

Review Article

Structure-preserving model reduction of interconnected dynamical systems[☆]

N. van de Wouw^{a,*,}, B. Besselink^b, M.F. Shakib^a, L.A.L. Janssen^a, L. Poort^a, G. Scarcitti^c,
R.H.B. Fey^a

^a Department of Mechanical Engineering, Eindhoven University of Technology, Eindhoven, The Netherlands

^b Bernoulli Institute for Mathematics, Computer Science and Artificial Intelligence, University of Groningen, Groningen, The Netherlands

^c Department of Electrical and Electronic Engineering, Imperial College London, London, UK

ARTICLE INFO

Keywords:

Model reduction
Interconnected dynamical systems
Lur'e-type systems
Accuracy guarantees
Stability
Passivity
Structure preservation

ABSTRACT

This paper presents model reduction techniques for two classes of interconnected dynamical systems: firstly, feedback interconnections between large-scale linear systems and static nonlinearities and, secondly, interconnections between (many) linear dynamical systems. For the first class we provide both balancing-based and moment matching approaches that are applicable to large-scale systems, by reducing the linear part only and leaving the nonlinearity intact. Hence, the original feedback interconnection structure is preserved as well. Moreover, we provide a reduction error bound that expresses reduction accuracy depending on the level of reduction of the linear part and the properties of the nonlinearities. In addition, these methods preserve (global and incremental) stability properties. For the second class of systems, we present an approach to link reduction accuracy specifications on the level of the sub-systems to related specifications on the level of the interconnected system. This allows for a modular approach that preserves the structure of the high-order, interconnected system. In turn, this promotes the interpretability of the reduced-order system. In addition, we introduce the concept of abstracted reduction for interconnected linear systems, which enables the modular reduction of the sub-systems, while taking into account the dynamics of the rest of the interconnected system in a computationally tractable way. Finally, these methods also provide an error bound and preserve stability and (optionally) passivity.

Contents

1. Introduction	2
1.1. Model reduction for nonlinear systems	3
1.2. Model reduction for linear interconnected systems	3
1.3. Main contributions and outline	3
1.4. Notation	4
2. Model reduction of feedback interconnections of linear and nonlinear systems	4
2.1. Background and motivation	4
2.2. Incremental system properties	6
2.2.1. Input-to-state convergence	6
2.2.2. Incrementally dissipative systems	6
2.2.3. Lur'e-type systems with incremental properties	7
2.3. Error analysis using incremental system properties	8
2.4. Model reduction by balanced truncation	9
2.4.1. Balanced model reduction for LTI systems	10
2.4.2. Dissipativity-preserving balanced model reduction	11
2.4.3. Balanced model reduction for Lur'e-type systems	12
2.5. Model reduction by moment matching	12

[☆] This work was supported by the project Digital Twin with project number P18-03 of the research programme Perspectief which is (mainly) financed by the Dutch Research Council (NWO) and by Holland High tech, top sector High-Tech Systems and Materials, with a PPP innovation subsidy public private partnerships for research and development.

* Corresponding author.

E-mail address: n.v.d.wouw@tue.nl (N. van de Wouw).

<https://doi.org/10.1016/j.arcontrol.2025.101016>

Received 27 January 2025; Received in revised form 28 August 2025; Accepted 29 August 2025

Available online 1 October 2025

1367-5788/© 2025 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

2.5.1.	Interconnection-based notion of a (nonlinear) moment.....	13
2.5.2.	Time domain moment matching for LTI systems.....	13
2.5.3.	Structure-preserving moment matching for Lur'e-type systems.....	14
2.5.4.	Error bound minimization using moment matching.....	15
2.6.	Numerical example.....	17
2.6.1.	System description.....	17
2.6.2.	Loop transformation.....	17
2.6.3.	Linear model reduction.....	18
2.6.4.	Time-domain simulations.....	18
2.7.	Summary.....	19
3.	Model reduction of interconnected, linear dynamical systems.....	19
3.1.	Background and motivation.....	19
3.2.	System class and problem setting.....	20
3.3.	Distributing accuracy requirements from system level to sub-system level.....	21
3.4.	Structure-preserving reduction using abstracted environments.....	23
3.5.	Preservation of stability and passivity.....	26
3.6.	Summary.....	27
4.	Conclusions and outlook.....	27
4.1.	Conclusions.....	27
4.2.	Extensions and outlook.....	27
	Declaration of competing interest.....	28
	Data availability.....	28
	References.....	28

1. Introduction

The modeling of dynamical systems has been a cornerstone in understanding complex systems, both in the natural and engineering domains. In engineering systems, models are an indispensable tool to support the design, optimization, diagnostics, and control of such systems. Several factors have exacerbated the complexity of such engineering systems. Firstly, the ever-increasing demands on functionality and performance drive up design complexity. An example is complex semiconductor equipment such as wafer scanners used for chip manufacturing. These high-tech systems require nanometer-level positioning accuracy and sub-milliKelvin thermal conditioning to meet the increasing demand for ever smaller and faster chips. Lithography equipment consists of hundreds of components and modules that need to function in harmony to realize such performance specifications. Secondly, the design process of such systems is highly multi-disciplinary and the resulting models are typically of a multi-physics, or even cyber-physical, nature. Thirdly, as the above example shows, (controlled) engineering systems are increasingly consisting of a multitude of interconnected dynamical systems. Other examples of such complex networked systems are power networks, traffic systems, robotic networks, water management systems, manufacturing systems, logistic networks, etc.

Therefore, model complexity reduction (also known as model order reduction) is an essential methodological tool in any model-based design workflow. Techniques for model order reduction have been developed for dynamical systems within different domains. Within the mechanical engineering domain, component mode synthesis techniques have been developed to reduce the complexity of interconnected structural dynamics models (Craig Jr., 2000; de Klerk, Rixen, & Voormeeren, 2008). Within the numerical mathematics community, Krylov-based reduction techniques (Bai, 2002; de Klerk, Rixen, & Voormeeren, 2003) and reduced basis methods (Chaturantabut & Sorensen, 2010; Hesthaven, Rozza, & Stamm, 2016) have been established. In addition, time-domain moment matching techniques (Astolfi, 2010) and balancing-based reduction techniques (Enns, 1984; Moore, 1981) have emerged from the systems and control community together with data-driven methods (Burohman, Besselink, Scherpen, & Camlibel, 2023, 2024; Nakatsukasa, Sète, & Trefethen, 2018; Scariotti & Astolfi, 2017b). In this paper, we focus on approaches based on balancing and moment matching techniques and we refer to Antoulas (2005a), Besselink et al. (2013), Gallivan, Grimme, and Van Dooren

(1999), Schilders, van der Vorst, and Rommes (2008) for (comparative) overviews of the different approaches that arose from different communities. These references focus on reduction methods for large-scale linear systems, while the current paper focuses on the problem of model reduction for interconnected linear systems and the class of (nonlinear) Lur'e-type systems.

As mentioned above, complex engineering systems typically consist of multiple (often many) sub-systems that are interconnected. This further increases the complexity (and order) of the resulting interconnected system dynamics and further increases the need for complexity reduction. In this paper, we will consider two classes of interconnected systems for which model complexity reduction approaches will be presented. Firstly, many engineering systems consist of large-scale linear time-invariant dynamics interconnected with (relatively low-order) nonlinear system elements. Examples are large-scale structural dynamics models interconnected to local nonlinearities associated with, e.g., friction, hysteresis, nonlinear stiffness or damping, nonlinear journal bearings, cutting processes, and nonlinear actuator dynamics. Such model structure arises in high-precision positioning systems for high-tech equipment, manufacturing (e.g., milling) processes (Faassen, van de Wouw, Oosterling, & Nijmeijer, 2003; Insperger & Stépán, 2000), drilling processes (Aarsnes & van de Wouw, 2019; Gernay, Denoël, & Detournay, 2009), civil engineering applications (Suy, Fey, Galanti, & Nijmeijer, 2007), and many more. Also, models for large-scale electric circuits with diodes may be cast in Lur'e-type form. The above applications give rise to a class of models consisting of the (feedback) interconnection of a large-scale linear system with a (relatively low-order, or static) nonlinear system. Secondly, we will consider a class of systems consisting of multiple (potentially many) interconnected linear time-invariant sub-systems. As said, such models may arise in diverse domains: for example, mechanical or thermal models of high-tech equipment consisting of many interconnected sub-systems (Egelmeers, Janssen, Fey, Gerritsen, & van de Wouw, 2024; Poort, Besselink, Fey, & van de Wouw, 2024), or complex electric circuit models (Benner, Hinze, & ter Maten, 2011) and power systems (Scariotti, 2016; Sturk, Vanfretti, Chompoobutrgool, & Sandberg, 2014).

Generic challenges that arise when pursuing model reduction for both classes of interconnected models are:

- How can the structure of the interconnected system model be preserved after reduction? In other words, how can we ensure that, after reduction, the same sub-systems and their interconnection can be readily recognized? This serves to retain the interpretability of the reduced-order model.

- A related challenge is how to develop methods that allow the reduction of sub-systems in a modular fashion, i.e., that support the reduction of the system at sub-system level. This promotes the feasibility of the reduction of systems consisting of many sub-systems and it allows us to pursue reduction of sub-systems in an (at least partially) independent fashion.
- How can the preservation of stability properties and/or passivity properties be guaranteed in the reduction process?
- How can we provide quantitative guarantees on the accuracy of the reduced-order model and can we design the trade-off between complexity and accuracy?
- How can scalability towards increasingly large-scale systems be promoted? In other words, how can we make the model reduction of increasingly large-scale interconnected systems (either in terms of the dimension of the sub-systems or the number of sub-systems) computationally feasible?

The remainder of this introduction includes a concise literature review on model reduction for nonlinear systems and linear interconnected systems, a summary of this paper's contributions, and a section on notation.

1.1. Model reduction for nonlinear systems

In the scope of nonlinear systems, results on model reduction have been obtained in the scope of balancing-based approaches (Scherpen, 1993), moment matching techniques (Astolfi, 2010), reduced-basis methods (Chaturantabut & Sorensen, 2010; Hesthaven et al., 2016), and data-based surrogate modeling approaches (Abbasi, Iapichino, Schilders, & van de Wouw, 2022; Berkooz, Holmes, & Lumley, 1993; Brunton, Proctor, & Kutz, 2016; Filanova, Pontes Duff, Goyal, & Benner, 2023). In this paper, we focus on the first two methodological categories.

For linear systems, balanced truncation (Moore, 1981) guarantees the preservation of stability (Pernebo & Silverman, 1982) and provides an error bound (Enns, 1984; Glover, 1984). Alternative balanced truncation methods, such as positive real balancing (Desai & Pal, 1984; Harshavardhana, Jonckheere, & Silverman, 1984) and bounded real balancing (Opdenacker & Jonckheere, 1988), can be used to preserve passivity (positive realness) and contractivity (bounded realness) of the system, respectively. For an overview, see Antoulas (2005b). For nonlinear systems, existing model reduction techniques typically do not satisfy the properties of preservation of stability and the existence of an error bound. Note that, contrary to the linear setting, in the nonlinear context both internal stability (i.e., stability of an equilibrium point for zero input) as well as input-output stability (i.e., a bounded gain) have to be considered separately. Balanced truncation for nonlinear systems as developed in Fujimoto and Scherpen (2010), Scherpen (1993) does preserve (local) stability of the equilibrium point, but it does not provide input-output stability guarantees nor a reduction error bound. Extensions based on incremental stability notions (Angeli, 2002) (including convergence (Pavlov, Pogromsky, van de Wouw, & Nijmeijer, 2004) and contraction (Lohmiller & Slotine, 1998)), which can provide forms of (input-output) stability and accuracy guarantees, can be found in Besselink, van de Wouw, Scherpen, and Nijmeijer (2013), Besselink, van de Wouw, Scherpen, and Nijmeijer (2014), Kawano (2022), Kawano and Scherpen (2015, 2017), Sarkar and Scherpen (2022). Nevertheless, all of these approaches are computationally challenging and are not applicable to truly large-scale systems.

Moment matching for linear time-invariant models (Astolfi, 2010) fits in the class of interpolation approaches (Gallivan, Vandendorpe, & Van Dooren, 2004). Moments of linear time-invariant models are defined as the coefficients of the Laurent series expansion of the transfer function at a complex interpolation point, see Antoulas (2005a). Moment matching aims to match (make equal) the moments of the reduced-order model to those of the full-order model. The time-domain

interpretation of moment matching as proposed in Astolfi (2010) has formed the basis for the extension of the method to nonlinear models, enabling the construction of reduced-order nonlinear models that achieve moment matching (Astolfi, 2010; Scarcioiti & Astolfi, 2017a, 2017c). Herein, the definition of moments relies on the center manifold theorem and is, consequently, defined only locally in the neighborhood of the origin. Hence, the formulated reduced-order model is by definition only an approximation of the full-order model in the neighborhood of the origin. In addition, these results do not provide an error bound and do not preserve the model structure of the full-order model.

1.2. Model reduction for linear interconnected systems

Important challenges also arise in the complexity reduction of linear systems consisting of multiple sub-systems. While the reduction of the interconnected system as a whole can be expected to be accurate, two important challenges may arise in such an approach. Firstly, it is typically computationally infeasible to reduce the entire system for truly large-scale systems and, secondly, such reduction destroys the interconnection structure (Lutowska, 2012), thus compromising interpretability. These challenges gave rise to an increasing interest in the development of alternative reduction techniques for such systems (Reis & Stykel, 2008). Here, one can think of a modular approach (Vaz & Davison, 1990) in which the sub-system models are reduced independently while retaining the interconnection structure. However, in general, this approach does not lead to accurate reduced-order models if the dynamics of the total interconnected system is not taken into account in such independent reduction of the sub-systems. With these problems in mind, so-called structure-preserving methods have been developed for interconnected systems (Lutowska, 2012; Monshizadeh, Trentelman, & Camlibel, 2013; Sandberg & Murray, 2009; Schilders & Lutowska, 2014; Vandendorpe & Van Dooren, 2008). Furthermore, when a system model is reduced modularly, i.e., individually at a sub-system level, it is challenging to quantify how the stability and accuracy of the reduced-order interconnected system model are affected by a loss of accuracy induced by the reduction of an individual sub-system model. Some a priori error bounds are available in the literature (Ishizaki, Kashima, Imura, & Aihara, 2014; Reis & Stykel, 2008). However, these are typically conservative and, as such, less equipped for application to real-life engineering applications. In addition, it is in general highly challenging to translate requirements on accuracy and stability of the interconnected reduced-order system model to requirements on the reduced-order sub-system models. Finally, modular, structure-preserving model reduction methods are based on, in one form or another, taking into account the dynamics of the 'environment' (the rest of the interconnected system) of the sub-system under reduction. Given the large-scale nature of the system as a whole, this again raises concerns about the computational feasibility of such approaches in truly large-scale applications.

1.3. Main contributions and outline

In this paper, we will give an overview of recent developments on model order reduction techniques for both nonlinear systems consisting of a feedback interconnection between linear time-invariant dynamics and multi-variable, static nonlinearities (Lur'e-type systems) and systems consisting of interconnected linear time-invariant dynamical systems. In particular, we will give a unified presentation of these model reduction problems, the most recent results, and opportunities for future research.

In the scope of model reduction approaches for Lur'e-type systems, see Section 2, we warrant scalability to systems of large order by taking an approach in which only the linear dynamics are reduced while retaining the exact nonlinearities in the reduced-order nonlinear system. We present reduction approaches, based on balancing and moment matching, with the following key properties:

- We preserve the feedback interconnection structure between the linear dynamics and the nonlinearities in the reduced-order model. This provides interpretability of the involved feedback signals.
- We provide conditions under which (incremental) stability properties can be preserved for the reduced-order model. This provides global guarantees for entire classes of time-varying inputs.
- We formulate computable, quantitative bounds for the reduction error. For the proposed moment matching approach, we, additionally, present an approach to minimize the reduction error.

In the scope of model reduction for linear interconnected systems, see Section 3, we provide the following main contributions:

- A computational framework for translating accuracy specifications for the reduced-order sub-systems to those for the reduced-order interconnected system, and vice versa. This framework supports a truly modular reduction approach, where reduction can be performed independently for the sub-systems while still guaranteeing both stability and accuracy for the reduced-order interconnected system.
- A so-called abstracted reduction framework, which makes modular, structure-preserving model reduction of interconnected systems, based on balancing, computationally feasible, while still accounting for the dynamics of the interconnected system as a whole. This is achieved by reducing sub-systems while taking an abstracted (i.e., reduced-complexity) representation of their environment (the rest of the system) into account. This approach warrants both modularity and scalability of structure-preserving reduction approaches to large-scale systems. In addition, a variant of this approach is presented, which preserves not only stability, but also passivity properties of the original dynamics.

Section 4 presents concluding remarks and an outlook towards future research.

1.4. Notation

The symbols $\mathbb{R}$, $\mathbb{R}_+$, $\mathbb{C}$, $\mathbb{C}^0$, $\mathbb{C}^-$, and $\mathbb{N}$ denote the set of real numbers, positive real numbers, complex numbers, complex numbers with zero real parts, complex numbers with negative real parts, and natural numbers (excluding 0), respectively. We denote by e_i the canonical vector with appropriate dimension and all its elements zero except the i th element, which is 1. The symbols I_n and 0_n denote the $n \times n$ identity and zero matrices, respectively, and $0_{n \times v}$ denotes the zero matrix of dimensions $n \times v$. For $A \in \mathbb{R}^{n \times n}$, $\lambda(A)$ denotes its spectrum and $A^\top$ denotes its transpose, while for $A \in \mathbb{C}^{n \times m}$, A^* denotes its Hermitian transpose. For a symmetric matrix $A \in \mathbb{R}^{n \times n}$, $A > 0$ ($A < 0$) denotes that A is positive (negative) definite, and $A \geq 0$ ($A \leq 0$) denotes that A is positive (negative) semi-definite. Furthermore, $A = \text{diag}(A_1, A_2)$ denotes a block-diagonal matrix of submatrices A_1 and A_2 . In symmetric block matrices, the element $\star$ is induced by transposition. For matrices $\mathcal{X} \in \mathbb{R}^{n \times n}$, $A \in \mathbb{R}^{n \times n}$, $B \in \mathbb{R}^{n \times m}$, $C \in \mathbb{R}^{p \times n}$, and the constant $\gamma \geq 0$, we use the short-hand notations

$$\text{He}(A) := A + A^\top, \quad (1a)$$

$$\mathcal{N}_\gamma(A, B, C, \mathcal{X}) := \begin{bmatrix} \text{He}(\mathcal{X}A) & \mathcal{X}B & C^\top \\ \star & -\gamma I_m & 0_{m \times p} \\ \star & \star & -\gamma I_p \end{bmatrix}. \quad (1b)$$

A continuous function $\alpha : [0, a) \rightarrow [0, +\infty)$ is said to belong to class $\mathcal{K}$ if it is strictly increasing and $\alpha(0) = 0$. It is said to belong to class $\mathcal{K}_\infty$ if $a = +\infty$ and $\alpha(r) \rightarrow +\infty$ as $r \rightarrow +\infty$. A continuous function $\beta : [0, a) \times [0, +\infty) \rightarrow [0, +\infty)$ is said to belong to class $\mathcal{KL}$ if, for each fixed s , the mapping $\beta(\cdot, s)$ belongs to class $\mathcal{K}$ and, for each fixed r , the mapping $\beta(r, \cdot)$ is decreasing and $\beta(r, s) \rightarrow 0$ as $s \rightarrow \infty$. Given two mappings, $f : Y \rightarrow Z$ and $g : X \rightarrow Y$, with $f \circ g : X \rightarrow Z$ we denote the composite function that maps all $x \in X$ to $f(g(x)) \in Z$ as $(f \circ g)(x) = f(g(x))$.

For a vector x , its Euclidean norm is denoted by $|x|$. The symbol $\Re(s)$ indicates the real part of the complex number s . For a transfer-function matrix Φ defined on $\mathbb{C}$, with all its poles in $\mathbb{C}^-$, its $\mathcal{H}_\infty$ -norm is defined by $\|\Phi\|_{\mathcal{H}_\infty} := \sup_{\omega \in [0, \infty)} \bar{\rho}(\Phi(i\omega))$ with $\bar{\rho}$ the maximum singular value of Φ and $i := \sqrt{-1}$. For some interval $\mathcal{T} \subset \mathbb{R}$, we denote by $\mathcal{L}_2^n(\mathcal{T})$ the class of functions that are square integrable on $\mathcal{T}$, i.e.,

$$\mathcal{L}_2^n(\mathcal{T}) = \left\{ x : \mathcal{T} \rightarrow \mathbb{R}^n \mid \int_{\mathcal{T}} |x(t)|^2 dt < \infty \right\}. \quad (2)$$

Next, for $T > 0$, the set $\mathcal{P}_2^n(T)$ denotes the class of functions that are periodic with period T and whose restriction to $[0, T]$ is in $\mathcal{L}_2^n([0, T])$. Explicitly

$$\mathcal{P}_2^n(T) = \left\{ x : \mathbb{R} \rightarrow \mathbb{R}^n \mid \begin{array}{l} x(t) = x(t+T) \\ \text{and } x|_{[0, T]} \in \mathcal{L}_2^n([0, T]) \end{array} \right\}. \quad (3)$$

The norms on these signal spaces are denoted by $\|\cdot\|_{\mathcal{L}_2}$ and $\|\cdot\|_{\mathcal{P}_2}$, respectively, and are defined as

$$\|x\|_{\mathcal{L}_2}^2 = \int_{\mathcal{T}} |x(t)|^2 dt, \quad \|x\|_{\mathcal{P}_2}^2 = \frac{1}{T} \int_0^T |x(t)|^2 dt, \quad (4)$$

where $\mathcal{T}$ and T are clear from context. Finally, $\mathcal{L}_\infty^n$ denotes the class of essentially bounded piecewise-continuous functions $x : \mathbb{R} \rightarrow \mathbb{R}^n$.

2. Model reduction of feedback interconnections of linear and nonlinear systems

2.1. Background and motivation

This section motivates the model reduction problem for nonlinear systems with incremental system properties. Then, it presents the class of so-called Lur'e-type models. Finally, it provides a high-level problem description and an outline for solving this problem in the subsequent sections.

In the context of model reduction for nonlinear systems, the complexity of generic nonlinear behavior, such as the presence of multiple attractors or chaotic dynamics, can complicate the interpretation of reduced-order models. For example, when a reduced model exhibits a different number of attractors than the full-order system, it becomes challenging to assess the behavior of the error dynamics and, consequently, provide guarantees on the accuracy of the reduced model.

In contrast to generic nonlinear systems, nonlinear systems with *global incremental properties*, such as, e.g., an incremental $\mathcal{L}_2$ -gain, incremental stability, incremental passivity, or convergence, provide a system-theoretic understanding of their behavior (e.g., guaranteeing a single steady-state attractor in case of the convergence property). As we will demonstrate in Section 2.3, the incremental properties of such systems can be used to derive (a priori) error bounds between the system responses and the responses of the reduced-order model. Such an error bound is of crucial importance as it provides an accuracy certificate for the reduced-order model. Moreover, systems with incremental properties are inherently robust to input perturbations. For example, small perturbations in the input to systems with an incremental $\mathcal{L}_2$ -gain lead to small perturbations in the output.

In many branches of engineering and science, system behavior can be described by the interaction of (potentially high order) linear time-invariant (LTI) dynamics with a limited number of static nonlinearities or low-order nonlinear dynamics, as depicted in Fig. 1. Let us motivate this via two examples. Structural dynamics, e.g., of buildings and bridges, are typically modeled using structural finite elements with *linear* behavior, while nonlinearities appear only at the interfaces with other structures or with the ground (Suy et al., 2007). Electrical circuit networks can typically be described by LTI dynamics, while nonlinear elements, such as diodes and transistors, only appear locally. Intrinsically, the underlying modeling methods provide large-scale LTI dynamics, while their nonlinearities are typically only a few and only of a low order.

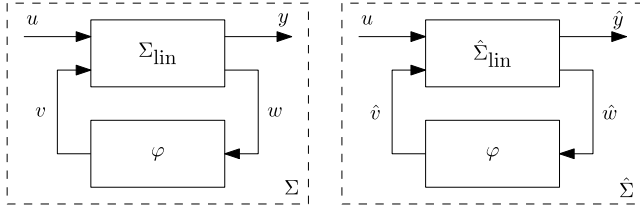


Fig. 1. A full-order (left) and a reduced-order (right) nonlinear model that consists of an LTI block placed in feedback with a nonlinear block. Only the order of the LTI block is reduced, while the nonlinear block is preserved.

