2024 IEEE 63rd Conference on Decision and Control (CDC)
December 16-19, 2024. MiCo, Milan, Italy

Forty Plus Years of Model Reduction and Still Learning

A. Astolfi[1,2], C. L. Beck[3], D. Bhattacharjee[1], Y. Kawano[4],
A. Moreschini[1], H. Sandberg[5], J. M. A. Scherpen[6]

[1]Department of Electrical and Electronic Engineering, Imperial College London, United Kingdom
[2]Dipartimento di Ingegneria Civile e Ingegneria Informatica, Università di Roma "Tor Vergata", Italy
[3]Univ. of Illinois at Urbana-Champaign, Illinois, USA
[4]Graduate School of Advanced Science and Engineering, Hiroshima University, Japan
[5]KTH Royal Institute of Technology, Stockholm, Sweden
[6]Jan C. Willems Center for Systems and Control, Engineering and Technology Institute Groningen,
Faculty of Science and Engineering, University of Groningen, the Netherlands

a.astolfi@imperial.ac.uk beck3@illinois.edu d.bhattacharjee22@imperial.ac.uk
ykawano@hiroshima-u.ac.jp a.moreschini@imperial.ac.uk hsan@kth.se
j.m.a.scherpen@rug.nl

Abstract— **The approximation of complex dynamical systems models by reduced order models has been considered an important research problem for over four decades, not only in the field of control, but also in economics, image processing, circuit analysis, statistical mechanics, aircraft structures, and more recently in hybrid energy systems, to name just a sampling of fields. In this paper, we provide an overview of the development of balanced truncation and interpolation approaches for reducing linear and non-linear dynamical systems models for the purpose of control analysis and design.**

I. Introduction

The construction of reduced order models for dynamical systems has long been considered an important research problem in the field of control, beginning perhaps with the linear system realization problems proposed by Ho and Kalman for learning minimal models from data [1]. Model reduction, however, transcends the field of control, and has also garnered interest in the fields of economics [2], image processing [3], circuit analysis and simulation [4], and statistical mechanics [5], [6], to cite just a few specific examples. The classic problem formulation is this: given a difference or differential equation model for a dynamical system, determine a relevant simplified model that facilitates tractable system analysis and controller synthesis where none may have existed before. Interest in this problem grew with the interest in optimal controller synthesis when it was realized that the growth in computational complexity for an optimal control design is faster than $O(n^3)$, where n represents the state dimension of the model or the order of the differential equations describing the dynamical behavior of the system. Any particular reduction approach has since been judged, generally speaking, on the level of complexity reduction achieved, how closely the reduced model captures the original system behavior, and the complexity of the reduction process itself.

The classic formulation of the model reduction problem in the controls community, following the foundation in realization theory, began more directly with the balancing and principal component analysis perspective elegantly stated by B. Moore [7]. With time, the mathematical models considered relevant in system and control problems have increased in both dimension and complexity, involving distributed and interconnected systems of dynamical systems, switching systems, and complex nonlinear dynamics. At this point, the main model reduction approaches that have been extensively pursued for the purpose of control are balanced truncation, and interpolation or moment matching methods (see, for example [8], [9, Chapters 7, 8], [10, Chapters 9, 10]. While the development of these methods typically has origins in linear systems theory, they have been fully extended to nonlinear systems [11]–[16], with data-based approaches more recently being pursued [17]. In this tutorial paper, we provide an overview of the development of these methods for linear and nonlinear control systems models.

A. A Brief Overview of Model Reduction from Classic to Current

Since its introduction to the control community around 40 years ago, balanced truncation has been a cornerstone of model reduction, especially for the reduction of linear systems. So-called balanced realizations were introduced in 1976 [18], to minimize round-off errors in digital implementations of linear filters. In 1981 [7], B. Moore proposed to truncate such realizations to construct reduced order models, motivated by the principal component analysis of the controllability and observability operators; E. Verriest provided the first extension of these notions to analytic time-varying systems [19]. The key features of balanced

This work has been partially supported by the European Union's Horizon 2020 Research and Innovation Programme under grant agreement No. 739551 (KIOS CoE); by the Italian Ministry for Research in the framework of the 2020 Program for Research Projects of National Interest (PRIN), Grant 2020RTWES4; by the EPSRC grants EP/W005557 and EP/X033546; by NSF grant ECCS 20-32321, the Swedish Research Council under Grant 2016-00861; and by the JST FOREST Program Grant Number JPMJFR222E and JSPS KAKENHI Grant Number JP24K00910.

truncation, including guaranteed preservation of model stability [20], [21] and the a priori error bounds established in the mid-80's [21]–[23], make it an attractive first method of choice. The extension of these results to the discrete time case in 1990 introduced the use of Lyapunov *inequalities* [24], opening the door to Linear Matrix Inequality (LMI) approaches to model reduction.

Based on the LMI framework, notable extensions of balanced truncation methods were then developed for linear parameter varying (LPV) systems [25]; uncertain and multi-dimensional systems [26]; linear time-varying (LTV) systems [27], [28]; non-stationary LPV systems [29]; jump-linear systems [30]; and interconnected and structured systems [31], with some of the more recent work including extensions to nonstationary LPV distributed systems [32].

Additional related model reduction results have been established for optimal $\mathscr{H}_\infty$ model reduction [33]; Krylov and projection-based reduction methods [34]; and stochastic model reduction [35]. Algorithms that improved the computational efficiency of balanced model reduction early on also played a main role in establishing the popularity of the method, with key examples including Schur and Cholesky factorization based algorithms [36]–[38].

Remark 1. *Mainstream balancing-based model reduction methods assume the underlying linear systems of interest are stable. This constraint was addressed in 1988 by Meyer, who proposed the use of coprime factorizations for unstable system reduction [39]. Extensions of the coprime factors approach to more complex unstable systems followed [40]–[42].*

Remark 2. *Direct connections of the principal component analysis concepts underlying balanced truncation methods clearly exist with the subspace identification methods first proposed in the 1980s [43]–[45]. Specifically, the principal components of linear systems are those that are both strongly controllable and observable, or equivalently, those components with large associated Hankel singular values. Both balanced truncation and subspace identification are based on determining these components. Details are provided in Section II. Interestingly, recent results on learning low-order models from data based on finite-size sample data sets have brought the field full-circle back to consideration of the original Ho-Kalman realization algorithm; see for example [46]–[48].*

Extensions of the balanced truncation approaches for model reduction of nonlinear systems was first developed in [11], where "local" results were established, that is results on a neighborhood of an equilibrium admitting smooth solutions to Lyapunov type equations for the nonlinear state space. These local results were shown to be applicable to larger regions than would be obtained by simply linearizing the nonlinear system. The relation to balancing of the linearized system was also established.

Data-based reduction of nonlinear systems in a balanced system framework were established in [49], where empirical balanced realizations for nonlinear systems are constructed from both simulated and observed data, which coincides with balanced realizations when considering linear systems. A specific projection is then applied (the Galerkin projection) to the balanced realization to construct low-dimensional nonlinear models. Related empirical methods based in proper orthogonal decomposition (POD)s were further developed [50], which generalized some of these results.

An alternative approach to model reduction is that based on rational interpolation theory. Over the past two to three decades, substantial progress has been made in interpolation based reduction methods, leading to these methods emerging as another main choice for model reduction of large dynamical systems and nonlinear systems. A projection framework for interpolatory model reduction was introduced by Skelton and collaborators [51], [52]. For SISO systems these methods are sometimes referred to as rational Krylov methods, but for MIMO systems these are known more widely as interpolation or moment-matching methods. These methods are based on the use of projections, and thus have some connections to the earlier mentioned POD methods; for details see [53]–[56].

In the remainder of this tutorial paper, we present an overview of the classic model reduction problem formulation and review the timeline of developments in linear, nonlinear, and interpolation-based model reduction methods. We provide in-depth discussions of linear balanced-truncation based methods in Section II, including results for linear time-varying systems; data-based balanced reduction methods for nonlinear systems in Section III; and moment matching reduction methods based on closed-loop interpolation schemes for nonlinear systems in Section IV-A. Open problems in model reduction will be highlighted throughout.

II. Balanced Truncation of Linear Systems: Fundamentals and Extensions

In this section, we will first review the fundamentals of balanced truncation. We will discuss the balancing of controllability and observability Gramians and how this leads to classes of reduced-order models with desirable properties. We will highlight the definition and importance of Hankel singular values for model-order selection and both upper and lower bounds on the $\mathscr{H}_\infty$-approximation error [22], [23]. After this, we will mention some of the extensions to balanced truncation that have appeared since its inception. There are several excellent books available that cover balanced truncation and its extensions in much greater detail; see [8]–[10], [57], [58]

A. The Fundamentals

Consider a linear state-space system $\mathbf{G}$ with realization

$$\begin{aligned} \dot{x} &= Ax + Bu \\ y &= Cx + Du, \end{aligned} \tag{1}$$

with *state* $x(t) \in \mathbb{R}^n$, *input* $u(t) \in \mathbb{R}^m$, and *output* $y(t) \in \mathbb{R}^p$. Balanced truncation applies in continuous and discrete time with only minor differences[1]. We will focus on continuous time here for simplicity. Often we assume A is a Hurwitz matrix, and (1) is asymptotically stable. We denote its transfer function by $G(s) := C(sI-A)^{-1}B+D$, which lies in the space $\mathscr{H}_\infty$ ($\|G\|_\infty := \sup_{s\in\mathbb{C}_+} \bar{\sigma}(G(s)) < \infty$). As mentioned in the previous section, the goal of model reduction is to find a reduced-order model $\mathbf{G}_r$ with realization

$$\begin{aligned} \dot{z} &= A_r z + B_r u \\ y_r &= C_r z + D_r u, \end{aligned} \tag{2}$$

with reduced *state* $z(t) \in \mathbb{R}^r$, same *input* $u(t) \in \mathbb{R}^m$, and same *output* $y_r(t) \in \mathbb{R}^p$. As previously noted, model reduction seeks to find $\mathbf{G}_r$ such that (i) $r \ll n$; and (ii) error$(\mathbf{G},\mathbf{G}_r)$ is small. The success of balanced truncation is based on the fact that it provides suggestions for reasonable dimensions r, together with (A_r,B_r,C_r,D_r) and simple a priori bounds on the error in the $\mathscr{H}_\infty$-norm. We will review the required steps next.

The first step involves quantifying the controllability and observability properties of (1). The great insight of B. Moore in [7] was that states that are simultaneously hard to influence from u *and* difficult to see from y are good candidates for truncation from the model. Hence, we need to choose coordinates of the state space x that reflect both these properties. The controllability Gramian P and observability Gramian Q satisfy the Lyapunov equations

$$\begin{aligned} AP + PA^\top + BB^\top &= 0, \\ A^\top Q + QA + C^\top C &= 0. \end{aligned}$$

When A is Hurwitz, Gramians P and Q are unique and always positive semidefinite (positive definite in case the system realization is minimal). The eigenvectors of P and Q characterize the principal components of the controllability and observability operators of (1), respectively; see [7]. Thus, the controllability and observability of state elements should not be viewed as merely binary concepts; instead, they can be quantitatively assessed.

