly accessible and locally observable.

The proof is given in Appendix C.

Remark 3.6: In Theorem 3.5, the triangular structure assumption for $\hat{f}$ is to guarantee incremental global asymptotic stability of the reduced order model $\Sigma_{\hat{g}\hat{h}}^r$. In the case that we know that $\Sigma_{\hat{g}\hat{h}}^r$ is incrementally globally asymptotically stable, we can remove this assumption. $\triangleleft$

Remark 3.7: In [8], under the asymptotic stability, controllability, and observability assumptions for the linearized system, it is shown that if all Hankel singular values of the linearized system are distinct, the Hankel singular values of the traditional nonlinear balancing of [6] locally satisfy a similar condition to that in Theorem 3.5. It is possible to

extend this to the differential balancing case, i.e., then locally $\sigma_i(\varphi^{-1}(z_a, z_b, t), t) \equiv \sigma_i(\varphi^{-1}(z_a, 0, t), t)$, and the assumption on the differential singular value function in Theorem 3.5 can be removed. Moreover, the reduced order nonlinear system by differential balancing at the origin is the reduced system of the linearized system by linear balancing, which is known to be an asymptotically stable linear system. Therefore, the reduced order nonlinear system is locally incrementally asymptotically stable. Then with Remark 3.6, we do not need to assume $\hat{f}_a(z_a, z_b, t) \equiv \hat{f}_a(z_a, 0, t)$.

B. On the Differential Hankel Operator for Time-Invariant Systems

For linear systems, the σ'_i s of differential balancing in (8) are nothing but the Hankel singular values. We now extend these results to the time-invariant differential balancing case. Hence, we now assume that Σ_{gh} is *time-invariant*. For this, we study the differential Hankel operator, as is introduced in [7]. In order to be self-contained, we first introduce the Fréchet derivative and the various operators.

Consider a nonlinear operator $\Sigma_{gh}(x_0, u) : \mathbb{R}^n \times L_2[a, b] \ni (x_0, u) \mapsto y \in L_2[a, b]$ defined by the system Σ_{gh} , where $[a, b] \subset \mathbb{R}^n$ is finite or infinite. A linear operator $d\Sigma_{gh}(x_0, u)(\delta x_0, \delta u)$ is said to be the Fréchet derivative if for each $x_0 \in \mathbb{R}^n$ and $u \in L_2[a, b]$, the following limit exists

$$d\Sigma_{gh}(x_0, u)(\delta x_0, \delta u) := \lim_{s \rightarrow 0} \frac{1}{s} (\Sigma_{gh}(x_0 + s\delta x_0, u + s\delta u) - \Sigma_{gh}(x_0, u))$$

for all $\delta x_0 \in \mathbb{R}^n$, $\delta u \in L_2[a, b]$. From its definition, the Fréchet derivative is given by the differential dynamics $d\Sigma_{gh}$.

To define the nonlinear Hankel operator, we define the controllability and observability operators [7], [27] of Σ_{gh} as follows

$$\begin{aligned} \mathcal{C} : L_2^m[0, \infty) &\rightarrow \mathbb{R}^n, \quad x_0 = \mathcal{C}(u), \\ \begin{cases} \dot{x}(t) = -f(x(t)) - g(x(t))u(t), & x(\infty) = 0 \\ x_0 = x(0) \end{cases}, \\ \mathcal{O} : \mathbb{R}^n &\rightarrow L_2^p[0, \infty), \quad y = \mathcal{O}(x_0), \\ \begin{cases} \dot{x}(t) = f(x(t)), & x(0) = x_0 \\ y(t) = h(x(t)) \end{cases}. \end{aligned}$$

Assume that $\mathcal{C}$ and $\mathcal{O}$ are Fréchet differentiable, whose state space representations are denoted by $d\mathcal{C}$ and $d\mathcal{O}$, respectively, and the linear adjoints of $d\mathcal{C}$ and $d\mathcal{O}$ are denoted by $(d\mathcal{C})^*$ and $(d\mathcal{O})^*$, respectively. They are obtained as follows [7], [27].

$$\begin{aligned} d\mathcal{C} : L_2^m[0, \infty) \times L_2^m[0, \infty) &\rightarrow \mathbb{R}^n, \quad v_0 = d\mathcal{C}(u)(u_p), \\ \begin{cases} \dot{x}(t) = -f(x(t)) - g(x(t))u(t), & x(\infty) = 0 \\ \dot{v}(t) = -\frac{\partial(f(x(t)) + g(x(t))u(t))}{\partial x} v(t) - g(x(t))u_p(t), \\ v(\infty) = 0, \quad v_0 = v(0) \end{cases}, \\ d\mathcal{O} : \mathbb{R}^n \times \mathbb{R}^n &\rightarrow L_2^p[0, \infty), \quad y_p = d\mathcal{O}(x_0)(v_0), \\ \begin{cases} \dot{x}(t) = f(x(t)), & x(0) = x_0 \\ \dot{v}(t) = \frac{\partial f(x(t))}{\partial x} v(t), & v(0) = v_0, \\ y_p(t) = \frac{\partial h(x(t))}{\partial x} v(t) \end{cases}, \end{aligned}$$

$$(d\mathcal{C})^* : L_2^m[0, \infty) \times \mathbb{R}^n \rightarrow L_2^m[0, \infty), \quad y_a = (d\mathcal{C}(u))^*(p_0),$$

$$\begin{cases} \dot{x}(t) = -f(x(t)) - g(x(t))u(t), & x(\infty) = 0 \\ \dot{p}(t) = \frac{\partial^T(f(x(t)) + g(x(t))u(t))}{\partial x} p(t), & p(0) = p_0, \\ y_a(t) = g^T(x(t))p(t) \end{cases},$$

$$(d\mathcal{O})^* : \mathbb{R}^n \times L_2^p[0, \infty) \rightarrow \mathbb{R}^n, \quad p_0 = (d\mathcal{O}(x_0))^*(u_a),$$

$$\begin{cases} \dot{x}(t) = f(x(t)), & x(0) = x_0 \\ \dot{p}(t) = -\frac{\partial^T f(x(t))}{\partial x} p(t) - \frac{\partial^T h(x(t))}{\partial x} u_a(t), \\ p(\infty) = 0, \quad p_0 = p(0) \end{cases}.$$

The Hankel operator associated to Σ_{gh} [7] is defined as $\mathcal{H} : L_2^m[0, \infty) \rightarrow L_2^p[0, \infty)$, $\mathcal{H}(u) := \mathcal{O} \circ \mathcal{C}(u)$. That is, the Hankel operator is the observability operator with the initial state $x_0 := \mathcal{C}(u)$. Throughout this subsection, we use this initial condition. Since $\mathcal{C}$ and $\mathcal{O}$ are Fréchet differentiable, $\mathcal{H}$ is also, which we denote by $d\mathcal{H}$. According to [7], $d\mathcal{H}$ satisfies

$$d\mathcal{H}(u)(u_p) = d\mathcal{O}(\mathcal{C}(u)) \circ (d\mathcal{C}(u)(u_p)), \quad (12)$$

In this subsection, our aim is to study relationships between singular values of $d\mathcal{H}(u)(u_p)$ and differential balancing. For that, we need to define a differential Hankel singular value in relation to the corresponding differential Hankel operator.

Definition 3.8: Function $\lambda : \mathbb{R}^n \rightarrow \mathbb{R}_{\geq 0}$ and vector valued function $v : \mathbb{R}^n \rightarrow \mathbb{R}^n$ are respectively said to be a differential Hankel singular value and Hankel singular vector of the system Σ_{gh} if

$$(d\mathcal{H}(u))^* \circ d\mathcal{H}(u) \circ v(\mathcal{C}(u)) = \lambda^2(\mathcal{C}(u)) \cdot v(\mathcal{C}(u)).$$

Related to this, we follow [7] to define the differential Hankel norm.

Definition 3.9: The differential Hankel norm of the system Σ_{gh} is

$$\|d\Sigma\|_{d\mathcal{H}} := \sup_{\substack{\delta u \in L_2^m[0, \infty) \\ \delta u \neq 0}} \frac{\|d\mathcal{H}(u)(\delta u)\|_{L_2^p[0, \infty)}}{\|\delta u\|_{L_2^m[0, \infty)}}.$$

In general, the differential Hankel norm depends on the input trajectory u . As will be shown in Theorem 3.15, the differential Hankel norm as well as the differential controllability function does not depend on u if the solution $P(x, t)$ to (1) and (2) exists.

We use some properties of the controllability, observability and Hankel operators obtained from [7], i.e.,

Lemma 3.10: [7] For the controllability, observability and Hankel operators it holds that

$$\langle d\mathcal{C}(u)(u_p), p_0 \rangle_{\mathbb{R}^n} = \langle u_p, (d\mathcal{C}(u))^*(p_0) \rangle_{L_2^m}, \quad (13)$$

$$\langle d\mathcal{O}(x_0)(v_0), u_a \rangle_{L_2^p} = \langle v_0, (d\mathcal{O}(x_0))^*(u_a) \rangle_{\mathbb{R}^n}, \quad (14)$$

$$(d\mathcal{H}(u))^*(u_a) = (d\mathcal{C}(u))^* \circ (d\mathcal{O}(\mathcal{C}(u)))^*(u_a). \quad (15)$$

In order to establish a relation between the differential Hankel operator and the differential singular values, we need to have some boundedness properties, i.e.,

Assumption 3.11: The Fréchet derivatives of controllability and observability operators $d\mathcal{C}(u)(u_p)$ and $d\mathcal{O}(x_0)(v_0)$ are bounded.

Now we can establish a relation between the controllability operator and the differential controllability Gramian as follows.