Lur'e-type systems, see Fig. 1, consist of an interconnection between LTI dynamics and *static* nonlinearities, and are described by

$$\Sigma_{\text{lin}} : \begin{cases} \dot{x} &= Ax + B_u u + B_v v, \\ y &= C_y x, \\ w &= C_w x, \end{cases} \quad (5a)$$

$$v = \varphi(w), \quad (5b)$$

where $x(t) \in \mathbb{R}^n$ is the state of the LTI dynamics, $u(t) \in \mathbb{R}^m$ is the input, $y(t) \in \mathbb{R}^p$ is the output, $w(t) \in \mathbb{R}^q$ is the input to the nonlinear block, and $v(t) \in \mathbb{R}^r$ is the output of the nonlinear block. The LTI system is characterized by the matrices $A \in \mathbb{R}^{n \times n}$, $B_u \in \mathbb{R}^{n \times m}$, $B_v \in \mathbb{R}^{n \times r}$, $C_y \in \mathbb{R}^{p \times n}$, and $C_w \in \mathbb{R}^{q \times n}$. Using the Laplace variable $s \in \mathbb{C}$, we denote by $G(s)$ the transfer function of the LTI part of system (5), which is defined as

$$G(s) := C(sI_n - A)^{-1}B. \quad (6)$$

Here, for notational convenience, we have defined

$$B := \begin{bmatrix} B_u & B_v \end{bmatrix}, \quad C := \begin{bmatrix} C_y \\ C_w \end{bmatrix}. \quad (7)$$

Throughout this section, we will consider static, continuous nonlinearities $\varphi : \mathbb{R}^q \rightarrow \mathbb{R}^r$ with $\varphi(0) = 0$ satisfying so-called *incremental* properties. The simplest example, for a scalar nonlinearity $\varphi : \mathbb{R} \rightarrow \mathbb{R}$ and some finite $\mu > 0$, is the incremental sector condition

$$|\varphi(w) - \varphi(w')| \leq \mu |w - w'|, \quad \forall w, w' \in \mathbb{R}, \quad (8)$$

which can be regarded as a bound on the slope of φ , see Fig. 2 for an example. This is a Lipschitz condition with Lipschitz constant μ , and ensures that φ is globally Lipschitz continuous. If φ is differentiable, this is equivalent to

$$\left| \frac{d\varphi}{dw}(w) \right| \leq \mu, \quad \forall w \in \mathbb{R}. \quad (9)$$

A natural extension of (8) for multivariable nonlinearities is

$$|\varphi(w) - \varphi(w')| \leq \mu |w - w'|, \quad \forall w, w' \in \mathbb{R}^q, \quad (10)$$

where we recall that $|\cdot|$ denotes the Euclidean norm. We will discuss broader classes of incremental properties in Section 2.2. The Lur'e-type system resulting from the interconnection of the linear dynamics Σ_{lin} in (5a) with the static nonlinearity φ in (5b) will be denoted as $\Sigma = \mathcal{I}(\Sigma_{\text{lin}}, \varphi)$.

Remark 1. As we will explain in the following sections, most of the reduction results presented in this article are not restricted to Lur'e-type systems; they can also be applied when the nonlinear block includes dynamic nonlinearities. Lur'e-type systems are commonly used to model physical systems (see Shakib, Scarcioiti, Pogromsky, Pavlov, and van de Wouw (2023a), Zamani and Arcak (2017) and references therein), and they frequently arise from system identification methods (Shakib, Pogromsky, Pavlov, & van de Wouw, 2022) and nonlinear control designs (Daniel Andreas et al., 2021; Pavlov, Hunnekens, van de Wouw, & Nijmeijer, 2013; van de Wouw, Pastink, Heertjes, Pavlov, & Nijmeijer,

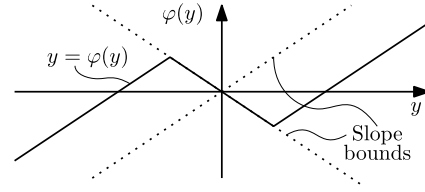


Fig. 2. An example of a single-input, single-output piece-wise linear static nonlinearity (solid line) that satisfies $\varphi(0) = 0$ and the condition (8), where the dotted lines represent the bounds on the slope of φ . Note that the nonlinearity is not required to be continuously differentiable.

2008). Furthermore, Lur'e-type models benefit from a wide range of established analysis and design tools (see, e.g., Khalil (1996)), which are compatible with the resulting Lur'e-type reduced-order models.

As a reduced-order model for (5), we consider the model described by

$$\hat{\Sigma}_{\text{lin}} : \begin{cases} \dot{\hat{x}} &= \hat{A}\hat{x} + \hat{B}_u u + \hat{B}_v \hat{v}, \\ \hat{y} &= \hat{C}_y \hat{x}, \\ \hat{w} &= \hat{C}_w \hat{x}, \end{cases} \quad (11a)$$

$$\hat{v} = \varphi(\hat{w}), \quad (11b)$$

where $\hat{x}(t) \in \mathbb{R}^{\hat{n}}$ is the state of the *reduced-order* LTI dynamics, $u(t) \in \mathbb{R}^m$ is the input, $\hat{y}(t) \in \mathbb{R}^p$ is the output, $\hat{w}(t) \in \mathbb{R}^q$ is the input to the nonlinear block, and $\hat{v}(t) \in \mathbb{R}^r$ is the output of the nonlinear block. The LTI part of the reduced-order model is characterized by the matrices $\hat{A} \in \mathbb{R}^{\hat{n} \times \hat{n}}$, $\hat{B}_u \in \mathbb{R}^{\hat{n} \times m}$, $\hat{B}_v \in \mathbb{R}^{\hat{n} \times r}$, $\hat{C}_y \in \mathbb{R}^{p \times \hat{n}}$, and $\hat{C}_w \in \mathbb{R}^{q \times \hat{n}}$. The transfer function of the LTI part of (11) is denoted by $\hat{G}$ and is defined similarly to (6). Note that the LTI block of this reduced-order model has a reduced dimension, i.e., $\hat{n} \ll n$, while the nonlinearity φ in (11b) is *inherited* from the system dynamics (5). Hence, by construction, (11) is a structure-preserving reduced-order model in the sense that the interconnection structure between the LTI block and the nonlinear block remains the same as the structure in the system (5). Consequently, the signals $\hat{v}$ and $\hat{w}$ preserve the physical meaning of the signals v and w , respectively. We denote the reduced-order model (11) by $\hat{\Sigma} = \mathcal{I}(\hat{\Sigma}_{\text{lin}}, \varphi)$.

In the next section, we will introduce different notions of incremental system properties for system (5), e.g., the notion of *convergence* (Yakubovich, 1964) or *incremental $\mathcal{L}_2$ gain* (Zames, 1966). Using these notions, in this article, we aim to find the matrices $\hat{A}, \dots, \hat{C}_w$, such that

- the reduced-order model (11) preserves the relevant incremental system property of system (5), e.g., convergence or incremental $\mathcal{L}_2$ gain, and
- the error between the responses of the system (5) and the reduced-order model (11) is bounded by a computable (a priori) error bound.

As a consequence, the reduced-order model enjoys the same incremental property and provides interpretability in terms of its interconnection structure, while the (bound on the) error between its responses and the system responses can be (a priori) computed.

In Section 2.2, we will introduce the considered notions of incremental system properties as well as their characterizations. Subsequently, in Section 2.3, we will demonstrate how these properties enable the derivation of error bounds between system responses and the responses of the reduced-order model. In Sections 2.4 and 2.5, we will recall LTI reduction approaches based on balanced truncation and moment matching, respectively, which are subsequently modified to address the problem statements in this article. Finally, we will provide numerical examples in Section 2.6.

2.2. Incremental system properties

In this section, we introduce incremental system properties for generic nonlinear systems and recall conditions for system (5) to enjoy such properties. These notions are introduced for general nonlinear systems of the form

$$\dot{x} = f(x, u), \quad (12)$$

where $x(t) \in \mathbb{R}^n$ is the state, $u(t) \in \mathbb{R}^m$ is the input, and $f : \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}^n$ is assumed to satisfy $f(0, 0) = 0$ and be locally Lipschitz in x and continuous in u . We will sometimes also consider an associated output equation

$$y = h(x, u) \quad (13)$$

with output $y(t) \in \mathbb{R}^p$ and $h : \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}^p$ continuous satisfying $h(0, 0) = 0$.

First, we will introduce the notion of *input-to-state convergence* in Section 2.2.1. In Section 2.2.2, we will present the notion of *incremental dissipativity*, before providing characterizations for these notions in Section 2.2.3.

2.2.1. Input-to-state convergence

The notion of convergence ensures the existence of a unique, well-defined, bounded, and globally asymptotically stable steady-state solution $\bar{x}_u$ for any bounded input u . Therefore, independent of the initial condition, all solutions of a convergent system converge to the uniquely defined, bounded steady-state solution $\bar{x}_u$. The convergence property has proven instrumental in many systems and control problems such as the output regulation problem (Pavlov, van de Wouw, & Nijmeijer, 2006), frequency-domain analysis of nonlinear systems (Pavlov, van de Wouw, & Nijmeijer, 2007), the performance optimization of nonlinear control systems (Pavlov et al., 2013), the observer design problem (Khalin, Efimov, & Ushirobira, 2023), synchronization (Pavlov, , Steur, & van de Wouw, 2022), extremum seeking control (Hazeleger, Haring, & van de Wouw, 2020) and the system identification problem (Amiri, Shakib, Lopez Arteaga, & van de Wouw, 2025; Shakib, Pogromsky, Pavlov, & van de Wouw, 2019; Shakib et al., 2022). We formally define the notion of convergence for the nonlinear system (12) as follows.

Definition 2 (Pavlov et al., 2006). The system (12) is said to be *globally uniformly convergent* if, for every input $u \in \mathcal{L}_\infty^m$, there exists a solution $\bar{x}_u$ to (12) that satisfies the following conditions:

- $\bar{x}_u$ is defined and bounded on $\mathbb{R}$.
- $\bar{x}_u$ is globally uniformly asymptotically stable.

The solution $\bar{x}_u$ is called the *steady-state* solution.

As shown in Property 2.15 in Pavlov et al. (2006), for a globally uniformly convergent system, this time-varying steady-state solution $\bar{x}_u$ is unique (the only solution with the above properties). This steady-state solution depends on the input u , i.e., if u is changed, then so is $\bar{x}_u$. We next introduce a stronger notion of convergence, namely input-to-state convergence, which bounds the change in $\bar{x}_u$ as a function of the change in u .

Definition 3 (Pavlov et al., 2006). The system (12) is said to be *input-to-state convergent* if it is globally uniformly convergent and, for every input $u \in \mathcal{L}_\infty^m$, system (12) is input-to-state stable with respect to the steady-state solution $\bar{x}_u$, i.e., there exists a $\mathcal{KL}$ -function β and a $\mathcal{K}_\infty$ -function γ such that any solution x of system (12) corresponding to some input $\hat{u}(t) := u(t) + \Delta u(t)$ satisfies

$$|x(t) - \bar{x}_u(t)| \leq \beta(|x(t_0) - \bar{x}_u(t_0)|, t - t_0) + \gamma \left(\sup_{t_0 \leq \tau \leq t} |\Delta u(\tau)| \right) \quad (14)$$

for all $t, t_0 \in \mathbb{R}$, $t \geq t_0$. The functions β and γ may depend on the particular input u .

As evidenced from (14), input-to-state convergent systems are robust against input variations. Namely, firstly, the bound on the steady-state difference $|\bar{x}_u(t) - \bar{x}_u(t)|$ increases monotonically with $|\Delta u(t)|$ given the fact that the function γ is of class $\mathcal{K}_\infty$, and, secondly, this property of the gain function γ also implies (since it is continuous and zero at zero argument) that a small change in the input, i.e., a small $|\Delta u(t)|$, results in a small steady-state difference $|\bar{x}_u(t) - \bar{x}_u(t)|$. Throughout this article, we call an input-to-state convergent system simply a convergent system.

2.2.2. Incrementally dissipative systems

While the notion of convergence of Section 2.2.1 characterizes behavior of the *state trajectory* of a nonlinear system (12), we consider *input-output behavior* in the current section.

Input-output behavior of nonlinear systems is most conveniently described using *dissipativity theory*, originally introduced in Willems (1972a, 1972b).

Definition 4 (Willems, 1972a). The system (12)–(13) is said to be *dissipative with respect to the supply rate* $s : \mathbb{R}^m \times \mathbb{R}^p \rightarrow \mathbb{R}$ if there exists a function $V : \mathbb{R}^n \rightarrow [0, \infty)$ such that

$$V(x(t_1)) - V(x(t_0)) \leq \int_{t_0}^{t_1} s(u(t), y(t)) dt \quad (15)$$

for any trajectory (x, u, y) satisfying (12)–(13), and all $t_0, t_1 \in \mathbb{R}$ with $t_0 \leq t_1$. Such function V is called a *storage function*.

The dissipation inequality (15) allows for a nice physical interpretation when thinking of the supply rate s as characterizing the “power” supplied to the nonlinear system, in the sense that $s(u(t), y(t))$ is the supplied power at time t . Then, $V(x(t))$ can be regarded as the “energy” stored in the nonlinear system at time t and the dissipation inequality essentially states that any change in (internally) stored energy can never exceed the energy supplied to the system through the (external) input and output. By relating external (input–output) to internal (state) system behavior, dissipativity theory has become a cornerstone of systems theory, see, e.g., the books (Arcak, Meissen, & Packard, 2016; Brogliato, Lozano, Maschke, & Egeland, 2020; van der Schaft, 2017).

Remark 5. In case the storage function V in Definition 4 is assumed to be differentiable, it follows immediately from the dissipation inequality (15) that the system (12)–(13) is dissipative with respect to the supply rate s if the infinitesimal dissipation inequality

$$\frac{\partial V}{\partial x}(x) f(x, u) \leq s(u, h(x, u)), \quad \forall x \in \mathbb{R}^n, u \in \mathbb{R}^m. \quad (16)$$

is satisfied for some (differentiable) function $V : \mathbb{R}^n \rightarrow [0, \infty)$. As this is a purely algebraic condition that does not require computation of trajectories, this gives a feasible approach towards verifying dissipativity.

Commonly used supply rates are those that define a bounded $\mathcal{L}_2$ gain and passivity, see, e.g., van der Schaft (2017).

Definition 6. The system (12)–(13) is said

1. to have an $\mathcal{L}_2$ gain bounded by $\gamma > 0$ if it is dissipative with respect to

$$s(u, y) = \gamma^2 |u|^2 - |y|^2. \quad (17)$$

It is (strictly) *contractive* if this holds with $\gamma \leq 1$ ($\gamma < 1$).

2. to be *passive* if $m = p$ and there exists $\delta \geq 0$ such that it is dissipative with respect to

$$s(u, y) = u^T y - \delta |y|^2. \quad (18)$$

It is *output-strictly passive* if this holds with $\delta > 0$.

We are interested in characterizing *incremental* input–output behavior, through an incremental version of dissipativity. This relies on the auxiliary system

$$\begin{aligned}\dot{x} &= f(x, u), & \dot{x}' &= f(x', u'), \\ y &= h(x, u), & y' &= h(x', u'),\end{aligned}\quad (19)$$

which comprises nothing more than two copies of the nonlinear system (12)–(13).

Definition 7. The system (12)–(13) is said to be incrementally dissipative with respect to the supply rate s if the auxiliary system (19) is dissipative with respect to the supply rate s_{aux} defined as

$$s_{\text{aux}}(u, u', y, y') = s(u - u', y - y'). \quad (20)$$

Following this definition, we emphasize that the storage function for the auxiliary system is a function of the state (x, x') , which is nonnegative but without any further structure. The application of Definition 7 to the supply rates of Definition 6 immediately leads to the definitions of a bounded incremental $\mathcal{L}_2$ gain, incremental contractivity, and incremental (output-strict) passivity.

Remark 8. When specializing (12)–(13) to linear systems

$$\begin{aligned}\dot{x} &= Ax + Bu, \\ y &= Cx + Du,\end{aligned}\quad (21)$$

it follows immediately from linearity that (21) is dissipative with respect to a supply rate s if and only if it is incrementally dissipative with respect to the same supply rate. More interestingly, the notions of bounded $\mathcal{L}_2$ gain and passivity can equivalently be expressed in terms of the transfer function

$$G(s) = C(sI - A)^{-1}B + D. \quad (22)$$

Namely, assuming that the realization (21) is minimal, we have that the linear system has an $\mathcal{L}_2$ gain bounded by γ if and only if G has no poles in $\{s \mid \Re(s) \geq 0\}$ and

$$\|G\|_{H_\infty} \leq \gamma. \quad (23)$$

If $\gamma \leq 1$, such a transfer function is said to be bounded real. Similarly, the linear system is passive if and only if G is positive real, i.e., G has no poles in $\{s \mid \Re(s) > 0\}$ and

$$G(s) + G(s)^* \geq 0 \text{ for all } s \text{ with } \Re(s) > 0. \quad (24)$$

These classical results are referred to as the bounded real and positive real lemma, respectively, see, e.g., Anderson and Vongpanitlerd (2006), Brogliato et al. (2020) for details.

While the notions of (incremental) dissipativity in Definitions 4 and 7 are defined for dynamical systems as in (12)–(13), they are readily extended to static nonlinearities.

Definition 9. The static nonlinearity $\varphi : \mathbb{R}^q \rightarrow \mathbb{R}^r$ is said to be *incrementally dissipative* with respect to the supply rate s if

$$s(w - w', \varphi(w) - \varphi(w')) \geq 0 \quad (25)$$

for all $w, w' \in \mathbb{R}^q$.

We note that this definition is a direct consequence of (15) (or (16)) in the absence of a storage function.

If (25) holds with s as in (17), we simply say that the nonlinearity has an incremental gain bounded by γ , and it is incrementally contractive if $\gamma \leq 1$. Note that, in this case, (25) is equivalent to

$$|\varphi(w) - \varphi(w')| \leq \gamma |w - w'|, \quad (26)$$

thus providing a direct generalization of the sector bound (8) from scalar nonlinearities to multivariable nonlinearities and recovering (10)

when $\gamma = \mu$. A static nonlinearity is called *incrementally* (output-strictly) *passive* if there exists a constant $\delta \geq 0$ ($\delta > 0$) such that

$$(w - w')^\top (\varphi(w) - \varphi(w')) \geq \delta |\varphi(w) - \varphi(w')|^2, \quad (27)$$

holds for all $w, w' \in \mathbb{R}^q$, which is again a direct consequence of Definition 9.

2.2.3. Lur'e-type systems with incremental properties

Having introduced the incremental properties of (input-to-state) convergence in Section 2.2.1 and incremental dissipativity in Section 2.2.2, we now turn to characterizations of these properties for Lur'e-type systems as in (5). Subsequently, we will present the formal problem statement at the end of this section.

We first present a set of conditions guaranteeing both the convergence property and a bounded incremental $\mathcal{L}_2$ gain for system (5).

Theorem 10. Consider the Lur'e-type system $\Sigma = I(\Sigma_{\text{lin}}, \varphi)$ in (5). Suppose that

- A1. A is Hurwitz, i.e., $\lambda(A) \subset \mathbb{C}^-$;
- A2. the nonlinearity φ satisfies $\varphi(0) = 0$ and the incremental sector condition (10) holds for some $\mu > 0$;
- A3. the small-gain condition

$$\gamma_{wv}\mu < 1 \quad (28)$$

holds, where $\gamma_{wv} = \|G_{(w,v)}\|_{H_\infty}$, $G_{(w,v)}(s) = C_w(sI - A)^{-1}B_v$.

Then, the system (5)

1. is globally uniformly convergent and input-to-state convergent for the class of inputs $\mathcal{L}_{\infty}^m$ and
2. has a bounded incremental $\mathcal{L}_2$ gain.

Proof. The proof of the (input-to-state) convergence property is similar to the proof of Shakib, Scarcioiti, Jungers, et al. (2024, Theorem 1). The bounded incremental $\mathcal{L}_2$ gain follows along the same lines of Besselink, van de Wouw, and Nijmeijer (2013, Lemma 12), after noting that the assumption that A is Hurwitz guarantees that Σ_{lin} has a bounded $\mathcal{L}_2$ gain. $\square$

A key observation from Theorem 10 is that the assumptions A1–A3 simultaneously guarantee convergence and incremental $\mathcal{L}_2$ gain properties. The fact that the nonlinearity satisfies an *incremental* sector bound is crucial in this.

Next, we stress that the assumption that A is Hurwitz guarantees that Σ_{lin} has a bounded $\mathcal{L}_2$ gain. In case the realization (5a) is both controllable and observable, the converse holds as well (i.e., a bounded $\mathcal{L}_2$ gain implies that A is Hurwitz). This can be further specialized to the case of contractive Σ_{lin} , which leads to the following corollary of Theorem 10.

Corollary 11. Consider the Lur'e-type system $\Sigma = I(\Sigma_{\text{lin}}, \varphi)$ in (5). Suppose that

- (A, B) is controllable, (A, C) is observable, and Σ_{lin} is contractive with input (u, v) and output (y, w) ;
- the nonlinearity φ satisfies $\varphi(0) = 0$ and is strictly incrementally contractive, i.e., (26) holds with $\gamma < 1$.

Then, the system (5)

1. is globally uniformly convergent and input-to-state convergent for the class of inputs $\mathcal{L}_{\infty}^m$ and
2. is incrementally contractive.

A counterpart to this result, exploiting (incremental) passivity rather than (incremental) contractivity, is given next. It can be concluded from standard results on passive interconnections, e.g., van der Schaft (2017).

Theorem 12. Consider the Lur'e-type system $\Sigma = I(\Sigma_{\text{lin}}, \varphi)$ in (5). Suppose that

- (A, B) is controllable, (A, C) is observable, and Σ_{lin} is (output-strictly) passive with input (u, v) and output (y, w) ;
- the nonlinearity φ satisfies $\varphi(0) = 0$ and is incrementally passive, i.e., (27) holds with $\delta \geq 0$.

Then, the system (5) is incrementally (output-strictly) passive.

At this point, we recall that we are interested in approximating the Lur'e-type system (5) with one of reduced order, see Fig. 1. On the basis of Theorem 10, we can give the main problem statement of Section 2.

Problem 13. Assume that the Lur'e-type system (5) satisfies assumptions A1–A3 in the statement of Theorem 10, i.e., it has incremental stability properties. Given $\hat{n} < n$, find a reduced-order Lur'e-type system (11) and (a priori or a posteriori) a constant $\epsilon \in \mathbb{R}_+$ such that

- the reduced-order model enjoys the same incremental stability properties, i.e., items A1 and A3 of Theorem 10 hold for the reduced-order Lur'e-type system;
- for each $T > 0$, the error bound

$$\|\bar{y}_u - \hat{y}_u\|_{\mathcal{P}_2} \leq \epsilon \|u\|_{\mathcal{P}_2} \quad (29)$$

holds for any periodic, bounded input $u \in \mathcal{P}_2^m(T)$, where $\bar{y}_u$ and $\hat{y}_u$ are the unique steady-state outputs of the high-order and reduced-order Lur'e-type systems;

- the error bound

$$\|y - \hat{y}\|_{\mathcal{L}_2} \leq \epsilon \|u\|_{\mathcal{L}_2} \quad (30)$$

holds for any bounded input $u \in \mathcal{L}_2^m(\mathbb{R}_+)$, where y and $\hat{y}$ are the outputs of the high-order and reduced-order Lur'e-type systems, starting from zero initial conditions.

Hence, we ask for the reduced-order Lur'e-type system to not only preserve relevant incremental system (stability) properties, but also to satisfy error bounds on two different classes of signals. First, (29) asks for an error bound on the (periodic) steady-state outputs for periodic input functions (defined on $\mathbb{R}$), for which, by the grace of the convergence property, the steady-state solutions exist and are bounded for any bounded input u (see Definition 2), hence also for T -periodic inputs $u \in \mathcal{P}_2^m(T)$. Second, (30) asks for an $\mathcal{L}_2$ error bound for input functions in $\mathcal{L}_2^m$ (defined on $\mathbb{R}_+$ and for zero initial conditions).

2.3. Error analysis using incremental system properties

Problem 13 asks for the distinct error bounds (29) and (30), which we will refer to as the *steady-state* and $\mathcal{L}_2$ perspective, respectively, to hold. In this section, we will show that these two perspectives admit a unified analysis of the reduction order.

To present this unified error analysis, it will be useful to view the linear subsystem (5a) of the Lur'e-type system as a general linear operator

$$\mathcal{G} : \mathcal{U} \times \mathcal{V} \rightarrow \mathcal{Y} \times \mathcal{W}, \quad (31)$$

with $\mathcal{U}$, $\mathcal{V}$, $\mathcal{Y}$, and $\mathcal{W}$ the signal spaces in which u , v , y , and w live, respectively. The above perspectives specialize this operator as follows.

- For the *steady-state* perspective, we consider periodic signals and take

$$\begin{aligned} \mathcal{U} &= \mathcal{P}_2^m(T), & \mathcal{V} &= \mathcal{P}_2^r(T), \\ \mathcal{Y} &= \mathcal{P}_2^p(T), & \mathcal{W} &= \mathcal{P}_2^q(T). \end{aligned} \quad (32)$$

Correspondingly, $\mathcal{G}$ represents the steady-state operator

$$\mathcal{G} : \begin{bmatrix} u \\ v \end{bmatrix} \mapsto \begin{bmatrix} \bar{y}_{(u,v)} \\ \bar{w}_{(u,v)} \end{bmatrix}. \quad (33)$$

We emphasize that, if A is Hurwitz, this steady-state operator is well-defined and indeed maps periodic inputs to (bounded) periodic (steady-state) outputs.

- For the $\mathcal{L}_2$ perspective, we take

$$\begin{aligned} \mathcal{U} &= \mathcal{L}_2^m(\mathbb{R}_+), & \mathcal{V} &= \mathcal{L}_2^r(\mathbb{R}_+), \\ \mathcal{Y} &= \mathcal{L}_2^p(\mathbb{R}_+), & \mathcal{W} &= \mathcal{L}_2^q(\mathbb{R}_+), \end{aligned} \quad (34)$$

and $\mathcal{G}$ corresponds to the usual input–output operator for zero initial conditions. Thus,

$$\mathcal{G} : \begin{bmatrix} u \\ v \end{bmatrix} \mapsto \begin{bmatrix} y \\ w \end{bmatrix}, \text{ with } \begin{bmatrix} y(t) \\ w(t) \end{bmatrix} = \int_0^t C e^{A(t-\tau)} B \begin{bmatrix} u(\tau) \\ v(\tau) \end{bmatrix} d\tau. \quad (35)$$

Again, this operator is well-defined (and maps $\mathcal{L}_2$ input signals to $\mathcal{L}_2$ output signals) if A is Hurwitz.