If we perform the coordinate transformation $x \leftarrow T^{-1}x$, for invertible matrices T, the controllability and observability Gramians transform as

$$P \leftarrow T^{-1}PT^{-\top}, \quad Q \leftarrow T^\top QT. \tag{3}$$

Interestingly, we see that the eigenvalues

$$\sigma_i := \sqrt{\lambda_i(PQ)}, \quad i = 1,2,\ldots,n,$$

are invariant under coordinate transformations of (1). It is shown in [22] that these eigenvalues are the singular values of the Hankel operator of the system $\mathbf{G}$. Hence, σ_i are called the *Hankel singular values* of $\mathbf{G}$, and are system invariants with many interesting properties.

More importantly, it was shown in [7], [18] that there exists a specific transformation T that *balances* the state coordinates[2] and achieves

$$P = Q = \Sigma, \tag{4}$$

$$\Sigma = \mathrm{diag}\{\sigma_1,\sigma_2,\ldots,\sigma_n\}, \quad \sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_n > 0. \tag{5}$$

Hence, for such a *balanced realization* of $\mathbf{G}$, the Gramians are identical and diagonal, and the elements coincide with the Hankel singular values. The state coordinate x_n, corresponding to σ_n, is the least simultaneously controllable and observable state direction. Intuitively, it is the state coordinate that participates the least in the mapping from u to y in $\mathbf{G}$, and appears suitable for truncation. The balanced Gramian is partitioned as

$$\Sigma = \begin{bmatrix} \Sigma_1 & 0 \\ 0 & \Sigma_2 \end{bmatrix},$$

where

$$\Sigma_1 = \begin{bmatrix} \sigma_1 I_{r_1} & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \sigma_l I_{r_l} \end{bmatrix}, \quad \Sigma_2 = \begin{bmatrix} \sigma_{l+1} I_{r_{l+1}} & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \sigma_N I_{r_N} \end{bmatrix},$$

where $n = r_1 + \ldots + r_N$, $r = r_1 + \ldots + r_l$, and the singular values are re-arranged in strictly decreasing order such that $\sigma_i \neq \sigma_j$, $i \neq j$. This notation is introduced to strengthen the results when singular values have a multiplicity greater than one, i.e., $r_i > 1$ for some i. Conformably to Σ_1 and Σ_2, the balanced realization (1) of $\mathbf{G}$ is partitioned into

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}, \quad B = \begin{bmatrix} B_1 \\ B_2 \end{bmatrix}, \quad C = \begin{bmatrix} C_1 & C_2 \end{bmatrix}.$$

As already hinted at, balanced truncation now proceeds by truncating all the states $x_{r+1},\ldots,x_n$ to obtain the realization (2) of $\mathbf{G}_r$ as

$$(A_r,B_r,C_r,D_r) = (A_{11},B_1,C_1,D).$$

Intuitively, we expect a small approximation error if the elements in Σ_2 are much smaller than those in Σ_1. Remarkable and powerful results have been proven to support this intuition. First, L. Pernebo and L. Silverman [20] showed that $A_r = A_{11}$ is guaranteed to be asymptotically stable if no elements in Σ_1 and Σ_2 coincide. The assured maintenance of stability throughout the reduction process has consistently been a significant consideration in model reduction, and the result has been rightfully celebrated. Two years later, K. Glover [22] and D. F. Enns [23] independently showed an elegant upper a priori error bound that has been equally celebrated. We state the bound in the following theorem, which closely follows the statement in [23].

Theorem 1. *Suppose (A,B,C,D) is a balanced realization and that (A_{11},B_1,C_1,D) is a balanced truncation such that Σ_1 and Σ_2 have no entries in common. Then A_{11} is Hurwitz,*

[1]Stronger results on the stability of reduced models are available in discrete time [24].

[2]Numerically efficient methods for computing such coordinate transformations are discussed in, for example, [8]. However, computing and balancing the Gramians remain the main computational bottleneck for applying balanced truncation to very high-dimensional systems.

and (A_{11}, B_1, C_1, D) is a minimal and balanced realization of $\mathbf{G}_r$ with Gramian Σ_1. Furthermore,

$$\|G - G_r\|_\infty \leq 2 \sum_{i=l+1}^{N} \sigma_i.$$

When $l = N-1$, equality holds and $\bar{\sigma}(G(0) - G_r(0)) = 2\sigma_N$ if r_N is odd.

The upper error bound in Theorem 1 should be compared with the following fundamental lower bound on the approximation error [22],

$$\inf_{G_r \in \mathscr{H}_\infty(r)} \|G - G_r\|_\infty \geq \sigma_{r+1}, \tag{6}$$

where $\mathscr{H}_\infty(r)$ denotes transfer functions in $\mathscr{H}_\infty$ with McMillan degree less than or equal to r. The lower bound cannot always be attained with equality, but Theorem 1 shows that balanced truncation is no worse than a factor of two when $l = m-1$. The lower bound (6) also illustrates that the Hankel singular values reveal what approximation orders r are possible for a given threshold on the $\mathscr{H}_\infty$-norm error.

Since $D_r = D$ for balanced truncation, a perfect model fit is achieved at infinite frequency, $G(\infty) = G_r(\infty)$. It is often more desirable to have a good approximation at smaller frequencies. To this end, a singular perturbation approximation of a balanced realization of $\mathbf{G}$,

$$\begin{aligned}(A_r, B_r, C_r, D_r) = (&A_{11} - A_{12}A_{22}^{-1}A_{21}, B_1 - A_{12}A_{22}^{-1}B_2,\\ &C_1 - C_2A_{22}^{-1}A_{21}, D - C_2A_{22}^{-1}B_2),\end{aligned}$$

achieves $G(0) = G_r(0)$ and was shown to satisfy the same error bounds as balanced truncation in [59].

B. Time-Varying and Other Extensions

Early extensions of these classic results included balanced truncation of frequency-weighted [23], time-varying [19], [27], [28], [60], uncertain [26], parameter-varying [25], [61], and stochastic systems [35], [62], [63].

To provide some detail, consider the linear time-varying system $\mathbf{G}$ with realization

$$\begin{aligned}\dot{x} &= A(t)x + B(t)u\\ y &= C(t)x + D(t)u.\end{aligned}$$

In the works [19], [60], it was proven that it is possible to balance the coordinates using time-varying transformations (3), under appropriate assumptions on the smoothness of the realization $(A(t), B(t), C(t), D(t))$. (See [64] for the discrete-time case.) This balancing process leads to diagonal, time-varying Gramians,

$$\begin{aligned}&P(t) = Q(t) = \Sigma(t),\\ &\Sigma(t) = \operatorname{diag}\{\sigma_1(t), \sigma_2(t), \ldots, \sigma_n(t)\},\end{aligned}$$

which satisfy time-varying differential Lyapunov equations, or LMIs more generally, that is,

$$\begin{aligned}A(t)\Sigma(t) + \Sigma(t)A^\top(t) - \dot{\Sigma}(t) + B(t)B^\top(t) &\leq 0\\ \Sigma(t)A(t) + A^T\Sigma(t) + \dot{\Sigma}(t) + C^\top(t)C(t) &\leq 0.\end{aligned} \tag{7}$$

The time-varying Hankel singular values $\sigma_i(t)$, $i = 1, 2, \ldots, n$, have a similar meaning here as in the time-invariant case, for each time t. However, one now needs to take some care in specifying the boundary conditions of $\Sigma(t)$ and the considered time interval $[t_0, t_\mathrm{f}]$. After balancing, one can proceed to truncate states to obtain a time-varying reduced system $\mathbf{G}_r$ with realization $(A_{11}(t), B_1(t), C_1(t), D(t))$. In the works [27], [28] error bounds generalizing those given in Theorem 1 and (6) were obtained. For example, in [28] it was shown that

$$\max_t \sigma_{r+1}(t) \leq \|\mathbf{G} - \mathbf{G}_r\| \leq 2 \sum_{i=l+1}^{N} \mathscr{S}_{T_i}(\sigma_i), \tag{8}$$

where $\mathscr{S}_{T_i}$ is the so-called max-min ratio (see (9) below). Here we understand $\mathbf{G}$ and $\mathbf{G}_r$ as operators on $\mathscr{L}_2[t_0, t_\mathrm{f}]$, and $\|\cdot\|$ is the induced norm on $\mathscr{L}_2[t_0, t_\mathrm{f}]$. The lower bound in (8) assumes the boundary conditions $\Sigma(t_0) = \Sigma(t_\mathrm{f}) = 0$ and applies to any r-th order linear time-varying system. The upper bound, on the other hand, applies to systems arising from truncated balanced realizations, and any solution $\Sigma(t)$ to (7) can be used. If we find a solution $\Sigma(t)$ where the singular values truncated are monotonic, the max-min ratio $\mathscr{S}_{T_i}$ simplifies to a maximum. In particular, if a constant solution $\Sigma(t) = \Sigma$ to the LMIs is found, the bound in Theorem 1 is recovered. The existence of constant solutions is further discussed in [27].

The max-min ratio $\mathscr{S}_{T_i}$ appearing in (8) is defined as

$$\mathscr{S}_{[t_0, t_\mathrm{f}]}(\sigma) := \sigma(t_0) \prod_i \frac{\sigma(t_i^{\max})}{\sigma(t_i^{\min})}, \tag{9}$$

where the local minima and maxima of σ over the time interval $[t_0, t_\mathrm{f}]$ have been ordered as

$$t_0 \leq t_1^{\min} < t_1^{\max} < t_2^{\min} < t_2^{\max} \cdots \leq t_\mathrm{f}.$$

The magnitude and the time-variation of the singular values affect the bound. An interesting feature of model reduction of time-varying systems is that some states may be of importance to the system $\mathbf{G}$ only over some subset of time $T_i \subset [t_0, t_\mathrm{f}]$. It is then possible to construct a reduced model $\mathbf{G}_r$ of time-varying order, where state x_i is only truncated over time intervals T_i where $\mathscr{S}_{T_i}(\sigma_i)$ is small [28].

Further extensions considered generalized balanced truncation, which is based on generalized Gramians that satisfy Lyapunov LMIs. Attractive reduction properties using generalized Gramians include potentially better error bounds [24][3] and that underlying structural constraints in the system of interest can be enforced in the reduction process [40], [66]. Generalized Gramians are particularly applicable in the model reduction of multi-dimensional and uncertain systems [26], [41], [67]. The so-called extended balanced truncation approach [31] was later introduced to handle interconnection constraints between connected subsystems in the model reduction process.

[3] Generalized Gramians also appear in the characterization of the solutions to the optimal $\mathscr{H}_\infty$-model approximation problem [26], [65].

There are numerous related open research problems, with perhaps those receiving the greatest focus at this time being related to finite sample learning of low-order LTI models [48]. Along these lines, related finite sample analyses of subspace system identification methods have very recently been considered [68]. Future directions include improved sample complexity bounds, and extending non-asymptotic low-order LTI model learning to bilinear and nonlinear system models.