Lemma 3.12: Under the assumptions in Theorem 2.5 and Assumption 3.11,

$$d\mathcal{C}(u) \circ (d\mathcal{C}(u))^*(p_0) = P(x_0)p_0. \quad (16)$$

Proof: First, we introduce the observability function of $(d\mathcal{C})^*$.

$$2L_{\mathcal{C}}^*(x_0, p_0) := \|y_a(t)\|_{L_2^p[0, \infty)}, \quad (17)$$

where $x(0) = x_0 = \mathcal{C}(u) \in \mathbb{R}^n$, and $p(0) = p_0 \in \mathbb{R}^n$. We next aim to show that

$$2L_{\mathcal{C}}^*(x_0, p_0) = p_0^T P(x_0)p_0 = \langle P(x_0)p_0, p_0 \rangle_{\mathbb{R}^n}. \quad (18)$$

Let $\bar{L}_{\mathcal{C}}^*(x(t), p(t)) := (1/2)p^T(t)P(x(t))p(t)$. By differentiating $\bar{L}_{\mathcal{C}}^*$ with respect to t along the trajectories of $(d\mathcal{C})^*$, we have, from (1) and (2), $\frac{d\bar{L}_{\mathcal{C}}^*}{dt} = -\frac{1}{2}p^T g(x)g^T(x)p$. By integration over time, we have

$$\begin{aligned} L_{\mathcal{C}}^*(x_0, p_0) &= \frac{1}{2} \int_0^\infty p^T(t)g(x(t))g^T(x(t))p(t)dt \\ &= - \int_0^\infty \frac{d\bar{L}_{\mathcal{C}}^*(x(t), p(t))}{dt} dt = \bar{L}_{\mathcal{C}}^*(x_0, p_0), \end{aligned}$$

where $p(\infty) = 0$ is used, which holds as a consequence of (13) and Assumption 3.11. Therefore, (18) holds.

Next, from (17), we have $2L_{\mathcal{C}}^*(\mathcal{C}(u), p_0) = \langle (d\mathcal{C}(u))^*(p_0), (d\mathcal{C}(u))^*(p_0) \rangle_{L_2^p}$. By substituting $u_p = (d\mathcal{C}(u))^*(p_0)$ into (13), $2L_{\mathcal{C}}^*(\mathcal{C}(u), p_0) = \langle d\mathcal{C}(u) \circ (d\mathcal{C}(u))^*(p_0), p_0 \rangle_{\mathbb{R}^n}$. From $x_0 = \mathcal{C}(u)$ and (18), we have (16). ■

Similarly, or dually, a relationship between the observability operator and the differential observability Gramian is obtained as follows.

Lemma 3.13: Under the assumptions in Theorem 2.7 and Assumption 3.11,

$$(d\mathcal{O}(x_0))^* \circ d\mathcal{O}(x_0)(v_0) = Q(x_0)v_0. \quad (19)$$

Proof: From Definition 2.3, we have $2L_{\mathcal{O}}(x_0, v_0) = \langle d\mathcal{O}(x_0)(v_0), d\mathcal{O}(x_0)(v_0) \rangle_{L_2^p}$. By substituting $u_a = d\mathcal{O}(x_0)(v_0)$ into (14), $2L_{\mathcal{O}}(x_0, v_0) = \langle v_0, (d\mathcal{O}(x_0))^* \circ d\mathcal{O}(x_0)(v_0) \rangle_{\mathbb{R}^n}$. On the other hand, from Theorem 2.7, $2L_{\mathcal{O}}(x_0, v_0) = \langle v_0, Q(x_0)v_0 \rangle_{\mathbb{R}^n}$. Thus, we have (19). ■

Based on the above lemmas, we now are able to obtain the relationship between the Hankel operator and the singular values.

Theorem 3.14: Suppose that the assumptions of Theorems 2.5 and 2.7 and Assumption 3.11 hold. Then, $\sigma_i(x_0)$, the nonzero eigenvalues of matrix $(P(x_0)Q(x_0))^{1/2}$ and the differential Hankel singular values of Definition 3.8 are the same.

Proof: Let $\lambda(x_0)$ be a nonzero differential singular value with a nonzero differential singular vector $v(x_0)$. For ease of notation, from now on we omit the argument $x_0 = \mathcal{C}(u)$ of λ and v . By using (12) and (15), the definition of the singular value can be rewritten as

$$(d\mathcal{C}(u))^* \circ (d\mathcal{O}(x_0))^* \circ d\mathcal{O}(x_0) \circ d\mathcal{C}(u) \circ v = \lambda^2 v. \quad (20)$$

By substituting (20) into u_p in $d\mathcal{C}(u)(u_p)$, we have, from linearity of the Fréchet derivative $d\mathcal{C}(u)(\cdot)$,

$$\begin{aligned} &(d\mathcal{C}(u) \circ (d\mathcal{C}(u))^* \circ ((d\mathcal{O}(x_0))^* \circ d\mathcal{O}(x_0)) \circ d\mathcal{C}(u)(v) \\ &= d\mathcal{C}(u)(\lambda^2 v) = \lambda^2 d\mathcal{C}(u)(v). \end{aligned}$$

Note that $\lambda^2 \cdot d\mathcal{C}(u)(v) \neq 0$. Otherwise, $d\mathcal{C}(u)(\lambda^2 \cdot v) \equiv 0$ implies $\lambda^2 v \equiv 0$, which is not possible. From (19), (16) and $x_0 = \mathcal{C}(u)$, we obtain $P(x_0)Q(x_0) \cdot d\mathcal{C}(u)(v) = \lambda^2 \cdot d\mathcal{C}(u)(v)$. Denote $\hat{v}(x_0) := d\mathcal{C}(u)(v(x_0))$, then we have $P(x_0)Q(x_0)\hat{v}(x_0) = \lambda^2(x_0)\hat{v}(x_0)$. Therefore, $\lambda^2(x_0)$ is a nonzero eigenvalue of $P(x_0)Q(x_0)$.

On the other hand, let $\lambda^2(x_0)$ and $\hat{v}(x_0)$ be respectively a nonzero eigenvalue and eigenvector of $P(x_0)Q(x_0)$. Again, for ease of notation, from now on we omit the argument $x_0 = \mathcal{C}(u)$ of λ and $\hat{v}$. From (19), (16) and $x_0 = \mathcal{C}(u)$, we have $(d\mathcal{C}(u) \circ (d\mathcal{C}(u))^* \circ ((d\mathcal{O}(\mathcal{C}(u)))^* \circ d\mathcal{O}(\mathcal{C}(u)))) \circ \hat{v} = \lambda^2 \hat{v}$. By substituting this into u_p of $(d\mathcal{C}(u))^* \circ (d\mathcal{O}(\mathcal{C}(u)))^* \circ d\mathcal{O}(\mathcal{C}(u))(u_p)$, we have from (12) and (15)

$$\begin{aligned} &((d\mathcal{H}(u))^* \circ d\mathcal{H}(u)) \\ &\quad \circ ((d\mathcal{C}(u))^* \circ (d\mathcal{O}(\mathcal{C}(u)))^* \circ d\mathcal{O}(\mathcal{C}(u))(\hat{v})) \\ &= \lambda^2((d\mathcal{C}(u))^* \circ (d\mathcal{O}(\mathcal{C}(u)))^* \circ d\mathcal{O}(\mathcal{C}(u))(\hat{v})), \end{aligned}$$

where we used linearity of the Fréchet derivatives. Denote $v(\mathcal{C}(u)) := (d\mathcal{C}(u))^* \circ (d\mathcal{O}(\mathcal{C}(u)))^* \circ d\mathcal{O}(\mathcal{C}(u)) \circ \hat{v}(\mathcal{C}(u))$. Then, $\lambda^2(\mathcal{C}(u))$ and $v(\mathcal{C}(u))$ are respectively a differential Hankel singular value and vector. Note that $v(\mathcal{C}(u)) \neq 0$. Otherwise, as above noted, $\hat{v}(\mathcal{C}(u)) \equiv d\mathcal{C}(u)(v(\mathcal{C}(u))) \equiv 0$, which is not possible. ■

Using the above theorem, and similar to the linear case, e.g., [4], [5], as well as the nonlinear balancing case of [7], [8], we are now able to establish an explicit expression for the differential Hankel norm as follows.

Theorem 3.15: Suppose that the assumptions of Theorem 2.5 and 2.7 and Assumption 3.11 hold. The differential Hankel norm is the largest nonzero eigenvalue of $(P(x_0)Q(x_0))^{1/2}$.