Due to linearity, we can introduce a partitioning of $\mathcal{G}$ as

$$\mathcal{G} \left(\begin{bmatrix} u \\ v \end{bmatrix} \right) = \begin{bmatrix} \mathcal{G}_{yu}(u) + \mathcal{G}_{yv}(v) \\ \mathcal{G}_{wu}(u) + \mathcal{G}_{wv}(v) \end{bmatrix}, \quad (36)$$

and let $\|\cdot\|$ denote the standard operator norm. For example,

$$\|\mathcal{G}_{yu}\| = \sup \left\{ \frac{\|\mathcal{G}_{yu}(u)\|_{\mathcal{Y}}}{\|u\|_{\mathcal{U}}} \mid u \in \mathcal{U}, u \neq 0 \right\}. \quad (37)$$

An important observation is that, regardless whether $\mathcal{G}$ takes the form (33) or (35), this operator norm is related to the $\mathcal{H}_\infty$ norm of the transfer function G of (5a).

Lemma 14. Consider the linear dynamics (5a) with associated transfer function G as in (6). Assume that A is Hurwitz, let $\mathcal{G}$ denote either (33) or (35), and consider the partitioning (36). Then, for any $i \in \{u, v\}$ and $j \in \{y, w\}$,

$$\|\mathcal{G}_{ji}\| = \gamma_{ji}, \quad \text{with } \gamma_{ji} = \|G_{(j,i)}\|_{\mathcal{H}_\infty} \quad (38)$$

and $G_{(j,i)} = C_j(sI - A)^{-1}B_i$.

In the steady-state perspective, γ_{ji} is the worst-case, i.e., the maximum amplification of periodic inputs to the steady-state response, see, e.g., Pavlov et al. (2013). In the $\mathcal{L}_2$ perspective, this statement is nothing more than (a scaled version of) the bounded real lemma (Scherer, 1990, Theorem 2.2) (see also Remark 8).

Next, shifting attention to the local nonlinearity φ of the Lur'e-type system, we have that the incremental sector condition (10) immediately translates to an (incremental) gain on the corresponding signals w and v .

Lemma 15. Consider the nonlinearity φ as in (5b), assume that it satisfies $\varphi(0) = 0$ and the incremental sector condition (10) holds for some $\mu > 0$. Let the pair $(\mathcal{V}, \mathcal{W})$ satisfy either (32) (for some $T > 0$) or (34) and, for a given $w \in \mathcal{W}$, let v denote the signal defined as $v(t) = \varphi(w(t))$. Then, the following holds:

1. if $w \in \mathcal{W}$, then $v \in \mathcal{V}$ and $\|v\|_{\mathcal{V}} \leq \mu \|w\|_{\mathcal{W}}$;
2. if $w, w' \in \mathcal{W}$, then $w - w' \in \mathcal{W}$, $v - v' \in \mathcal{V}$, and $\|v - v'\|_{\mathcal{V}} \leq \mu \|w - w'\|_{\mathcal{W}}$.

After these preparatory observations, we are in a position to derive a general error bound for the reduction of Lur'e-type systems as in Fig. 1. Here, we make the following assumptions.

Assumption 16. The Assumptions A1–A3 in the statement of Theorem 10 hold and, moreover,

- A4. $\hat{A}$ is Hurwitz, i.e., $\lambda(\hat{A}) \subset \mathbb{C}^-$;
- A5. the small-gain condition

$$\hat{g}_{wv} \mu < 1 \quad (39)$$

holds, where $\hat{g}_{wv} = \|\hat{G}_{(w,v)}\|_{\mathcal{H}_\infty}$, $\hat{G}_{(w,v)} = \hat{C}_w(sI - \hat{A})^{-1}\hat{B}_v$.

In other words, we assume that the conditions of [Theorem 10](#) hold for the reduced-order Lur'e-type system as well. As $\hat{A}$ is Hurwitz, we can define the input-output operator $\hat{G}$ for the reduced-order linear dynamics and partition it in a similar way as in [\(36\)](#). Then, the result of [Lemma 14](#) holds for the reduced-order linear dynamics as well, and we denote the corresponding gains by $\hat{\gamma}_{ji}$.

Moreover, we assume that we have access to error bounds on the reduction of the linear dynamics, i.e., we know constants $\epsilon_{ji} > 0$ such that, for any $i \in \{u, v\}$ and $j \in \{y, w\}$,

$$\|G_{(j,i)} - \hat{G}_{(j,i)}\|_{\mathcal{H}_\infty} \leq \epsilon_{ji}. \quad (40)$$

Now, we proceed to the analysis of the reduction error between the original and reduced-order Lur'e-type systems. Taking $u \in \mathcal{U}$, we have for any $v, \hat{v} \in \mathcal{V}$ that

$$\begin{aligned} w - \hat{w} &= \mathcal{G}_{wu}(u) + \mathcal{G}_{wv}(v) - \hat{\mathcal{G}}_{wu}(u) - \hat{\mathcal{G}}_{wv}(\hat{v}) \\ &= \mathcal{G}_{wv}(v - \hat{v}) + (\mathcal{G}_{wu} - \hat{\mathcal{G}}_{wu})(u) + (\mathcal{G}_{wv} - \hat{\mathcal{G}}_{wv})(\hat{v}), \end{aligned} \quad (41)$$

where linearity of the operator $\mathcal{G}$ is exploited. Application of the triangle inequality, together with the bounds [\(38\)](#) and [\(40\)](#), leads to

$$\|w - \hat{w}\|_{\mathcal{W}} \leq \gamma_{wv}\|v - \hat{v}\|_{\mathcal{V}} + \epsilon_{wu}\|u\|_{\mathcal{U}} + \epsilon_{wv}\|\hat{v}\|_{\mathcal{V}}. \quad (42)$$

Next, noting that v and w (and $\hat{v}$ and $\hat{w}$) are related through the static nonlinearity φ , it follows from the incremental sector condition and [Lemma 15](#) that

$$\|v - \hat{v}\|_{\mathcal{V}} \leq \mu\|w - \hat{w}\|_{\mathcal{W}}. \quad (43)$$

Substitution of [\(43\)](#) in [\(42\)](#) and collecting terms leads to

$$\|w - \hat{w}\|_{\mathcal{W}} \leq \frac{1}{1 - \gamma_{wv}\mu} (\epsilon_{wu}\|u\|_{\mathcal{U}} + \epsilon_{wv}\|\hat{v}\|_{\mathcal{V}}), \quad (44)$$

where we note that the denominator is positive due to [\(28\)](#).

Focusing on the reduced-order Lur'e-type system, note that

$$\hat{w} = \hat{\mathcal{G}}_{wu}(u) + \hat{\mathcal{G}}_{wv}(\hat{v}), \quad (45)$$

leading to the bound

$$\|\hat{w}\|_{\mathcal{W}} \leq \hat{\gamma}_{wu}\|u\|_{\mathcal{U}} + \hat{\gamma}_{wv}\|\hat{v}\|_{\mathcal{V}}. \quad (46)$$

Exploiting again the incremental sector condition and $\varphi(0) = 0$, it follows from [Lemma 15](#) that $\|\hat{v}\|_{\mathcal{V}} \leq \mu\|\hat{w}\|_{\mathcal{W}}$, after which substitution in [\(46\)](#) yields

$$\|\hat{w}\|_{\mathcal{W}} \leq \frac{\hat{\gamma}_{wu}}{1 - \hat{\gamma}_{wv}\mu} \|u\|_{\mathcal{U}} \quad (47)$$

and

$$\|\hat{v}\|_{\mathcal{V}} \leq \frac{\hat{\gamma}_{wu}\mu}{1 - \hat{\gamma}_{wv}\mu} \|u\|_{\mathcal{U}}. \quad (48)$$

Using the same line of reasoning as in the derivation of the bound [\(42\)](#), we find

$$\|y - \hat{y}\|_{\mathcal{Y}} \leq \gamma_{yv}\|v - \hat{v}\|_{\mathcal{V}} + \epsilon_{yu}\|u\|_{\mathcal{U}} + \epsilon_{yv}\|\hat{v}\|_{\mathcal{V}}. \quad (49)$$

Finally, subsequently substituting [\(43\)](#), [\(44\)](#), and [\(48\)](#) in [\(49\)](#) leads to

$$\|y - \hat{y}\|_{\mathcal{Y}} \leq \epsilon\|u\|_{\mathcal{U}}, \quad (50)$$

with

$$\epsilon = \epsilon_{yu} + \frac{\gamma_{yv}\mu}{1 - \gamma_{wv}\mu} \epsilon_{wu} + \left(\epsilon_{yv} + \frac{\gamma_{yv}\mu}{1 - \gamma_{wv}\mu} \epsilon_{wv} \right) \frac{\hat{\gamma}_{wu}\mu}{1 - \hat{\gamma}_{wv}\mu}. \quad (51)$$

Thus, we have proven the following result.

Theorem 17. Consider the Lur'e-type system $\Sigma = \mathcal{I}(\Sigma_{\text{lin}}, \varphi)$ in [\(5\)](#) and the reduced-order Lur'e-type system $\hat{\Sigma} = \mathcal{I}(\hat{\Sigma}_{\text{lin}}, \varphi)$ in [\(11\)](#). Let [Assumption 16](#) hold and suppose the constants $\epsilon_{ji} > 0$ are such that [\(40\)](#) holds for any $i \in \{u, v\}$ and $j \in \{y, w\}$. Then, the following error bounds hold for $\epsilon = \epsilon$ in [\(51\)](#):

1. for each $T > 0$, and any $u \in \mathcal{P}_2^m(T)$, we have

$$\|\bar{y}_u - \hat{y}_u\|_{\mathcal{P}_2} \leq \epsilon\|u\|_{\mathcal{P}_2} \quad (52)$$

with $\bar{y}_u \in \mathcal{P}_2^p(T)$ and $\hat{y}_u \in \mathcal{P}_2^p(T)$ the steady-state outputs of Σ and $\hat{\Sigma}$, respectively;

2. for any $u \in \mathcal{L}_2^m(\mathbb{R}_+)$, we have

$$\|y - \hat{y}\|_{\mathcal{L}_2} \leq \epsilon\|u\|_{\mathcal{L}_2} \quad (53)$$

with $y \in \mathcal{L}_2^p(\mathbb{R}_+)$ and $\hat{y} \in \mathcal{L}_2^p(\mathbb{R}_+)$ the outputs of Σ and $\hat{\Sigma}$, respectively, for zero initial conditions.

In other words, the desired error bounds of [Problem 13](#) hold for $\epsilon = \epsilon$ in [\(51\)](#). We note that $\epsilon = 0$ when the reduction of the linear dynamics does not introduce any error, i.e., when $\epsilon_{ji} = 0$, as desired.

Remark 18. The error bounds of [Theorem 17](#) are a *posteriori* error bounds. Namely, they rely on the availability of $\hat{\gamma}_{wv}$ and $\hat{\gamma}_{wu}$, which can only be computed after the reduced-order linear dynamics are obtained. However, assuming that a reduction procedure on the linear dynamics is used with an a priori error bound, i.e., the bounds ϵ_{ji} in [\(40\)](#) can be obtained before computing $\hat{\Sigma}_{\text{lin}}$, we can provide bounds on $\hat{\gamma}_{wv}$ and $\hat{\gamma}_{wu}$ as

$$\hat{\gamma}_{wv} \leq \gamma_{wv} + \epsilon_{wv}, \quad \hat{\gamma}_{wu} \leq \gamma_{wu} + \epsilon_{wu}. \quad (54)$$

Assuming that $(\gamma_{wv} + \epsilon_{wv})\mu < 1$ to guarantee that A5 in [Assumption 16](#) holds, $\hat{\gamma}_{wv}$ and $\hat{\gamma}_{wu}$ in [\(51\)](#) can be replaced by these bounds. In this case, the resulting expression can be evaluated *a priori*. Of course, this comes at the cost of conservatism, in the sense that the resulting bound is larger than that in [\(51\)](#).

Remark 19. The expression for the constant ϵ in [\(51\)](#) can be further simplified at the cost of conservatism by noting that $\epsilon_{ji} \leq \epsilon_{\text{lin}}$ with

$$\epsilon_{\text{lin}} := \|G - \hat{G}\|_{\mathcal{H}_\infty}, \quad (55)$$

where G is the transfer function of Σ_{lin} , defined in [\(6\)](#), and $\hat{G}$ is the transfer function of $\hat{\Sigma}_{\text{lin}}$. If we use $\epsilon_{ji} \leq \epsilon_{\text{lin}}$, we find that

$$\epsilon \leq \epsilon_{\text{lin}} \left(1 + \frac{\gamma_{yv}\mu}{1 - \gamma_{wv}\mu} \right) \left(1 + \frac{\hat{\gamma}_{wu}\mu}{1 - \hat{\gamma}_{wv}\mu} \right). \quad (56)$$

A key consequence of [Theorem 17](#) is that an error bound is guaranteed as long as the reduced-order linear dynamics satisfy A4 and A5 in [Assumption 16](#) and an error bound [\(40\)](#) is available on the reduction of the linear dynamics. The next sections discuss reduction methods for the linear dynamics targeting these properties.

Remark 20. The steady-state output error bound [\(52\)](#) is derived specifically for T -periodic excitation signals $u \in \mathcal{P}_2^m(T)$. A related result for the steady-state output error bound is provided in [Besselink, van de Wouw, and Nijmeijer \(2012\)](#), which considers essentially bounded, piecewise-continuous signals $u \in \mathcal{L}_\infty^m$ that are not necessarily periodic. In that case, the bound is expressed in terms of the $\mathcal{L}_\infty$ signal norm, with a different constant ϵ .

2.4. Model reduction by balanced truncation

Balanced truncation is one of the most popular methods for model reduction of (linear) systems. It aims at approximating the input-output behavior of a high-dimensional system by preserving state components that are simultaneously easily controlled and easily observed. The most important feature of balanced truncation is that it provides guarantees on the reduced-order approximant. Namely, the stability of the reduced-order model is guaranteed if the high-order model is stable and, moreover, an a priori bound on the reduction error is available.

In this section, we will first review balanced truncation for linear time-invariant systems, see [Section 2.4.1](#). In [Section 2.4.2](#), we will

discuss extensions aimed at the preservation of dissipativity properties. Finally, in Section 2.4.3 we will employ balanced truncation in the scope of Lur'e-type systems. We will also apply balanced truncation in Section 3 on model reduction of interconnected, linear dynamical systems.

2.4.1. Balanced model reduction for LTI systems

Consider the linear time-invariant system

$$\begin{aligned}\dot{x} &= Ax + Bu, \\ y &= Cx + Du,\end{aligned}\quad (57)$$

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

The method of balanced truncation for model reduction of systems of the form (57) was first introduced by Moore in Moore (1981), inspired by methods in filtering (Mullis & Roberts, 1976). Desirable properties of balanced truncation such as stability preservation and the availability of an a priori error bound were subsequently shown in Pernebo and Silverman (1982) and Enns (1984), Glover (1984), respectively. The material in this section is based on texts such as Antoulas (2005a), Zhou, Doyle, and Glover (1996).

Balanced truncation relies on the characterization of controllability and observability properties of (57) through so-called energy functions. The controllability function is the first such function.

Definition 21. For the linear system (57) with (A, B) controllable, its *controllability function* is the function $L_c : \mathbb{R}^n \rightarrow \mathbb{R}$ defined as

$$L_c(x_0) = \inf \left\{ \int_{-\infty}^0 |u(t)|^2 dt \mid \begin{array}{l} u \in L_2^m((-\infty, 0]) \text{ such that} \\ \lim_{t \rightarrow -\infty} x(t) = 0 \text{ and } x(0) = x_0 \end{array} \right\}. \quad (58)$$

The controllability function provides a measure of controllability by characterizing the least amount of “energy”, in terms of the $\mathcal{L}_2$ norm of the input function, that is required to reach a state $x_0 \in \mathbb{R}^n$. We can think of a state x_0 to be easily reachable when $L_c(x_0)$ is small, and difficult to reach if $L_c(x_0)$ is large.

Similarly, a measure of observability is provided through the observability function.

Definition 22. For the linear system (57) with A Hurwitz, its *observability function* is the function $L_o : \mathbb{R}^n \rightarrow \mathbb{R}$ defined as

$$L_o(x_0) = \int_0^\infty |y(t)|^2 dt \quad \text{for } u = 0 \text{ and } x(0) = x_0. \quad (59)$$

Thus, the observability function measures the “energy” in the output for a given initial condition $x_0 \in \mathbb{R}^n$ (and zero input). We can think of a state x_0 to be easily observable when $L_o(x_0)$ is large and difficult to observe if $L_o(x_0)$ is small.

The controllability and observability functions admit a clear characterization.

Theorem 23. Consider the linear system (57) with (A, B) controllable, (A, C) observable, and A Hurwitz. Then,

1. its controllability function (58) satisfies

$$L_c(x_0) = x_0^\top P^{-1} x_0, \quad (60)$$

with P the unique symmetric solution to

$$AP + PA^\top + BB^\top = 0. \quad (61)$$

This matrix P is positive definite.

2. its observability function (59) satisfies

$$L_o(x_0) = x_0^\top Q x_0, \quad (62)$$

with Q the unique symmetric solution to

$$A^\top Q + QA + C^\top C = 0. \quad (63)$$

This matrix Q is positive definite.

The matrices P and Q in items 1. and 2. above are known as the *controllability Gramian* and *observability Gramian*, respectively. They play a crucial role in balancing theory as they fully characterize the controllability and observability functions and can be obtained as solutions to the Lyapunov Eqs. (61) and (63), for which effective computational tools exist.

To evaluate the effect of a change of coordinates on the energy functions, let $R \in \mathbb{R}^{n \times n}$ be a nonsingular matrix defining a change of coordinates $x = R\bar{x}$, with $\bar{x}(t) \in \mathbb{R}^n$ the new coordinates. This change of coordinates affects the system matrices as

$$(A, B, C, D) \mapsto (R^{-1}AR, R^{-1}B, CR, D), \quad (64)$$

and it can easily be verified that the controllability and observability Gramian transform as

$$P \mapsto R^{-1}PR^{-\top} = (R^\top P^{-1}R)^{-1}, \quad Q \mapsto R^\top QR. \quad (65)$$

A key result in balanced truncation is that a change of coordinates can be found in which the Gramians are equal. This can be concluded from the following result on simultaneous diagonalization, see, e.g., Horn and Johnson (2013, Theorem 7.6.1).

Lemma 24. Let $X \in \mathbb{R}^{n \times n}$ and $Y \in \mathbb{R}^{n \times n}$ be positive definite symmetric matrices. Then, there exists a nonsingular matrix $R \in \mathbb{R}^{n \times n}$ such that

$$R^\top X R = (R^\top Y R)^{-1} = \Sigma, \quad (66)$$

with Σ of the form

$$\Sigma = \text{diag}(\sigma_1 I_{m_1}, \sigma_2 I_{m_2}, \dots, \sigma_\ell I_{m_\ell}), \quad (67)$$

where σ_i^2 are the distinct eigenvalues of the product XY^{-1} , with multiplicities m_i , for $i \in \{1, 2, \dots, \ell\}$, and with $\sum_{i=1}^\ell m_i = n$. In addition, $\sigma_1 > \sigma_2 > \dots > \sigma_\ell > 0$.

In particular, the application of Lemma 24 for $X = Q$ and $Y = P^{-1}$ shows that there exist coordinates in which the Gramians are equal. This motivates the following definition.

Definition 25. A realization (A, B, C, D) is called *balanced* if the controllability and observability Gramians satisfy $P = Q = \Sigma$, with Σ of the form (67).

Note that, for a balanced realization, the energy functions take the form

$$L_c(x_0) = x_0^\top \Sigma^{-1} x_0, \quad L_o(x_0) = x_0^\top \Sigma x_0, \quad (68)$$

allowing for a clear interpretation. Namely, the Gramians are *diagonal*, which admits an analysis of controllability and observability properties on the basis of individual state components. In particular, as the Gramians are *ordered*, we have that the first state requires the least energy to reach, i.e., $L_c(e_1) = \sigma_1^{-1} < \sigma_2^{-1} < \dots < \sigma_\ell^{-1} = L_c(e_n)$. Crucially, the notions of controllability and observability are aligned as the Gramians are *equal*. In other words, the first state component is simultaneously easy to reach ($L_c(e_1)$ is small) and easy to observe ($L_o(e_1)$ is large).

Remark 26. The constants $\sigma_1, \dots, \sigma_\ell$, which equal the square roots of the eigenvalues of the product QP , are called *Hankel singular values*. Namely, they equal the singular values of the so-called Hankel operator associated with (57), which is the operator mapping past inputs to future outputs, see, e.g., Glover (1984).

The above motivates maintaining the first state components in the balanced realization (associated with the largest Hankel singular values) in defining a reduced-order model.

From now on, we assume that the original realization (57) is balanced (and note that this can always be achieved after a change of coordinates). For this balanced realization, we introduce the partitioning

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}, \quad B = \begin{bmatrix} B_1 \\ B_2 \end{bmatrix}, \quad C = \begin{bmatrix} C_1 & C_2 \end{bmatrix} \quad (69)$$

and choose the reduced-order model as

$$\begin{aligned}\dot{\xi} &= A_{11}\xi + B_1u, \\ \hat{y} &= C_1\xi + Du.\end{aligned}\quad (70)$$

Here, $\xi(t) \in \mathbb{R}^{\hat{n}}$ is the state of the reduced-order system with $\hat{n} < n$, and $\hat{y}(t) \in \mathbb{R}^p$ is the approximation of the output. We emphasize that (70) is obtained by a two-step procedure referred to as *balanced truncation*: first, a change of coordinates is performed such that (A, B, C, D) in (57) is a balanced realization (balancing); and, second, the reduced-order model (70) is obtained by discarding the states that are simultaneously difficult to reach and observe (truncation).

The reduced-order model (70) has the following properties.

Theorem 27. Consider the linear system (57) with (A, B) controllable, (A, C) observable, and A Hurwitz. Let (70) be obtained by balanced truncation of (57). Assume in addition that $\hat{n} = \sum_{i=1}^k m_i$ for some $k < \ell$. Then:

1. (A_{11}, B_1) is controllable, (A_{11}, C_1) is observable, and (A_{11}, B_1, C_1, D) is a balanced realization;
2. A_{11} is Hurwitz;
3. the error bound

$$\|G - \hat{G}\|_{H_\infty} \leq 2 \sum_{i=k+1}^{\ell} \sigma_i \quad (71)$$

holds with G and $\hat{G}$ the transfer functions of (57) and (70), respectively, and σ_i the Hankel singular values of G .

The last two items state that balanced truncation preserves asymptotic stability and guarantees a bound on the reduction error, respectively. In particular, the error bound depends on the Hankel singular values σ_i that are truncated (not counting multiplicities). As they can be computed before the actual reduction is performed, this is an *a priori* error bound.

2.4.2. Dissipativity-preserving balanced model reduction

Balanced truncation, discussed in Section 2.4.1, preserves relevant system properties such as controllability, observability, and stability. In this section, we briefly discuss balancing methods for linear systems (57) aimed at preserving *dissipativity* properties. We focus on the preservation of contractivity, which leads to a method referred to as bounded real balancing. This was originally introduced in [Opdenacker and Jonckheere \(1988\)](#), see also [Ober \(1991\)](#), [Trentelman \(2009\)](#).

In dissipativity-preserving balanced model reduction, a crucial role is played by two “special” storage functions. The first, referred to as the *required supply*, is defined as

$$V_r(x_0) = \inf \left\{ \int_{-T}^0 s(u(t), y(t)) dt \mid \begin{array}{l} T \geq 0, u : [-T, 0] \rightarrow \mathbb{R}^m \\ \text{s.t. } x(-T) = 0, x(0) = x_0 \end{array} \right\},$$

where we note that the infimum is taken over both T and u . Recalling the energy interpretation of dissipativity theory, the required supply can be thought of as quantifying the least amount of energy (measured through the supply rate) that needs to be supplied to the system to reach the state $x_0 \in \mathbb{R}^n$ from the origin. Loosely speaking, the required supply thus plays a similar role as the controllability function of [Definition 21](#), although it uses a different measure of “energy”.

The second storage function is the *available storage*

$$V_a(x_0) = \sup \left\{ -\int_0^T s(u(t), y(t)) dt \mid \begin{array}{l} T \geq 0, u : [0, T] \rightarrow \mathbb{R}^m \\ \text{with } x(0) = x_0 \end{array} \right\},$$

which gives the maximum amount of energy that can be extracted from the system when starting from the initial condition $x_0 \in \mathbb{R}^n$. Thus, the available storage plays a similar role as the observability function of [Definition 22](#). It can be shown that, for a dissipative system, the available storage and required supply give the smallest and largest storage function, respectively, in the sense that $V_a(x) \leq V(x) \leq V_r(x)$ for all $x \in \mathbb{R}^n$ and any storage function V .

In the remainder of this section, we focus on contractive linear systems, i.e., systems that are dissipative with respect to the supply rate (17) for $\gamma \leq 1$. The fact that such systems admit quadratic storage functions, together with [Remark 5](#), contributes to the following result (see also, e.g., [Anderson & Vongpanitlerd, \(2006\)](#)).