III. Data-Based Model Reduction for Non-Linear Systems Based on Differential Balancing

For linear systems, the controllability Gramian P and observability Gramian Q can be computed using impulse and initial state responses, respectively, e.g., [8]. This allows for balanced truncation of LTI systems using empirical data. The application of linear data-driven methods to nonlinear systems has garnered significant research interest, e.g., [49], [69]–[73]. These methods aim to reduce the computational complexity of nonlinear controller design, such as model predictive control, e.g., [74], [75]. However, these data-driven methods have traditionally been proposed only around a steady-state. To capture nonlinear behavior, data-driven methods explicitly taking nonlinearity into account have been developed for controllability in a stochastic setting [76] and for observability [77], [78]. However, neither method addresses both controllability and observability Gramians simultaneously, and there is no direct connection between these two approaches.

In this section, we discuss a data-driven balancing method for a nonlinear system by utilizing its variational system. This theoretical framework is referred to as differential balancing theory [13], [79], [80]. Since the variational system can be viewed as an LTV system along the trajectory of the nonlinear system, one can extend the concept of the controllability and observability Gramians of the LTV system [27], [28], [81], [82]. We call these extensions the differential reachability and observability Gramians, respectively. These Gramians depend on the state trajectory of the nonlinear system, and their values at each fixed trajectory can be computed from the impulse and initial state responses of the variational system along this fixed trajectory. The obtained trajectory-wise Gramians are constant matrices, allowing for the computation of balanced coordinates and a reduced-order model in a manner similar to the LTI case.

A. The Fundamentals

Consider a nonlinear system with constant input vector fields, described by

$$\begin{cases} \dot{x} = f(x) + Bu \\ y = h(x) \end{cases} \tag{10}$$

with state $x(t) \in \mathbb{R}^n$, input $u(t) \in \mathbb{R}^m$, and output $y(t) \in \mathbb{R}^p$, where $f : \mathbb{R}^n \to \mathbb{R}^n$ and $h : \mathbb{R}^n \to \mathbb{R}^p$ are of class C^2 and $B \in \mathbb{R}^{n\times m}$. Let $\varphi_{t-t_0}(x_0,u)$ denote the state trajectory $x(t)$ of the system (10) starting from $x(t_0) = x_0 \in \mathbb{R}^n$ for each continuous and bounded $u : \mathbb{R} \to \mathbb{R}^m$. Note that since f is of class C^2, if u is also of class C^2, then the solution $\varphi_{t-t_0}(x_0,u)$ is a class C^2 function of (t,x_0) as long as it exists.

The variational system of (10) along its trajectory $\varphi_{t-t_0}(x_0,u)$ is

$$\begin{cases} \dot{\delta x} = \dfrac{\partial f(\varphi_{t-t_0})}{\partial \varphi_{t-t_0}} \delta x + B\delta u \\ \delta y = \dfrac{\partial h(\varphi_{t-t_0})}{\partial \varphi_{t-t_0}} \delta x \end{cases} \tag{11}$$

with state $\delta x(t) \in \mathbb{R}^n$, input $\delta u(t) \in \mathbb{R}^m$, and output $\delta y(t) \in \mathbb{R}^p$. As long as $\varphi_{t-t_0}(x_0,u)$ exists, the solution $\delta x(t)$ exists for any $\delta x(t_0) \in \mathbb{R}^n$ and continuous and bounded $\delta u : \mathbb{R} \to \mathbb{R}^m$ because the variational system is an LTV system for fixed (x_0,u).

Mimicking the LTI and LTV cases [81], [82], the differential reachability and observability Gramians are defined as Gramians of the variational systems [79].

Definition 1. *Given $x_0 \in \mathbb{R}^n$ and class C^2 bounded $u : [t_0,t_f] \to \mathbb{R}^m$, the differential reachability Gramian is defined by*

$$G_{\mathscr{R}}(t_0,t_f,x_0,u) := \int_{t_0}^{t_f} \frac{\partial \varphi_{t-t_0}}{\partial x_0} B \left(\frac{\partial \varphi_{t-t_0}}{\partial x_0} B \right)^\top dt \tag{12}$$

and the differential observability Gramian is defined by

$$\begin{aligned} &G_{\mathscr{O}}(t_0,t_f,x_0,u) \\ &:= \int_{t_0}^{t_f} \left(\frac{\partial h(\varphi_{t-t_0})}{\partial \varphi_{t-t_0}} \frac{\partial \varphi_{t-t_0}}{\partial x_0} \right)^\top \frac{\partial h(\varphi_{t-t_0})}{\partial \varphi_{t-t_0}} \frac{\partial \varphi_{t-t_0}}{\partial x_0} dt, \end{aligned} \tag{13}$$

where the arguments of φ_{t-t_0} are (x_0,u).

The differential Gramians exist in $[t_0,t_f]$ because the solution $\varphi_{t-t_0}(x_0,u)$ exists and is a class C^2 function of (t,x) in $[t_0,t_f] \times \mathbb{R}^n$ from the assumption. In the LTI case, the Gramians defined by (12) and (13) respectively are the controllability Gramian P and observability Gramian Q. The differential reachability Gramian does not coincide with the differential controllability function [13] whereas the differential observability Gramian does with the differential observability function. The motivation of introducing the reachability Gramian is to develop its numerical computational method by using trajectories forward in time.

Similarly to the LTI case, one can define a balanced realization between the differential reachability and observability Gramians. Since these differential Gramians depend on a trajectory $\varphi(x_0,u)$, we define our balanced realization trajectory-wise as follows [79].

Definition 2. *Given $\varphi_{t-t_0}(x_0,u)$, suppose that the differential reachability Gramian $G_{\mathscr{R}}(t_0,t_f,x_0,u) \in \mathbb{R}^{n\times n}$ and differential observability Gramian $G_{\mathscr{O}}(t_0,t_f,x_0,u) \in \mathbb{R}^{n\times n}$ are positive definite. A realization of the system (10) is said to be a differentially balanced realization along $\varphi_{t-t_0}(x_0,u)$ if there exists a constant diagonal matrix*

$$\Lambda = \mathrm{diag}\{\sigma_1,\dots,\sigma_n\}, \;\; \sigma_1 \geq \cdots \geq \sigma_n > 0$$

such that $G_{\mathscr{R}}(t_0,t_f,x,u) = G_{\mathscr{O}}(t_0,t_f,x,u) = \Lambda$.

It is possible to show that there always exists a differentially balanced realization along $\varphi_{t-t_0}(x_0,u)$ if the differential Gramians are positive definite [79].

Theorem 2. *Suppose that the differential Gramians $G_{\mathscr{R}}(t_0,t_f,x_0,u)$ and $G_{\mathscr{O}}(t_0,t_f,x_0,u)$ are positive definite at fixed $\varphi_{t-t_0}(x_0,u)$. Then, there exists a non-singular matrix $T_\varphi \in \mathbb{R}^{n\times n}$ which achieves*

$$\begin{aligned} &T_\varphi G_{\mathscr{R}}(t_0,t_f,x_0,u)T_\varphi^\top \\ &= T_\varphi^{-\top} G_{\mathscr{O}}(t_0,t_f,x_0,u)T_\varphi^{-1} = \Lambda. \end{aligned} \tag{14}$$

Note that T_φ in Theorem 2 is a constant matrix. Therefore, as in the LTI case, a reduced-order model can be constructed by applying a change of coordinates $z = T_\varphi x$. The obtained reduced order model is expected to have similar input-output behaviour as that of the original system (10) around the trajectory $\varphi_{t-t_0}(x_0,u)$.

B. Data-Based Computations of Differential Gramians

It is computationally challenging to compute the differential Gramians analytically as typically the implementation of nonlinear balancing theory requires solving nonlinear partial differential equations (PDEs), e.g., [13], [83]–[85]. However, computing the values of the differential Gramians along a fixed trajectory $\varphi_{t-t_0}(x_0,u)$ is numerically tractable as shown in this subsection.

First, we show that the differential reachability Gramian $G_{\mathscr{R}}(t_0,t_f,x_0,u)$ along a fixed trajectory $\varphi_{t-t_0}(x_0,u)$ can be computed by using an impulse response of the variational system. Let $\delta_D(\cdot)$ be Dirac's delta function, and let $\delta x_{\mathrm{Imp,i}}(t)$ be the impulse response of the variational system (11) along the trajectory $\varphi_{t-t_0}(x_0,u)$ with $\delta u(t) = e_i^m \delta_D(t-t_0)$, where $e_i^m \in \mathbb{R}^m$ is the standard basis. Then, substituting $\delta x_0 = 0$ and $u(t) = e_i^m \delta_D(t-t_0)$ into (11) yields

$$\delta x_{\mathrm{Imp},i}(t) = \frac{\partial \varphi_{t-t_0}(x_0,u)}{\partial x} B_i, \tag{15}$$

where B_i is the ith column vector of B. Note that $\delta x_{\mathrm{Imp},i}(t)$ exists as long as $\varphi_{t-t_0}(x_0,u)$ exists. From (12), we obtain

$$\begin{aligned} G_{\mathscr{R}}(t_0,t_f,x_0,u) &= \int_{t_0}^{t_f} \delta x_{\mathrm{Imp}}(t)\delta x_{\mathrm{Imp}}^\top(t)dt, \\ \delta x_{\mathrm{Imp}}(t) &:= \begin{bmatrix} \delta x_{\mathrm{Imp},1}(t) & \cdots & \delta x_{\mathrm{Imp},m}(t) \end{bmatrix}. \end{aligned}$$

Thus, for each $\varphi_{t-t_0}(x_0,u)$, the value of the differential reachability Gramian $G_{\mathscr{R}}(t_0,t_f,x_0,u)$ is obtained by using the impulse response of the variational system (11). It is worth mentioning that (15) does not hold if B is not constant in general, which is the main reason of focusing on constant input vector fields.

Next, we show that the differential observability Gramian $G_{\mathscr{O}}(t_0,t_f,x_0,u)$ along a fixed trajectory $\varphi_{t-t_0}(x_0,u)$ can be computed by using initial state responses. Substituting $\delta x_0 = e_i^n$ and $\delta u = 0$ into (11), we have the following initial output response of the variational system (11) along $\varphi_{t-t_0}(x_0,u)$,

$$\delta y_{\mathrm{Is},i}(t) = \frac{\partial h(\varphi_{t-t_0}(x_0,u))}{\partial x}\frac{\partial \varphi_{t-t_0}(x_0,u)}{\partial x} e_i^n, \tag{16}$$

From (13), we obtain

$$\begin{aligned} G_{\mathscr{O}}(t_0,t_f,x_0,u) &= \int_{t_0}^{t_f} \delta y_{\mathrm{Is}}^\top(t)\delta y_{\mathrm{Is}}(t)dt, \\ \delta y_{\mathrm{Is}}(t) &:= \begin{bmatrix} \delta y_{\mathrm{Is},1}(t) & \cdots & \delta y_{\mathrm{Is},n}(t) \end{bmatrix}. \end{aligned}$$

Therefore, for each $\varphi_{t-t_0}(x_0,u)$, the value of the differential observability Gramian $G_{\mathscr{O}}(t_0,t_f,x_0,u)$ is obtained by using the initial state responses of the variational system (11).