Proof: From the definitions of $L_{\mathcal{O}}$ and $d\mathcal{O}$, we have

$$\begin{aligned} 2L_{\mathcal{O}}(\mathcal{C}(u), d\mathcal{C}(u)(\delta u)) &= \|\delta y(t)\|_{L_2^p[0, \infty)} \\ &= \|d\mathcal{O}(\mathcal{C}(u))(d\mathcal{C}(u)(\delta u))\|_{L_2^p[0, \infty)}, \end{aligned}$$

and consequently from (12),

$$2L_{\mathcal{O}}(\mathcal{C}(u), d\mathcal{C}(u)(\delta u)) = \|d\mathcal{H}(u)(\delta u)\|_{L_2^p[0, \infty)}. \quad (21)$$

From the definitions of $L_{\mathcal{C}}$, $\mathcal{C}$ and $d\mathcal{C}$, we have

$$L_{\mathcal{C}}(\mathcal{C}(u), d\mathcal{C}(u)(\delta u)) = \frac{1}{2} \inf_{\delta u \in L_2^m[0, \infty)} \int_0^\infty \|\delta u(t)\|^2 dt. \quad (22)$$

Therefore, from (21) and (22), we obtain

$$\begin{aligned} \|\delta\Sigma\|_{\mathcal{H}} &= \sup_{\substack{\delta u \in L_2^m[0, \infty) \\ \delta u \neq 0}} \sqrt{\frac{L_{\mathcal{O}}(\mathcal{C}(u), d\mathcal{C}(u)(\delta u))}{L_{\mathcal{C}}(\mathcal{C}(u), d\mathcal{C}(u)(\delta u))}} \\ &= \sup_{d\mathcal{C}(u)(\delta u) \neq 0} \sqrt{\frac{L_{\mathcal{O}}(\mathcal{C}(u), d\mathcal{C}(u)(\delta u))}{L_{\mathcal{C}}(\mathcal{C}(u), d\mathcal{C}(u)(\delta u))}}. \end{aligned}$$

From Theorems 2.5 and 2.7 and initial conditions $x_0 = \mathcal{C}(u)$ and $\delta x_0 = d\mathcal{C}(u)(\delta u)$, we obtain

$$\begin{aligned} &\sup_{d\mathcal{C}(u)(\delta u) \neq 0} \frac{L_{\mathcal{O}}(\mathcal{C}(u), d\mathcal{C}(u)(\delta u))}{L_{\mathcal{C}}(\mathcal{C}(u), d\mathcal{C}(u)(\delta u))} \\ &= \sup_{\delta x_0 \neq 0} \frac{L_{\mathcal{O}}(x_0, \delta x_0)}{L_{\mathcal{C}}(x_0, \delta x_0)} = \sup_{\delta z_0 \neq 0} \frac{\delta z_0^T \Lambda(x_0) \delta z_0}{\delta z_0^T \Lambda^{-1}(x_0) \delta z_0} = \sigma_1^2(x_0), \end{aligned} \quad (23)$$

where $\sigma_1^2(x_0)$ is the largest eigenvalue of $P(x_0)Q(x_0)$. That completes the proof. ■

Remark 3.16: In this section we have assumed the system to be time-invariant. However, following a similar reasoning, equality (23) also holds for time-varying systems, i.e., we have

$$\begin{aligned} \sup_{\delta x_0 \neq 0} \frac{L_O(x_0, t_0, \delta x_0)}{L_C(x_0, t_0, \delta x_0)} &= \sup_{\delta z_0 \neq 0} \frac{\delta z_0^T \Lambda(x_0, t_0) \delta z_0}{\delta z_0^T \Lambda^{-1}(x_0, t_0) \delta z_0} \\ &= \sigma_1^2(x_0, t_0) \end{aligned} \quad (24)$$

Another expression of this inequality is

$$\|\delta y(t)\|_{L_2^p[t_0, \infty)} \leq \sigma_1(x_0, t_0) \inf_{\delta u \in L_2^m(-\infty, t_0]} \|\delta u(t)\|_{L_2^m(-\infty, t_0]}, \quad (25)$$

where $u(t) = 0, \delta u(t) = 0$ for all $t \geq t_0$, $x(t_0) = x_0$, $\delta x(-\infty) = 0$, $\delta x(t_0) = \delta x_0$, $\delta x(\infty) = 0$ and $u \in L_2^m(-\infty, t_0]$. We however do not have an appropriate analysis of a Hankel operator for the time-variant case, and thus we are unable to establish the full time-variant counterparts of Theorem 3.14 and 3.15. ◁

Remark 3.17: By taking a path integral, we also obtain another inequality from (25). Namely, similar to [16], we have the following

$$\begin{aligned} &\|y_1(t) - y_2(t)\|_{L_2^p[t_0, \infty)} \\ &\leq \sigma_1(x_0, t_0) \inf_{u_1 - u_2 \in L_2^m(-\infty, t_0]} \|u_1(t) - u_2(t)\|_{L_2^m(-\infty, t_0]}, \end{aligned}$$

where $u_1(t) = u_2(t) = 0$ for all $t \geq t_0$, $x_1(t_0) = x_0$ and $x_1(-\infty) = x_2(-\infty)$. This inequality implies that differential balancing evaluates the effect of $y_1(t) - y_2(t)$ from $u_1(t) - u_2(t)$. A similar effect is studied in incremental balancing [3] but it evaluates $u_1(t) + u_2(t)$, i.e., differential and incremental balancing are based on different energy functions. Moreover, incremental balancing relies on systems having specific drift and input vector fields whose elements are odd functions. This is in contrast to differential balancing. ◁

Theorems 3.14 and 3.15 are generalizations of relationships between linear Hankel operators and balancing [3]. In the more recent nonlinear balancing methods such as incremental balancing, these relationships have not been generalized. In traditional balancing as defined in [6], [7] it is generalized, but in contrast to differential balancing in the traditional case the pseudo inverse of the controllability operator is needed. Furthermore, the use of the controllability function in [6], [7] means that we do not deal with a Lyapunov type of equation, but merely with an Hamilton-Jacobi equation, while in the linear case the controllability Gramian is characterized by a Lyapunov equation. Thus, in this respect differential balancing naturally extends the linear case.

IV. GENERALIZED DIFFERENTIAL BALANCING

A. Generalized Differential Energy Functions

In previous sections, balancing theory in the contraction framework has been established, which is a natural extension of linear balancing theory. From the application point of view, it is worth constructing a computationally more feasible method. Inspired by generalized incremental balancing as presented in [3], we have developed generalized differential balancing, which is more oriented towards computation. These results are published in [28], but since they are a natural computationally attractive extension, we summarize the results here. The results are illustrated by a new physically relevant example. The proofs can be

found in [28]. Here we consider systems of the form

$$\Sigma_{BC} : \begin{cases} \dot{x}(t) = f(x(t), t) + B(t)u(t), \\ y(t) = C(t)x(t), \end{cases}$$

with the only difference w.r.t. Σ_{gh} that the input vector field and the output map are not depending on the state. The generalized differential Gramians are defined as follows.

Definition 4.1: If there exists a uniformly positive definite matrix $\bar{P}(t) = \bar{P}^T(t)$ such that

$$-\dot{\bar{P}}(t) + \frac{\partial f(x, t)}{\partial x} \bar{P}(t) + \bar{P}(t) \frac{\partial^T f(x, t)}{\partial x} \leq -B(t)B^T(t) \quad (26)$$

for all $x \in \mathbb{R}^n, t \in \mathbb{R}$ then the function $\bar{L}_C(\delta x_0, t_0) := (1/2)\delta x_0^T \bar{P}^{-1}(t_0)\delta x_0^T$, is said to be a generalized differential controllability function.

Definition 4.2: If there exists a uniformly positive definite matrix $\bar{Q}(t) = \bar{Q}^T(t)$ such that

$$\dot{\bar{Q}}(t) + \bar{Q}(t) \frac{\partial f(x, t)}{\partial x} + \frac{\partial^T f(x, t)}{\partial x} \bar{Q}(t) \leq -C^T(t)C(t) \quad (27)$$

for all $x \in \mathbb{R}^n, t \in \mathbb{R}$ then the function $\bar{L}_O(\delta x_0, t_0) := (1/2)\delta x_0^T \bar{Q}(t_0)\delta x_0^T$, is said to be a generalized differential observability function.

It has been shown in [28] that existence of generalized differential controllability and observability functions respectively imply boundedness of trajectories and convergence of the output of the associated system of differential dynamics. Clearly, generalized differential controllability and observability functions are not unique, and they provide lower bounds for the differential controllability function and upper bounds for the differential observability function. That is, $\bar{L}_C(\delta x_0, t_0) \leq L_C(x_0, \delta x_0, t_0)$ and $\bar{L}_O(\delta x_0, t_0) \geq L_O(x_0, \delta x_0, t_0)$ for all $x_0 \in \mathbb{R}^n, \delta x_0 \in \mathbb{R}^n, t_0 \in \mathbb{R}$ if these four functions exist [28]. Therefore, it is desirable to find larger generalized differential controllability functions and smaller generalized differential observability functions. Hereafter, we assume existence of these four differential functions.

We can partially generalize the results of Theorems 2.13 and 2.17, establishing relationships between existence, and positivity of generalized differential energy functions and properties of the time-invariant system Σ_{BC} such as strong accessibility, observability, and incremental stability. We refer to [28] for details.

B. Generalized Differentially Balanced Realization and Model Reduction

Based on the generalized differential controllability and observability functions, we now are able to define a generalized differentially balanced realization.

Definition 4.3: A realization of the associated system to Σ_{BC} is said to be a generalized differentially balanced realization on an open subset $D \subset \mathbb{R}$ if there exists a diagonal matrix

$$\bar{\Lambda}(t) = \text{diag}\{\bar{\sigma}_1(t), \bar{\sigma}_2(t), \dots, \bar{\sigma}_n(t)\}, \quad (28)$$

where $\bar{\sigma}_1(t) \geq \bar{\sigma}_2(t) \geq \dots \geq \bar{\sigma}_n(t) > 0$ on D holds, and $\bar{P}(t) = \bar{\Lambda}(t)$ and $\bar{Q}(t) = \bar{\Lambda}(t)$.

Based on this definition, the following generalized differential balancing theorem is a natural consequence.