Theorem 28. Consider the linear system (57) with (A, B) controllable and (A, C) observable. The system is contractive if and only if there exists a positive definite symmetric matrix K satisfying

$$\begin{bmatrix} A^\top K + KA & KB \\ B^\top K & 0 \end{bmatrix} - \begin{bmatrix} 0 & I \\ C & D \end{bmatrix}^\top \begin{bmatrix} I & 0 \\ 0 & -I \end{bmatrix} \begin{bmatrix} 0 & I \\ C & D \end{bmatrix} \leq 0. \quad (72)$$

In this case, A is Hurwitz. Moreover, the required supply and available storage satisfy

$$V_r(x_0) = x_0^\top K_r x_0, \quad (73)$$

$$V_a(x_0) = x_0^\top K_a x_0, \quad (74)$$

with K_r and K_a the maximum and minimum solution to (72), in the sense that $K_a \leq K \leq K_r$ for any solution K to (72). Finally, we have $K_a > 0$.

Inequality (72) can alternatively be derived from the bounded real lemma ([Scherer, 1990](#), Theorem 2.2), by substituting the gain bound $\gamma = 1$. The core idea of dissipativity-preserving model reduction is to simultaneously diagonalize the matrices K_a and K_r (instead of the observability and controllability Gramians). Namely, noting that these matrices transform under the change of coordinates (64) as $K_a \mapsto R^\top K_a R$ and $K_r \mapsto R^\top K_r R$, it follows from [Lemma 24](#) (with $X = K_a$ and $Y = K_r$) that there exists a change of coordinates such that the contractive system is bounded real balanced.

Definition 29. A realization (A, B, C, D) of a contractive system is called *bounded real balanced* if the required supply and available storage satisfy $K_a = K_r^{-1} = Y$, with Y of the form

$$Y = \text{diag}(v_1 I_{m_1}, v_2 I_{m_2}, \dots, v_\ell I_{m_\ell}), \quad (75)$$

with $\sum_{i=1}^{\ell} m_i = n$ and $v_1 > v_2 > \dots > v_\ell > 0$.

This definition is motivated by the observation that this realization admits a similar interpretation as the balanced realization of Section 2.4.1. In particular, in the R -transformed coordinates, the first state component requires a small amount of energy to reach ($V_r(e_1)$ is small), while much energy can be extracted from it ($V_a(e_1)$ is large).

The method of bounded real balanced truncation amounts to, first, performing the above change of coordinates such that (A, B, C, D) is a bounded real balanced realization and, second, obtaining the reduced-order model by truncation (as in (70)). This reduced-order model is guaranteed to be contractive as stated next.

Theorem 30. Consider the contractive linear system (57) with (A, B) controllable and (A, C) observable. Let (70) be obtained by bounded real balanced truncation of (57). Assume in addition that $\hat{n} = \sum_{i=1}^k m_i$ for some $k < \ell$. Then:

1. (A_{11}, B_1) is controllable, (A_{11}, C_1) is observable, and (A_{11}, B_1, C_1, D) is a bounded real balanced realization;
2. A_{11} is Hurwitz;
3. the reduced-order model (70) is contractive;
4. the error bound

$$\|G - \hat{G}\|_{H_\infty} \leq 2 \sum_{i=k+1}^{\ell} v_i \quad (76)$$

holds, with G and $\hat{G}$ the transfer functions of (57) and (70), respectively.

Remark 31. Similar to the reduction of contractive systems, a balancing-based reduction method for passive systems can be stated.

This method, known as positive real balancing (recall Remark 8) was first proposed in Desai and Pal (1984) and Harshavardhana et al. (1984), see also Gugercin and Antoulas (2004), Phillips, Daniel, and Silveira (2003). In an analogous way to Theorem 28, we have that a (minimal realization of a) system is passive if and only if there exists a positive definite symmetric matrix K such that

$$\begin{bmatrix} A^\top K + KA & KB \\ B^\top K & 0 \end{bmatrix} - \begin{bmatrix} 0 & I \\ C & D \end{bmatrix}^\top \begin{bmatrix} 0 & I \\ I & -2\delta I \end{bmatrix} \begin{bmatrix} 0 & I \\ C & D \end{bmatrix} \preceq 0, \quad (77)$$

with $\delta \geq 0$ as in Definition 6. The required supply V_r and available storage V_a are again given as in (73) and (74), respectively, with K_r and K_a the maximum and minimum solution to (77). Positive real balanced truncation then proceeds in an analogous way as bounded real balanced truncation, guaranteeing the preservation of passivity. We note, however, that passivity does not imply asymptotic stability and, as a result, an error bound on $\|G - \hat{G}\|_{\mathcal{H}_\infty}$ can only be found under additional assumptions, see Gugercin and Antoulas (2004).

We refer to Section 3.5 for an application of a passive system's required supply function and available storage function in the scope of model reduction.

2.4.3. Balanced model reduction for Lur'e-type systems

We now return to the main problem of model reduction of Lur'e-type systems, following the approach of Fig. 1 in which only the linear dynamics is reduced (and the nonlinearity φ is maintained). The results in this section are derived from Besselink, van de Wouw, and Nijmeijer (2013) and Besselink, van de Wouw, and Nijmeijer (2009), see also Reis and Heinkenschloss (2009) for a similar idea.

The main objectives of Problem 13 are the preservation of incremental stability properties and the availability of an error bound. However, in Section 2.3 we have concluded that these properties are guaranteed when the reduced linear dynamics satisfy A4 and A5 of Assumption 16.

When applying balanced truncation to obtain this reduced linear dynamics, we get the following result.

Theorem 32. Consider the Lur'e-type system $\Sigma = I(\Sigma_{\text{lin}}, \varphi)$ in (5) and suppose that A1–A3 in the statement of Theorem 10 hold for some $\gamma_{wv} \geq 0$ and $\mu \geq 0$. Let $\hat{\Sigma}_{\text{lin}}$ in (11a) be obtained by applying balanced truncation to Σ_{lin} , with reduced order $\hat{n} = \sum_{i=1}^k m_i$ for some $k < \ell$ (recall Definition 25). Denote with ϵ_{ji} error bounds such that (40) holds. Furthermore, assume that

$$(\gamma_{wv} + \epsilon_{wv})\mu < 1. \quad (78)$$

Then, the reduced-order Lur'e-type system in (11)

1. is globally uniformly convergent and input-to-state convergent for the class of inputs $\mathcal{L}_\infty^m$;
2. has a bounded incremental $\mathcal{L}_2$ gain;
3. satisfies the error bounds (29) and (30), with

$$\begin{aligned} \epsilon &= \epsilon_{yu} + \frac{\gamma_{yv}\mu}{1 - \gamma_{wv}\mu} \epsilon_{wu} \\ &\quad + \left(\epsilon_{yv} + \frac{\gamma_{yv}\mu}{1 - \gamma_{wv}\mu} \epsilon_{wv} \right) \frac{(\gamma_{wu} + \epsilon_{wu})\mu}{1 - (\gamma_{wv} + \epsilon_{wv})\mu}. \end{aligned} \quad (79)$$

The proof of Theorem 32 follows immediately from earlier results. Namely, the fact that balanced truncation preserves stability (Theorem 27) guarantees that A4 in Assumption 16 holds. Moreover, this guarantees the existence of the bounds in (40). Next, A5 in Assumption 16 is implied by (78). Then, items 1. and 2. follow from Theorem 10, whereas the error bound in item 3. follows from Theorem 17 and Remark 18. This theorem thus shows that Problem 13 can be solved.

The error bound (79) relies on knowledge of the bounds in (40) on the reduction errors for the different input–output pairs of the partitioned linear system Σ_{lin} . It is clear from Theorem 27, however, that these can be bounded as

$$\epsilon_{ji} \leq 2 \sum_{r=k+1}^{\ell} \sigma_r, \quad i \in \{u, v\}, j \in \{y, w\}. \quad (80)$$

Replacing ϵ_{ji} in (79) with this bound leads to an *a priori* error bound for the reduction of Lur'e-type systems.

Remark 33. The statement and result of Theorem 32 relies on the gains γ_{ji} of the high-order linear dynamics and the error bounds ϵ_{ji} of the reduction, as in (40). In case knowledge of the gains $\hat{\gamma}_{ji}$ of the reduced-order linear dynamics is available, the small-gain condition (78) and the error bound (79) can be replaced by $\hat{\gamma}_{wv}\mu < 1$ and (51) (for $\epsilon = \epsilon$), respectively. Generally, this leads to less conservative conditions.

As a special case of Theorem 32, we consider Lur'e-type systems with contractive linear dynamics and incrementally strictly contractive nonlinearity. This leads to the following result.

Theorem 34. Consider the Lur'e-type system $\Sigma = I(\Sigma_{\text{lin}}, \varphi)$ in (5) and assume that the conditions of Corollary 11 hold. Denote by $\mu < 1$ the incremental gain of φ . Let $\hat{\Sigma}_{\text{lin}}$ in (11a) be obtained by applying bounded real balanced truncation to Σ_{lin} , with reduced order $\hat{n} = \sum_{i=1}^k m_i$ for some $k < \ell$ (recall Definition 29). Denote by ϵ_{ji} the error bounds such that (40) holds. Then, the reduced-order Lur'e-type system in (11)

1. is globally uniformly convergent and input-to-state convergent for the class of inputs $\mathcal{L}_\infty^m$;
2. is incrementally contractive;
3. satisfies the error bounds (29) and (30), with

$$\epsilon = \epsilon_{yu} + \frac{\mu}{1 - \mu} \epsilon_{wu} + \left(\epsilon_{yv} + \frac{\mu}{1 - \mu} \epsilon_{wv} \right) \frac{\mu}{1 - \mu}. \quad (81)$$

The proof is similar to that of Theorem 32, noting that $\gamma_{ji} \leq 1$ due to contractivity, and $\hat{\gamma}_{ji} \leq 1$ as a consequence of applying bounded real balancing (recall Theorem 30). The latter also guarantees that Theorem 34 does not require a counterpart of (78).

Remark 35. The reduction approach for Lur'e-type systems in this section relies on the application of balanced truncation to the linear dynamics, whereas the preservation of stability properties and the availability of the error bound crucially rely on incremental stability properties. Interestingly, extensions of balanced truncation to general nonlinear systems, i.e., not specifically in Lur'e-type form, also rely on incremental properties to guarantee relevant stability properties and an error bound, see, e.g., Besselink et al. (2014), Kawano and Scherpen (2017).

Remark 36. The reduction approach for Lur'e-type systems in this and the next section is presented for continuous nonlinearities φ . Still, we conjecture that extensions to particular classes of Lur'e-type systems with set-valued, or even impulsive, nonlinearities are possible. In this case, the nonlinearities should exhibit appropriate monotonicity properties and the linear dynamical sub-system should be passive to at least ensure that the required incremental system properties (which we use as basis for our model reduction approach) are satisfied, see, e.g., Brogliato and Tanwani (2020), van de Wouw and Leine (2010), Yakubovich, Leonov, and Gelig (2004).

2.5. Model reduction by moment matching

Moment matching is a projection-based model reduction technique that approximates the behavior of a high-dimensional LTI system by “matching” key system properties, known as “moments” (Antoulas, 2005a). Moments have been traditionally defined in terms of the value of the transfer function (and possibly its derivatives) of the system, at regular points of the function (called interpolation points) (Antoulas, 2005a). In this classical view, a reduced-order model is such that its transfer function (and possibly its derivatives) takes the same values of the transfer function (and possibly its derivatives) of the system at

the same interpolation points. However, instead of focusing on this approach, in this article, we shift our attention to the *interconnection-based* notion of moments, as introduced in Astolfi (2010), see also (Scarciotti & Astolfi, 2024) for a recent treatise. This notion extends naturally beyond linear systems to encompass nonlinear systems and provides a parameterized family of the reduced-order models for which additional properties can be enforced. In the scope of model reduction for Lur'e-type systems, this parametrization can be exploited to guarantee the preservation of the incremental system properties and to reduce error bounds of the form (29) and (30). Since there is a strong relation between moments of a nonlinear system and its steady-state response, we focus on convergent systems in this section. Namely, the convergence property guarantees the existence of a well-defined steady-state solution. Where possible, we will also present connections to incremental system properties.

We start in Section 2.5.1 by presenting the notion of a moment of a general nonlinear system of the form (12)–(13) and will show its intimate connection with the steady-state system response. Then, in Section 2.5.2, we will turn to the interconnection-based moment matching technique for LTI systems, which is the core technique used for the reduction of the LTI block of the Lur'e-type system (5). Subsequently, the LTI moment matching technique will be extended to provide families of convergence-preserving reduced-order models for the Lur'e-type system in Section 2.5.3. Finally, in Section 2.5.4, we will propose optimization problems to reduce the error bound by searching within the family of the convergence-preserving reduced-order models.

2.5.1. Interconnection-based notion of a (nonlinear) moment

The notion of moment of a *smooth* nonlinear dynamical system was presented in Astolfi (2010), Scarciotti and Astolfi (2017c) and relies on the solution of a (nonlinear) partial differential equation. In Astolfi (2010), Scarciotti and Astolfi (2017c), the existence of such a solution was locally guaranteed by the so-called center manifold theorem (Carr, 1981). The nonlinearities in Lur'e-type systems considered in this article are, however, not necessarily smooth, but only Lipschitz continuous, see, e.g., (10). Therefore, based on Shakib et al. (2023a), we recall the *global* notion of moment for generic, possibly non-smooth, nonlinear systems (12)–(13). These results are based on a global invariance theorem developed in Pavlov et al. (2006). Furthermore, in this section, we recall sufficient conditions for this moment to exist.

Consider a signal generator described by the equations

$$\dot{\tau} = s(\tau), \quad u = \ell(\tau) \quad (82)$$

with $\tau(t) \in \mathbb{R}^{\hat{n}}$, and the mappings $s : \mathbb{R}^{\hat{n}} \rightarrow \mathbb{R}^{\hat{n}}$ and $\ell : \mathbb{R}^{\hat{n}} \rightarrow \mathbb{R}^m$ locally Lipschitz in τ . We assume that the solutions of (82) exist on $\mathbb{R}$, which is the case, e.g., if s is globally Lipschitz in τ . The state dimension $\hat{n}$ of the signal generator plays a crucial role in the model reduction process, as it directly determines the state dimension of the reduced-order model. Consequently, the choice of $\hat{n}$ has a significant impact on the complexity of the reduced model.

The moment of the system at the signal generator (s, ℓ) is related to an invariant manifold of the interconnection between system (12)–(13) and the signal generator (82), defined as follows:

$$\dot{\tau} = s(\tau), \quad \dot{x} = f(x, \ell(\tau)), \quad y = h(x). \quad (83)$$

The notion of moment is defined next.

Definition 37 (Shakib et al., 2023a). Consider the interconnected system (83) and suppose that the pair (s, ℓ) is observable.¹ Suppose there exists a unique function $\pi : \mathbb{R}^{\hat{n}} \rightarrow \mathbb{R}^n : \tau \mapsto \pi(\tau)$ such that the graph

$$\mathcal{M} := \{(\tau, x) : x = \pi(\tau), \tau \in \mathbb{R}^{\hat{n}}\} \quad (84)$$

is invariant with respect to the interconnected system (83). Then, the function $h \circ \pi$ is called the *moment of the system* (12)–(13) at (s, ℓ) .

¹ See Scarciotti and Astolfi (2017c, Definition 2.10) for a definition of observability of the pair (s, ℓ) .

Definition 37 characterizes the moment through the function π in (84). While it can be challenging to prove the existence of π for generic nonlinear systems, it was shown in Pavlov et al. (2006) that the input-to-state convergence property plays a key role in the existence and uniqueness of π . Let us first introduce conditions on the signal generator (82) and the system (12)–(13).

Assumption 38. The signal generator (82) is observable and neutrally stable² and the system (12)–(13) is input-to-state convergent.

Under these assumptions, the mapping π exists and is unique, as presented next.

Lemma 39 (Shakib et al., 2023a). Under Assumption 38, there exists a unique, continuous function π as in Definition 37, such that the graph $\mathcal{M}$ in (84) is invariant with respect to the interconnected system (83). Moreover, for every input $u(t)$ generated by (82), the corresponding steady-state solution of (83) is given by $\bar{x}_u(t) = \pi(\tau(t))$ and is globally uniformly asymptotically stable. Furthermore, if $\pi(\tau)$ is continuously differentiable, then $\pi(\tau)$ solves the system of n partial differential equation

$$\frac{\partial \pi(\tau)}{\partial \tau} s(\tau) = f(\pi(\tau), \ell(\tau)), \quad \tau(t) \in \mathbb{R}^{\hat{n}}. \quad (85)$$

Lemma 39 is based on a global invariance theorem presented in Pavlov et al. (2006). It guarantees the *global* asymptotic stability of the invariant manifold $\mathcal{M}$ in Definition 37. Hence, this invariant manifold is closely related to the steady-state response of the interconnected system (83), namely, the moment $h \circ \pi$ at (s, ℓ) evaluated along a particular trajectory τ coincides with the steady-state output response $\bar{y}_u$ of the interconnected system (83). We will show later that the error bound in the steady-state perspective, i.e., (29), naturally bounds the difference between the moment of the high-order system and the reduced-order model. For the class of Lur'e-type systems (5) (and (11)), steady-state solutions on the manifold $\mathcal{M}$ can be computed efficiently using the so-called mixed-time-frequency simulation algorithm (Pavlov & van de Wouw, 2008; Pavlov et al., 2013).

2.5.2. Time domain moment matching for LTI systems

In this article, we focus on *linear* moment matching of the LTI block of the Lur'e-type system (5). In this section, we recall LTI interconnection-based moment matching as presented in Astolfi (2010) for SISO systems and in Shakib, Scarciotti, Pogromsky, Pavlov, and van de Wouw (2023b) for MIMO systems. In Section 2.5.3, we will adapt this approach to preserve the convergence property of Lur'e-type systems and show that the error bound (29) also bounds the mismatch in the *nonlinear* moment as in Definition 37. Therefore, although we only achieve linear moment matching, we do provide a measure for the mismatch in the nonlinear moment.

Consider the LTI part Σ_{lin} in (5a) described by the real matrices A, B , and C . Recall that this LTI system has $m + r$ inputs and $p + q$ outputs. In the linear case, the signal generator (82) reduces to

$$s(\tau) = S\tau, \quad \ell(\tau) = L\tau, \quad (86)$$

where $S \in \mathbb{R}^{\hat{n} \times \hat{n}}$ and $L \in \mathbb{R}^{(m+r) \times \hat{n}}$ are an observable pair. Provided that the non-resonance condition $\lambda(A) \cap \lambda(S) = \emptyset$ holds, there exists an invariant manifold $\mathcal{M} = \{(x, \tau) \in \mathbb{R}^{n+\hat{n}} : x = \Pi\tau\}$, where $\Pi \in \mathbb{R}^{n \times \hat{n}}$ is the unique solution of the Sylvester equation

$$A\Pi + BL = \Pi S. \quad (87)$$

Based on Π , the moment of the LTI system (5a) at (S, L) is denoted by $C\Pi$, consistent with the nonlinear counterpart in Definition 37. The matrix $C\Pi$ has a well-defined meaning in terms of the transfer function G of the LTI part of system (5). It is similar (by a change of coordinates)

² See Pavlov et al. (2006, Definition 3.1) for a definition of neutral stability.

to the evaluation of the transfer function G at the interpolation points characterized by the eigenvalues of the matrix S and along tangential directions characterized by the columns of the matrix L , see Astolfi (2010), Shakib et al. (2023b) for details.

Let us denote the moment of the LTI part of the reduced-order model (11) at (S, L) by $\hat{C}\hat{\Pi}$, where $\hat{\Pi} \in \mathbb{R}^{\hat{n} \times \hat{n}}$ is the unique solution to the Sylvester equation $\hat{A}\hat{\Pi} + \hat{B}L = \hat{\Pi}S$ with matrices $\hat{A} \in \mathbb{R}^{\hat{n} \times \hat{n}}$ and $\hat{B} \in \mathbb{R}^{\hat{n} \times (m+r)}$. The reduced-order model (11a) is said to achieve *moment matching* at (S, L) if it matches the moment of the system (5a) at (S, L) , i.e., $\hat{C}\hat{\Pi} = C\Pi$. We next recall a family of LTI models that achieve moment matching at (S, L) .

Theorem 40 (Shakib et al., 2023b). *Consider the system (5a) and the observable pair (S, L) in (86). Suppose $\lambda(S) \cap \lambda(A) = \emptyset$. Then, for any $\hat{B} \in \mathbb{R}^{\hat{n} \times (m+r)}$ such that $\lambda(S) \cap \lambda(S - \hat{B}L) = \emptyset$, the model (11a) with matrices*

$$\hat{A} = S - \hat{B}L, \quad \hat{C} = C\Pi, \quad (88)$$

where $\Pi \in \mathbb{R}^{n \times \hat{n}}$ is the unique solution to (87), achieves moment matching at (S, L) .

Theorem 40 provides a family of models that is parameterized in the matrix $\hat{B}$ and corresponds to the choice $\hat{\Pi} = I_{\hat{n}}$. As shown in Shakib, Scariotti, Jungers, et al. (2024, Lemma 2), the family of models (88) contains all models that achieve moment matching up to a similarity transformation. The freedom in $\hat{B}$ can be exploited to solve various problems, such as matching additional moments, placing the poles of the reduced-order model, and assigning a prescribed $\mathcal{L}_2$ gain, see Scariotti and Astolfi (2024) for a recent survey. Note that the condition $\lambda(S) \cap \lambda(S - \hat{B}L) = \emptyset$ is often trivially satisfied, e.g., by the observability of (S, L) , the eigenvalues of $S - \hat{B}L$ can be placed at any arbitrarily chosen location different from $\lambda(S)$.

In contrast to the balanced truncation model reduction approaches for LTI systems (see Section 2.4), moment matching approaches are not equipped with an error bound. In fact, depending on the choice of $\hat{B}$, the reduced-order LTI model might not even be globally exponentially stable. As we detail in the subsequent sections, we exploit the parametrization of $\hat{B}$ for (i) preserving incremental system properties, in particular, convergence, for the reduced-order Lur'e-type model and (ii) reducing error bounds to obtain a more accurate reduced-order model. Importantly, under (i), the preservation of convergence ensures that error bounds remain bounded, enabling their reduction.

2.5.3. Structure-preserving moment matching for Lur'e-type systems

To preserve the LTI structure of Σ_{lin} in (5a), we apply *linear* moment matching with a *linear* signal generator, as in Theorem 40, to Σ_{lin} to obtain a family of reduced-order Lur'e-type models $\hat{\Sigma} = I(\hat{\Sigma}_{\text{lin}}, \varphi)$. However, to ensure the existence and uniqueness of the function π in the definition of the *nonlinear* moment of the Lur'e-type system (5), see Definition 37, we require a counterpart of Assumption 38.

Assumption 41. The pair (S, L) in (86) is observable and the eigenvalues of S are simple and located on the imaginary axis. The Lur'e-type system $\Sigma = I(\Sigma_{\text{lin}}, \varphi)$ in (5) satisfies the conditions A1–A3 in the statement of Theorem 10 for some $\gamma_{wv} \geq 0$ and $\mu \geq 0$.

Let us show that Assumption 41 implies the satisfaction of Assumption 38. Firstly, the *linear* signal generator, as in (86), is neutrally stable since S has simple eigenvalues on the imaginary axis. Secondly, the pair (S, L) is observable by assumption. Finally, the Lur'e-type system is input-to-state convergent by the conditions of Theorem 10. Hence, by application of Lemma 39, the function π in Definition 37 exists and is unique. Consequently, the moment of the nonlinear Lur'e-type system (5) at (S, L) and the unique steady-state response $\bar{y}_u$ of system (5) are well defined.

Let the moment of the Lur'e-type system (5) be denoted by $C_y\pi$, and the moment of the reduced-order Lur'e-type model (11) by $\hat{C}_y\theta$, where

$\theta : \mathbb{R}^{\hat{n}} \rightarrow \mathbb{R}^{\hat{n}}$ plays the same role as π in Definition 37. Model reduction by moment matching seeks reduced-order models that have the *same* moment as the original system (Astolfi, 2010), i.e., $\hat{C}_y\theta$ equals $C_y\pi$. However, achieving moment matching in the context of Fig. 1, where the nonlinearity φ is inherited from the full-order system, presents two challenges, namely:

1. since φ introduces an infinite number of harmonics, achieving exact moment matching (via a linear signal generator) generally requires an infinite number of interpolation points ($\hat{n} = \infty$), as shown in Shakib et al. (2023a);
2. obtaining an explicit analytical expression for the mapping π in Definition 37 is highly complex or even infeasible.

To address these issues, the approaches in Shakib, Scariotti, Jungers, et al. (2024), Shakib et al. (2023a) focus on achieving only *linear* moment matching between Σ_{lin} and $\hat{\Sigma}_{\text{lin}}$. In addition, Shakib, Scariotti, Jungers, et al. (2024), Shakib et al. (2023a) also derive a bound for the mismatch between the *nonlinear* moments of the Lur'e-type systems $I(\Sigma_{\text{lin}}, \varphi)$ and $I(\hat{\Sigma}_{\text{lin}}, \varphi)$, which is then minimized.

This simplification is justified as follows. Under Assumption 41, the signal generator output u is a harmonic signal ($u \in \mathcal{P}_2^m(T)$) with the period T depending on the chosen interpolation points in S . The steady-state output $\bar{y}_u$ of the Lur'e-type system is in $\mathcal{P}_2^p(T)$, with the same period T but typically containing higher harmonics due to φ (Pavlov et al., 2006). It has been shown that the Fourier coefficients of these higher harmonics in $\bar{y}_u$ vanish for large frequencies, see Shakib et al. (2023a, Property 12). Thus, achieving linear moment matching at a finite number of interpolation points, while neglecting sufficiently high harmonics, has a negligible impact on the nonlinear moment.