In summary, the values of the differential reachability/observability Gramian at given (x_0,u) is obtained by computing impulse/initial state responses of the variational system (11) along $\varphi_{t-t_0}(x_0,u)$. Hence, trajectory-wise differential balanced truncation is doable based on empirical data.

C. Approximation without Variational Dynamics

The data-based approach in the previous subsection requires the variational system model in addition to the original system model. Computing the variational system model may need an effort, which motivates us to develop an approximation method based on trajectories of the original system only.

The main idea for the approximation is based on the fact that the variational system (11) coincides with the Fréchet derivative of a nonlinear operator $(x_f,y) = \Sigma(x_0,u)$ induced by the system (10), where the Fréchet derivative is the following linear operator

$$\begin{aligned} &d\Sigma_{(x_0,u)}(\delta x_0,\delta u) \\ &:= \lim_{s\to 0} \frac{\Sigma(x_0+s\delta x_0, u+s\delta u) - \Sigma(x_0,u)}{s}. \end{aligned}$$

Its simple approximation is

$$\begin{aligned} &d\Sigma_{(x_0,u)}(\delta x_0,\delta u) \\ &\approx d\Sigma^{\mathrm{app}}_{(x_0,u)}(\delta x_0,\delta u) \\ &:= \frac{\Sigma(x_0+s\delta x_0, u+s\delta u) - \Sigma(x_0,u)}{s}. \end{aligned}$$

Since the nonlinear operator $\Sigma(x_0,u)$ is given by the system (10), a state space representation of the discretized approximation $d\Sigma^{\mathrm{app}}_{(x_0,u)}(\delta x_0,\delta u)$ is

$$\begin{cases} \dot{x} = f(x)+Bu, \quad x(t_0)=x_0,\ u=u \\ \dot{x}' = f(x')+Bu',\ x'(t_0)=x_0+s\delta x_0,\ u'=u+s\delta u \\ x_{vf} = \dfrac{x'(t_f)-x(t_f)}{s},\ y_v = \dfrac{h(x')-h(x)}{s}. \end{cases}$$

Therefore, we have the following approximations

$$\begin{aligned} \delta x(t) &\approx \frac{x'(t)-x(t)}{s} \\ \delta y(t) &\approx y_v(t), \end{aligned}$$

where δx_0 and δu coincide with the differences of a pair of the initial states $(x_0'-x_0)/s$ and a pair of inputs $(u'-u)/s$, respectively.

Consequently, an approximation of the impulse response (15) is computed by

$$\begin{aligned} &\delta x_{\mathrm{Imp,i}}(t) \approx \frac{x'(t)-x(t)}{s} \\ &\qquad \delta x_0 = 0,\ \delta u = e_i^m \delta_D(t-t_0),\ i=1,2,\ldots,m. \end{aligned}$$

Similarly, an approximation of the initial state response (16) is computed by

$$\begin{aligned}\delta y_{\mathrm{Is},i}(t) &\approx y_v(t)\\ &\delta x_0 = e_i^n,\ \delta u = 0,\ i = 1,\dots,n.\end{aligned}$$

One notices that these computations only require trajectories of the system (10). If these data are available, one can compute a change of coordinates $z = T_\varphi x$ for a differential balanced realization along $\varphi_{t-t_0}(x_0,u)$ without a system model.

The remaining challenge is to construct a reduced order model having a good approximation for all trajectories, because this essentially requires solving nonlinear PDEs. A potential approach is to employ the basic idea of proper orthogonal decomposition [8], [86]. First, we compute the summation of differential reachability and observability Gramians,

$$\begin{aligned}G_{\mathscr{R}}(t_0,t_f) &:= \frac{1}{r}\sum_{i=1}^{r} G_{\mathscr{R}}(t_0,t_f,x_i,u_i)\\ G_{\mathscr{O}}(t_0,t_f) &:= \frac{1}{r}\sum_{i=1}^{r} G_{\mathscr{O}}(t_0,t_f,x_i,u_i)\end{aligned}$$

for different choices of trajectories $\varphi_{t-t_0}(x_i,u_i)$, $i = 1,\dots,r$. Then, we construct a linear change of coordinates which simultaneously diagonalizes $G_{\mathscr{R}}(t_0,t_f)$ and $G_{\mathscr{O}}(t_0,t_f)$.

D. Remarks

Related nonlinear balanced realizations are found in flow balancing [83]–[85]. In flow balancing, variational systems represent small perturbations, with a primary focus on local analysis. In contrast, differential balancing [13], motivated by contraction theory [87], [88], emphasizes global analysis. Leveraging contraction theory, differential balancing can preserve the incremental exponential stability of the original model through generalized or extended differential balancing [80], [89]. In this section, we have tailored differential balancing to a data-based approach. For approximating the Fréchet derivative, we utilized a pair of trajectories of the original system, reminiscent of incremental balancing, which is based on a pair of trajectories [14]. Unlike differential balancing, incremental balancing lacks a data-based approach. Investigating this is an interesting topic in itself. On the other hand, balanced truncation has been explored for preserving the structure of positive or monotone systems [90], [91]. These methods have strong connections with steady-state responses to constant inputs, such as DC gains, making them compatible with data-based approaches. For monotone systems, model reduction methods based on nonlinear DC gains have been investigated [92]. Since DC gains can be viewed as zero-moments, this suggests a possible link among balanced truncation, moment matching, and data-based approaches for monotone systems.

IV. Moment Matching of Nonlinear Systems: Fundamentals and Extensions

A. The Fundamentals

In this section we introduce the building blocks of model reduction by moment matching for nonlinear systems. Before we dive into the theory, we take a brief journey into the notion of moment for linear systems, making the connection between the classical (rational) interpolation theory [8], [93] and the interconnection-based interpretation [15] that has paved the way for the extension of the method to nonlinear systems.

1) Linear Systems [94]: Consider a linear, single-input, single-output, continuous-time, system described by the equations[4]

$$\begin{aligned}\dot{x} &= Ax + Bu,\\ y &= Cx,\end{aligned} \tag{17}$$

with state $x(t) \in \mathbb{R}^n$, input $u(t) \in \mathbb{R}$, output $y(t) \in \mathbb{R}$, and matrices, A, B, and C of appropriate dimensions. Let

$$W(s) = C(sI - A)^{-1}B$$

for all $s \in \mathbb{C}$ be the (rational) transfer function associated with (17) and assume that the state-space realization (17) is minimal, *i.e.* reachable and observable.

Definition 3 (See [95])**.** *Let $s_i \in \mathbb{C} \setminus \sigma(A)$ and $k \in \mathbb{N}$. The 0-moment of system* (17) *at s_i is given by $\eta_0(s_i) = W(s_i)$. For $k \geq 1$, the k-moment of the system* (17) *at s_i is given by*

$$\eta_k(s_i) = \frac{(-1)^k}{k!}\frac{d^k W}{ds^k}(s_i).$$

From the definition, it results that the moments of system (17) at s_i determine the coefficients of the Laurent series expansion [95] of the transfer function in the neighborhood of s_i.

With this notion of moment at hand, the moment matching problem can be formulated as an interpolation problem at operating points s_i's on the complex plane [95].

Moment Matching Problem: *Given $\{s_i\}_{i=1}^{\nu}$ and $\{k_i\}_{i=1}^{\nu}$ and $k_i \in \mathbb{N}$, find a (rational) transfer function of order ρ such that the associated k_i-moment at s_i, which is defined as $\overline{\eta}_{k_i}(s_i)$, verifies the interpolation condition*

$$\overline{\eta}_{k_i}(s_i) = \eta_{k_i}(s_i). \tag{18}$$

It is clear that the notion of moment, as per Definition 3, is limited to linear systems. The turning point result, which allows going beyond the frequency domain, has been recognized in [54], [55], where the moments $\eta_0(s_i),\dots,\eta_k(s_i)$ were linked to a Sylvester equation.

Lemma 1 (See [96])**.** *Consider system* (17) *with $s_i \in \mathbb{C} \setminus \sigma(A)$. Then the moments $\eta_0(s_i),\dots,\eta_k(s_i)$, for $i = 1,\cdots,N$,*

[4]The discussion in this section directly extends to DAE multi-input, multi-output (MIMO) systems, regardless of the time domain (discrete-time or continuous-time systems).

are in one-to-one relation with the matrix $C\Pi$*, where* Π *is the (unique) solution of the Sylvester equation*

$$A\Pi + BL = \Pi S, \tag{19}$$

with S, any non-derogatory real matrix with characteristic polynomial given by $\det(sI - S) = \prod_{i=1}^{N}(s - s_i)^{k+1}$*, and L is such that the pair* (S, L) *is observable.*

This intimate connection between $\eta_0(s_i), \ldots, \eta_k(s_i)$ and $C\Pi$ inspired [94] to revisit the moment matching problem for linear systems from a time-domain perspective. Indeed, from a geometric viewpoint, it can be shown that the cascade interconnected system

$$\begin{aligned} \dot{\omega} &= S\omega, \\ \dot{x} &= Ax + BL\omega, \\ y &= Cx, \end{aligned} \tag{20}$$

has a well-defined invariant set described by the equation $x = \Pi\omega$ and, accordingly, the restriction of (20) to the invariant set reproduces the behaviour of $S\omega$. With this in mind, the moment of system (17) can be redefined from a property of a point in the complex plane to a (linear) mapping on a real vector space. In particular, the moment of system (17) can be reinterpreted by means of the invariant set described by the equations $x = \Pi\omega$ with interpolation points generated by $\sigma(S) \in \mathbb{C} \setminus \sigma(A)$.

Definition 4 (See [96])**.** *The moment of system* (17) *at* (S, L) *is defined as* $C\Pi$*.*

The key fact stemming from the algebraic characterization of the moments is that, under certain assumptions, the moments boil down to the mapping characterizing the steady-state output response (if any) of (20). By leveraging these arguments, the concept of moments has been extended in [96] to nonlinear systems by taking advantage of the steady-state property of the cascade interconnected system.

2) Nonlinear Systems [96]: Consider a nonlinear, single-input, single-output, continuous-time, system described by the equations

$$\begin{aligned} \dot{x} &= f(x, u), \\ y &= h(x), \end{aligned} \tag{21}$$

with state $x(t) \in \mathbb{R}^n$, input $u(t) \in \mathbb{R}$, output $y(t) \in \mathbb{R}$, and mappings[5] f and h of appropriate dimension such that $f(0,0) = 0$ and $h(0) = 0$. We assume that (21) is (locally) observable and (locally) accessible. For the notion of moment to be well-defined, we need a (locally) observable signal generator described by the equations

$$\begin{aligned} \dot{\omega} &= s(\omega), \\ u &= \ell(\omega), \end{aligned} \tag{22}$$

where $\omega(t) \in \mathbb{R}^{\nu}$ and the mappings s and ℓ are of appropriate dimension and such that $s(0) = 0$ and $\ell(0) = 0$. The role of the signal generator is to excite the system (21) with signals

[5]All functions and mappings are assumed sufficiently smooth.