Theorem 4.4: Let $L_C(\delta x_0, t_0)$ and $L_O(\delta x_0, t_0)$ be generalized differential controllability and observability functions, respectively. For every system Σ_{BC} , there exists a coordinate transformation $z = T(t)x$ which make $d\Sigma_{BC}$ into a generalized differentially balanced realization on a domain $D \subset \mathbb{R}$. Also $\bar{\sigma}_i^2(t)$ ($i = 1, \dots, n$) in (28) are the eigenvalues of $\bar{P}(t)\bar{Q}(t)$. ■

For a generalized balanced realized system a relation with the Hankel operator type of analysis can be performed as follows,

$$\begin{aligned} \sup_{\delta x_0 \neq 0} \frac{L_O(x_0, t_0, \delta x_0)}{L_C(x_0, t_0, \delta x_0)} &\leq \sup_{\delta x_0 \neq 0} \frac{\bar{L}_O(x_0, t_0, \delta x_0)}{\bar{L}_C(x_0, t_0, \delta x_0)} \\ &= \sup_{\delta z_0 \neq 0} \frac{\delta z_0^T \bar{\Lambda}(t_0) \delta z_0}{\delta z_0^T \bar{\Lambda}^{-1}(t_0) \delta z_0} = \bar{\sigma}_1^2(t_0). \end{aligned}$$

This can be reformulated as

$$\|\delta y(t)\|_{L_2^p[t_0, \infty)} \leq \bar{\sigma}_1(t_0) \inf_{\delta u \in L_2^m(-\infty, t_0]} \|\delta u(t)\|_{L_2^m(-\infty, t_0]},$$

where $u(t) = 0, \delta u(t) = 0$ for all $t \geq t_0$ and $x(t_0) = x_0, \delta x(-\infty) = 0, \delta x(t_0) = \delta x_0, \delta x(\infty) = 0, u \in L_2^m(-\infty, 0]$. Moreover, from a similar discussion as provided in Remark 3.17, for a given feasible pair $(u_2(t), y_2(t))$, we have

$$\begin{aligned} \|y_1(t) - y_2(t)\|_{L_2^p[t_0, \infty)} \\ \leq \bar{\sigma}_1(t_0) \inf_{u_1 - u_2 \in L_2^m(-\infty, t_0]} \|u_1(t) - u_2(t)\|_{L_2^m(-\infty, t_0]}, \end{aligned}$$

where $y_1(t)$ is the output trajectory corresponding to an admissible input $u_1(t)$, and $u_1(t) = u_2(t) = 0$ for all $t \geq t_0$, and $x_1(t_0) = x_0, x_1(-\infty) = x_2(-\infty)$. Thus, there is a relation with the induced H_∞ norm of the system, see [28] for more information about this result.

Similar to Section III.A we are able to provide a model reduction procedure for a generalized differentially balanced realization. The error bound of this model reduction is studied [28]. Let $\bar{\sigma}_k(t) > \bar{\sigma}_{k+1}(t)$ in equation (28). We partition the system in the z -coordinates as follows:

$$\begin{aligned} \bar{f}(z, t) &= \begin{bmatrix} \bar{f}_a(z_a, z_b, t) \\ \bar{f}_b(z_a, z_b, t) \end{bmatrix} := T(t)f(T^{-1}(t)z(t), t), \\ \bar{B}(t) &= \begin{bmatrix} \bar{B}_a(t) \\ \bar{B}_b(t) \end{bmatrix} := T(t)B(t), \\ \bar{C}(t) &= [\bar{C}_a(t) \quad \bar{C}_b(t)] := C(t)T^{-1}(t), \end{aligned}$$

where $z_a := [z_1, \dots, z_k]^T$ and $z_b := [z_{k+1}, \dots, z_n]^T$. If $\bar{f}_b(z_a, 0, t) = 0$, and the system is truncated, i.e., we set $z_b = 0$, then the reduced system is the generalized differentially balanced realization with singular value functions $\bar{\sigma}_1(t) \geq \dots \geq \bar{\sigma}_k(t)$.

In contrast to the differential energy functions, the generalized differential controllability and observability functions are independent from the state variables x . Thanks to this fact, an error bound in the H_∞ -norm of the error of the prolonged system and the reduced order prolonged system in terms of the sum of the truncated generalized singular values can be given for model reduction by generalized differential balancing. We refer to [28] for more details.

C. Academic Example

We compute the differential Hankel singular values and achieve model reduction for the system Σ_{gh} by differential and generalized differential balancing, given by

$$f = \begin{bmatrix} -x_1 + x_2 - x_2^2 \\ -x_2 \end{bmatrix}, \quad g = \begin{bmatrix} 1 + 2x_2 \\ 1 \end{bmatrix}, \quad h = x_1.$$

It is possible to show that this system is incrementally globally asymptotically stable. By using Maple 18, we compute solutions to (1), (2) and (7),

$$\begin{aligned} P(x) &= \begin{bmatrix} 5/4 + 3x_2 + 2x_2^2 & 3/4 + x_2 \\ 3/4 + x_2 & 1/2 \end{bmatrix}, \\ Q(x) &= \begin{bmatrix} 1/2 & 1/4 - x_2/3 \\ 1/4 - x_2/3 & 1/4 - 5x_2/9 + x_2^2/3 \end{bmatrix}. \end{aligned}$$

Both $P(x)$ and $Q(x)$ are positive definite for all $x \in \mathbb{R}^2$. Therefore, from Corollary 2.20, the system is locally strongly accessible and locally observable.

According to Theorem 3.14, the differential Hankel singular values σ_i^2 and eigenvalues λ_i of PQ are equivalent. The eigenvalues are

$$\begin{aligned} \lambda_1 &= \sigma_1^2 = 9/16 + 11x_2/18 + x_2^2/4 + a/36, \\ \lambda_2 &= \sigma_2^2 = 9/16 + 11x_2/18 + x_2^2/4 - a/36, \\ a &:= \sqrt{81x_2^4 + 396x_2^3 + 844x_2^2 + 900x_2 + 405}. \end{aligned}$$

Note that $\lambda_1 > \lambda_2 > 0$ for all $x_2 \in \mathbb{R}^2$. From Theorem 3.15, σ_1 is the differential Hankel norm.

Next, we compute the reduced order model. Since the system is two dimensional, there exist integral factors $c_1(x_1, x_2)$ and $c_2(x_1, x_2)$ such that

$$\frac{\partial \varphi(x_1, x_2)}{\partial (x_1, x_2)} = \begin{bmatrix} c_1 T_{11} & c_1 T_{12} \\ c_2 T_{21} & c_2 T_{22} \end{bmatrix}$$

holds for some vector valued function $\varphi(x_1, x_2) : \mathbb{R}^2 \rightarrow \mathbb{R}^2$. By using this coordinate transformation, we obtain some sort of weighted balanced coordinates such that controllability and observability Gramians are $\text{diag}\{c_1^2 \lambda_1, c_1^2 \lambda_2\}$ and $\text{diag}\{c_1^{-2} \lambda_1, c_1^{-2} \lambda_2\}$, respectively.

The coordinate transformation $z = \varphi(x)$ in the original coordinates can be obtained, which consists of numerous terms of x . For convenience, here, the third order approximation around the origin is shown.

$$\begin{aligned} \varphi_1 &\simeq -0.232x_1 - 0.143x_2 + 0.552 \times 10^{-3}x_1^2 \\ &\quad + 0.682 \times 10^{-3}x_1x_2 + 0.110x_2^2 \\ &\quad - 0.523 \times 10^{-3}x_1x_2^2 - 0.0214x_2^3, \\ \varphi_2 &\simeq 0.214x_1 - 0.347x_2 - 0.552 \times 10^{-3}x_1^2 \\ &\quad + 0.179 \times 10^{-2}x_1x_2 - 0.209x_2^2 \\ &\quad + 0.107 \times 10^{-2}x_1x_2^2 - 0.273 \times 10^{-2}x_2^3. \end{aligned}$$

After the coordinate transformation, by substituting $z_1 = \hat{z}_1, z_2 = 0$, the reduced system is obtained, whose

Fig. 1. Output trajectories of original system and reduced-order models.

approximation up to the second order is

$$\begin{aligned}\dot{z}_1 &\simeq -0.553z_1 + 0.640z_1^2 + (-0.375 + 0.463z_1)u, \\ \hat{y} &\simeq -3.12z_1 + 2.30z_1^2.\end{aligned}$$

Finally, we consider generalized differential balancing. For this system, it is possible to prove that there does not exist a generalized differential controllability or observability function but there exist $\bar{P}$ and $\bar{Q}$ locally satisfying (26) and (27). For instance,

$$\bar{P} = \begin{bmatrix} 9 & 1 \\ 1 & 1 \end{bmatrix}, \quad \bar{Q} = \begin{bmatrix} 1 & 1/2 \\ 1/2 & 10 \end{bmatrix}$$

satisfy (26) and (27) for $x_2^2 \leq 1$. Moreover, $\bar{P} - P$ and $\bar{Q} - Q$ are positive definite for $x_2^2 \leq 1$.

Fig. 1 shows output trajectories of the original system and reduced-order models given by differential and generalized differential balancing, starting from zero initial states and input $u(t) = \sin t$. It can be observed that the response of the reduced order model by differential balancing follows the trajectory of the original model well. The reduced model by generalized differential balancing performs worse, but approximates the period of the oscillation well.