Applying *linear* moment matching with a *linear* signal generator, as in Theorem 40, to Σ_{lin} , gives a family of reduced-order Lur'e-type models $\hat{\Sigma} = I(\hat{\Sigma}_{\text{lin}}, \varphi)$. Note here that the condition $\lambda(A) \cap \lambda(S) = \emptyset$ in Theorem 40 is trivially satisfied since A is Hurwitz while the $\lambda(S)$ are located on the imaginary axis. This family, parameterized in $\hat{B}$, achieves linear moment matching, i.e., the moment of $\hat{\Sigma}_{\text{lin}}$ matches the moment of Σ_{lin} . To preserve the convergence property for the Lur'e-type reduced-order model $\hat{\Sigma}$, conditions A4–A5 of Assumption 16 should be satisfied for the reduced-order model $\hat{\Sigma}$ by an appropriate selection of the matrix $\hat{B}$. To this extent, we decompose the matrices $\hat{B}$ and $\hat{C}$ of the reduced-order model as follows:

$$\hat{B} = [\hat{B}_u \quad \hat{B}_v], \quad \hat{C} = \begin{bmatrix} \hat{C}_y^\top & \hat{C}_w^\top \end{bmatrix}^\top, \quad (89)$$

where $\hat{B}_u \in \mathbb{R}^{\hat{n} \times m}$, $\hat{B}_v \in \mathbb{R}^{\hat{n} \times r}$, $\hat{C}_y \in \mathbb{R}^{p \times \hat{n}}$, and $\hat{C}_w \in \mathbb{R}^{q \times \hat{n}}$. Next, we encode the conditions A4–A5 of Assumption 16 in the set of models B , defined as

$$B := \left\{ \hat{B} \in \mathbb{R}^{\hat{n} \times (m+r)} \left| \left\| \hat{C}_w (sI - (S - \hat{B}L))^{-1} \hat{B}_v \right\|_{\mathcal{H}_\infty} < \mu^{-1} \right. \right\}. \quad (90)$$

This set has several properties as stated in the next result.

Theorem 42. *Consider the Lur'e-type system $\Sigma = I(\Sigma_{\text{lin}}, \varphi)$ in (5) and the signal generator (86). Suppose Assumption 41 holds. Then, the following statements are true for any order $\hat{n} \in \mathbb{N}$:*

1. Any reduced-order model (11) characterized by $(\hat{A}, \hat{C})$ in (88) and $\hat{B} \in B$
 - (a) is globally uniformly convergent and input-to-state convergent for the class of inputs $\mathcal{L}_\infty^m$;
 - (b) has a bounded incremental $\mathcal{L}_2$ gain;
 - (c) satisfies the error bounds (29) and (30) with $\varepsilon = \varepsilon$ with ε in (51).
2. The set of reduced-order models (11) characterized by $(\hat{A}, \hat{C})$, defined in (88), and $\hat{B} \in B$

- (a) contains all models (up to a similarity transformation) that achieve linear moment matching at (S, L) in the linear block and satisfy the conditions A4–A5 of [Assumption 16](#);
 (b) is non-empty.

The proof of [Theorem 42](#) follows immediately from earlier results and is similar to the proof of [Shakib, Scarciotti, Jungers, et al. \(2024, Theorem 3\)](#). Firstly, item 1. is true by virtue of [Theorem 10](#) after noting that A4–A5 are satisfied for any model in the set $\mathcal{B}$. Moreover, the error bounds (29) and (30) are true by [Theorem 17](#), after noting that the constants ϵ_{ji} in (40) exist and are finite since A and $\hat{A}$ are both Hurwitz matrices.

Secondly, item 2. (a) is true since [Theorem 40](#) parameterized all models that achieve linear moment matching and since the conditions A4–A5 are directly encoded in $\mathcal{B}$. This statement shows that any model that achieves moment matching and preserves input-to-state convergence under the conditions A4–A5 is characterized by the given set of reduced-order models. The last statement of [Theorem 42](#), i.e., item 2. (b), shows that there exists at least one $\hat{B}$ such that the corresponding reduced-order Lur'e-type model satisfies the conditions A4–A5 of [Assumption 16](#). Therefore, there exists a family of models, parameterized by $\hat{B} \in \mathcal{B}$, that solves [Problem 13](#). Moreover, this result is independent of the order $\hat{n}$. In fact, such a specific selection of $\hat{B}$ is given in [Shakib, Scarciotti, Jungers, et al. \(2024, Lemma 3\)](#), which is recalled next with small adaptations.

Lemma 43. Consider the Lur'e-type system (5) and the signal generator (86). Suppose [Assumption 41](#) holds. Denote by $\mathcal{X} > 0_n$ any solution to

$$\mathcal{N}_{\mu-1}(A, B_v, C_w, \mathcal{X}) < 0_{n+q+r}, \quad (91)$$

where $\mathcal{N}_{\mu-1}$ is defined in (1b). Then, the reduced-order model (11) with matrices $(\hat{A}, \hat{C})$ in (88) and

$$\hat{B} = (\Pi^\top \mathcal{X} \Pi)^{-1} \Pi^\top \mathcal{X} B, \quad (92)$$

where Π is the solution to the Sylvester equation (87), satisfies conditions A4–A5 of [Assumption 16](#).

[Lemma 43](#) provides a constructive way of computing a matrix $\hat{B}$ that is in the set $\mathcal{B}$, i.e., that satisfies A4–A5. Note that the linear matrix inequality (91) is feasible for some $\mathcal{X} > 0_n$ by the gain condition A3 of [Theorem 10](#), which is assumed to be satisfied. The found $\hat{B}$ can be seen as a weighted projection³ along Π with a positive-definite weighting matrix $\mathcal{X}$. Finally, we note that under [Assumption 41](#), we have $\mathcal{X} > 0_n$ and Π full column rank, which guarantee that $(\Pi^\top \mathcal{X} \Pi)^{-1}$ in (92) exists. For a detailed proof, we refer to the proof of [Shakib, Scarciotti, Jungers, et al. \(2024, Lemma 3\)](#). Thus, for any order $\hat{n}$, there exists at least one reduced-order Lur'e-type model that achieves linear moment matching at (S, L) and solves [Problem 13](#).

Remark 44. The steady-state error bound (29) also bounds the mismatch between the nonlinear moment of the reduced-order model and the nonlinear moment of the original model, both evaluated along solutions τ of the signal generator, i.e., $C_y \pi(\tau) - \hat{C}_y \theta(\tau)$. Note that the evaluation of the moment $C_y \pi$ along τ is nothing more than the steady-state system response $\bar{y}_u$, since $\bar{x}_u = \pi(\tau)$ at (S, L) , see [Definition 37](#) and [Lemma 39](#). Similarly, we have $\hat{x}_u = \theta(\tau)$ and $\hat{C}_y \theta(\tau) = \hat{y}_u$ at (S, L) . Therefore, we have

$$\|C_y \pi(\tau(t)) - \hat{C}_y \theta(\tau(t))\|_{p_2} = \|\bar{y}_u(t) - \hat{y}_u(t)\|_{p_2}, \quad (93)$$

³ This expression is also reminiscent of the two-sided moment matching formula, see [Scarciotti and Astolfi \(2024, Eq. \(15\)\)](#), and plays a role in dissipativity-preserving moment matching, see [Shakib, Moreschini, and Scarciotti \(2025\)](#).

which, using the steady-state error bound (29), can be bounded by

$$\|C_y \pi(\tau(t)) - \hat{C}_y \theta(\tau(t))\|_{p_2} \leq \epsilon \|u(t)\|_{p_2} = \epsilon \|L \tau(t)\|_{p_2} \quad (94)$$

with ϵ given in [Theorem 42](#). We thus conclude that the bound on the steady-state mismatch $\|\bar{y}_u(t) - \hat{y}_u(t)\|_{p_2}$ also bounds the mismatch between the nonlinear moments evaluated along the signal generator solutions.

Remark 45. The nonlinear moment of the Lur'e-type system (5) is well defined under [Assumption 41](#). Specifically, this assumption requires that the eigenvalues of S are simple and located on the imaginary axis. However, this assumption can be relaxed at the cost of the nonlinear moment no longer being well defined or the connection with the steady-state solution being lost. For example:

1. If S has repeated eigenvalues, the signal generator may produce unbounded responses.
2. If S includes eigenvalues in the complex right-half plane, the signal generator becomes unstable and produces unbounded responses.

In such cases, [Lemma 39](#) does not apply because the function π may either not exist or fail to be unique. Nonetheless, the results of [Theorems 40](#) and [42](#) remain valid as long as S satisfies the condition $\lambda(A) \cap \lambda(S) = \emptyset$. It is important to stress that, under these relaxed conditions, the connection between the nonlinear moment and the steady-state response $\bar{y}_u$ no longer holds, see also ([Astolfi, 2010](#)). Consequently, the analysis provided in [Remark 44](#) is no longer applicable.

Remark 46. Independent of the choice of S and L (and thus $\hat{n}$), [Theorem 42](#) shows that there exists at least one model that achieves linear moment matching and preserves the incremental system property. This is in line with the bounded-real balanced truncation method, which also preserves the incremental system property independent of the order $\hat{n}$ of the reduced-order model, see [Theorem 34](#). However, it is in contrast with the properties obtained by (standard) balanced truncation (see [Theorem 32](#)), where the preservation of incremental system properties has to be checked a priori or a posteriori for a given reduction order.

2.5.4. Error bound minimization using moment matching

[Theorem 42](#) provides a family of reduced-order models that satisfy an error bound. In this error bound, the constant ϵ plays a central role. Furthermore, [Remark 44](#) shows that this error bound also has an interpretation in terms of the matching of the nonlinear moment. In this section, we present methods to minimize ϵ , thus minimizing the error bounds (29) and (30), by searching within the family of models parameterized by $\hat{B} \in \mathcal{B}$.

The error bounds (29) and (30) are characterized through the constant $\epsilon = \epsilon$ with ϵ in (51). In turn, ϵ depends on the constants ϵ_{ji} with $i \in \{u, v\}$ and $j \in \{y, w\}$. For clarity of exposition, we use the simplified notation in [Remark 19](#), where $\epsilon_{ji} \leq \epsilon_{\text{lin}}$ has been used, and we note that the error bounds are proportional to ϵ_{lin} . Recall that ϵ_{lin} is defined as

$$\epsilon_{\text{lin}} = \|G(s) - \hat{G}(s)\|_{\mathcal{H}_\infty}, \quad (95)$$

where G and $\hat{G}$ are the transfer functions of the LTI part of system (5) and reduced-order model (11), respectively. In the scope of moment matching, $\hat{G}$ can be written as follows

$$\hat{G} = \hat{C} (sI - \hat{A})^{-1} \hat{B} = C \Pi (sI - (S - \hat{B}L))^{-1} \hat{B}, \quad (96)$$

where the expressions for $\hat{A}$ and $\hat{C}$ in (88) are used. Therefore, we aim to minimize ϵ_{lin} by minimizing γ , which represents a bound on ϵ_{lin} , while constraining $\hat{B} \in \mathcal{B}$:

$$\min_{\hat{B}, \gamma} \gamma \quad (97a)$$

$$\text{subject to } \hat{B} \in \mathcal{B}. \quad (97b)$$

$$\|G(s) - C\Pi(sI - (S - \hat{B}L))^{-1}\hat{B}\|_{\mathcal{H}_\infty} < \gamma. \quad (97c)$$

We propose two approaches to solve this problem. The first approach directly minimizes the $\mathcal{H}_\infty$ -norm using linear matrix inequalities and is based on [Shakib, Scarciotti, Jungers, et al. \(2024\)](#). The second approach minimizes a proxy for the $\mathcal{H}_\infty$ -norm in the frequency-domain and is based on [Shakib et al. \(2023a\)](#).

A linear matrix inequality (LMI) approach based on [Shakib, Scarciotti, Jungers, et al. \(2024\)](#): The approach proposed by [Shakib, Scarciotti, Jungers, et al. \(2024\)](#) relies on the so-called bounded real lemma ([Scherer, 1990](#)) and directly minimizes (97a). Using the bounded real lemma, for any $\gamma > 0$, the constraint (97c) can equivalently be written as the matrix inequality:

$$\mathcal{N}_\gamma(A_e, B_e, C_e, \mathcal{X}) < 0_{n+\hat{n}+m+r+p+q}, \quad \mathcal{X} > 0_{n+\hat{n}}, \quad (98)$$

where $\mathcal{N}_\gamma$ is defined in (1b) and

$$A_e := \begin{bmatrix} A & 0_{n \times \hat{n}} \\ 0_{\hat{n} \times n} & S - \hat{B}L \end{bmatrix}, \quad B_e := \begin{bmatrix} B \\ \hat{B} \end{bmatrix}, \quad C_e = \begin{bmatrix} C & -C\Pi \end{bmatrix}.$$

The matrix inequality (98) is bilinear in the variables $\mathcal{X}$ and $\hat{B}$ due to products between these two variables.

Using the matrix inequality characterization of constraint (97c) in (98), we present the following set of models:

$$\mathcal{B}_\gamma := \left\{ \hat{B} \in \mathbb{R}^{\hat{n} \times (m+r)} \mid \begin{array}{l} \exists \mathcal{X}_1 > 0, \exists \mathcal{X}_2 > 0 : \\ \mathcal{N}_\gamma(A_e, B_e, C_e, \mathcal{X}_1) < 0, \\ \mathcal{N}_{\mu^{-1}}(S - \hat{B}L, \hat{B}_v, C_w\Pi, \mathcal{X}_2) < 0 \end{array} \right\}. \quad (99)$$

Here, the constraint $\mathcal{N}_{\mu^{-1}}(S - \hat{B}L, \hat{B}_v, C_w\Pi, \mathcal{X}_2) < 0$ together with $\mathcal{X}_2 > 0$ ensures that $\|\hat{C}_w(sI - \hat{A})^{-1}\hat{B}_v\|_{\mathcal{H}_\infty} < \mu^{-1}$ and that $\hat{A}$ is Hurwitz, i.e., satisfaction of the conditions A4–A5 of [Assumption 16](#), and thus (97b). Note that the set $\mathcal{B}_\gamma$ is a subset of $\mathcal{B}$, i.e., $\mathcal{B}_\gamma \subseteq \mathcal{B}$. Therefore, any model from this set achieves linear moment matching, is input-to-state convergent, has a bounded incremental $\mathcal{L}_2$ gain, and respects the error bounds (29) and (30). The constraint $\mathcal{N}_\gamma(A_e, B_e, C_e, \mathcal{X}_1) < 0$ ensures that $\epsilon_{\text{lin}} < \gamma$.

We note that the variables $\mathcal{X}_1, \mathcal{X}_2$ and $\hat{B}$ appear nonlinearly in the matrix inequalities in (99). These matrix inequalities can be written in a so-called primal form $\mathcal{R}_p(\mathcal{X}_1, \mathcal{X}_2, \hat{B}, \gamma) < 0$, where

$$\mathcal{R}_p(\mathcal{X}_1, \mathcal{X}_2, \hat{B}, \gamma) := \mathcal{M}_1(\mathcal{X}_1, \mathcal{X}_2, \gamma) + \text{He}(\mathcal{M}_2(\mathcal{X}_1, \mathcal{X}_2)\mathcal{M}_3^\top(\hat{B})), \quad (100)$$

for some matrix functions $\mathcal{M}_1, \mathcal{M}_2$, and $\mathcal{M}_3$ defined in [Shakib, Scarciotti, Jungers, et al. \(2024, Equation \(23\)\)](#), where $\mathcal{M}_1, \mathcal{M}_2$, and $\mathcal{M}_3$ are linear in their arguments, and where He is defined in (1a). Notice that (100) still contains a product between the variables $\hat{B}$ and $\mathcal{X}_1, \mathcal{X}_2$.

Using Finsler's lemma ([Boyd, Ghaoui, Feron, & Balakrishnan, 1994](#)), the matrix inequality $\mathcal{R}_p < 0$ in (100) can be written as $\mathcal{R}_F(\mathcal{X}_1, \mathcal{X}_2, \mathcal{Q}, \hat{B}, \gamma) < 0$ with $\mathcal{Q} \in \mathbb{R}^{(2n+4\hat{n}) \times (4n+8\hat{n}+2q+2r+m+p)}$ an additional decision variable. This form is called the Finsler's form, where

$$\mathcal{R}_F(\mathcal{X}_1, \mathcal{X}_2, \mathcal{Q}, \hat{B}, \gamma) := \mathcal{M}_4(\mathcal{X}_1, \mathcal{X}_2, \gamma) + \text{He} \left(\begin{bmatrix} \mathcal{M}_3(\hat{B}) \\ -I_{2n+4\hat{n}} \end{bmatrix} \mathcal{Q}^\top \right), \quad (101)$$

and $\mathcal{M}_4$, defined in [Shakib, Scarciotti, Jungers, et al. \(2024, Equation \(25\)\)](#), is a matrix function linear in its arguments. Notice that the Finsler's form (101) does not contain a product between the matrices $\hat{B}$ and $\mathcal{X}_1, \mathcal{X}_2$, but it introduces a product between the variables $\hat{B}$ and $\mathcal{Q}$. The reason for the introduction of the Finsler's form will be explained later, but we note now that, by the Finsler's Lemma ([Boyd et al., 1994](#)), the constraints are equivalent in primal and Finsler's form, i.e., for given $\mathcal{X}_1, \mathcal{X}_2, \hat{B}, \gamma$, the inequality $\mathcal{R}_p(\mathcal{X}_1, \mathcal{X}_2, \hat{B}, \gamma) < 0$ holds if and only if there exists an additional matrix $\mathcal{Q}$ such that $\mathcal{R}_F(\mathcal{X}_1, \mathcal{X}_2, \mathcal{Q}, \hat{B}, \gamma) < 0$.

Optimization problems with bilinear matrix inequality constraints can be solved iteratively using a so-called coordinate-descent algorithm (CDA). In each iteration, a CDA fixes one of the bilinear variables and solves the resulting problem with LMI constraints for the other variables to minimize γ . For example, suppose we are given an initial $\hat{B} \in \mathcal{B}_\gamma$, possibly for a large γ . Then, one can fix $\hat{B}$ in the primal form of the constraints (100) and solve the resulting optimization problem with LMI constraints in the matrix variables $\mathcal{X}_1, \mathcal{X}_2$, using, e.g., a bisection search over γ . Then, $\mathcal{X}_1, \mathcal{X}_2$ can be fixed to its new value and (100) can be solved with the variable $\hat{B}$ appearing linearly. This iterative process can be repeated and guarantees a non-increasing sequence of γ .

Similarly, one can do iterations on the matrix inequalities in the Finsler's form of the constraints (101), where either $\hat{B}$ or $\mathcal{Q}$ is fixed iteratively (notice that $\mathcal{X}_1$ and $\mathcal{X}_2$ are not fixed in any of these iterations). Thanks to the primal and Finsler's form of the constraints being equivalent, ([Shakib, Scarciotti, Jungers, et al., 2024](#)) proposes switching between the primal and Finsler's form when γ can no longer be reduced. This combined CDA is formalized in [Shakib, Scarciotti, Jungers, et al. \(2024, Algorithm 1\)](#). An initial solution, for a possibly large γ , can be found using [Lemma 43](#), where the dimension of the LMIs (91) is proportional to the state dimension n of the high-order system. More efficiently, an initial solution can be computed via ([Shakib, Scarciotti, Jungers, et al., 2024, Lemma 3](#)), where the dimension of the involved LMIs is only proportional to the state dimension $\hat{n}$ of the reduced-order model.

The drawback of this method is that it can (currently) only be applied to relatively small-sized systems (roughly up to a hundred states). This is due to the fact that the dimension of the LMI constraints in (99), in particular (98), is proportional to the state dimension n of the high-order model. We refer to [Shakib, Scarciotti, Jungers, et al. \(2024\)](#) for detailed numerical examples. A preliminary version of this method was presented in [Shakib et al. \(2021\)](#) and focused on the $\mathcal{H}_\infty$ -optimal model reduction of LTI systems.

A frequency-gridding approach based on [Shakib et al. \(2023a\)](#): To address the numerical complexity of solving large-scale LMI problems, we present an alternative, frequency-gridding approach, which is a modification from the approach proposed in [Shakib et al. \(2023a\)](#). The approach we present here minimizes a proxy for $\|G(s) - \hat{G}(s)\|_{\mathcal{H}_\infty}$ over a finite number of frequency points.

We define the set of $R \in \mathbb{N}$ frequencies Ω as follows:

$$\Omega := \{\omega_k\}_{k=1}^R. \quad (102)$$

We do not pose any conditions on the selection of the frequencies $\omega_i \geq 0$, $i = 1, \dots, R$. For example, these can be taken equidistantly spaced or logarithmically spaced over a finite range. Next, we define a cost function

$$J(\hat{B}) = \sum_{k=1}^R \left[\bar{\rho} \left(G(i\omega_k) - C\Pi(i\omega_k I_{\hat{n}} - (S - \hat{B}L))^{-1}\hat{B} \right) \right], \quad (103)$$

which measures the distance between the transfer function G and $\hat{G}$ for a given $\hat{B}$ at the set of frequencies Ω , and where $\bar{\rho}$ is the maximum singular value. The cost function J is a proxy for $\|G(s) - \hat{G}(s)\|_{\mathcal{H}_\infty}$. This cost function can be readily used to define an optimization problem in which $\hat{B}$ is constrained to the set $\mathcal{B}$ in (90). We recall that, after solving such an optimization problem, the results of [Theorem 42](#) hold for any $B \in \mathcal{B}$, e.g., the resulting reduced-order model achieves linear moment matching and preserves the input-to-state convergence property.

The constraints in the set $\mathcal{B}$ in (90) can be characterized through LMIs (by virtue of the bounded real lemma), which only depend on the size of the reduced-order model $\hat{n}$ (not on n). Alternatively, we propose an LMI-free, unconstrained method for numerical efficiency. The conditions in $\mathcal{B}$ in (90), i.e.,

$$\|\hat{G}_{(w,v)}\|_{\mathcal{H}_\infty} = \|\hat{C}_w(sI - (S - \hat{B}L))^{-1}\hat{B}_v\|_{\mathcal{H}_\infty} < \mu^{-1} \quad (104)$$

and $S - \hat{B}L$ is Hurwitz, can be efficiently checked by computing $\|\hat{G}_{(w,v)}\|_{\mathcal{H}_\infty}$, e.g., using ([Bruinsma & Steinbuch, 1990](#)), where we use the

convention that $\|\hat{G}_{(w,v)}\|_{\mathcal{H}_\infty} = \infty$ for $S - \hat{B}L$ not Hurwitz. Furthermore, we propose to use a barrier function to turn the problem into an unconstrained problem. The barrier function is defined as

$$\psi(\hat{B}) = \begin{cases} -\log\left(\mu^{-1} - \|\hat{G}_{(w,v)}\|_{\mathcal{H}_\infty}\right), & \text{if } \|\hat{G}_{(w,v)}\|_{\mathcal{H}_\infty} < \mu^{-1}, \\ +\infty, & \text{else,} \end{cases} \quad (105)$$

which tends to $+\infty$ when $\|\hat{G}_{(w,v)}\|_{\mathcal{H}_\infty}$ tends to μ^{-1} or when $S - \hat{B}L$ is not Hurwitz, thereby rejecting any $\hat{B}$ that is not in $\mathcal{B}$. Finally, we solve the *unconstrained* optimization problem

$$\hat{B} = \arg \min_{B \in \mathbb{R}^{n \times (m+r)}} J(B) + \alpha \psi(B), \quad (106)$$

where $\alpha > 0$ is a hyper parameter which can generally be chosen small, e.g., $\alpha = 10^{-5}$. The initialization method described in the LMI-based approach can also be used to initialize this optimization problem.

2.6. Numerical example

This section examines a mechanical system as a case study for the reduction techniques presented in this article. The system is a chain of N masses connected through a combination of linear and nonlinear elements, with the dynamics incorporating features such as softening and hardening springs. The structure and dynamics of this system make it an ideal candidate to showcase the interplay between linear and nonlinear components, as well as to illustrate the efficacy of model reduction techniques in preserving key dynamic and incremental system properties.

2.6.1. System description

We apply the reduction results of this section to a chain of N mass-spring-damper systems. The first and last carts are connected to the fixed earth through a combination of a linear and nonlinear spring and a linear damper. All carts are connected with adjacent carts through linear springs and dampers. The output y corresponds to the position of the first cart, while the input u applies a force on the first cart, resulting in a collocated input–output configuration. Fig. 3 depicts this system schematically. Note that this example is adapted from Kawano and Scherpen (2015), where all carts were connected to adjacent carts through nonlinear springs.