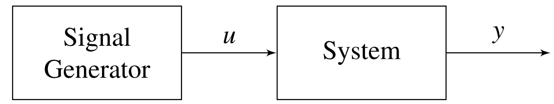

Fig. 1: Diagrammatic illustration of the cascade interconnection between the signal generator (22) and the nonlinear system (21).

of interest so that one can analyze (21) at specific operating conditions associated to the interconnected system

$$\begin{aligned} \dot{\omega} &= s(\omega), \\ \dot{x} &= f(x, \ell(\omega)), \\ y &= h(x). \end{aligned} \tag{23}$$

However, to extend the notion of moment to nonlinear systems we need to assume that there exists a (unique) mapping $\pi(\cdot)$, locally defined in a neighborhood of $\omega = 0$, which is the unique (analytic) solution of the partial differential equation

$$f\big(\pi(\omega), \ell(\omega)\big) = \frac{\partial \pi}{\partial \omega} s(\omega), \quad \pi(0) = 0. \tag{24}$$

Note that, the partial differential equation (24) is the nonlinear enhancement of the Sylvester equation (19).

Definition 5 (See [96])**.** *The moment of system* (21) *at* (s, ℓ) *is defined as* $h(\pi(\cdot))$*.*

Analogous to the linear case, the existence of $\pi(\cdot)$ implies that the interconnected system (23) has a well-defined invariant manifold described by the equation $x = \pi(\omega)$ and the restriction of (23) to the invariant manifold reproduces the behaviour of $\dot{\omega} = s(\omega)$.

Theorem 3 (See [96])**.** *Consider the system* (21) *and the signal generator* (22)*. Assume that* $x = 0$ *is a locally exponentially stable equilibrium for* $\dot{x} = f(x, 0)$ *and that* $\omega = 0$ *is a neutrally stable*[6] *equilibrium for* $\dot{\omega} = s(\omega)$*. Then the moment of system* (21) *at* (s, ℓ) *has a one-to-one relation with the (locally well-defined) steady-state output response of* (23)*.*

Stated another way, the moment $h(\pi(\cdot))$ is in one-to-one relation with the (locally well-defined) steady-state response of the output of the interconnected system (23) [96]. With this notion of moment at hand, the moment matching problem can be formulated as an interpolation problem at an operating signal in the time domain.

Moment Matching Problem (Revisited): *Given the signal generator* (22) *and the solution of* (24)*, find a (nonlinear) system such that the associated moment at* (s, ℓ)*, which is defined as* $\kappa(p(\cdot))$*, verifies the matching condition*

$$h(\pi(\cdot)) = \kappa(p(\cdot)). \tag{25}$$

Specializing the revisited moment matching problem to a state-space realization of order $\rho < n$ of the form

$$\dot{\xi} = \phi(\xi, u), \tag{26a}$$

$$\psi = \kappa(\xi), \tag{26b}$$

[6]$\omega = 0$ is a neutrally stable equilibrium if it is stable (in the sense of Lyapunov) and each $\omega(0)$ is Poisson stable, see [97, Sec. 8.1].

with $\xi(t) \in \mathbb{R}^\rho$, we say that (26) is a reduced order model of (21) if $\kappa(p(\cdot))$ is a (well-defined) moment at (s, ℓ) which satisfies (25), where $p : \mathbb{R}^\rho \to \mathbb{R}^\nu$ is the unique analytic solution of the partial differential equation

$$\phi(p(\omega), l(\omega)) = \frac{\partial p}{\partial \omega} s(\omega), \qquad p(0) = 0. \tag{27}$$

In [96] it is also shown that, for $\rho = \nu$, there exists a particular state-space realization of the reduced-order model (26), which is given by selecting

$$\begin{aligned} \phi(\cdot, u) &= s(\cdot) - \delta(\cdot, \ell(\cdot)) + \delta(\cdot, u), \\ \kappa(\cdot) &= h(\pi(\cdot)), \end{aligned} \tag{28}$$

where $\delta(\cdot, \cdot)$ is a free mapping that can be chosen arbitrarily to achieve moment matching with specific properties.

3) Bibliographical Notes: The literature on model reduction by moment matching from a time-domain perspective is rich and has evolved over the last decade. Exploiting (28), the problem of constructing reduced order models by preserving properties of interest has been studied in various works, such as passivity, zero dynamics, port-Hamiltonian structure, general second-order structure [98]–[100]. The nonlinear Petrov–Galerkin projection given in [101] is based on [102], whereas a general family of parameterized models achieving moment matching has been studied in [103]. The problem of model reduction by moment matching for other classes of systems has also received significant attention. This includes systems of differential–algebraic equations [103]–[106], systems with delays, stochastic systems, Lur'e type systems, hybrid systems [107]–[113]. Moment matching with an explicit signal generator (which generates a richer class of input signals) has been studied in [114], [115]. While most moment matching approaches rely on the availability of the model of the underlying system, data-driven moment matching techniques have gained popularity due to their ability to derive reduced-order models without the knowledge of the full-order model [116]–[119]. The interconnection-based framework was extended to the Loewner framework for linear and nonlinear systems [120]–[124]. We refer the reader to [56] for a comprehensive survey on the interconnection-based moment matching problem.

B. The Extensions

In this section, we focus on two main challenges: first, the design of a data-driven procedure for computing reduced-order models for an unknown system driven by an unknown implicit signal generator [118], [119]; second, the development of a model-based procedure for computing reduced-order models for a system without requiring internal stability of the system to be reduced [110], [111].

1) Signal Generator Agnostic Moment Matching [119]: While the majority of the literature on moment matching has focused on deriving reduced-order models from known models of the underlying system driven by a known signal generator, a breakthrough was achieved in [116], which provided a way of constructing reduced-order models given data obtained from an unknown underlying system. Even

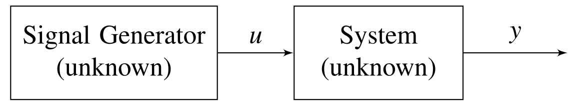

Fig. 2: Diagrammatic illustration of the interconnection between the (unknown) signal generator (35) and the (unknown) nonlinear system (21) in [119].

though this procedure serves as a point of departure in the realm of moment matching by not requiring the knowledge of the underlying system, it is still limited by its dependence on the knowledge of the structure of the signal generator and its internal state. In this respect, [118] has proposed a method to construct reduced-order models from input-output data in the case in which the models of the underlying system and of the signal generator driving it are unknown. Recall that the output response of the full-order model (21) is given by

$$y(t) = h(\pi(\omega(t))) + \varepsilon(t). \tag{29}$$

In a similar fashion, the output response of the reduced-order model (26) is given by

$$\psi(t) = \kappa(\xi(t)) + \varepsilon(t). \tag{30}$$

With this in mind, we can state the result of [118] that allows the definition of the moment of the underlying system in terms of the state of the reduced-order model.

Lemma 2. *Consider the system (21), the system (26), and the signal generator (22). Then system (26) matches the moment of system (21) at (s, l) and*

$$\lim_{t \to \infty} (h(\pi(\omega(t))) - \kappa(\xi(t))) = 0. \tag{31}$$

The result stated in Lemma 2 is particularly useful as it allows equation (29) to be transformed into a form in which the output of the underlying system is related to the state of the reduced-order model. To construct the reduced-order model, the mapping $\kappa(\cdot)$ must be identified. However, an exact identification may not be feasible and an approximation should be considered. Thus, we consider a standard assumption for approximating a mapping by a family of basis functions [125]. Typically, the family of basis functions representing the mapping can be implemented by a trial and error procedure by using a polynomial expansion or an expansion based on the class of input signals. It is worthwhile to note that there are results of "universal" approximation for some families of basis functions [126], [127]. In light of this, we now introduce an assumption on the mapping $\kappa(p(\cdot))$.

Assumption 1. *The mapping $\kappa(p(\cdot))$ belongs to the function space described by the family of continuous basis functions $\varphi : \mathbb{R}^\rho \to \mathbb{R}$, with $i = 1, \ldots, M$, i.e., there exist constants $\Gamma_i \in$ with $i = 1, \ldots, M$, such that $\kappa(\xi) = \sum_{i=1}^{M} \Gamma_i \varphi_i(\xi)$. If $M = \infty$, then we assume that the series is uniformly convergent.*

While it may be true that a specific family of basis functions may be able to exactly represent a certain mapping $\kappa(p(\cdot))$

with a finite basis, this is not the case generally. Therefore, consider the following approximation

$$\overline{\Gamma} = \begin{bmatrix}\overline{\Gamma}_1 & \overline{\Gamma}_2 & \dots & \overline{\Gamma}_N\end{bmatrix},$$
$$\Omega_N(\xi) = \begin{bmatrix}\varphi_1(\xi) & \varphi_2(\xi) & \cdots & \varphi_N(\xi)\end{bmatrix},$$

such that $N \leq M$. Using a weighted sum of these basis functions, equation (30) can be rewritten as

$$\begin{aligned} y(t) &= \sum_{i=1}^{N} \overline{\Gamma}_i \varphi_i(\xi(t)) + e(t) + \varepsilon(t) \\ &= \overline{\Gamma}\Omega_N(\xi(t))^\top + e(t) + \varepsilon(t), \end{aligned}$$

where $e(t) = \sum_{i=N+1}^{M} \overline{\Gamma}_i \varphi_i(\xi(t))$ is the error due to the termination of the summation at N. Finally, since ε is an exponentially decaying signal, consider the approximation

$$\tilde{y} = \sum_{i=1}^{N} \overline{\Gamma}_i \varphi_i(\xi) = \overline{\Gamma}\Omega_N(\xi)^\top, \tag{32}$$

which neglects the truncation error e and the transient error ε. Since ε is not known, $\overline{\Gamma}_k$ needs to be approximated. This approximation can be determined from data as follows.

Theorem 4. *Consider the system (21), the system (26), and the signal generator (22). Let the time snapshots* $U_k \in^{w\times N}$ *and* $\Xi_k \in^{w}$*, with* $w \geq N > 0$*, be defined as*

$$\tilde{U}_k = \begin{bmatrix}\Omega(\xi(t_{k-w+1})) \\ \vdots \\ \Omega(\xi(t_{k-1})) \\ \Omega(\xi(t_k))\end{bmatrix}, \quad \tilde{\Xi}_k = \begin{bmatrix} y(t_{k-w+1}) - \tilde{e}(t_{k-w+1}) \\ \vdots \\ y(t_{k-1}) - \tilde{e}(t_{k-1}) \\ y(t_k) - \tilde{e}(t_k)\end{bmatrix},$$

respectively, where $\tilde{e}(t_k) = \varepsilon(t_k) + e(t_k)$ *and* $k \geq w-1$*. If* $\tilde{U}_k$ *is full column rank, then*

$$\mathrm{vec}(\hat{\Gamma}_k) = (\tilde{U}_k^\top \tilde{U}_k)^{-1}\tilde{U}_k^\top \tilde{\Xi}_k, \tag{33}$$

is an estimate of $\overline{\Gamma}_k$*.*

$\overline{\Gamma}_k$ is the estimate of the matrix $\overline{\Gamma}$ at T_k^w, which is a moving window of sample times $T_k^w := \{t_{k-w+1}, \dots, t_{k-1}, t_k\}$, such that $0 \leq t_0 < \dots < t_{k-w} < \dots < t_k$, and $k > w-1 \geq 0$. For the approximation $\overline{\Gamma}_k$, defined in (33), to be well-defined for all k, the set of sample times T_k^w needs to be selected in a way that the matrix $\tilde{U}_k^\top \tilde{U}_k$ is full column rank. This condition on T_k^w highlights a property of persistence of excitation of the signal generator (22), which is guaranteed by the following assumption, see [128].