D. Example: 50 Mass-Spring-Damper Systems

Based on generalized differential balancing, we tackle the model reduction of a system composed by 50 mass-spring-damper systems with nonlinear springs in Fig. 2, where k_l and k_n are respectively spring constants of linear and nonlinear springs, and $m = k_l = d = 1$ and $k_n = 2$. Its state space representation Σ_{BC} is a 100 dimensional system whose f , B and C are

$$\begin{aligned}f_{2i-1} &= x_{2i} \quad (i = 1, \dots, 50), \\ f_2 &= -x_1 + x_3 - 2(x_1 - x_3)^3 - x_2 + x_4, \\ f_{2i} &= -x_{2i-1} + x_{2i-3} - 2(x_{2i-1} - x_{2i-3})^3 \\ &\quad - x_{2i-1} + x_{2i+1} - 2(x_{2i-1} - x_{2i+1})^3 \\ &\quad - x_{2i} + x_{2i-2} - x_{2i} + x_{2i+2} \quad (i = 2, \dots, 49), \\ f_{100} &= -x_{99} + x_{97} - 2(x_{99} - x_{97})^3 - x_{100} + x_{98}, \\ B &= [0 \quad \dots \quad 0 \quad 1]^T, \quad C = [0 \quad \dots \quad 0 \quad 1 \quad 0],\end{aligned}$$

where x_{2i-1} and x_{2i} ($i = 1, \dots, 100$) are respectively position and velocity of the i th mass-spring damper subsystems. By solving both (26) and (27), we obtain positive definite solutions, and consequently the system can be transformed into the generalized differentially balanced realization. We can estimate the

Fig. 2. Mass-spring-damper systems with nonlinear springs.

Fig. 3. Error bound versus order of reduced model.

Fig. 4. Output trajectories of 200-dimensional original system and 10, 12, and 20-dimensional reduced-order models.

error bound of the model reduction based on the results of [28], which is shown in Fig. 3, and 4 shows output trajectories of the original system and reduced-order model starting from zero initial states and with input $u(t) = \sin t$. It can be observed from Fig. 4 that the response of the 20th order model follows the trajectory of the original model really well. The 12th order model is still quite a good approximation, but shows some small errors at the peaks of the oscillation. The 10th order reduced model still approximates the period of the oscillation well, but does not make some significant error in the amplitude.

V. CONCLUSION

In this paper, we propose a model reduction method based on the differential controllability and observability functions defined via prolonged systems. In contrast to flow and incremental balancing, this model reduction can always be achieved by taking line integrals, and the differential balanced realization has been studied in relation to the nonlinear Hankel operator. Our method can be seen a generalization of linear frameworks rather than nonlinear frameworks, e.g., the differential controllability function is characterized by a generalization of the Lyapunov equation, and not of the Riccati equation as is used for most nonlinear balancing methods. We have summarized a related more computationally feasible method, which is illustrated by a physical example.

APPENDIX A

PROOFS FOR CONTROLLABILITY ANALYSIS

The following is an important proposition for incremental stability analysis, which is a specific case of the LaSalle invariance principle for contraction [16], i.e., when $V = \delta x^T X \delta x$ and $\alpha = \delta x^T Y \delta x$ in Theorem 2 in [16].

Proposition A.1: Consider the system Σ_h , and a matrix $Y(x) = Y^T(x) \geq 0$ ($\forall x \in \mathbb{R}^n$). Suppose that there exists a solution $X(x) = X^T(x) > 0$ ($\forall x \in \mathbb{R}^n$) to

$$\delta_f(X) + X \frac{\partial f}{\partial x} + \frac{\partial^T f}{\partial x} X = -Y. \quad (29)$$

Then any pair of solutions of the systems Σ_h and $d\Sigma_h$ converge to the largest invariant set contained in

$$\{(x, \delta x) \in \mathbb{R}^n \times \mathbb{R}^n : \delta x^T Y \delta x = 0\}.$$

If the largest invariant set is $\mathbb{R}^n \times \{0\}$, the system Σ_h is incrementally globally asymptotically stable. ■

To complete the proof of Theorem 2.13, we present the following lemmas.

Lemma A.2: Suppose that Assumption 2.11 holds, the system Σ_h is incrementally globally asymptotically stable, and the system Σ_h has a nonempty α limit set L^- (which is fulfilled if Σ_h has an equilibrium point). Then, the inverse-time system Σ_g^- is incrementally globally asymptotically stabilizable if $0 < P(x) < \infty$ at each $x \in \mathbb{R}^n$.

Proof: Consider the inverse time system of Σ_h denoted by Σ_h^- . Since Σ_h has a nonempty α limit set, there exists a trajectory $x^{*-}(t)$ of the inverse-time system Σ_h^- starting from initial state $x_0^{*-} \in \mathbb{R}^n$ such that $\lim_{t \rightarrow \infty} x^{*-}(t) \in L^-$. Let us consider the inverse time system Σ_{gh}^- , which contains Σ_h^- as a special case $u \equiv 0$. Consider a minimizing path $\gamma^-(\cdot)$ connecting x_0^{*-} and x_0^- parametrized $s \in [0, 1]$ with respect to the metric $P^{-1}(x)$. That is, $\gamma^-(0) = x_0^{*-}$ and $\gamma^-(1) = x_0^-$. Also, consider path $\nu^-(t, \cdot)$ connecting 0 and $u^-(t)$ parametrized by $s \in [0, 1]$. We denote the state trajectory $\psi^-(t, \gamma^-(s), \nu^-(t, s))$ of Σ_{gh}^- starting from $\gamma^-(s)$ with input $\nu^-(t, s)$ by $\gamma_t^-(s)$ ($s \in [0, 1] \subset \mathbb{R}$). Note that $\gamma_t^-(\cdot)$ is a path connecting $x^{*-}(t) = \psi^-(t, t_0, x_0^{*-}, 0)$ and $x^-(t) = \psi^-(t, t_0, x_0^-, u^-)$. Here, we show that the system

$$\begin{aligned} \dot{x}^-(t) = & -f(x^-(t)) \\ & - g(x^-(t)) \int_0^1 g(\gamma_t^-(s))^T P^{-1}(\gamma_t^-(s)) \frac{\partial \gamma_t^-(s)}{\partial s} ds, \end{aligned} \quad (30)$$

is incrementally globally asymptotically stable. This is a system Σ_g^- with input

$$u^- = \int_0^1 g(\gamma_t^-(s))^T P^{-1}(\gamma_t^-(s)) \frac{\partial \gamma_t^-(s)}{\partial s} ds. \quad (31)$$

Then, the path ν^- connecting 0 and u^- is obtained as

$$\nu^-(t, s) = \int_0^s g(\gamma_t^-(r))^T P^{-1}(\gamma_t^-(r)) \frac{\partial \gamma_t^-(r)}{\partial r} dr, \quad (32)$$

which is the feedback controller when the initial condition is $\gamma(s)$. Denote $\chi^-(t, x_0^-) := \psi(t, x_0^-, u^-)$ by a state trajectory of (30) starting from $x^-(0) = x_0^- \in \mathbb{R}^n$. Then, a path $\gamma_t(s)$ connecting $x^{*-}(t)$ and $x^-(t)$ is $\chi^-(t, \gamma^-(s)) = \psi(t, \gamma(s), \nu^-(t, s))$.

Next, we show that the associated differential system of (30) satisfies (3). The state trajectory of (30) starting from $\gamma^-(s)$ denoted by $\chi^-(t, \gamma^-(s))$ satisfies, from (32)

$$\begin{aligned} \dot{\chi}^- = & -f(\chi^-) \\ & - g(\chi^-) \int_0^s g(\chi^-)^T P^{-1}(\chi^-) \frac{\partial \chi^-(t, \gamma^-(r))}{\partial r} dr. \end{aligned} \quad (33)$$

Note that we use a small abuse of notation, since we leave out the argument of χ^- . In the integral, arguments of χ^- are $(t, \gamma^-(r))$. Note that $\gamma_t^-(s) = \chi^-(t, \gamma^-(s))$ for ν^- with u^- in (31). By using this and (33), compute

$$\begin{aligned} & \frac{\partial}{\partial t} \frac{\partial \chi^-(t, \gamma^-(s))}{\partial s} \\ &= \frac{\partial}{\partial s} \frac{\partial \chi^-(t, \gamma^-(s))}{\partial t} \\ &= \frac{\partial}{\partial s} \left(-f(\chi^-) - g(\chi^-) \int_0^s g(\chi^-)^T P^{-1}(\chi^-) \frac{\partial \chi^-(t, \gamma^-(r))}{\partial r} dr \right) \\ &= - \left(\frac{\partial(f(\chi^-) + g(\chi^-) \int_0^s g(\chi^-)^T P^{-1}(\chi^-) \frac{\partial \chi^-(t, \gamma^-(r))}{\partial r} dr)}{\partial x} \right. \\ & \quad \left. + g(\chi^-) g(\chi^-)^T P^{-1}(\chi^-) \right) \frac{\partial \chi^-(t, \gamma^-(s))}{\partial s}. \end{aligned}$$

Thus, the associated differential system of (30) is (3) with input (31).