We consider identical masses, i.e., $m_i = m$ for $i = 1, \dots, N$, while all linear springs have the same stiffness k and all dampers have the same damping constant d . The nonlinear spring attached to the fixed earth and the first cart is a so-called softening spring, which is modeled as a piecewise linear function $\tilde{\varphi}_1 : \mathbb{R} \rightarrow \mathbb{R}$ of the displacement $w_1 = x_1$. The nonlinear spring attached to the last cart and the fixed earth is a so-called hardening spring and is modeled as a piecewise linear function $\tilde{\varphi}_2 : \mathbb{R} \rightarrow \mathbb{R}$ of the displacement $w_2 = x_N$. The functions $\tilde{\varphi}_1$ and $\tilde{\varphi}_2$ are given by:

$$\tilde{v}_1 = \tilde{\varphi}_1(w_1) = \begin{cases} -k_1 \delta_1 & \text{if } w_1 < -\delta_1, \\ k_1 w_1 & \text{if } |w_1| \leq \delta_1, \\ k_1 \delta_1 & \text{if } w_1 > \delta_1, \end{cases} \quad (107)$$

$$\tilde{v}_2 = \tilde{\varphi}_2(w_2) = \begin{cases} k_2(w_2 + \delta_2) & \text{if } w_2 < -\delta_2, \\ 0 & \text{if } |w_2| < \delta_2, \\ k_2(w_2 - \delta_2) & \text{if } w_2 > \delta_2, \end{cases}$$

where k_1 and k_2 are stiffness parameters and δ_1 and δ_2 are non-negative parameters. For $k_1 = k_2 = 4$ and $\delta_1 = \delta_2 = 0.5$, the left subplot of Fig. 4 depicts the nonlinearities $\tilde{\varphi}_1$ and $\tilde{\varphi}_2$. The nonlinearity $\tilde{\varphi} : \mathbb{R}^2 \rightarrow \mathbb{R}^2$, i.e., $q = r = 2$, is defined as

$$\tilde{\varphi}(w) := \begin{bmatrix} \tilde{\varphi}_1(w_1) \\ \tilde{\varphi}_2(w_2) \end{bmatrix}. \quad (108)$$

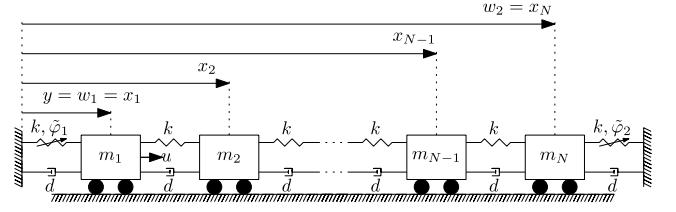


Fig. 3. Schematic overview of a mass-spring-damper system with N degrees of freedom. The function $\tilde{\varphi}_1$ and $\tilde{\varphi}_2$ are nonlinear spring elements.

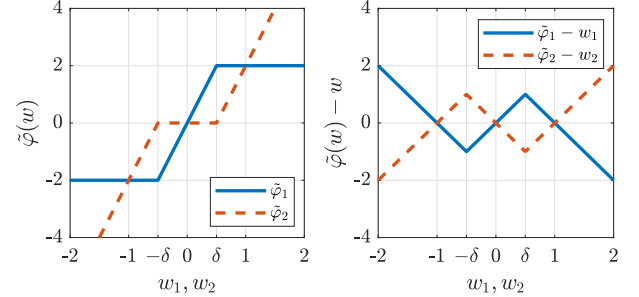


Fig. 4. Left: the function $\tilde{\varphi}_1$ and $\tilde{\varphi}_2$ for $\delta_1 = \delta_2 = 0.5$. Right: the functions after performing the loop transformation.

The dynamics of the mass-spring-damper system can be written as a Lur'e-type system $\Sigma = I(\Sigma_{\text{lin}}, \varphi)$ in (5) with state dimension $n := 2N$. Hereto, let us define the variable $W \in \mathbb{R}^{N \times N}$ as

$$W := \begin{bmatrix} 2 & -1 & & 0 \\ -1 & 2 & -1 & \\ & \ddots & \ddots & \ddots \\ 0 & & -1 & 2 \end{bmatrix}. \quad (109)$$

Using W , we define the stiffness matrix built from the linear springs as $K := kW$ and the damping matrix as $D := dW$. Noting that an external force u acts on the first mass, then, the matrices of (5) read as

$$\tilde{A} = \begin{bmatrix} 0 & I_N \\ -\frac{1}{m}K & -\frac{1}{m}D \end{bmatrix}, \quad B_u = \frac{e_{N+1}}{m}, \quad B_v = \frac{-1}{m} [e_{N+1} \quad e_{2N}],$$

$$C_y = e_1^T, \quad C_w = \begin{bmatrix} e_1^T \\ e_N^T \end{bmatrix},$$

where e_i are canonical vectors of dimension $n = 2N$. The dynamics then read as

$$\Sigma_{\text{lin}} : \begin{cases} \dot{x} = \tilde{A}x + B_u u + B_v \tilde{v}, \\ y = C_y x, \\ w = C_w x, \end{cases} \quad (110a)$$

$$\tilde{v} = \tilde{\varphi}(w). \quad (110b)$$

In the remainder of this example, we take the parameters $N = 100$, $m = 1$, $k = 2$, and $d = 0.8$ for the linear part of model (5), and we take $k_1 = 4$, $k_2 = 4$, $\delta_1 = 0.5$, and $\delta_2 = 0.05$ as the parameters of the nonlinear springs. Consequently, the state dimension $n = 2N = 200$. Note that δ_2 was taken 0.5 in the graphs in Fig. 4 for the sake of visualization, but it is taken 0.05 in the time-domain simulations presented below.

2.6.2. Loop transformation

Fig. 4 illustrates that the slope of both $\tilde{\varphi}_1$ and $\tilde{\varphi}_2$ lies between 0 and 4 (note that these are independent of δ_1 and δ_2). Consequently, the nonlinearity $\tilde{\varphi}$ satisfies the condition (10) for $\mu = 4$. However, condition (10) for $\mu = 4$ would also allow any nonlinearity with slope between -4 and 4 , which is not of interest and would then lead to a conservative result.

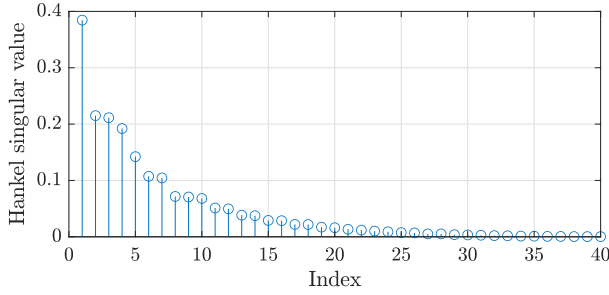


Fig. 5. The 40 largest Hankel singular values of the LTI system Σ_{lin} .

To reduce the conservatism of condition (10) and without changing the dynamics of the model, we introduce a loop transformation so that the slope of the transformed nonlinearity lies within -2 and 2 , i.e., $\mu = 2$ in (10). Hereto, consider the nonlinearity $\varphi : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ defined as

$$\varphi(w) := \tilde{\varphi}(w) - 2w = \begin{bmatrix} \tilde{\varphi}_1(w_1) - 2w_1 \\ \tilde{\varphi}_2(w_2) - 2w_2 \end{bmatrix}, \quad (111)$$

which is depicted in the right subplot of Fig. 4 (for $\delta_2 = 0.5$). It can be confirmed that it satisfies the condition (10) with its slope between -2 and 2 . Changing the nonlinearity also results in a modification in the LTI part of the system. In particular, denote by $\tilde{v} = \tilde{\varphi}(w) - 2w + 2w = \varphi(w) + 2C_w x$. Now, substituting $\tilde{v}$ into the LTI dynamics (110a) results in the system (5) with $A = (\tilde{A} + 2B_v C_w)$. This loop transformation can be interpreted as shifting a part of the incremental slope from the nonlinearity to the linear part of the system, thus tightening the slope bounds of the nonlinear feedback and enabling a less conservative application of Theorem 10. We refer the interested reader to Khalil (1996) for more details on loop transformations.

It can be verified that the transformed system satisfies conditions A1–A3 of Theorem 10 for $\gamma_{uv} = 0.43$ and $\mu = 2$. Hence, the transformed system is globally uniformly convergent and input-to-state convergent, see Definitions 2 and 3, respectively. Moreover, it has a bounded incremental $\mathcal{L}_2$ gain. The goal is to preserve all these *incremental system properties* during the reduction of $\Sigma = \mathcal{I}(\Sigma_{\text{lin}}, \varphi)$.

2.6.3. Linear model reduction

As already mentioned, the state dimension n of Σ_{lin} is 200. We apply balanced truncation and time domain moment matching techniques to reduce this dimension to 11, i.e., $\hat{n} = 11$. The model obtained through balanced truncation, as described in Section 2.4, is denoted by $\hat{\Sigma}_{\text{lin,BT}}$. The model obtained through moment matching, as described in Section 2.5, is denoted by $\hat{\Sigma}_{\text{lin,MM}}$.

Balanced truncation: The standard balancing method of Section 2.4.1 is applied to obtain $\hat{\Sigma}_{\text{lin,BT}}$. The 40 largest Hankel singular values are depicted in Fig. 5, where it can be observed that the Hankel singular values decay to zero quickly. For $\hat{n} = 11$, the a priori H_∞ -norm error bound is found to be $\|G - \hat{G}\|_{H_\infty} \leq 0.80$, see item 3 of Theorem 27. However, it can be verified that with $\gamma_{uv} = 0.43$ and $\mu = 2$, the small-gain condition (78) of Theorem 32 is not satisfied. Hence, the preservation of incremental system properties cannot be guaranteed a priori. In fact, using the a priori error bound derived from the Hankel singular values, incremental system properties can only be ensured for $\hat{n} \geq 26$, where $\|G - \hat{G}\|_{H_\infty} \leq 0.06$. For $\hat{n} < 26$, the preservation of incremental system properties is not guaranteed a priori but can still be verified a posteriori, as discussed in Remark 33. For the purposes of comparison with the moment-matching, as described next, we reduce the system dimension to $\hat{n} = 11$, and verify the incremental system properties a posteriori.

Moment matching: The moment matching technique for model reduction requires several user choices. Firstly, we define the matrix

S in (86) as $S = \text{diag}(0, \Psi(\omega_1), \dots, \Psi(\omega_5))$, where

$$\Psi(\omega) := \begin{bmatrix} 0 & -\omega \\ \omega & 0 \end{bmatrix} \quad (112)$$

is a matrix with eigenvalues $\pm i\omega$ on the imaginary axis. We select the interpolation frequencies (in rad/sec) as follows:

$$[\omega_1 \quad \dots \quad \omega_5] = [0.044 \quad 0.0876 \quad 0.1 \quad 0.131 \quad 10]. \quad (113)$$

The frequencies correspond to the zero frequency, the frequency of the first three resonance peaks in the transfer function from u to y , and the frequencies 0.1 and 10 rad/sec. The choice of the matrix $L \in \mathbb{R}^{3 \times 11}$ in (86) can influence the accuracy of the resulting reduced-order model. The matrix L is related to the so-called tangential directions, which are non-trivial to select. In this example (for experimental purposes only) we took the following approach. Ten realizations of L were generated, i.e., $L^{[1]}, \dots, L^{[10]}$, by drawing its elements from a zero-mean normal distribution with unit variance such that each pair $(S, L^{[k]})$, $k = 1, \dots, 10$, is observable. Hence, each pair $(S, L^{[k]})$, $k = 1, \dots, 10$, meets Assumption 41.

Having the system and the pairs $(S, L^{[k]})$, $k = 1, \dots, 10$, at hand, we solve the Sylvester equation (87) for the matrix $\Pi \in \mathbb{R}^{n \times \hat{n}}$, resulting in $\Pi^{[k]}$, $k = 1, \dots, 10$, which are subsequently used to define the sets of models $\mathcal{B}^{[k]}$ as per Theorem 42. Each set contains all models that achieve linear moment matching at $(S, L^{[k]})$ and satisfy the conditions A4–A5 of Assumption 16, i.e., these models preserve the incremental system properties of the high-order system. Next, we compute an initial matrix $\hat{B}^{[k]}$, $k = 1, \dots, 10$, using Lemma 43. Finally, for each $k \in \{1, \dots, 10\}$, we compute ϵ_{lin} in (95), resulting in $\epsilon_{\text{lin}}^{[k]}$.

We select the initial reduced-order model corresponding to the smallest $\epsilon_{\text{lin}}^{[k]}$, $k \in \{1, \dots, 10\}$, and drop the superscript $[k]$ from here onward. The corresponding $\hat{B}$ acts as an initial solution for the frequency-gridding optimization method detailed in Section 2.5.4. Specifically, we select the frequency grid Ω as a logarithmically spaced array of 1000 frequencies between 0.01 rad/sec and 10 rad/sec, and we select $\alpha = 10^{-5}$ in (106). To solve the unconstrained optimization problem (106), we employ MATLAB's `fminunc` command, which is based on the quasi-Newton optimization approach (Shanno, 1970).

Results: The frequency-response functions (FRF) of the obtained linear models are depicted in Fig. 6, alongside with that of the LTI part of the original system, i.e., Σ_{lin} . It can be observed that both reduced-order models capture the FRF of the system accurately between the first and third resonance peaks. Beyond this, the balanced truncation model typically offers a more precise fit at higher frequencies, whereas at lower frequencies, the moment matching model proves more accurate. The a posteriori error bounds (29) and (30) depend on the constant ϵ , for which, for the sake of simplicity, we take the right-hand side of (56). The model obtained through balanced truncation has an associated $\epsilon = 4.00$. The initial moment matching reduced-order model has an associated $\epsilon = 3.50$ with $\hat{B}$ obtained through Lemma 43. Then, by performing the optimization in the moment matching approach, the corresponding ϵ is reduced to 3.16, showing the efficacy of the optimization problem in (106). The value of $\hat{\gamma}_{uv}$ for the reduced-order models is 0.44 and 0.43 for $\hat{\Sigma}_{\text{lin,BT}}$ and $\hat{\Sigma}_{\text{lin,MM}}$, respectively. Therefore, it can be verified that both models preserve the incremental system properties, e.g., the input-to-state convergence property, since they satisfy conditions A4–A5 of Assumption 16.

2.6.4. Time-domain simulations

We perform numerical time integration simulations to assess the error between the time responses of the reduced-order models and the time responses of the original system. Consider the input signal depicted in the top plot of Fig. 7, which is a sum of two square wave signals. Such block-wave inputs generally excite the dynamics over a wide frequency range.

Starting from zero initial conditions, the responses of the system and the reduced-order models are depicted in the middle plot of Fig.

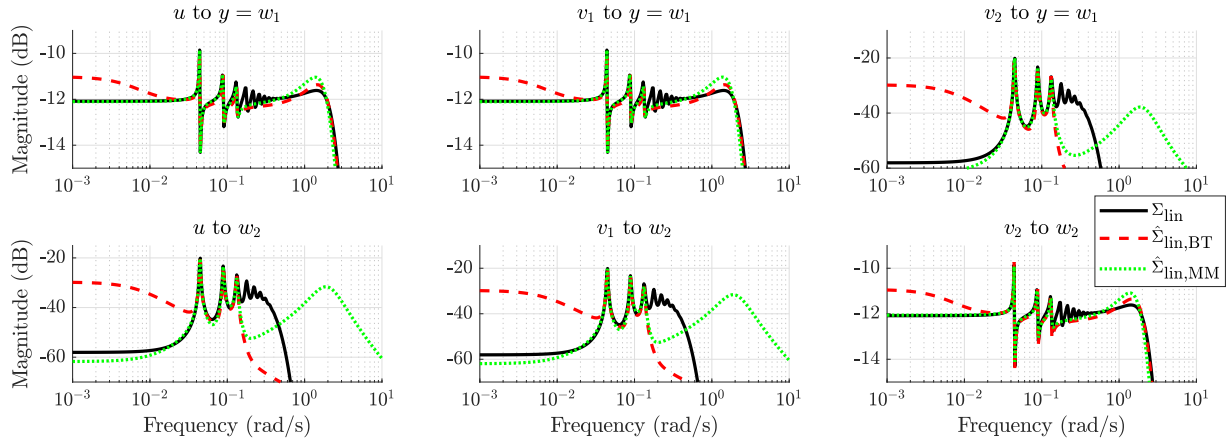


Fig. 6. Bode magnitude plot of the dynamics Σ_{lin} (solid black), $\hat{\Sigma}_{\text{lin,BT}}$ (dashed red), and $\hat{\Sigma}_{\text{lin,MM}}$ (dotted green). Note that the limits on the vertical axis are different in all plots. Furthermore, the transfer functions from u to y and from u to w_1 coincide as $y = w_1$. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

7. The errors between the full-order model and the two reduced-order model responses are depicted in the bottom plot of Fig. 7. It can be observed that both reduced-order models yield an error that is roughly one order of magnitude smaller than the signal y . Moreover, the model obtained by balanced truncation generally offers a slightly more accurate transient response compared to the model obtained by moment matching, see also Fig. 8 for a zoomed-in comparison. This is thanks to a more accurate high-frequency match of the FRFs. However, the reduced-order model obtained by balanced truncation requires more time to converge to the static constant value of y after a step and it has a larger error in this region, see Fig. 8. The latter is due to a larger mismatch at low frequencies in the FRFs.

Finally, we consider the error bound in (30). Recall that ϵ is 4.00 for the model obtained by balanced truncation and 3.16 for the model obtained by moment matching. Numerically integrating the input u depicted in Fig. 8 over $\mathcal{T} = [0, 300]$ results in $\|u\|_{\mathcal{L}_2} = 38.73$. Furthermore, numerically integrating the errors $y - \hat{y}_{\text{BT}}$ and $y - \hat{y}_{\text{MM}}$ over $\mathcal{T}$ results in $\|y - \hat{y}_{\text{BT}}\|_{\mathcal{L}_2} = 0.21$ and $\|y - \hat{y}_{\text{MM}}\|_{\mathcal{L}_2} = 0.19$, respectively. Therefore, we can observe that the error bound given in (30) holds, that is, $\|y - \hat{y}\|_{\mathcal{L}_2} \leq \epsilon \|u\|_{\mathcal{L}_2}$, whether $\hat{y}$ refers to $\hat{y}_{\text{BT}}$ or $\hat{y}_{\text{MM}}$. Similar to the observations in Besselink, van de Wouw, and Nijmeijer (2013), Shakib, Scariotti, Jungers, et al. (2024), Shakib et al. (2023a), these error bounds tend to be conservative due to various approximation steps taken in their derivation. However, as indicated in Shakib, Scariotti, Jungers, et al. (2024), Shakib et al. (2023a), reducing the error bounds generally results in a corresponding decrease in the discrepancy between the outputs of the full-order model and the outputs of the reduced-order model.

2.7. Summary

In this section, the model reduction problem for nonlinear systems with incremental system properties has been motivated, and a framework has been presented to derive reduced-order models that preserve these properties while ensuring interpretability, robustness, and accuracy. Specifically, the focus has been on Lur'e-type systems, which combine LTI dynamics with static nonlinearities, as these widely occur in engineering and science and benefit from established analysis and design tools. By leveraging the incremental system properties of the original system, such as convergence and incremental $\mathcal{L}_2$ -gain, we demonstrated how to derive a priori error bounds between the responses of the full-order model and the reduced-order models. Furthermore, adapted versions of balanced truncation and moment matching are presented which ensure that the reduced-order models

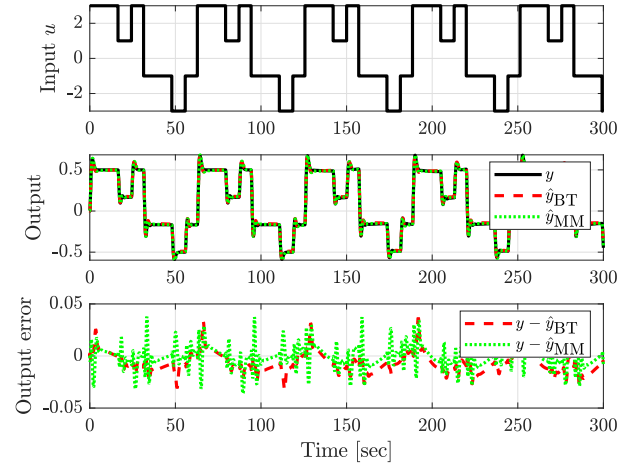


Fig. 7. Top: the input signal u as a function of time. Middle: the output response of the original system (solid black) and of the reduced-order models obtained by balanced truncation (dashed red) and by moment matching (dotted green). Bottom: the error between the output response of the original system and of the reduced-order models obtained by balanced truncation (dashed red) and by moment matching (dotted green). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

preserve the interconnection structure and incremental system properties of the original, full-order model. Numerical examples are provided to illustrate the accuracy guarantees of these structure-preserving reduced-order models.

3. Model reduction of interconnected, linear dynamical systems

3.1. Background and motivation

The overall dynamic behavior of many complex engineering systems results from the interaction of multiple sub-systems interconnected through (physical or controlled) feedback loops. In model-based engineering, interconnected system models arise from the modular development of highly complex systems in the field of structural or thermal dynamics (or more generally multi-physics systems) and are common in application domains such as aerospace, robotics, and high-tech mechatronics and manufacturing equipment.

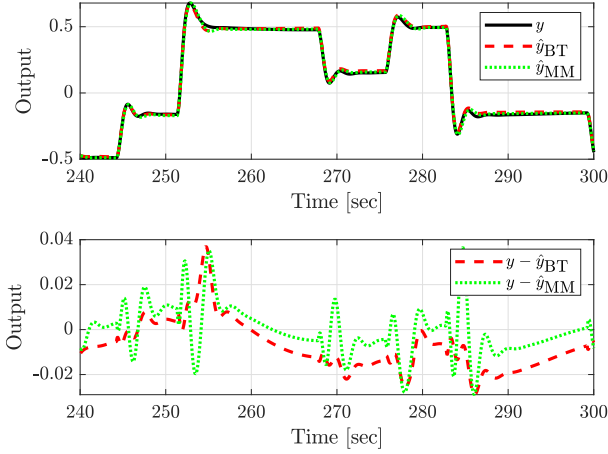


Fig. 8. Top: the output response of the full-order model (solid black) and of the reduced-order models obtained by balanced truncation (dashed red) and by moment matching (dotted green). This figure depicts the last 60 s of the middle plot of Fig. 7. Bottom: the error between the output response of the original system and of the reduced-order models obtained by balanced truncation (dashed red) and by moment matching (dotted green). This figure depicts the last 60 s of the bottom plot of Fig. 7. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

In this scope, linear models are commonly used to describe the dynamics of such systems. When the number of sub-systems grows, the complexity of the overall model increases, making analysis, simulation, and real-time diagnostics and control challenging or even impossible. This gives rise to the need for model reduction techniques tailored to linear interconnected systems, which can simplify the model while retaining essential system dynamics. In addition, key system properties, such as stability and passivity, are often required to be preserved in the reduced-order model. Moreover, in the models that are used in model-based engineering workflows, it is typically required that the interconnection structure is preserved and that model reduction can be pursued in a modular fashion.

Unlike single system models, the coupling and feedback interaction between multiple sub-system models must be handled with care during the reduction process. Direct application of standard model reduction techniques on individual subsystems, such as balanced truncation (see Section 2.4) or moment matching (see Section 2.5), can result in instability or model accuracy degradation in the overall reduced-order system model. Such detrimental effects can be caused by the fact that these methods typically do not account for the dynamics that arise from the interactions between sub-systems. Therefore, model reduction techniques for interconnections of linear systems are needed that explicitly take into account the interconnected nature of these systems.

3.2. System class and problem setting

We represent models of interconnected systems by linear sub-system models that are interconnected through their inputs and outputs. Specifically, in this general framework, each of the k sub-systems is represented by the following equation

$$v^j = G^j(s)w^j, \quad (114)$$

for $j = 1, 2, \dots, k$, where w^j and v^j are the m_j inputs and p_j outputs of the j th sub-system, respectively, and G^j is the corresponding transfer function.

To enable a systematic approach for connecting the sub-systems, these k transfer functions are collected in a block diagonal transfer function, given by

$$G_B(s) = \text{diag}(G^1(s), \dots, G^k(s)). \quad (115)$$

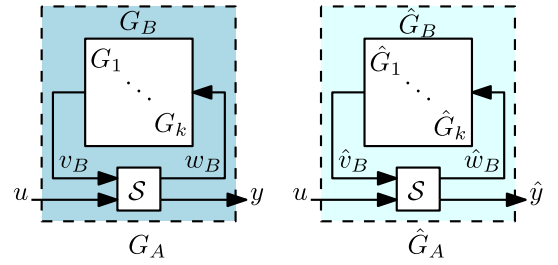


Fig. 9. Block diagram representations of, on the left, the high-order interconnected system G_A with high-order sub-system models $G^1, \dots, G^k$, and the interconnection matrix S , and, on the right, the reduced-order interconnected system $\hat{G}_A$ with reduced-order sub-system models $\hat{G}^1, \dots, \hat{G}^k$, and the same interconnection matrix S (Janssen, Besselink, Fey, & van de Wouw, 2024a).

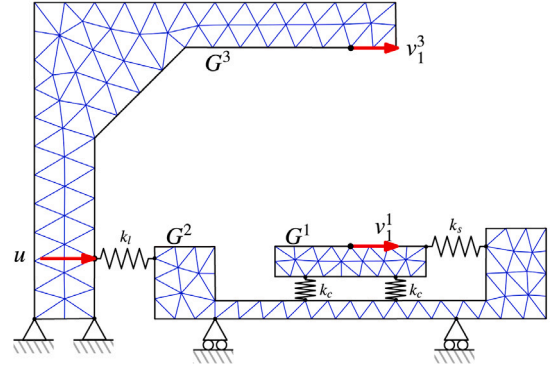


Fig. 10. Schematic illustration of the 2D wafer stage model.

The interconnection structure now relates the inputs and outputs of the sub-systems, denoted by $w_B = [(w^1)^T \dots (w^k)^T]^T$ and $v_B = [(v^1)^T \dots (v^k)^T]^T$, respectively, to m external inputs u and p external outputs y . Particularly, the interconnection structure S is given by

$$\begin{bmatrix} w_B \\ y \end{bmatrix} = S \begin{bmatrix} v_B \\ u \end{bmatrix}, \quad \text{with } S = \begin{bmatrix} S_{BB} & S_{BA} \\ S_{AB} & 0 \end{bmatrix}. \quad (116)$$

Here, we refer to S_{BB} as the interface matrix. Furthermore, the external inputs and outputs are connected to the sub-systems via matrices S_{BA} and S_{AB} , respectively.