Assumption 2. *The initial condition* $\omega(0)$ *of the signal generator (22) is almost periodic and all the solutions of the system are analytic. Furthermore, the signal generator (22) satisfies the excitation rank condition [128] at* $\omega(0)$*.*

We are now ready to state a result that allows relating the output map of the reduced-order model to the steady-state output response of the underlying system.

Theorem 5. *Consider the system (21), the system (26), and the signal generator (22). There exist sequences* $\{t_k\}$ *such that*

$$\lim_{k\to\infty}\left(y(t_k) - \lim_{N\to M}\overline{\Gamma}_k\Omega_N(\xi(t_k))^\top\right) = 0. \tag{34}$$

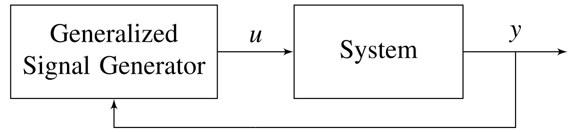

Fig. 3: Diagrammatic illustration of the cascade interconnection between the (unknown) signal generator (22) and the (unknown) nonlinear system (21) in [111].

In essence, through this result, the model reduction by moment matching problem is transformed into the estimation of the output mapping for a partially constructed reduced-order model (26). Indeed, the choice of basis functions may affect the speed of convergence of this estimation. We refer the interested reader to [129, Eq. 2] for more on the art of choosing basis functions. We conclude this part of the tutorial, by noting that this procedure is applicable to a very general class of systems, including systems with delay, and is robust to variations in the signal generator.

2) Stability Agnostic Moment Matching [111]: We have observed that moments are closely linked to the steady-state response of the nonlinear system driven by a signal generator. We highlight "if any" because a steady-state output (if any exists) response is ensured only when the origin is a locally exponentially stable equilibrium. However, this condition might be too restrictive for general nonlinear systems that exhibit complex behaviors. For these scenarios, the method of closed-loop interpolation has been recently introduced in [110], [111]. The point of departure of this closed-loop concept is the notion of *generalized signal generator*, which provides an extension of the signal generator (22). Specifically, it is described by the equations

$$\dot{\omega} = s(\omega), \tag{35a}$$
$$\dot{z} = \eta\left(z, \omega, y\right), \tag{35b}$$
$$u = \theta(z) + \ell(\omega), \tag{35c}$$

with $z(t) \in \mathbb{R}^r$, η and θ of appropriate dimension and such that $\eta(0,0,0) = 0$ and $\theta(0) = 0$. The state ω and the mappings s and ℓ are related to the signal generator (22). The generalized signal generator (35) must be (locally) observable, the subsystem (35a) must be neutrally stable, and the mappings $\eta(\cdot,\cdot,\cdot)$ and $\theta(\cdot)$ should be such that the equilibrium $(z,x) = (0,0)$ is locally exponentially stable. Moreover, the mapping $\eta(\cdot,\cdot,\cdot)$ should be such that

$$\eta\left(0, \omega, h(\pi(\omega))\right) = 0, \ \forall \omega.$$

With this generalized signal generator at hand, we need to redefine the associated notion of moment.

Definition 6. *The moment of system* (21) *at* (s,η,θ,ℓ) *is defined as* $h(\pi(\cdot))$*, where* $\pi(\cdot)$ *is the unique solution of the partial differential equation* (24)*.*

It is readily seen that the moment of (21) at (s,η,θ,ℓ) is equivalent to the moment of (21) at (s,ℓ). However, employing the generalized signal generator the internal stability condition of (21) is replaced by the local exponential stability of $(z,x)=(0,0)$. Under these premises, the following results can be proved.

Lemma 3. *Consider the system* (21) *and the generalized signal generator* (35)*. Then, for every* $(x(0),z(0),\omega(0))$ *in some neighborhood of* $(0,0,0)$*, the steady-state response of the (closed-loop) interconnected system*

$$\dot{\omega} = s(\omega), \tag{36a}$$
$$\dot{z} = \eta\,(z,\omega,h(x))\,, \tag{36b}$$
$$\dot{x} = f\big(x,\theta(z)+\ell(\omega)\big), \tag{36c}$$
$$y = h(x) \tag{36d}$$

exists, is unique, and it is such that

$$\lim_{t\to\infty}(z(t),x(t)-\pi(\omega(t))) = (0,0).$$

Theorem 6. *Consider the system* (21) *and the generalized signal generator* (35)*. Then, the steady-state output response of the (closed-loop) interconnected system* (36) *has a one-to-one relation with the moment at* (s,η,θ,ℓ) *of* (21)*.*

While the standard moment matching framework – in the time domain – is limited by the strong stability requirements of the system to be reduced, the closed-loop interpolation scheme overcomes this limitation by designing a signal generator that guarantees internal stability. This allows for the relaxation of stability conditions, even when the system's state is not fully known or measured. We can construct a family of nonlinear reduced-order models that achieves moment matching (21). Specifically, consider the family of nonlinear systems described by the equations

$$\dot{\xi}_a = \overline{f}_1\big(\xi_a,\xi_b,u\big), \tag{37a}$$
$$\dot{\xi}_b = \overline{f}_2\big(\xi_a,\xi_b,u\big), \tag{37b}$$
$$\overline{y} = \overline{h}\big(\xi_a,\xi_b\big), \tag{37c}$$

with $\xi_a(t)\in\mathbb{R}^{\nu}$ and $\xi_b(t)\in\mathbb{R}^{\rho-\nu}$, $u(t)\in\mathbb{R}$, $\overline{y}(t)\in\mathbb{R}$, and mappings $\overline{f}_1$, $\overline{f}_2$, and $\overline{h}$ of appropriate dimensions such that $\overline{f}_1(0,0,0)=0$, $\overline{f}_2(0,0,0)=0$, and $\overline{h}(0,0)=0$. Let $\overline{f}_1$ and $\overline{f}_2$ be such that the zero equilibrium of the (closed-loop) interconnected system (35)-(37) is locally exponentially stable, and such that

$$\overline{f}_1\big(\xi_a,0,\ell(\xi_a)\big) = s(\xi_a),$$
$$\overline{f}_2\big(\xi_a,0,\ell(\xi_a)\big) = 0,$$
$$\overline{h}(\xi_a,0) = h(\pi(\xi_a)).$$

Theorem 7. *Consider the system* (37) *and the generalized signal generator* (35)*. Let* $h(\pi(\cdot))$ *be the moment of the system* (21) *at* (s,η,θ,ℓ)*. Then* (37) *is a model of order* $\rho\geq\nu$ *which matches the moment of* (21) *at* (s,η,θ,ℓ)*.*

The parameterized model (37) boils down from the parameterization of all moment matching models discussed in [103], [106]. Hence, following [103], one can further prove that if (26) is a model of order $\rho\geq\nu$ achieving moment matching at (s,η,θ,ℓ), then the family of systems (37) defines a parameterization of (26), and thus it contains all the nonlinear moment matching interpolants with the same structure of the underlying system. Finally, it is worth mentioning that if $\omega(t)\equiv 0$ the generalized signal generator plays the role of a stabilizer of order r for any system of order ρ which is contained in the family of systems 37. Hence, the family of systems (37) defines also the set of all admissible plants stabilized by a given stabilizing controller. This leads to a moment-matching interpretation of the *dual Youla-Kučera parameterization* for nonlinear systems, which again provides strong connections to results in linear systems, namely via coprime-factorizations and associated reduction approaches.

V. Summary

A brief yet broad overview of the state of the model reduction problem for the purpose of control has been presented. The focus of this overview runs from classic balanced truncation methods through more recent moment-matching methods, including some discussions of error bounds, computability, and closed-loop reduction.

References

[1] B.Ho and R. Kalman, "Effective construction of linear state variable models from input/output functions," *Automatisierungstechnik*, no. 1/12, p. 545–548, 1966.

[2] H. Simon and A. Ando, "Aggregation of variables in dynamic systems," *Econometrica*, pp. 111–138, 1961.

[3] T. Zhou and D. Tao, "Double shrinking sparse dimension reduction," *IEEE Transactions on Image Processing*, vol. 22, no. 1, pp. 244–257, 2013.

[4] "Model reduction for circuit simulation," in *Lecture Notes in Electrical Engineering 74* (P. Benner, M. Hinze, and E. W. ter Maten, eds.), Springer Dordrecht, 2011.

[5] Y. Zhu and H. Lei, "Effective Mori-Zwanzig equation for the reduced order modeling of stochastic systems," *Journal of the American Institute of Mathematical Sciences*, pp. 959–982, 2022.

[6] C. Beck, S. Lall, T. Liang, and M. West, "Model reduction, optimal prediction, and the Mori-Zwanzig representation of markov chains," in *Proceedings of the IEEE Conference on Decision and Control*, 2009.

[7] B. Moore, "Principal component analysis in linear systems: Controllability, observability, and model reduction," *IEEE Transactions on Automatic Control*, vol. 26, pp. 17–32, Feb. 1981.

[8] A. C. Antoulas, *Approximation of Large-Scale Dynamical Systems*. Philadelphia: SIAM, 2005.

[9] K. Zhou, J. C. Doyle, and K. Glover, *Robust and Optimal Control*, vol. 40. New Jersey: Prentice Hall, 1996.

[10] M. Green and D. Limebeer, *Linear Robust Control*. Prentice Hall, 1995.

[11] J. M. A. Scherpen, "Balancing for nonlinear systems," *Sys. Cont. Lett.*, vol. 21, no. 2, pp. 143–153, 1993.

[12] K. Fujimoto and J. M. A. Scherpen, "Nonlinear input-normal realizations based on the differential eigenstructure of Hankel operators," *IEEE Transactions on Automatic Control*, vol. 50, no. 1, pp. 2–18, 2005.

[13] Y. Kawano and J. M. A. Scherpen, "Model reduction by differential balancing based on nonlinear Hankel operators," *IEEE Transactions on Automatic Control*, vol. 62, no. 7, pp. 3293–3308, 2017.

[14] B. Besselink, N. van de Wouw, J. M. A. Scherpen, and H. Nijmeijer, "Model reduction for nonlinear systems by incremental balanced truncation," *IEEE Transactions on Automatic Control*, vol. 59, no. 10, pp. 2739 – 2753, 2014.