By differentiating $\delta x^{*-T}(t) P^{-1}(x^-(t)) \delta x^-(t)$ with respect to t along the trajectories of (30) and (3) with input (31), we have, from (1), (2), and (4) that $\frac{d}{dt} (\delta x^{*-T} P^{-1} \delta x^-) = -\|g^T P^{-1} \delta x^-\|^2 \leq 0$. According to Proposition A.1, it remains to show that the maximal invariant set in $V := \{(x^-, \delta x^-) \in \mathbb{R}^n \times \mathbb{R}^n : g^T P^{-1} \delta x^- = 0\}$ is $\mathbb{R}^n \times \{0\}$. The maximal invariant set in V of the vector field of (30) is equivalent to that of $-f(x^-)$ and thus, that of Σ_h . From (1), V can be expressed as

$$\begin{aligned} V = & \left\{ (x^-, \delta x^-) \in \mathbb{R}^n \times \mathbb{R}^n : \right. \\ & \left. \delta x^{*-T} \left(\delta_f(P^{-1}) + P^{-1} \frac{\partial f}{\partial x^-} + \frac{\partial^T f}{\partial x^-} P^{-1} \right) \delta x^- = 0 \right\} \\ = & \left\{ (x, \delta x) \in \mathbb{R}^n \times \mathbb{R}^n : d(\delta x^T P^{-1} \delta x)/dt = 0 \right\}, \end{aligned}$$

where d/dt is computed along the trajectories of the systems Σ_h and $d\Sigma_h$. Since the system Σ_h is incrementally globally asymptotically stable, and since $P^{-1}(x)$ is positive definite for all $x \in \mathbb{R}^n$, we conclude that the maximal invariant set in V is $\mathbb{R}^n \times \{0\}$. ■

Remark A.3: In the above proof, we take a minimizing path $\gamma^-(s)$ connecting x_0^{*-} and x_0^- with respect to the metric $P^{-1}(x)$. This is due to using Proposition 1.1 with $X := P^{-1}$. For more details, see the proofs of Theorems 1 and 2 in [16]. Based on Proposition 1.1, we proved incremental asymptotic stability of the closed loop system (30). One of the trajectories of (30) is $x^{*-}(t)$, which implies that the distance between $x^{*-}(t)$ and any trajectory of (30) converges to the zero, i.e. any trajectory converges to $x^{*-}(t)$. If the system Σ_h has an equilibrium, x_0^{*-} can be chosen as the equilibrium, and consequently $x^{*-}(t) \equiv x_0^- \in L^-$. ◀

In Lemma A.2, we assume that Σ_h is incremental globally asymptotically stable. Then, its reverse time system Σ_h^- is unstable. Thus, incremental stabilizability of Σ_g^- can be viewed as a kind of controllability of Σ_g . Hence, the existence of a positive definite solution $P(x)$ implies controllability of Σ_g . This is an analogy of the relationship between the controllability Gramian and controllability in the linear case. We proceed further under Assumption 2.12.

Lemma A.4: Suppose that Assumption 2.12 holds, and the system Σ_h is incrementally globally asymptotically stable. Then, the system Σ_{gh} is locally strongly accessible if the reverse time system Σ_{-g} is incrementally globally asymptotically stabilizable.

Proof: We prove this by contraposition. If the system Σ_{gh} is not locally strongly accessible, Σ_{-g} is also not locally strongly ac-

cessible. According to [24], by a suitable coordinate transformation $z^- = \varphi(x^-)$, the system Σ_{-g} can be transformed into

$$\begin{aligned}\dot{z}_1^- &= -\hat{f}_1(z_1^-), \\ \dot{z}_2^- &= -\hat{f}_2(z_1^-, z_2^-) - \hat{g}_2(z_1^-, z_2^-)u^-, \end{aligned}$$

where z_2^- is locally strongly accessible, and z_1^- is *not*. Since the system Σ_h is incrementally globally asymptotically stable, its time reverse system Σ_h^- is not. Hence, $\dot{z}_1^- = -\hat{f}_1(z_1^-)$ is not incrementally globally asymptotically stable either, and thus, Σ_{-g}^- is not incrementally globally asymptotically stable. ■

Finally, we give the proof of Theorem 2.14.

Proof of Theorem 2.14: Necessity follows from Lemma A.4 and Theorem 2.13. We prove the sufficiency. By differentiating $\delta x^T(t)P^{-1}(x(t))\delta x(t)$ with respect to t along the trajectories of the systems Σ_h and $d\Sigma_h$, we have $\frac{d}{dt}(\delta x^T P^{-1} \delta x) = -\|g^T P^{-1} \delta x\|^2$ from (1). According to Proposition A.1, it remains to show that the maximal invariant set in $V := \{(x, \delta x) \in \mathbb{R}^n \times \mathbb{R}^n : g^T P^{-1} \delta x = 0\}$ is $\mathbb{R}^n \times \{0\}$. Notice that the maximal invariant set in V of Σ_h is equivalent to that of $-f(x^-)$ and thus, from the proof of Lemma A.2, that of (30). Since the system (30) is incrementally globally asymptotically stable, and since $P^{-1}(x)$ is positive definite, we conclude that the maximal invariant set in V is $\mathbb{R}^n \times \{0\}$. ■

APPENDIX B

PROOFS FOR OBSERVABILITY ANALYSIS

Here, we consider the autonomous system Σ_h . Local observability of Σ_h is defined by considering when $u \equiv 0$. Denote $\delta x(t; x_0, \delta x_0)$ and $\delta y(t; x_0, \delta x_0)$ as the state and output of $d\Sigma_h$ at time $t \geq 0$ starting from $x_0, \delta x_0 \in \mathbb{R}^n$, respectively.

Lemma B.1: Suppose that Assumption 2.15 holds, and $\delta y(t; x_0, \delta x_0) \equiv 0$ implies $\delta x(t; x_0, \delta x_0) \equiv 0$ for all initial states $x_0, \delta x_0 \in \mathbb{R}^n$. Then, system Σ_h is incrementally globally asymptotically stable if and only if $0 < Q(x) < \infty$ at each $x \in \mathbb{R}^n$.

Proof: (Sufficiency) Compute $\frac{d}{dt}(\delta x^T Q \delta x) = -\|\delta y\|^2$ with (7). Since $\delta y \equiv 0$ implies $\delta x \equiv 0$ for all $x_0 \in \mathbb{R}^n$, the maximal invariant set in $V := \{(x, \delta x) \in \mathbb{R}^n \times \mathbb{R}^n : \delta y = 0\}$ is $\mathbb{R}^n \times \{0\}$. Thus, from Proposition A.1 and $Q(x) > 0$ for all $x \in \mathbb{R}^n$, the system Σ_h is incrementally globally asymptotically stable.

(Necessity) From Theorem 2.7, for an incrementally globally asymptotically stable system we have

$$L_O(x_0, \delta x_0) = \frac{1}{2} \int_0^\infty \|\delta y(t)\|^2 dt = \frac{1}{2} \delta x_0^T Q(x_0) \delta x_0. \quad (34)$$

Assume that $x_0 \in \mathbb{R}^n, \delta x_0 \neq 0$. Then, $L_O(x_0, \delta x_0) \neq 0$. We prove this by contradiction. Assume that $L_O(x_0, \delta x_0) = 0$ for some $\delta x_0 \neq 0$, which implies $\delta y \equiv 0$ for this $\delta x_0 \neq 0$. From the assumption of this lemma, $\delta x \equiv 0$, and thus $\delta x_0 = 0$, which contradicts that $\delta x_0 \neq 0$. Thus, $L_O(x_0, \delta x_0) = \delta x_0^T Q(x_0) \delta x_0 > 0$ for all $x_0 \in \mathbb{R}^n$ and $\delta x_0 \neq 0$, which implies $Q(x) > 0$ at each $x \in \mathbb{R}^n$. ■

Lemma B.2: Suppose that Assumption 2.15 holds, and the system Σ_h is incrementally globally asymptotically stable. Then, $\delta y(t; x_0, \delta x_0) \equiv 0$ implies $\delta x(t; x_0, \delta x_0) \equiv 0$ for all initial state $x_0, \delta x_0 \in \mathbb{R}^n$ if and only if $0 < Q(x) < \infty$ at each $x \in \mathbb{R}^n$.

Proof: Since necessity is shown in Lemma B.1, we prove sufficiency by contraposition. Under incremental stability assumption, we have (34). Thus, $\delta y(t; x_0, \delta x_0) \equiv 0$ implies $Q(x_0) \delta x_0 = 0$ for some $\delta x_0 \neq 0$. Therefore, $Q(x_0)$ is nonsingular. ■

The assumption in Lemma B.1 relates to local observability of the system Σ_h .

Lemma B.3: Suppose that Assumption 2.16 holds. Then, the system Σ_h is locally observable if and only if $\delta y(t; x_0, \delta x_0) \equiv 0$ implies $\delta x(t; x_0, \delta x_0) \equiv 0$ for all $x_0, \delta x_0 \in \mathbb{R}^n$.

Proof: (Sufficiency) For analytic systems, $\delta y(t; x_0, \delta x_0) \equiv 0$ if and only if

$$\left. \frac{d^i}{dt^i} \left(\frac{\partial h(x(t))}{\partial x} \delta x(t) \right) \right|_{\substack{x(t)=x_0 \\ \delta x(t)=\delta x_0}} = \frac{\partial L_f^i h(x_0)}{\partial x} \delta x_0 = 0 \quad (35)$$

for all $i = 0, 1, \dots$, where $L_f^0 h := h$ and $L_f^{i+1} h := (\partial L_f^i h / \partial x) f$ ($i \geq 0$). By using (35), the condition of this theorem can be rearranged such that (35) implies $\delta x(t; x_0, \delta x_0) \equiv 0$. Note that $\delta x(t; x_0, \delta x_0) \equiv 0$ implies $\delta x_0 = 0$. In summary, (35) implies $\delta x_0 = 0$, i.e., the observability rank condition [24] holds for all $x_0 \in \mathbb{R}^n$. Therefore, Σ_h is locally observable.