With (114)–(116), as illustrated in the left panel of Fig. 9, we can define the system's high-order transfer function from u to y as

$$G_A(s) := F_u(S, G_B) \quad (117a)$$

$$:= S_{AB} G_B(s) (I - S_{BB} G_B(s))^{-1} S_{BA}, \quad (117b)$$

where F_u represents the upper linear fractional transformation (LFT), e.g., Zhou and Doyle (1998, Definition 9.1).

Illustrative example: 2D wafer stage. Throughout this section, we consider a motivating (and illustrative) industrial use case, involving a 2D wafer stage model, as visualized schematically in Fig. 10. A wafer stage is a mechatronic motion stage in lithography equipment for chip manufacturing. Here, we consider the structural (motion and vibrational) dynamics of this system. It consists of three components: the wafer (G^1) in the middle, the stage (G^2) on the bottom right, and the illuminator frame (G^3 , illuminator for short) on the left. In the model, the illuminator is fixed at its two bottom corners in both the horizontal and vertical directions, whereas the stage is fixed at two points in only the vertical direction, such that it can translate freely in the horizontal direction. The wafer is not connected to the fixed world.

We consider a disturbance attenuation problem, in which the horizontal external input force u [N] is a disturbance at the illuminator (see

Table 1Number of states, inputs m_j and outputs p_j per sub-system.

Symbol	Name	Order	m_j	p_j
G^1	Wafer	336	3	4
G^2	Stage	704	4	4
G^3	Illuminator	936	1	2
G_A	Interconnected system	1976	1	1

Fig. 10) and the external output y [m] is the horizontal misalignment (overlay) between wafer and illuminator $v_1^3 - v_1^1$. All interconnections are considered linear stiffnesses, representing the vertical suspension between the wafer and the stage (k_c), the short-stroke (k_s), and the long-stroke actuation (k_l). The structural dynamics of the components (sub-systems) themselves are modeled using triangular, three-node, linear elements, resulting in the model orders indicated in Table 1.

For this benchmark model, the interconnection structure S , as in (116), is given as

$$S = \begin{bmatrix} 0 & -k_s & 0 & 0 & k_s & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -k_c & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -k_c & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & k_s & 0 & 0 & -k_s & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -k_l & 0 & 0 & 0 & k_l & 0 & 0 \\ 0 & 0 & k_c & 0 & 0 & 0 & -k_c & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & k_c & 0 & 0 & 0 & -k_c & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & k_l & 0 & 0 & 0 & -k_l & 1 & 0 \\ -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \end{bmatrix}, \quad (118)$$

where $k_s = 10^{13}$ [N/m], $k_l = k_c = 10^{11}$ [N/m] are the short-stroke, long-stroke, and suspension spring constants, respectively (see also Fig. 10). For this application, it is crucial to enable model order reduction in a modular fashion, since component (sub-system) designs and models are governed by different, independent design teams.

In the model reduction of interconnected systems, given a high-order interconnected system model G_A , the aim is to compute a reduced-order interconnected system model $\hat{G}_A$ with a significantly reduced complexity (order). Herein, the accuracy of the transfer function of the reduced-order interconnected system from u to y is crucial, and is typically assessed by comparing the frequency response functions of the reduced-order interconnected system to those of the high-order system model.

Additional requirements for the reduction of the interconnected system are as follows:

1. **Modularity:** In many cases, the Reduced-Order Models (ROMs) of systems consisting of interconnected sub-systems should preserve the modular structure of the original model, in which the distinct sub-system ROMs, and their interpretable inputs and outputs, are still present in the ROM.
2. **Stability preservation:** In many cases, it is important that the stability of the interconnected system is preserved by the reduction. If that is warranted, we speak of a stability-preserving approach.
3. **Passivity preservation:** Passivity of the sub-system models can be essential in the design of robust controllers. In case the ROMs are used for passivity-based control, it is thus essential to retain the passivity of the sub-system models.
4. **Predictable accuracy of system ROM:** Reducing the complexity of (interconnected) models unavoidably leads to a loss of accuracy of the model. Therefore, there is a need for reduction methods that provide error bounds for the model reduction of interconnected systems.
5. **Efficient reduction:** Finally, as interconnected systems can lead to extremely high-order models, the reduction method typically needs to be computationally efficient enough to ensure the computational feasibility of the reduction.

Several approaches for model reduction of linear interconnected systems exist. One approach is to first combine all the linear sub-system models and then reduce the interconnected system as a whole, which in general leads to an accurate reduced system model. However, this approach is not structure-preserving and computationally inefficient (and potentially infeasible if the overall system order is too large).

Therefore, sub-system reduction, a computationally efficient approach in which the sub-systems are individually reduced before they are combined into a reduced-order interconnected system, is introduced next. With sub-system reduction, each of the k sub-systems G^j is reduced to obtain

$$\hat{v}^j = \hat{G}^j(s) \hat{u}^j, \quad (119)$$

for $j = 1, 2, \dots, k$, and collected in a block-diagonal transfer function, given by $\hat{G}_B(s) = \text{diag}(\hat{G}^1(s), \dots, \hat{G}^k(s))$. Then, the sub-system ROMs are connected using the same interconnection structure S defined in (116) to obtain the interconnected system ROM as follows:

$$\hat{G}_A(s) := S_{AB} \hat{G}_B(s) (I - S_{BB} \hat{G}_B(s))^{-1} S_{BA}, \quad (120)$$

as illustrated in the right panel of Fig. 9. By reducing the sub-system models individually, this approach is modular and preserves the structure of the interconnected system. However, reducing the sub-system models will typically introduce an error between the high-order and reduced-order sub-system models. When connecting the reduced sub-system models, it is difficult to predict how sub-system reduction errors propagate to errors in the total reduced-order, interconnected system. In addition, this might adversely affect the stability of the reduced-order, interconnected system.

In the following section, we will address requirements 1, 2, 4, and 5 by introducing a model order reduction framework based on sub-system reduction that is not only modular and computationally efficient, but also accurate (in a predictable manner) and stability preserving. The preservation of passivity, related to requirement 3, will be considered in Section 3.5. The subsequent model order reduction framework adopts a novel approach for the analysis of how sub-system reduction errors are related to reduction errors of the interconnected system. This will lead to a quantitative framework for the distribution of accuracy requirements from the system level to the sub-system level and vice versa as a basis for sub-system reduction.

3.3. Distributing accuracy requirements from system level to sub-system level

To predict and ensure the accuracy of linear interconnected system ROMs, we introduce the following specification.

Definition 47. For a high-order interconnected system transfer function $G_A(s)$ as in (117), let $\mathcal{E}_A$ denote the accuracy specification

$$\mathcal{E}_A := \left\{ \hat{G}_A \mid \|V_A(\hat{G}_A - G_A)W_A\|_{\mathcal{H}_\infty} < 1 \right\}. \quad (121)$$

We say that the interconnected system ROM $\hat{G}_A$ satisfies this specification, if $\hat{G}_A \in \mathcal{E}_A$. Furthermore, $\hat{G}_A \in \mathcal{E}_A$ implies that $\hat{G}_A$ is stable. In (121), $V_A(s)$ and $W_A(s)$ are user-selected weighting functions, which are taken to be stable proper transfer functions, with stable proper inverses.

Furthermore, we introduce a similar definition for a sub-system accuracy specification.

Definition 48. For the high-order sub-system transfer functions $G^j(s)$, let $\mathcal{E}^j$ denote the accuracy specification

$$\mathcal{E}^j := \left\{ \hat{G}^j \mid \|(W^j)^{-1}(\hat{G}^j - G^j)(V^j)^{-1}\|_{\mathcal{H}_\infty} \leq 1 \right\}. \quad (122)$$

We say that the sub-system ROM $\hat{G}^j$, $j \in \{1, 2, \dots, k\}$ satisfies this specification if $\hat{G}^j \in \mathcal{E}^j$. Furthermore, $\hat{G}^j \in \mathcal{E}^j$ implies that $\hat{G}^j$ is stable.

In (122), $V^j(s)$ and $W^j(s)$ are user-selected weighting functions, which are taken to be stable proper transfer functions, with stable proper inverses.

These system and sub-system ROM specifications impose a constraint on the accuracy of the models, by constraining the $\mathcal{H}_\infty$ -norm of a frequency-weighted model. The weighting in these specifications allows for flexibility in the requirements, which will be demonstrated later on in this section. However, first, we will introduce an approach that allows us to relate the sub-system and system accuracy requirements to each other. To achieve this, as follows from the robust control-based framework for model reduction of interconnected systems introduced in Janssen et al. (2024a), we introduce the transfer function of the so-called nominal system

$$N(s) := \begin{bmatrix} S_{BB}(I - G_B(s)S_{BB})^{-1} & (I - S_{BB}G_B(s))^{-1}S_{BA} \\ S_{AB}(I - G_B(s)S_{BB})^{-1} & 0 \end{bmatrix}, \quad (123)$$

and weighting functions

$$V(s) = \text{diag}(V^1(s), \dots, V^k(s), V_A(s)), \quad (124a)$$

$$W(s) = \text{diag}(W^1(s), \dots, W^k(s), W_A(s)). \quad (124b)$$

With these elements in place, we will introduce a methodology to relate sub-system accuracy specifications (122) to the accuracy specification of the interconnected system model (121). Specifically, the following theorem, introduced in Janssen, Besselink, Fey and van de Wouw (2023), addresses this issue.

Theorem 49. Consider the nominal system in (123) and the accuracy requirement-related weighting transfer function matrices in (124a) and (124b). Given

$$\begin{aligned} \mathbf{D} &:= \left\{ (D_\ell, D_r) \mid d^1, \dots, d^k, d_A \in \mathbb{R}_+, \right. \\ &\quad D_\ell = \text{diag}(d^1 I_{m_1}, \dots, d^k I_{m_k}, d_A I_p), \\ &\quad \left. D_r = \text{diag}(d^1 I_{p_1}, \dots, d^k I_{p_k}, d_A I_m) \right\}, \end{aligned} \quad (125)$$

if there exist

$$(D_\ell, D_r) \in \mathbf{D} \quad (126)$$

such that

$$\begin{bmatrix} W^{-2}(i\omega)D_r^{-1} & N^*(i\omega) \\ N(i\omega) & V^{-2}(i\omega)D_\ell \end{bmatrix} > 0 \quad (127)$$

holds for all $\omega \in \mathbb{R}$, then, satisfaction of the sub-system ROM accuracy requirements, i.e., $\hat{G}^j \in \mathcal{E}^j$ for all $j = 1, 2, \dots, k$, implies stability and satisfaction of the system accuracy requirements, i.e., $\hat{G}_A \in \mathcal{E}_A$.

With Theorem 49, a distribution of sub-system ROM accuracy specifications $\mathcal{E}^1, \dots, \mathcal{E}^k$ can then be found by maximizing

$$\text{tr}[V^{-2}(i\omega)] + \text{tr}[W^{-2}(i\omega)], \quad (128)$$

subject to (126) and (127) for all $\omega \in \mathbb{R}$. Here, the optimization is carried out over the parameters $V^{-2}(i\omega)$ and $W^{-2}(i\omega)$, defining the ROM accuracy bounds which can be trivially linearized for both (128) and (127) using parameter substitution. The matrix operator $|\cdot|$ in (128) denotes the element-wise magnitudes of the complex matrix. Note that in this optimization problem, $V(i\omega)$ and $W(i\omega)$ contain both the user-selected weighting functions that define the interconnected system ROM accuracy specification $V_A(i\omega)$ and $W_A(i\omega)$, respectively, as defined in (121), as well as the optimization parameters that define the sub-system ROM accuracy specifications $V^j(i\omega)$ and $W^j(i\omega)$, respectively, as defined in (122). Furthermore, the other optimization parameters are $d^1, \dots, d^k, d_A$, as defined in (125) while the model of the frequency-dependent high-order interconnected system $N(i\omega)$, as defined in (123), is fixed. For further discussion and implementation details, see Janssen, Besselink, Fey and van de Wouw (2023).

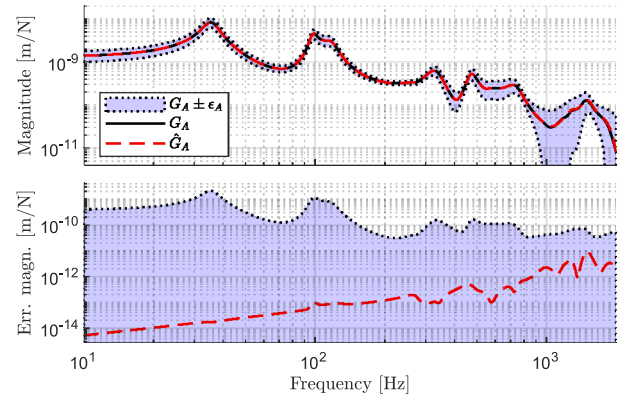


Fig. 11. Top plot: the magnitudes of the frequency responses of the interconnected system model G_A (black line) and reduced-order approximation $\hat{G}_A$ (dashed red line). Bottom plot: the magnitude of the error system $\hat{G}_A - G_A$. In both plots, the ROM system accuracy specifications $\mathcal{E}_A$ is indicated by the blue area. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Then, the obtained sub-system accuracy specifications $\mathcal{E}^1, \dots, \mathcal{E}^k$ can be used to individually reduce the sub-system models $G^1, \dots, G^k$ and obtain sub-system ROMs $\hat{G}^1, \dots, \hat{G}^k$ for which $\hat{G}^j \in \mathcal{E}^j$ is satisfied for all $j \in \{1, \dots, k\}$. Theorem 49 now guarantees that the interconnected system ROM, which follows from assembling the sub-system ROMs, is stable and satisfies the required accuracy requirement, i.e., $\hat{G}_A \in \mathcal{E}_A$.

Remark 50. Note that Theorem 49 provides a sufficient condition to guarantee stability and satisfaction of the system-level accuracy requirement. Consequently, failure to satisfy the constraint in (127) does not necessarily imply that the reduced interconnected system is unstable or inaccurate. Therefore, the condition may be conservative. In practice, however, as demonstrated in Janssen et al. (2024a), the conservativeness of the proposed approach is significantly lower than that of other existing methods.

Illustrative example: 2D wafer stage. The 2D wafer stage model introduced in Section 3.1 will be reduced to illustrate the framework for the distribution of accuracy specifications as introduced in Theorem 49. For this single-input-single-output example system, the (interconnected) system ROM accuracy requirement is prescribed by

$$|\hat{G}_A(i\omega) - G_A(i\omega)| \leq |\epsilon_A(i\omega)|, \quad (129)$$

with

$$\epsilon_A = \frac{1}{4}G_A + 5 \times 10^{-11}. \quad (130)$$

This accuracy requirement is formally described by accuracy specification $\mathcal{E}_A$ defined by $W_A = 1$ and $V_A = \epsilon_A^{-1}$ according to Definition 47.

This accuracy specification can be visualized in the frequency domain as a band around the frequency response of G_A , as shown in the top plot of Fig. 11, or as an upper bound on the error $\hat{G}_A - G_A$, as in the bottom plot of Fig. 11.

Then, the sub-system accuracy specification sets are obtained by maximizing (128) for a frequency grid of 100 logarithmically spaced points between 10 Hz and 2000 Hz, resulting in $\mathcal{E}^1, \mathcal{E}^2, \mathcal{E}^3$. Here, $V(s)$ and $W(s)$ are taken as diagonal matrices to facilitate the optimization. Next, each sub-system model G^j , for $j = 1, 2, 3$, is reduced to $\hat{G}^j$ by choosing balancing and residualization (defined as sub-system balanced reduction (ssBR), see Section 2.4) such that $\hat{G}^j \in \mathcal{E}^j$, i.e., $\hat{G}^j$ satisfies the sub-system ROM accuracy specification. The resulting reduced order models $\hat{G}^j$ are of order 8, 12, and 22, respectively, such that $\hat{G}_A$ is of order 42. The frequency response of each model G^j , the corresponding

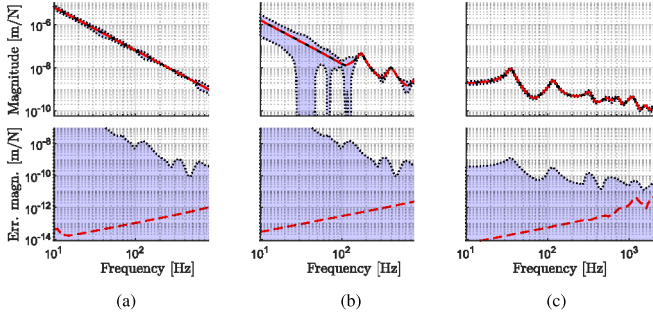


Fig. 12. Spectral norms of the MIMO transfer functions G^1 (in (a)), G^2 (in (b)), and G^3 (in (c)), their reduced counterparts $\hat{G}^1, \hat{G}^2, \hat{G}^3$, and the corresponding accuracy specifications $\mathcal{E}^1, \mathcal{E}^2$, and $\mathcal{E}^3$, respectively. The full-order models are indicated by black (—), the reduced models by (---), and the accuracy specifications by the blue areas.

requirement $\mathcal{E}^j$ and derived $\hat{G}^j$ are visualized in Fig. 12. The resulting reduced-order interconnected model $\hat{G}_A$, which is composed by interconnecting $\hat{G}^1, \hat{G}^2, \hat{G}^3$ with S as in (120), is guaranteed to satisfy $\hat{G}_A \in \mathcal{E}_A$ (as follows from Theorem 49). This can be verified visually by observing $\hat{G}_A$ in Fig. 11, where $\hat{G}_A$ lies completely within the blue area that represents the accuracy requirement $\mathcal{E}_A$.

3.4. Structure-preserving reduction using abstracted environments

Theorem 49 indicates that, as long as sub-system accuracy ROM requirements are fulfilled, the interconnected system ROM also satisfies its accuracy requirement. The actual order and accuracy of the sub-system ROMs depend on the selected reduction method, which can be chosen freely by the user (emphasizing the broad applicability of the proposed framework). However, in practice, any sub-system ROM $\hat{G}^j$ that satisfies (122) admits some measure of conservatism. Indeed, for $\hat{G}^j$ to satisfy (122) without conservatism, $\hat{G}^j$ should satisfy $\|(W^j(j\omega))^{-1}(\hat{G}^j(j\omega) - G^j(j\omega))(V^j(j\omega))^{-1}\|_2 = 1$ for all $\omega \in \mathbb{R}$, which will never occur in realistic scenarios. As, theoretically, a larger error allows more reduction, this conservatism of the framework of Section 3.3 prevents further reduction of sub-system models $\hat{G}^j$.

To enable further reduction of the sub-system models (for a given system accuracy specification $\mathcal{E}_A$), we consider *structure-preserving* reduction methods. Such methods reduce each of the component models G^j while explicitly aiming to approximate the external input–output behavior of the interconnected model G_A by a ROM $\hat{G}_A$ (i.e., the dynamics of the interconnected system is explicitly taken into account in the reduction of the sub-systems, not only in the formulation of their accuracy specifications, as done in the previous section) (Sandberg & Murray, 2009). Because of this, structure-preserving reduction methods typically yield interconnected system ROMs of significantly higher accuracy than with standard (separate) sub-system reduction as employed in Section 3.3.

A particular reduction method that considers the full interconnected system is introduced in Vandendorpe and Van Dooren (2008) as Interconnected System Balanced Truncation (ISBT). For the reduction using ISBT, consider a minimal state-space realization for each sub-system model G^j as

$$G^j : \begin{cases} \dot{x}^j = A^j x^j + B^j w^j, \\ v^j = C^j x^j + D^j w^j, \end{cases} \quad (131)$$

also denoted as $G^j : (A^j, B^j, C^j, D^j)$ and where $x^j \in \mathbb{R}^{n_j}$. Similarly, the interconnected system model admits a minimal state-space representation $G_A = (A_A, B_A, C_A, D_A)$ with state vector $x_A^T = [(x^1)^T, \dots, (x^k)^T] \in \mathbb{R}^{n_A}$.

With ISBT, individual sub-systems G^j are reduced based on the controllability and observability Gramians P_A and Q_A , respectively, of

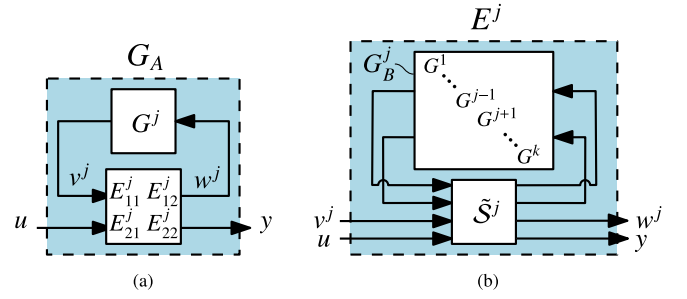


Fig. 13. (a) Upper LFT of E^j and G^j , constituting the interconnected model $G_A = F_u(E^j, G^j)$, and (b) upper LFT of $\tilde{S}^j$ and all G^l for $l \in \{1, \dots, k\} \setminus \{j\}$, constituting E^j .

the interconnected system model G_A , which is assumed to be asymptotically stable. These Gramians are the unique solutions to the Lyapunov equations

$$\begin{aligned} 0 &= A_A P_A + P_A A_A^T + B_A B_A^T, \\ 0 &= Q_A A_A + A_A^T Q_A + C_A^T C_A, \end{aligned} \quad (132)$$

which are partitioned according to the state partitioning of x_A as

$$P_A = \begin{bmatrix} P_{1,1} & \dots & P_{1,k} \\ \vdots & \ddots & \vdots \\ P_{k,1} & \dots & P_{k,k} \end{bmatrix}, \quad Q_A = \begin{bmatrix} Q_{1,1} & \dots & Q_{1,k} \\ \vdots & \ddots & \vdots \\ Q_{k,1} & \dots & Q_{k,k} \end{bmatrix}, \quad (133)$$

such that $P_{i,j}, Q_{i,j} \in \mathbb{R}^{n_i \times n_j}$ for $i, j \in \{1, \dots, k\}$. To reduce each G^j to $\hat{G}^j$, i.e., a balancing transformation (recall Lemma 24) is applied to balance $P_{j,j}$ and $Q_{j,j}$. In this way, each sub-system model is reduced based on the contribution it has to the interconnected system model, thereby resulting in a highly accurate interconnected system ROM.

However, solving the Lyapunov equations of (132) is computationally costly, as the order of the interconnected system is typically high. To overcome this, Poort, Janssen, Besselink, Fey, and van de Wouw (2025) takes a different perspective: upon the reduction of a single sub-system G^j , all remaining sub-systems act as an environment E^j in which the sub-system is embedded, as illustrated in Fig. 13(a). This environment thus determines how G^j influences the input–output behavior of G_A . Analogously, it determines how sub-system reduction errors $G^j - \hat{G}^j$ influence the total reduction error $G_A - \hat{G}_A$. Therefore, the environment model acts as a “weighting function” to indicate the most relevant dynamics of sub-system G^j in the interconnected system G_A .

As mentioned above, the equations of (132) are computationally costly to solve because this environment model is typically of high order. Therefore, Poort et al. (2025) proposes replacing the environment model with a low-order *abstraction*. The resulting approach of *abstracted reduction*, as further developed in Poort, Besselink, Fey, and van de Wouw (2025) and schematically visualized in Fig. 14, can drastically reduce the computational cost of structure-preserving reduction methods such as ISBT, without significantly compromising its accuracy benefits.

To formally introduce this approach, we first introduce the environment models E^j . Each E^j consists of the interconnection structure S and all sub-system models $G^1, \dots, G^k$ except for G^j , as visualized in Fig. 13, and is defined as follows.

Definition 51. Given S and $G^1(s), \dots, G^k(s)$, the environment $E^j(s)$ of each $G^j(s)$ is given as

$$E^j := F_u(\tilde{S}^j, G_B^j), \quad (134)$$

where G_B^j is the parallel connection of all sub-systems excluding G^j , such that

$$G_B^j = \text{diag}(G^1, \dots, G^{j-1}, G^{j+1}, \dots, G^k), \quad (135)$$

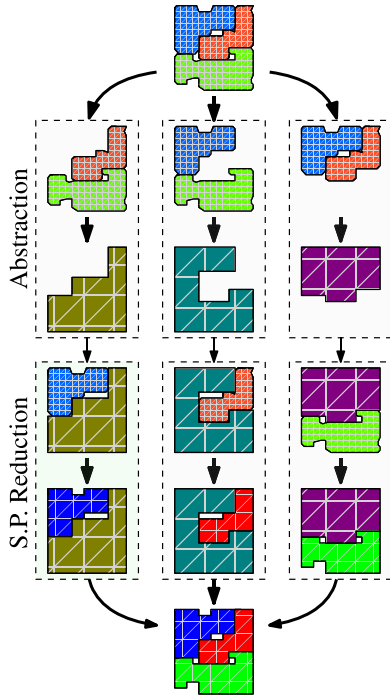


Fig. 14. Schematic representation of the framework of abstracted reduction, considering an interconnection of three component models. For the reduction of each sub-system, its environment, consisting of the remaining sub-systems, is first replaced by a low-order abstraction: the top part marked “Abstraction”. By considering only an abstraction of this environment in the subsequent structure-preserving (S.P.) reduction step of the sub-system (the bottom steps), the computational cost is significantly reduced.