[15] A. Astolfi, "Model reduction by moment matching for linear and nonlinear systems," *IEEE Transactions on Automatic Control*, vol. 55, no. 10, pp. 2321–2336, 2010.

[16] T. C. Ionescu and A. Astolfi, "Nonlinear moment matching-based model order reduction," *IEEE Transactions on Automatic Control*, vol. 61, no. 10, pp. 2837–2847, 2016.

[17] E. Qian, B. Kramer, B. Peherstorfer, and K. Willcox, "Lift & learn: Physics-informed machine learning for large-scale nonlinear dynamical systems," *Physica D: Nonlinear Phenomena*, 2020.

[18] C. Mullis and R. Roberts, "Synthesis of minimum roundoff noise fixed point digital filters," *IEEE Transactions on Circuits and Systems*, vol. 23, pp. 551–562, Sept. 1976.

[19] E. Verriest and T. Kailath, "On generalized balanced realizations," *IEEE Transactions on Automatic Control*, vol. 28, pp. 833–844, Aug. 1983.

[20] L. Pernebo and L. Silverman, "Model reduction via balanced state space representations," *IEEE Transactions on Automatic Control*, vol. 27, pp. 382–387, Apr. 1982.

[21] U. Al-Saggaf and G. Franklin, "An error bound for a discrete reduced order model of a linear multivariable system," *IEEE Transactions on Automatic Control*, vol. 32, pp. 815–819, Sept. 1987.

[22] K. Glover, "All optimal Hankel-norm approximations of linear multivariable systems and their L^∞-error bounds," *International Journal of Control*, vol. 39, pp. 1115–1193, June 1984.

[23] D. F. Enns, "Model reduction with balanced realizations: An error bound and a frequency weighted generalization," in *The 23rd IEEE Conference on Decision and Control*, pp. 127–132, Dec. 1984.

[24] D. Hinrichsen and A. Pritchard, "An improved error estimate for reduced-order models of discrete-time systems," *IEEE Transactions on Automatic Control*, vol. 35, pp. 317–320, Mar. 1990.

[25] G. Wood, P. J. Goddard, and K. Glover, "Approximation of linear parameter-varying systems," in *Proceedings, IEEE Conference on Decision and Control*, pp. 406–411, 1996.

[26] C. Beck, J. Doyle, and K. Glover, "Model reduction of multidimensional and uncertain systems," *IEEE Transactions on Automatic Control*, vol. 41, pp. 1466–1477, Oct. 1996.

[27] S. Lall and C. Beck, "Error-bounds for balanced model-reduction of linear time-varying systems," *IEEE Transactions on Automatic Control*, vol. 48, pp. 946–956, June 2003.

[28] H. Sandberg and A. Rantzer, "Balanced truncation of linear time-varying systems," *IEEE Transactions on Automatic Control*, vol. 49, pp. 217–229, Feb. 2004.

[29] M. Farhood and G. Dullerud, "Model reduction of nonstationary lpv systems," *IEEE Transactions on Automatic Control*, pp. 181–196, 2007.

[30] G. Kotsalis, A. Megretski, and M. Dahleh, "Balanced truncation for a class of stochastic jump linear systems and model reduction for hidden markov models," *IEEE Transactions on Automatic Control*, pp. 2543–2557, 2008.

[31] H. Sandberg, "An extension to balanced truncation with application to structured model reduction," *IEEE Transactions on Automatic Control*, vol. 55, pp. 1038–1043, Apr. 2010.

[32] D. Jaoude and M. Farhood, "Model reduction of distributed nonstationary lpv systems," *European Journal of Control*, pp. 27–39, 2018.

[33] K. Grigoriadis, "Optimal H_∞ model reduction via linear matrix inequalities: continuous and discrete-time cases," *Systems and Control Letters*, pp. 321–333, 1997.

[34] A. Antoulas and D. Sorensen, "Approximation of large-scale dynamical systems: an overview," *Int. J. Appl. Math. Comput. Sci.*, pp. 1093–1121, 2001.

[35] U. Desai and D. Pal, "A transformation approach to stochastic model reduction," *IEEE Transactions on Automatic Control*, vol. 29, pp. 1097–1100, Dec. 1984.

[36] A. Laub, M. T. Heath, C. C. Paige, and R. Ward, "Computation of system balancing transformations and other applications of simultaneous diagonalization algorithms," *IEEE Transactions on Automatic Control*, pp. 115–121.

[37] M. Safonov and R. Chiang, "A Schur method for balanced model reduction," in *Proceedings of the American Control Conference*, pp. 1036–1040, 1988.

[38] A. Varga, "Balancingfree square-root algorithm for computing singular perturbation approximations," in *Proceedings, IEEE Conference on Decision and Control*, 1991.

[39] D. Meyer, "Fractional balanced reduction: model reduction via fractional representation," *IEEE Transactions on Automatic Control*, pp. 1341–1345, 1990.

[40] L. Li and F. Paganini, "Structured coprime factor model reduction based on LMIs," *Automatica*, vol. 41, pp. 145–151, Jan. 2005.

[41] C. Beck, "Coprime factors reduction methods for linear parameter varying and uncertain systems," *Systems & Control Letters*, vol. 55, pp. 199–213, Mar. 2006.

[42] M. Farhood and C. L. Beck, "On the balanced truncation and coprime factors reduction of markovian jump linear systems," *Systems and Control Letters*, pp. 96–106, 2014.

[43] P. V. Overschee and B. D. Moor, "N4SID*: Subspace algorithms for the identification of combined deterministic-stochastic systems," *Automatica*, pp. 75–93, 1994.

[44] M. Verhaegen, "A novel non-iterative MIMO state space model identification technique," in *Proceedings, IFAC Identification and System Parameter Estimation Conference*, pp. 749–754, 1991.

[45] W. E. Larimore, "System identification, reduced-order filtering and modeling via canonical variate analysis," in *Proceedings of the American Control Conference*, pp. 445–451, 1983.

[46] I. Ziemann, A. Tsiamis, B. Lee, Y. Jedra, N. Matni, and G. Pappas, "A tutorial on the non-asymptotic theory of system identification," in *Proceedings of the IEEE Conference on Decision and Control*, 2023.

[47] S. Oymak and N. Ozay, "Revisiting Ho-Kalman based system identification: Robustness and finite sample analysis," *IEEE Transactions on Automatic Control*, pp. 1914–1928, 2022.

[48] T. Sarkar, A. Rakhlin, and M. Dahleh, "Finite time LTI system identification," *Journal of Machine Learning Research*, pp. 1–61, 2021.

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

[50] S. Chanturantabut and D. Sorenson, "Nonlinear model reduction via discrete empirical interpolation," *SIAM Journal on Scientific Computing*, pp. 2737–2764, 2010.

[51] C. D. Villemagne and R. Skelton, "Model reductions using a projection formulation," *Int. J. Control*, pp. 2141–2169, 1987.

[52] A. Yousuff, D. Wagie, and R. Skelton, "Linear system approximation via covariance equivalent realizations," *J. of Math. Analy. and Appl.*, pp. 91–115, 1985.

[53] C. Beattie and S. Gucercin, "Model reduction by rational interpretation," *Model Reduction and Approximation*, pp. 297–334, 2017.

[54] K. Gallivan, A. Vandendorpe, and P. Van Dooren, "Sylvester equations and projection-based model reduction," *Journal of Computational and Applied Mathematics*, vol. 162, no. 1, pp. 213–229, 2004.

[55] K. Gallivan, A. Vandendorpe, and P. Van Dooren, "Model reduction and the solution of Sylvester equations," *MTNS, Kyoto*, vol. 50, 2006.

[56] G. Scarciotti and A. Astolfi, "Interconnection-based model order reduction-a survey," *European Journal of Control*, p. 100929, 2023.

[57] G. E. Dullerud and F. Paganini, *A Course in Robust Control Theory*, vol. 36 of *Texts in Applied Mathematics*. New York, NY: Springer, 2000.

[58] G. Obinata and B. D. O. Anderson, *Model Reduction for Control System Design*. Communications and Control Engineering, London: Springer, 2001.

[59] Y. Liu and B. D. O. Anderson, "Singular perturbation approximation of balanced systems," *International Journal of Control*, vol. 50, pp. 1379–1405, Oct. 1989.

[60] S. Shokoohi, L. Silverman, and P. Van Dooren, "Linear time-variable systems: Balancing and model reduction," *IEEE Transactions on Automatic Control*, vol. 28, pp. 810–822, Aug. 1983.

[61] M. Farhood and G. E. Dullerud, "Model Reduction of Nonstationary LPV Systems," *IEEE Transactions on Automatic Control*, vol. 52, pp. 181–196, Feb. 2007. Conference Name: IEEE Transactions on Automatic Control.

[62] M. Green, "A relative error bound for balanced stochastic truncation," *IEEE Transactions on Automatic Control*, vol. 33, pp. 961–965, Oct. 1988.

[63] W. Wang and M. G. Safonov, "Relative-error bound for discrete balanced stochastic truncation," *International Journal of Control*, vol. 54, pp. 593–612, Sept. 1991.

[64] S. Shokoohi and L. M. Silverman, "Identification and model reduction of time-varying discrete-time systems," *Automatica*, vol. 23, pp. 509–521, July 1987.

[65] D. Kavranoğlu and M. Bettayeb, "Characterization of the solution to the optimal H_∞ model reduction problem," *Systems & Control Letters*, vol. 20, pp. 99–107, Feb. 1993.

[66] H. Sandberg and R. M. Murray, "Model reduction of interconnected linear systems," *Optimal Control Applications and Methods*, vol. 30, no. 3, pp. 225–245, 2009.

[67] L. Andersson, A. Rantzer, and C. Beck, "Model comparison and simplification," *International Journal of Robust and Nonlinear Control*, vol. 9, no. 3, pp. 157–181, 1999.

[68] J. He, I. Ziemann, C. Rojas, and H. Hjarmalsson, "Finite sample analysis for a class of subspace identification methods," *arXiv*, 2024.

[69] J. Hahn and T. F. Edgar, "An improved method for nonlinear model reduction using balancing of empirical Gramians," *Computers & Chemical Engineering*, vol. 26, no. 10, pp. 1379–1397, 2002.

[70] M. Condon and R. Ivanov, "Empirical balanced truncation of nonlinear systems," *Journal of Nonlinear Science*, vol. 14, no. 5, pp. 405–414, 2004.

[71] C. Himpe, "emgr—The empirical Gramian framework," *Algorithms*, vol. 11, no. 7, p. 91, 2018.

[72] J. Hahn and T. F. Edgar, "Balancing approach to minimal realization and model reduction of stable nonlinear systems," *Industrial & Engineering Chemistry Research*, vol. 41, no. 9, pp. 2204–2212, 2002.

[73] K. Willcox and J. Peraire, "Balanced model reduction via the proper orthogonal decomposition," *AIAA Journal*, vol. 40, no. 11, pp. 2323–2330, 2002.