(Necessity) If the analytic system is locally observable, then the observability rank condition holds for all $x_0 \in \mathbb{R}^n$ under Assumption 2.16 [24]. That is, (35) implies $\delta x_0 = 0$. From the definition of $d\Sigma_h$, we have $\delta x(t; x_0, 0) \equiv 0$. In summary, (35) implies $\delta x(t; x_0, \delta x_0) \equiv 0$. ■

APPENDIX C

PROOFS FOR THE DIFFERENTIALLY BALANCED REALIZATION

Proof of Theorem 3.3: Let $\hat{\psi}(\cdot, t_0, z_0, u)$ be the state trajectory of $\hat{\Sigma}_{\hat{g}\hat{h}}$. Then, we have $\hat{\psi}(t, t_0, \varphi(x_0, t_0), u) = \varphi(\psi(t, t_0, x_0, u), t)$, and consequently

$$\hat{\psi}(t, t_0, \varphi(\gamma(s), t_0), \nu(t, s)) = \varphi(\psi(t, t_0, \gamma(s), \nu(t, s)), t). \quad (36)$$

Thus, we have

$$\frac{\partial}{\partial t} \frac{\partial \hat{\psi}(t, t_0, \varphi(\gamma(s), t_0), \nu(t, s))}{\partial s} = \frac{\partial}{\partial t} \frac{\partial \varphi(\psi(t, t_0, \gamma(s), \nu(t, s)), t)}{\partial s}. \quad (37)$$

The left hand side of (37) can be computed as

$$\begin{aligned} & \frac{\partial}{\partial t} \frac{\partial \hat{\psi}(t, t_0, \varphi(\gamma(s), t_0), \nu(t, s))}{\partial s} \\ &= \frac{\partial}{\partial s} \frac{\partial \hat{\psi}(t, t_0, \varphi(\gamma(s), t_0), \nu(t, s))}{\partial t} \\ &= \frac{\partial}{\partial s} \left(\hat{f}(\hat{\psi}, t) + \hat{g}(\hat{\psi}, t) \nu(t, s) \right) \\ &= \frac{\partial(\hat{f}(\hat{\psi}, t) + \hat{g}(\hat{\psi}, t) \nu(t, s))}{\partial z} \frac{\partial \hat{\psi}}{\partial s} + \hat{g}(\hat{\psi}, t) \frac{\partial \nu(t, s)}{\partial s}. \end{aligned}$$

From (10), we have

$$\begin{aligned} & \varphi(\psi(\cdot, t_0, \gamma(s), \nu(t, s)), t) \\ &= \int_0^s T(\psi(\cdot, t_0, \gamma(r), \nu(t, r)), t) \frac{\partial \psi(\cdot, t_0, \gamma(r), \nu(t, r))}{\partial r} dr. \end{aligned}$$

By taking the partial derivative of both sides with respect to s , we have

$$\frac{\partial \varphi(\psi, t)}{\partial x} \frac{\partial \psi(\cdot, t_0, \gamma(s), \nu(t, s))}{\partial s} = T(\psi, t) \frac{\partial \psi(\cdot, t_0, \gamma(s), \nu(t, s))}{\partial s}. \quad (38)$$

By using (36) and (38), the left hand side of (37) can be rearranged as

$$\begin{aligned} & \frac{\partial}{\partial t} \frac{\partial \hat{\psi}(t, t_0, \varphi(\gamma(s), t_0), \nu(t, s))}{\partial s} \\ &= \frac{\partial(\hat{f}(\hat{\psi}, t) + \hat{g}(\hat{\psi}, t) \nu(t, s))}{\partial z} \frac{\partial \varphi(\psi, t)}{\partial x} \frac{\partial \psi}{\partial s} + \hat{g}(\hat{\psi}, t) \frac{\partial \nu(t, s)}{\partial s} \\ &= \frac{\partial(\hat{f}(\hat{\psi}, t) + \hat{g}(\hat{\psi}, t) \nu(t, s))}{\partial z} T(\psi, t) \frac{\partial \psi}{\partial s} + \hat{g}(\hat{\psi}, t) \frac{\partial \nu(t, s)}{\partial s}. \end{aligned} \quad (39)$$

Next, by using (38), the right hand side of (37) can be computed as

$$\begin{aligned}
& \frac{\partial}{\partial t} \frac{\partial \varphi(\psi(t, t_0, \gamma(s), \nu(t, s)), t)}{\partial s} \\
&= \frac{\partial}{\partial t} \left(\frac{\partial \varphi(\psi, t)}{\partial x} \frac{\partial \psi(t, t_0, \gamma(s), \nu(t, s))}{\partial s} \right) \\
&= \frac{\partial}{\partial t} \left(T(\psi, t) \frac{\partial \psi(t, t_0, \gamma(s), \nu(t, s))}{\partial s} \right) \\
&= \left(\delta_{f+g\nu}(T) + T(\psi, t) \frac{\partial(f(\psi, t) + g(\psi, t)\nu(s, t))}{\partial \psi} \right) \frac{\partial \psi}{\partial s} \\
&\quad + T(\psi, t) g(\psi, t) \frac{\partial \nu(t, s)}{\partial s} \quad (40)
\end{aligned}$$

From (37), (39), and (40), we now have

$$\begin{aligned}
& \frac{\partial}{\partial t} \frac{\partial \hat{\psi}}{\partial s} \\
&= \frac{\partial(\hat{f}(\hat{\psi}, t) + \hat{g}(\hat{\psi}, t)\nu(t, s))}{\partial z} T(\psi, t) \frac{\partial \psi}{\partial s} + \hat{g}(\hat{\psi}, t) \frac{\partial \nu(t, s)}{\partial s} \\
&= \left(\delta_{f+g\nu}(T) + T(\psi, t) \frac{\partial(f(\psi, t) + g(\psi, t)\nu(s, t))}{\partial x} \right) \frac{\partial \psi}{\partial s} \\
&\quad + T(\psi, t) g(\psi, t) \frac{\partial \nu(t, s)}{\partial s} = \frac{\partial}{\partial t} \frac{\partial \varphi(\psi, t)}{\partial s}.
\end{aligned}$$

Since T is nonsingular for all (x, t) , we have

$$\begin{aligned}
& \frac{\partial(\hat{f}(\hat{\psi}, t) + \hat{g}(\hat{\psi}, t)\nu(t, s))}{\partial z} T(\psi, t) \\
&= \left(\delta_{f+g\nu}(T) + T(\psi, t) \frac{\partial(f(\psi, t) + g(\psi, t)\nu(s, t))}{\partial x} \right), \\
&\quad \hat{g}(\hat{\psi}, t) = T(\psi, t) g(\psi, t).
\end{aligned}$$

Thus, $d\Sigma_{\hat{g}\hat{h}}$ and (9) have the same dynamics along $\psi(t, t_0, \gamma(\cdot), u(t, \cdot))$. From Theorem 3.2, the system is differentially balanced. ■

Proof of Theorem 3.4: Let $\hat{\psi}(\cdot, t_0, z_0, u)$ be the state trajectory of $\hat{\Sigma}_{\hat{g}\hat{h}}$. We partition the state trajectory $\hat{\psi}_a := [\hat{\psi}_1, \dots, \hat{\psi}_k]^T$ and $\hat{\psi}_b := [\hat{\psi}_{k+1}, \dots, \hat{\psi}_n]^T$. According to Theorem 3.3, (1), (2) and (7) along $\hat{\psi}(t, t_0, \varphi(\gamma(s), t_0), \nu(t, s))$ are

$$\begin{aligned}
& -\delta_{\hat{f}}(\hat{\Lambda}) + \hat{\Lambda} \begin{bmatrix} \frac{\partial \hat{f}_a}{\partial z_a} & \frac{\partial \hat{f}_a}{\partial z_b} \\ \frac{\partial \hat{f}_b}{\partial z_a} & \frac{\partial \hat{f}_b}{\partial z_b} \end{bmatrix}^T + \begin{bmatrix} \frac{\partial \hat{f}_a}{\partial z_a} & \frac{\partial \hat{f}_a}{\partial z_b} \\ \frac{\partial \hat{f}_b}{\partial z_a} & \frac{\partial \hat{f}_b}{\partial z_b} \end{bmatrix} \hat{\Lambda} \\
&= - \begin{bmatrix} \hat{g}_a \hat{g}_a^T & \hat{g}_a \hat{g}_b^T \\ \hat{g}_b \hat{g}_a^T & \hat{g}_b \hat{g}_b^T \end{bmatrix}, \\
& -\delta_{\hat{g}}(\hat{\Lambda}) + \hat{\Lambda} \begin{bmatrix} \frac{\partial \hat{g}_a}{\partial z_a} & \frac{\partial \hat{g}_a}{\partial z_b} \\ \frac{\partial \hat{g}_b}{\partial z_a} & \frac{\partial \hat{g}_b}{\partial z_b} \end{bmatrix}^T + \begin{bmatrix} \frac{\partial \hat{g}_a}{\partial z_a} & \frac{\partial \hat{g}_a}{\partial z_b} \\ \frac{\partial \hat{g}_b}{\partial z_a} & \frac{\partial \hat{g}_b}{\partial z_b} \end{bmatrix} \hat{\Lambda} = 0, \\
& \delta_{\hat{f}}(\hat{\Lambda}) + \begin{bmatrix} \frac{\partial \hat{f}_a}{\partial z_a} & \frac{\partial \hat{f}_a}{\partial z_b} \\ \frac{\partial \hat{f}_b}{\partial z_a} & \frac{\partial \hat{f}_b}{\partial z_b} \end{bmatrix}^T \hat{\Lambda} + \hat{\Lambda} \begin{bmatrix} \frac{\partial \hat{f}_a}{\partial z_a} & \frac{\partial \hat{f}_a}{\partial z_b} \\ \frac{\partial \hat{f}_b}{\partial z_a} & \frac{\partial \hat{f}_b}{\partial z_b} \end{bmatrix} \\
&= - \begin{bmatrix} \frac{\partial^T \hat{h}}{\partial z_a} \frac{\partial \hat{h}}{\partial z_a} & \frac{\partial^T \hat{h}}{\partial z_b} \frac{\partial \hat{h}}{\partial z_a} \\ \frac{\partial^T \hat{h}}{\partial z_a} \frac{\partial \hat{h}}{\partial z_b} & \frac{\partial^T \hat{h}}{\partial z_b} \frac{\partial \hat{h}}{\partial z_b} \end{bmatrix}.
\end{aligned}$$