[74] J. Hahn, U. Kruger, and T. F. Edgar, "Application of model reduction for model predictive control," *IFAC Proceedings Volumes*, vol. 35, no. 1, pp. 393–398, 2002.

[75] R. B. Choroszucha, J. Sun, and K. Butts, "Nonlinear model order reduction for predictive control of the diesel engine airpath," *Proc. 2016 American Control Conference*, pp. 5081–5086, 2016.

[76] K. Kashima, "Noise response data reveal novel controllability Gramian for nonlinear network dynamics," *Scientific Reports*, vol. 6, no. 27300, 2016.

[77] A. J. Krener and K. Ide, "Measures of unobservability," *Proc. 48th IEEE Conference on Decision and Control and the 28th Chinese Control Conference*, pp. 6401–6406, 2009.

[78] N. D. Powel and K. A. Morgansen, "Empirical observability Gramian rank condition for weak observability of nonlinear systems with control," *Proc. 54th IEEE Conference on Decision and Control*, pp. 6342–6348, 2015.

[79] Y. Kawano and J. M. Scherpen, "Empirical differential Gramians for nonlinear model reduction," *Automatica*, vol. 127, p. 109534, 2021.

[80] Y. Kawano, "Controller reduction for nonlinear systems by generalized differential balancing," *IEEE Transactions on Automatic Control*, vol. 67, no. 11, pp. 5856–5871, 2022.

[81] Y. Kawano and J. M. A. Scherpen, "Balanced model reduction for linear time-varying symmetric systems," *IEEE Transactions on Automatic Control*, vol. 64, no. 7, pp. 3060–3067, 2019.

[82] E. I. Verriest and T. Kailath, "On generalized balanced realizations," *IEEE Transactions on Automatic Control*, vol. 28, no. 9, pp. 833–844, 1983.

[83] E. I. Verriest and W. S. Gray, "Flow balancing nonlinear systems," *Proc. 14th International Symposium on Mathematical Theory of Networks and Systems*, 2000.

[84] E. I. Verriest and W. S. Gray, "Nonlinear balanced realizations," *Proc. 43rd IEEE Conference on Decision and Control*, pp. 1164–1169, 2004.

[85] E. I. Verriest, "Time variant balancing and nonlinear balanced realizations," *Model Order Reduction: Theory, Research Aspects and Applications*, pp. 213–250, 2008.

[86] P. Holmes, J. L. Lumley, G. Berkooz, and C. W. Rowley, *Turbulence, Coherent Structures, Dynamical Systems and Symmetry*. Cambridge: Cambridge University Press, 2012.

[87] W. Lohmiller and J.-J. E. Slotine, "On contraction analysis for nonlinear systems," *Automatica*, vol. 34, no. 6, pp. 683–696, 1998.

[88] F. Forni and R. Sepulchre, "A differential Lyapunov framework for contraction analysis," *IEEE Transactions on Automatic Control*, vol. 59, no. 3, pp. 614–628, 2014.

[89] A. Sarkar and J. M. A. Scherpen, "Extended differential balancing for nonlinear dynamical systems," *IEEE Control Systems Letters*, vol. 6, pp. 3170–3175, 2022.

[90] A. Sootla and A. Rantzer, "Scalable positivity preserving model reduction using linear energy functions," pp. 4285–4290, 2012.

[91] A. Sarkar, Y. Kawano, and J. M. Scherpen, "Model reduction of cooperative systems using separable energy functions," *Proc. 7th IFAC Conference on Analysis and Control of Nonlinear Dynamics and Chaos*, 2024. (to appear).

[92] Y. Kawano, B. Besselink, J. M. A. Scherpen, and M. Cao, "Data-driven model reduction of monotone systems by nonlinear DC gains," *IEEE Transactions on Automatic Control*, vol. 65, no. 5, pp. 2094–2167, 2020.

[93] P. Benner, M. Ohlberger, A. Cohen, and K. Willcox, *Model Reduction and Approximation: Theory and Algorithms*. SIAM, 2017.

[94] A. Astolfi, "A new look at model reduction by moment matching for linear systems," in *2007 46th IEEE Conference on Decision and Control*, pp. 4361–4366, IEEE, 2007.

[95] A. Antoulas, *Approximation of large-scale dynamical systems*. SIAM, 2005.

[96] A. Astolfi, "Model reduction by moment matching for linear and nonlinear systems," *IEEE Transactions on Automatic Control*, vol. 55, no. 10, pp. 2321–2336, 2010.

[97] A. Isidori, *Nonlinear Control Systems*. Springer London, 1995.

[98] T. C. Ionescu and A. Astolfi, "On moment matching with preservation of passivity and stability," in *49th IEEE Conference on Decision and Control (CDC)*, pp. 6189–6194, IEEE, 2010.

[99] T. C. Ionescu and A. Astolfi, "Families of moment matching based, structure preserving approximations for linear port Hamiltonian systems," *Automatica*, vol. 49, no. 8, pp. 2424–2434, 2013.

[100] J. D. Simard, A. Moreschini, and A. Astolfi, "Moment matching for nonlinear systems of second-order equations," in *2023 62nd IEEE Conference on Decision and Control (CDC)*, pp. 4978–4983, IEEE, 2023.

[101] T. C. Ionescu and A. Astolfi, "Nonlinear moment matching-based model order reduction," *IEEE Transactions on Automatic Control*, vol. 61, no. 10, pp. 2837–2847, 2015.

[102] A. Astolfi, "A note on model reduction by moment matching for nonlinear systems," *IFAC Proceedings Volumes*, vol. 43, no. 14, pp. 1244–1248, 2010.

[103] J. D. Simard, A. Moreschini, and A. Astolfi, "Parameterization of all moment matching interpolants," in *2023 European Control Conference (ECC)*, pp. 1–6, 2023.

[104] G. Scarciotti, "Model reduction by moment matching for linear singular systems," in *54th IEEE Conference on Decision and Control (CDC)*, pp. 7310–7315, IEEE, 2015.

[105] G. Scarciotti, "Steady-state matching and model reduction for systems of differential–algebraic equations," *IEEE Transactions on Automatic Control*, vol. 62, no. 10, pp. 5372–5379, 2017.

[106] J. Simard, A. Moreschini, and A. Astolfi, "Parameterization of all differential-algebraic moment matching interpolants," *IEEE Transactions on Automatic Control*, 2023. Submitted.

[107] G. Scarciotti and A. Astolfi, "Model reduction of neutral linear and nonlinear time-invariant time-delay systems with discrete and distributed delays," *IEEE Transactions on Automatic Control*, vol. 61, no. 6, pp. 1438–1451, 2015.

[108] G. Scarciotti and A. R. Teel, "On moment matching for stochastic systems," *IEEE Transactions on Automatic Control*, vol. 67, no. 2, pp. 541–556, 2021.

[109] M. F. Shakib, G. Scarciotti, A. Y. Pogromsky, A. Pavlov, and N. van de Wouw, "Model reduction by moment matching with preservation of global stability for a class of nonlinear models," *Automatica*, vol. 157, p. 111227, 2023.

[110] A. Moreschini and A. Astolfi, "Closing the loop in moment matching," in *IEEE Conference on Decision and Control*, 2024. Accepted.

[111] A. Moreschini and A. Astolfi, "Closed-Loop Interpolation by Moment Matching for Linear and Nonlinear systems," *IEEE Transaction on Automatic Control*. Submitted.

[112] G. Scarciotti and A. Astolfi, "Model reduction for hybrid systems with state-dependent jumps," *IFAC-PapersOnLine*, vol. 49, no. 18, pp. 850–855, 2016.

[113] S. Galeani and M. Sassano, "Model reduction by moment matching at discontinuous signals via hybrid output regulation," in *2015 European Control Conference (ECC)*, pp. 1189–1194, IEEE, 2015.

[114] G. Scarciotti and A. Astolfi, "Model reduction by matching the steady-state response of explicit signal generators," *IEEE Transactions on Automatic Control*, vol. 61, no. 7, pp. 1995–2000, 2015.

[115] D. Bhattacharjee and A. Astolfi, "Closed-loop model reduction by moment matching for linear systems," in *2023 62nd IEEE Conference on Decision and Control (CDC)*, pp. 4954–4959, IEEE, 2023.

[116] G. Scarciotti and A. Astolfi, "Data-driven model reduction by moment matching for linear and nonlinear systems," *Automatica*, vol. 79, pp. 340–351, 2017.

[117] A. Moreschini, M. Scandella, and T. Parisini, "Nonlinear data-driven moment matching in Reproducing Kernel Hilbert Spaces," in *2024 European Control Conference (ECC)*, pp. 3440–3445, 2024.

[118] D. Bhattacharjee and A. Astolfi, "Data-driven model reduction by moment matching for linear systems driven by an unknown implicit signal generator," in *2024 European Control Conference (ECC)*, pp. 3434–3439, 2024.

[119] D. Bhattacharjee, A. Moreschini, and A. Astolfi, "Signal generator agnostic moment matching," Submitted.

[120] J. D. Simard and A. Astolfi, "Nonlinear model reduction in the Loewner framework," *IEEE Transactions on Automatic Control*, vol. 66, no. 12, pp. 5711–5726, 2021.

[121] J. D. Simard and A. Moreschini, "Enforcing stability of linear interpolants in the loewner framework," *IEEE Control Systems Letters*, vol. 7, pp. 3537–3542, 2023.

[122] A. Moreschini, J. D. Simard, and A. Astolfi, "Data-driven model reduction for port-Hamiltonian and network systems in the Loewner framework," *Automatica*, vol. 169, p. 111836, 2024.

[123] A. Moreschini, J. D. Simard, and A. Astolfi, "Model reduction for linear port-Hamiltonian systems in the Loewner framework," *IFAC-PapersOnLine*, vol. 56, no. 2, pp. 9493–9498, 2023.

[124] A. Moreschini, J. D. Simard, and A. Astolfi, "Model reduction in the Loewner framework for second-order network systems on graphs," in *62nd IEEE Conference on Decision and Control (CDC)*, pp. 6713–6718, IEEE, 2023.

[125] R. Tóth, *Modeling and Identification of Linear Parameter-Varying Systems*, vol. 403. Springer, 2010.

[126] J. Park and I. W. Sandberg, "Universal approximation using radial-basis-function networks," *Neural Computation*, vol. 3, no. 2, pp. 246–257, 1991.

[127] H. Rocha, "On the selection of the most adequate radial basis function," *Applied Mathematical Modelling*, vol. 33, no. 3, pp. 1573–1583, 2009.

[128] A. Padoan, G. Scarciotti, and A. Astolfi, "A geometric characterization of the persistence of excitation condition for the solutions of autonomous systems," *IEEE Transactions on Automatic Control*, vol. 62, no. 11, pp. 5666–5677, 2017.

[129] S. L. Brunton, J. L. Proctor, and J. N. Kutz, "Discovering governing equations from data by sparse identification of nonlinear dynamical systems," *Proceedings of the National Academy of Sciences*, vol. 113, no. 15, pp. 3932–3937, 2016.