From the assumption of this theorem, by substituting $\hat{\psi}_b = 0$ into the first $k \times k$ matrix equations, we have

$$\begin{aligned}
& -\delta_{\hat{f}_a}(\hat{\psi}_a, 0, t)(\hat{\Lambda}_k) + \hat{\Lambda}_k \frac{\partial^T \hat{f}_a(\hat{\psi}_a, 0, t)}{\partial z_a} + \frac{\partial \hat{f}_a(\hat{\psi}_a, 0, t)}{\partial z_a} \hat{\Lambda}_k \\
&= -\hat{g}_a(\hat{\psi}_a, 0, t) \hat{g}_a^T(\hat{\psi}_a, 0, t), \\
& -\delta_{\hat{g}_a}(\hat{\psi}_a, 0, t)(\hat{\Lambda}_k) + \hat{\Lambda}_k \frac{\partial^T \hat{g}_a(\hat{\psi}_a, 0, t)}{\partial z_a} + \frac{\partial \hat{g}_a(\hat{\psi}_a, 0, t)}{\partial z_a} \hat{\Lambda}_k = 0, \\
& \delta_{\hat{f}_a}(\hat{\psi}_a, 0, t)(\hat{\Lambda}_k) + \frac{\partial^T \hat{f}_a(\hat{\psi}_a, 0, t)}{\partial z_a} \hat{\Lambda}_k + \hat{\Lambda}_k \frac{\partial \hat{f}_a(\hat{\psi}_a, 0, t)}{\partial z_a} \\
&= -\frac{\partial^T \hat{h}(\hat{\psi}_a, 0, t)}{\partial z_a} \frac{\partial \hat{h}(\hat{\psi}_a, 0, t)}{\partial z_a}.
\end{aligned}$$

These are nothing but (1), (2) and (7) for the reduced order system $\Sigma_{\hat{g}\hat{h}}^r$. Thus, in a similar manner with the proof of Theorem 3.3, it is possible to show that the reduced order system is differentially balanced. ■

Proof of Theorem 3.5: Incremental global asymptotic stability is preserved under coordinate transformations. Thus, under the assumption for $\hat{f}$, the reduced order model $\dot{z}_a = \hat{f}_a(z_a)$ is incrementally globally asymptotically stable.

In a similar manner as in the proof of Theorem 3.4, it is possible to show that the obtained equalities satisfy (1), (2) and (7) for $\Lambda_k(\varphi^{-1}(z_a, 0, t), t)$. According to Corollary 2.20, the reduced order model is locally strongly accessible and locally observable. ■

REFERENCES

- [1] Y. Kawano and J. M. A. Scherpen, "On differential balancing: Energy functions and balanced realization," *Proc. Eur. Control Conf.*, pp. 1159–1164, 2015.
- [2] Y. Kawano and J. M. A. Scherpen, "On differential energy functions: Controllability, observability and incremental stability," *Proc. IEEE Conf. Decision Control*, Osaka, Japan, pp. 4656–4661.
- [3] B. Besselink, N. Van De Wouw, J. M. A. Scherpen, and H. Nijmeijer, "Model reduction for nonlinear systems by incremental balanced truncation," *IEEE Trans. Autom. Control*, vol. 59, no. 10, pp. 2739–2753, 2014.
- [4] A. C. Antoulas, *Approximation of Large-Scale Dynamical Systems*. Philadelphia, PA: SIAM, 2005.
- [5] K. Zhou et al., *Robust and Optimal Control*. Prentice Hall: Englewood Cliffs, NJ, 1996, vol. 40.
- [6] J. M. A. Scherpen, "Balancing for nonlinear systems," *Syst. Control Lett.*, vol. 21, no. 2, pp. 143–153, 1993.
- [7] K. Fujimoto and J. M. A. Scherpen, "Nonlinear input-normal realizations based on the differential eigenstructure of Hankel operators," *IEEE Trans. Autom. Control*, vol. 50, no. 1, pp. 2–18, 2005.
- [8] K. Fujimoto and J. M. A. Scherpen, "Model reduction for nonlinear systems based on the balanced realization," *SIAM J. Control Optim.*, vol. 48, no. 7, pp. 4591–4623, 2010.
- [9] A. Astolfi, "Model reduction by moment matching for linear and nonlinear systems," *IEEE Trans. Autom. Control*, vol. 55, no. 10, pp. 2321–2336, 2010.
- [10] T. C. Ionescu and A. Astolfi, "Moment matching for nonlinear port Hamiltonian and gradient systems," in *Preprints 9th IFAC Symp. Nonlin. Control Syst.*, 2013.
- [11] E. I. Verriest and W. S. Gray, "Flow balancing nonlinear systems," in *Proc. 14th Int. Symp. Math. Theory Networks Syst.*, 2000.
- [12] E. I. Verriest and W. S. Gray, "Nonlinear balanced realizations," in *Proc. 43rd IEEE Conf. Decision Control*, vol. 2, 2004, pp. 1164–1169.
- [13] M. Sassano and A. Astolfi, "Dynamic generalized controllability and observability functions with applications to model reduction and sensor deployment," *Automatica*, vol. 50, no. 5, pp. 1349–1359, 2014.
- [14] W. Lohmiller and J.-J. E. Slotine, "On contraction analysis for non-linear systems," *Automatica*, vol. 34, no. 6, pp. 683–696, 1998.
- [15] D. Angeli, "A Lyapunov approach to incremental stability properties," *IEEE Trans. Autom. Control*, vol. 47, no. 3, pp. 410–421, 2002.
- [16] F. Forni and R. Sepulchre, "A differential Lyapunov framework for contraction analysis," *IEEE Trans. Autom. Control*, vol. 59, no. 3, pp. 614–628, 2014.

- [17] I. R. Manchester and J.-J. E. Slotine, "Control contraction metrics and universal stabilizability," in *Preprints 19th IFAC World Congress*, 2014.
- [18] I. R. Manchester and J.-J. E. Slotine, "Control contraction metrics: Differential L^2 gain and observer duality," arXiv Preprint arXiv:1403.5364v1, 2014.
- [19] F. Forni and R. Sepulchre, "On differentially dissipative dynamical systems," *Proc. 8th IFAC Symp. Nonlin. Control Syst.*, pp. 21–25, 2013.
- [20] A. J. van der Schaft, "On differential passivity," *Proc. 8th IFAC Symp. Nonlin. Control Syst.*, pp. 15–20, 2013.
- [21] W. S. Gray and J. M. A. Scherpen, "Hankel operators and gramians for nonlinear systems," in *Proc. 37th IEEE Conf. Decision Control*, vol. 2, 1998, pp. 1416–1421.
- [22] 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, 2000.
- [23] D. Bao, S.-S. Chern, and Z. Shen, *An Introduction to Riemann-Finsler Geometry*. Springer Science & Business Media, 2000, vol. 200.
- [24] H. Nijmeijer and A. van der Schaft, *Nonlinear Dynamical Control Systems*. New York: Springer-Verlag, 1990.
- [25] L. M. Silverman and H. Meadows, "Controllability and observability in time-variable linear systems," *SIAM J. Control*, vol. 5, no. 1, pp. 64–73, 1967.
- [26] L. Weiss, "The concepts of differential controllability and differential observability," *J. Math. Anal. Appl.*, vol. 10, no. 2, pp. 442–449, 1965.
- [27] T. C. Ionescu, K. Fujimoto, and J. M. Scherpen, "Singular value analysis of nonlinear symmetric systems," *IEEE Trans. Autom. Control*, vol. 56, no. 9, pp. 2073–2086, 2011.
- [28] Y. Kawano and J. M. A. Scherpen, "Model reduction by generalized differential balancing," in *Mathematical Control Theory I*, M. K. Camlibel, A. A. Julius, R. Pasumathy, and J. M. A. Scherpen, Eds. Springer-Verlag, 2015, pp. 349–362.

Yu Kawano (M'13) received the M.S. and Ph.D. degrees in engineering from Osaka University, Japan, in 2011 and 2013, respectively.

From August to November 2012, He was a Research Student at Tallinn University of Technology, Estonia. From April to October 2013, he served as a Research Fellow of the Japan Society for the Promotion Science. From October 2013 to November 2016, he was a Researcher with Kyoto University and JST CREST, Japan.

From April to November 2015, he was a Visiting Researcher at the University of Groningen, The Netherlands. Since November 2016, he has been a Researcher with the University of Groningen, at the Engineering and Technology institute Groningen (ENTEG) of the Faculty of Science and Engineering, Groningen, The Netherlands. His research interests include analysis and model reduction of nonlinear systems and large-scale systems.

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

She was with Delft University of Technology, The Netherlands, from 1994 to 2006. Since September 2006, she has been a Professor with the University of Groningen, at the Engineering and Technology institute Groningen (ENTEG) of the Faculty of Science and Engineering, Groningen, The Netherlands. She is currently scientific director of ENTÉG. She has held visiting research positions at The University of Tokyo, Japan, Université de Compiègne, SUPELEC, Gif-sur-Yvette, France, and the Old Dominion University, Norfolk, VA, USA. Her current research interests include nonlinear model reduction methods, nonlinear control methods, modeling and control of physical systems with applications to electrical circuits, electromechanical systems and mechanical systems, and distributed optimal control applications to smart grids.

Dr. Scherpen has been an Associate Editor of the IEEE TRANSACTIONS ON AUTOMATIC CONTROL, the International Journal of Robust and Nonlinear Control (IJRNC), and the IMA Journal of Mathematical Control and Information. She is on the Editorial Board of the IJRNC.