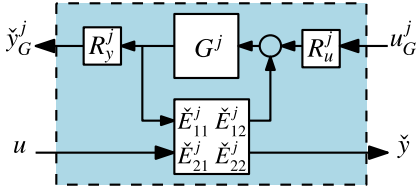


Fig. 15. Schematic representation of the j -th abstracted, augmented system $F_u(\tilde{E}^j, G^j)$ with $\tilde{E}^j$ as given in 137.

and where $\tilde{S}^j$ is obtained by permuting the rows and columns of S such that

$$\begin{bmatrix} w^1 \\ \vdots \\ w^{j-1} \\ w^{j+1} \\ \vdots \\ w^k \\ \hline w^j \\ u \end{bmatrix} = \tilde{S}^j \begin{bmatrix} v^1 \\ \vdots \\ v^{j-1} \\ v^{j+1} \\ \vdots \\ v^k \\ \hline v^j \\ y \end{bmatrix}. \quad (136)$$

The framework of abstracted reduction, as introduced in Poort et al. (2025), is formalized in Algorithm 1 below.

As shown by Algorithm 1, the framework consists of two reduction steps:

- The abstraction of the j^{th} environment model E^j to $\tilde{E}^j$. This abstraction can be accomplished by a relatively cheap reduction method, which is not required to be particularly accurate in the coupled context as $\tilde{E}^j$ mainly influences the structure-preserving reduction of sub-system j and only influences the accuracy of $\hat{G}_A$ indirectly and to a minor extent.

Algorithm 1 Abstracted reduction

1. **For all** $j = 1, \dots, k$,

- Construction of $E^j(s)$.** Construct the environments $E^j(s)$, according to Definition 51.
- Abstraction of $E^j(s)$.** Reduce each individual $E^j(s)$ to $\tilde{E}^j(s)$.
- Augmentation of $\tilde{E}^j(s)$.** Augment the set of inputs and outputs of $F_u(\tilde{E}^j, G^j)$ by incorporating G^j 's (weighted) inputs and outputs as in Fig. 15. This is equivalent to replacing $\tilde{E}^j(s)$ with $\tilde{F}^j(s)$, resulting in $F_u(\tilde{F}^j, G^j)$, where

$$\tilde{F}^j(s) = \begin{bmatrix} \tilde{F}_{11}^j & \tilde{F}_{12}^j \\ \tilde{F}_{21}^j & \tilde{F}_{22}^j \end{bmatrix} = \left[\begin{array}{c|cc} \tilde{E}_{11}^j & R_u^j & \tilde{E}_{12}^j \\ \hline R_y^j & 0 & 0 \\ \tilde{E}_{21}^j & 0 & \tilde{E}_{22}^j \end{array} \right], \quad (137)$$

for weighting matrices $R_y^j \in \mathbb{C}^{p_j \times p_j}$ and $R_u^j \in \mathbb{C}^{m_j \times m_j}$.

- Reduction of $G^j(s)$.** Use a structure-preserving reduction method to reduce $F_u(\tilde{F}^j, G^j)$ to $F_u(\tilde{F}^j, \hat{G}^j)$.

2. **Connection of $\hat{G}^j(s)$.** Compose the reduced-order, interconnected system $\hat{G}_A = F_u(S, \hat{G}_B)$.

- The structure-preserving reduction of $F_u(\tilde{F}^j, G^j)$ to $F_u(\tilde{F}^j, \hat{G}^j)$, as achieved by ISBT.⁴ By reducing $F_u(\tilde{F}^j, G^j)$ instead of G_A , this reduction step is computationally significantly cheaper.

The weighting matrices R_y^j and R_u^j , used to augment $\tilde{E}^j$ to $\tilde{F}^j$, provide additional control over the accuracy distribution of the input–output pairs depicted in Fig. 15 (Poort et al., 2025). The augmented input–output pair from u_G^j to y_G^j in Fig. 15 has the transfer function $R_y^j F_u(\tilde{E}^j, \hat{G}^j) R_u^j$ with

$$\tilde{E}^j(s) = \begin{bmatrix} \tilde{E}_{11}^j(s) & I \\ I & 0 \end{bmatrix}. \quad (138)$$

In the following, we will see that this augmented pair is key to relating the accuracy of $\hat{G}_A$ to the accuracy of $F_u(\tilde{F}^j, \hat{G}^j)$.

The approach in this section is similar to that in Section 3.3: we will formulate system-level accuracy specifications using Definition 47 and distribute this ‘room for reduction error’ over the reduction steps (b) and (d) in Algorithm 1 that introduce the reduction errors. The main difference is that, while in Section 3.3 the error sources are related to individual sub-systems, here the reduction errors are related to both the environment abstraction inaccuracies $\tilde{E}^j - E^j$ and the sub-system reduction errors in the abstracted interconnected system $F_u(\tilde{F}^j, \hat{G}^j) - F_u(\tilde{F}^j, G^j)$. We emphasize that the error $F_u(\tilde{F}^j, \hat{G}^j) - F_u(\tilde{F}^j, G^j)$ (for reasonable $\tilde{E}^j - E^j$, R_y^j and R_u^j) is generally a better indicator for the accuracy of $\hat{G}_A$ than $\hat{G}^j - G^j$, as $F_u(\tilde{E}^j, \hat{G}^j) - F_u(\tilde{E}^j, G^j) \approx \hat{G}_A - G_A$.

To quantify the abstraction and reduction errors $\tilde{E}^j - E^j$ and $F_u(\tilde{F}^j, \hat{G}^j) - F_u(\tilde{F}^j, G^j)$, respectively, we introduce the following accuracy specifications.

Definition 52. For an interconnected system with sub-system transfer functions $G^j(s)$, let $\mathcal{E}_E^j$ and $\mathcal{E}_F^j$ denote abstraction specifications and coupled ROM specifications, respectively, defined as

$$\mathcal{E}_E^j := \left\{ \tilde{E}^j \mid \left\| (W_E^j)^{-1} (\tilde{E}_{11}^j - E_{11}^j) (V_E^j)^{-1} \right\|_{H_\infty} \leq 1 \right\}, \quad (139a)$$

$$\mathcal{E}_F^j := \left\{ \hat{G}^j \mid \right\} \quad (139b)$$

⁴ Note that $\tilde{F}^j$ is an augmented version of $\tilde{E}^j$, see 137.

$$\left\| (W_F^j)^{-1} (F_u(\tilde{E}^j, \hat{G}^j) - F_u(\tilde{E}^j, G^j)) (V_F^j)^{-1} \right\|_{H_\infty} \leq 1 \}.$$

We say that any abstracted environment and sub-system ROM $j = 1, 2, \dots, k$ satisfies its specifications if $\tilde{E}^j \in \mathcal{E}_E^j$ and $\hat{G}^j \in \mathcal{E}_F^j$, respectively. In (139), $V_E^j(s)$, $V_F^j(s)$, $W_E^j(s)$, and $W_F^j(s)$ are user-selected weighting functions, which are taken to be stable proper transfer functions, with stable proper inverses.

To distribute the accuracy specification $\mathcal{E}_A$ for the interconnected system of Definition 47 over $\mathcal{E}_E^j$ (environment abstraction steps) and $\mathcal{E}_F^j$ (sub-system reduction steps) for $j = 1, \dots, k$, as given in Definition 52, Theorem 49 is adapted for the application of abstracted reduction. To this end, we redefine the nominal transfer function $N(s)$, the weighting matrices $W(s)$ and $V(s)$, and the set of scaling matrices $\mathbf{D}$ as in Poort et al. (2025). Firstly, the nominal transfer function $N(s)$ is defined as

$$N = \begin{bmatrix} M & I + MZ & M & MS_{BA} \\ -ZM - I & Z(I + MZ) & ZM & (I + ZM)S_{BA} \\ -M & MZ & M & MS_{BA} \\ -S_{AB}M & S_{AB}(I + MZ) & S_{AB}M & 0 \end{bmatrix}, \quad (140)$$

with S_{AB} , S_{BA} , and S_{BB} as in (116) and where

$$\begin{aligned} M(s) &= (I - G_B(s)S_{BB})^{-1}G_B(s), \\ Z(s) &= S_{BB} - E_{B,11}(s), \end{aligned} \quad (141)$$

with

$$E_{B,11} = \text{diag}(E_{11}^1, \dots, E_{11}^k). \quad (142)$$

Furthermore, the block-diagonal weighting transfer functions are defined as

$$W = \text{diag}(W_E^1, \dots, W_E^k, W_F^1, \dots, W_F^k, W_E^1, \dots, W_E^k, W_A), \quad (143a)$$

$$V = \text{diag}(V_E^1, \dots, V_E^k, V_F^1, \dots, V_F^k, V_E^1, \dots, V_E^k, V_A), \quad (143b)$$

whereas $\mathbf{D}$ reads

$$\mathbf{D} := \left\{ (D_\ell, D_r) \left| \begin{array}{l} d^1, \dots, d^k, d_A \in \mathbb{R}_+, S^j = (S^j)^* = \begin{bmatrix} s_{11}^j & s_{12}^j \\ s_{12}^j & s_{22}^j \end{bmatrix} \in \mathbb{C}^{2 \times 2}_{>0} \\ D_\ell = D(p_1, \dots, p_k, m_1, \dots, m_k, p) \\ D_r = D(m_1, \dots, m_k, p_1, \dots, p_k, m) \end{array} \right. \right\}, \quad (144)$$

with

$$D(a^1, \dots, a^k, b^1, \dots, b^k, c) = \begin{bmatrix} s_{11}^1 I_{a^1} & & & s_{12}^1 I_{a^1} & & \\ & \ddots & & & \ddots & \\ & & s_{11}^k I_{a^k} & & & s_{12}^k I_{a^k} \\ & & & d^1 I_{b^1} & & \\ & & & & \ddots & \\ & & & & & d^k I_{b^k} \\ s_{21}^1 I_{a^1} & & & s_{22}^1 I_{a^1} & & \\ & \ddots & & & \ddots & \\ & & s_{21}^k I_{a^k} & & & s_{22}^k I_{a^k} \\ & & & & & d_A I_c \end{bmatrix}. \quad (145)$$

Theorem 49 is then adapted to the problem setting of abstracted reduction, as explicated in the following corollary.

Corollary 53. Consider the nominal system in (140) and the weighting transfer function matrices in (143), related to the accuracy specifications. If there exist

$$(D_\ell, D_r) \in \mathbf{D} \quad (146)$$

with $\mathbf{D}$ as in (144), such that

$$\begin{bmatrix} W^{-2}(i\omega)D_r^{-1} & N^*(i\omega) \\ N(i\omega) & V^{-2}(i\omega)D_\ell \end{bmatrix} > 0 \quad (147)$$

holds for all $\omega \in \mathbb{R}$, then satisfaction of the accuracy specifications of both environment abstraction and sub-system reduction, i.e., respectively, $\tilde{E}^j \in \mathcal{E}_E^j$ and $\hat{G}^j \in \mathcal{E}_F^j$ for all $j = 1, 2, \dots, k$, implies satisfaction of the accuracy specification for the reduction of the interconnected system, i.e., $\hat{G}_A \in \mathcal{E}_A$.

The distribution of the accuracy specification for the environment abstractions and sub-system reductions, represented by the sets $\mathcal{E}_E^1, \dots, \mathcal{E}_E^k, \mathcal{E}_F^1, \dots, \mathcal{E}_F^k$, can then be obtained by maximizing (128) subject to (146) and (147). The resulting specifications can then be used to abstract the environment models $E^1, \dots, E^k$ and reduce the sub-system models $G^1, \dots, G^k$ and obtain $\tilde{E}^1, \dots, \tilde{E}^k$ and $\hat{G}^1, \dots, \hat{G}^k$ for which $\tilde{E}^j \in \mathcal{E}_E^j$ and $\hat{G}^j \in \mathcal{E}_F^j$ for all $j \in \{1, \dots, k\}$. Corollary 53 then guarantees that the interconnected system ROM, resulting from assembling the sub-system ROMs, satisfies the prescribed accuracy specification, i.e., $\hat{G}_A \in \mathcal{E}_A$, and is thereby also guaranteed to be stable.

Illustrative example: 2D wafer stage. To illustrate the framework of abstracted reduction, we consider again the benchmark system presented in Section 3.1 and reduce it using the abstracted reduction framework, such that $\hat{G}_A \in \mathcal{E}_A$, with $\mathcal{E}_A$ as prescribed by (129). To this end, $\mathcal{E}_E^1, \mathcal{E}_E^2, \mathcal{E}_E^3, \mathcal{E}_F^1, \mathcal{E}_F^2, \mathcal{E}_F^3$ are first determined by maximizing (128), after which the environment models E^j are constructed using Definition 51. These environment models are subsequently abstracted (typically a quite drastic reduction) to obtain the abstractions $\tilde{E}^j$, using the computationally inexpensive Hintz-Herting component mode synthesis reduction technique (Herting, 1985), such that $\tilde{E}^j \in \mathcal{E}_E^j$ for $j = 1, 2, 3$. Subsequently, the augmented models $F_u(\tilde{E}^j, G^j)$ are constructed following 137 using R_u^j and R_v^j as identity matrices (for further insight into their effect, see Poort et al. (2025)). Each augmented model $F_u(\tilde{E}^j, G^j)$ is then reduced to $F_u(\tilde{E}^j, \hat{G}^j)$ using closed-loop balanced reduction (CLBR) (Ceton, Wortelboer, & Bosgra, 1993; Wortelboer & Bosgra, 1994) via residualization, such that $\hat{G}^j \in \mathcal{E}_F^j$ for $j = 1, 2, 3$. Note that CLBR is simply a special case of ISBR, reducing a single sub-system model connected to another. This approach of abstracted reduction using CLBR is henceforth denoted aCLBR.

In addition, we also perform a structure-preserving reduction without abstracting the environments to determine the effect of using environment abstractions. To this end, an equivalent optimization is performed in which $\mathcal{E}_E^1, \mathcal{E}_E^2, \mathcal{E}_E^3$ are forced to be empty, thus maximizing $\mathcal{E}_F^1, \mathcal{E}_F^2, \mathcal{E}_F^3$. In this way, the environments are not abstracted and only G^1, G^2 , and G^3 are reduced using ISBR (equivalent to ISBT of Vandendorpe and Van Dooren (2008), but eliminating states using residualization, see, e.g., Antoulas, 2005a), such that $\hat{G}^j \in \mathcal{E}_F^j$ for $j = 1, 2, 3$. Recall that this approach, while generally resulting in a more accurate reduced-order model $\hat{G}_A$, is typically not feasible for high-order interconnected system models.

The resulting reduced-order models are compared both in terms of the orders required to meet $\hat{G}_A \in \mathcal{E}_A$ and their actual accuracy. To quantify this accuracy, let us define the relative error $\bar{\epsilon}_A := \|\hat{G}_A - G_A\|_2 / \|G_A\|_2$, where $\|\cdot\|_2$ represents the frequency-limited $\mathcal{L}_2$ -norm defined as

$$\|G\|_2 = \left(\frac{1}{2\pi} \int_{-\omega_b}^{\omega_b} G^*(j\omega)G(j\omega) d\omega \right)^{1/2}, \quad (148)$$

where $\omega_b = 2000 \times 2\pi$.

The resulting relative errors are displayed in Table 2 together with the required orders of (1) the reduced sub-system models $\hat{G}^j$, for $j = 1, 2, 3$, (2) the abstracted environment models $\tilde{E}^j$, and (3) the reduced interconnected system model $\hat{G}_A$. Here, we also included the subsystem balanced reduction (ssBR) method of Section 3.3 for comparison.

When comparing the three methods in terms of their reduced order and relative error $\bar{\epsilon}_A$, all three perform similarly. ISBR scores slightly better in terms of accuracy, aCLBR (surprisingly) slightly better in terms of the reduced order, and ssBR performs the worst, though only slightly.

Visual inspection of the three reduced-order models in Fig. 16 reflects the errors $\bar{\epsilon}_A$ in Table 2, showing that all errors are of similar magnitude. In addition, all reduced models $\hat{G}_A$ clearly comply with the specification, i.e. $\hat{G}_A \in \mathcal{E}_A$.

In conclusion, ssBR reduces each sub-system model G^j separately, ISBR reduces each sub-system model taking the full interconnected system G_A into account (which is computationally expensive), and

Table 2

Orders of the full-order model (FOM) and the reduced models (both of the sub-systems and the environments), resulting from the sub-system reduction with ssBR (Section 3.3), the structure-preserving reduction with ISBR (Vandendorpe & Van Dooren, 2008), and the abstracted reduction using CLBR (aCLBR) presented here. The relative, frequency-limited $\mathcal{L}_2$ -norm $\bar{\epsilon}_A$ is also included below.

System	FOM	ssBR	ISBR	aCLBR
G^1	336	8	8	8
G^2	704	12	15	13
G^3	936	22	19	20
E^1	1645	–	1645	26
E^2	1277	–	1277	30
E^3	1045	–	1045	6
G_A	1976	42	42	41
$\bar{\epsilon}_A$	–	3.8×10^{-3}	1.9×10^{-3}	3.1×10^{-3}

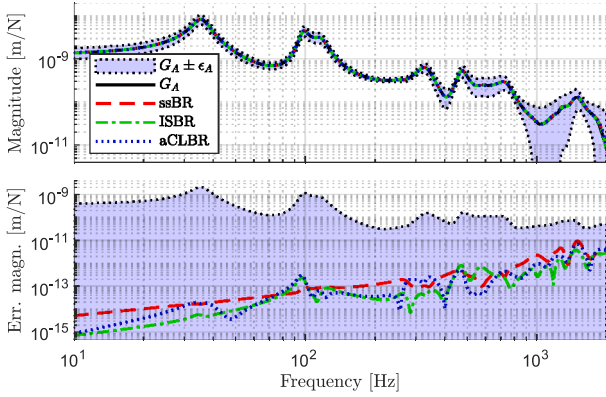


Fig. 16. Top plot: Magnitudes of the frequency responses of the interconnected model G_A and the three reduced-order approximations $\hat{G}_A$. Bottom plot: the magnitudes of the error systems $\hat{G}_A - G_A$ for the three approximate models. In both plots, the system accuracy specification ϵ_A is indicated by the blue area.

aCLBR takes only low-order abstracted environments into account. Comparing ISBR and aCLBR, the results shown in Table 2 signify that low-order approximate environment models, as used by aCLBR are sufficient to accurately reduce the sub-systems, which significantly reduces the computational cost of the structure-preserving reduction. For the considered use case, the subsystem reduction method ssBR performs quite well, being only slightly worse than ISBR. For problems where the environment is more important, however, ssBR is expected to perform significantly worse than ISBR and aCLBR.

3.5. Preservation of stability and passivity

In Sections 3.3 and 3.4, assembly-level accuracy specifications are distributed over the individual reduction (and abstraction) steps, thereby guaranteeing accuracy and stability. In case the sub-system models are passive according to Definition 6 (see also Remark 8), their passivity can also be preserved by employing passivity-preserving reduction methods. Then, if $-S$ is also passive, the full interconnected system G_A is passive, e.g., van der Schaft (2017). Consequently, the interconnected ROM $\hat{G}_A$ is passive if the passivity property is preserved for all sub-system ROMs $\hat{G}^j$ with $j = 1, \dots, k$. Preserving a model's passivity is particularly relevant in the context of control, where a passive ROM enables the use of passivity-based control strategies, which ensure robustness to system variability and uncertainty (Bao & Lee, 2007).

For individual sub-system reduction, as considered in Section 3.3, there exist many passivity-preserving reduction methods, such as Breiten and Unger (2022), Desai and Pal (1984), Jonckheere and Silverman

(1983), Knockaert, Dhaene, Ferranti, and de Zutter (2011), Muscato, Nunnari, and Fortuna (1997), Reis and Stykel (2010) (see also Remark 31). In case passivity preservation is preferred, the accuracy specifications can be distributed using Theorem 49, after which a passivity-preserving reduction method of choice can be used to reduce the sub-systems accordingly.

In case the environment needs to be taken into account explicitly, a passivity-preserving reduction method is required that also considers the full interconnected dynamics, as, e.g., ISBT in Section 3.4 does. To this end, ISBT was adapted in Poort et al. (2024) to combine the consideration of the interconnected input–output behavior, by considering the interconnected system's Gramians of (132) and assuming passivity characteristics of the sub-systems. Specifically, for a passive sub-system G^j , we denote by K_r^j and K_a^j the matrices characterizing the required supply function and available storage function, respectively, as in Remark 31.

To reduce each G^j to $\hat{G}^j$, it is then possible to use the balancing transformation of Lemma 24 to balance either $(P_{jj}$ and $K_a^j)$ or $(Q_{jj}$ and $K_r^j)$, after which the least important states are truncated. This approach, denoted as Passive Interconnected Balanced Truncation (PIBT) in Poort et al. (2024), thus combines the system-level accuracy, via the use of P_A or Q_A , with sub-system-level passivity via K_a^j or K_r^j . In this work, PIBT balances P_A and K_a^j , though balancing Q_A and K_r^j gives similar results. PIBT guarantees the passivity and stability of both the sub-systems and the interconnected system (asymptotic stability can be guaranteed by employing a stricter notion of passivity), see Poort et al. (2024) for more details.

Similar to ISBT, PIBT can also be used within the abstracted reduction framework of Algorithm 1 to reduce the computational complexity of solving the required Lyapunov equations.

Illustrative example: 2D wafer stage. To evaluate PIBT and its implementation within the abstracted reduction framework, we consider a similar use case as the one presented in Section 3.2 and Fig. 10. However, instead of position outputs, all sub-system models now have collocated velocity outputs, such that they are passive according to Definition 6. By considering only collocated outputs, $m_j = p_j$ for all $j \in \{1, 2, 3\}$, as shown in Table 1. To realize the spring interconnections, the new interconnection structure $\tilde{S}$ needs to integrate the sub-system velocity outputs to obtain the spring extensions. Therefore, $\tilde{S}$ becomes a dynamical system given by

$$\tilde{S} : \begin{cases} \dot{x}_S &= [B_1 \quad 0] \begin{bmatrix} v_B \\ u \end{bmatrix} \\ \begin{bmatrix} w_B \\ y \end{bmatrix} &= \begin{bmatrix} C_1 \\ C_2 \end{bmatrix} x_S + \begin{bmatrix} D_{11} & D_{12} \\ 0 & 0 \end{bmatrix} \begin{bmatrix} v_B \\ u \end{bmatrix}. \end{cases} \quad (149)$$

Now, the interconnected system model G_A can be represented as

$$G_A = F_u(\tilde{S}, G_B). \quad (150)$$

For the benchmark model of Fig. 10, B_1 is a signed Boolean matrix such that x_S represents the spring extensions and the overlay (displacement) output $y = v_1^3 - v_1^1$ (see Fig. 10), C_1 is a matrix of spring constants, C_2 is a signed Boolean matrix to output the overlay, and D_{12} is a signed Boolean matrix to map the external input u to w^3 .

The interconnected system model G_A is reduced to $\hat{G}_A$ by reducing G^j to $\hat{G}^j$, for $j = 1, 2, 3$, using three passivity-preserving reduction approaches: (i) reduction of the separate sub-systems using Mixed-Gramian Balanced Residualization (MGBR) of Unneland, Van Dooren, and Egeland (2007), (ii) reduction using Passive Interconnected Balanced Residualization (PIBR), and (iii) abstracted reduction, as in Algorithm 1, using PIBR for the reduction step 1.d (aPIBR). Note that PIBR is equivalent to PIBT besides the approach of state elimination (residualization instead of truncation, see, e.g., Antoulas, 2005b).

For all three approaches, the interconnected system accuracy specification ϵ_A , as prescribed by (129), is distributed over the individual

accuracy requirements $\mathcal{E}^j$ or $\mathcal{E}_E^j$ and $\mathcal{E}_F^j$, for $j = 1, 2, 3$. For approach (ii), using PIBR, Corollary 53 is used to translate $\mathcal{E}_A$ to $\mathcal{E}_F^1$, $\mathcal{E}_F^2$ and $\mathcal{E}_F^3$, while all $\mathcal{E}_E^j$ are fixed as empty sets. For approach (iii), using aPIBR, the abstraction and augmentation are performed in the same way as in the example treated in Section 3.4, using the Hintz-Herting reduction approach and R_u^j and R_y^j as identity matrices.

The model orders and relative errors, resulting from the three reduction approaches, are given in Table 3 and the frequency response functions of the (reduced) interconnected system models are visualized in Fig. 17. Table 3 clearly indicates the benefit of considering the (abstracted) environment models: PIBR and aPIBR can reduce significantly further than MGBR, while achieving a similar or even higher accuracy. Note that the focus is on reducing as far as possible while satisfying $\mathcal{E}_A$ and that PIBR only achieves a lower accuracy than MGBR because of the smaller order of the reduced interconnected model. Fig. 17 shows that all reduced-order models satisfy the prescribed accuracy requirement. In addition, all reduced-order models are stable and passive by considering each sub-system model's available storage in their reduction.

3.6. Summary

In this section, key challenges for model order reduction of interconnected LTI system models have been identified, and reduction methodologies have been presented that maintain and guarantee ROM accuracy and preserve ROM properties such as the interconnection structure (promoting interpretability), stability, and passivity. First, we have introduced a method to distribute the accuracy specifications of the interconnected system ROM to the accuracy specifications of the individual sub-system ROMs, ensuring that the sub-system model reduction aligns with the overall system model accuracy specifications. Then, to improve the efficiency of structure-preserving reduction methods, we introduced the framework of abstracted reduction, in which an abstraction of the sub-system's environment model is taken into account in the reduction process to improve the accuracy of the reduced interconnected system model. Indeed, by replacing high-order environment models with lower-order abstractions, high accuracy is achieved without significantly increasing computational cost compared to simply reducing the separate sub-systems independently. Finally, we described a passivity-preserving extension of the previously described reduction methods for interconnected sub-systems. Throughout this section, we used a motivating industrial use case of a 2D wafer stage model to demonstrate how the methods can be applied to obtain accurate, structure-preserving, stable, and passive reduced-order models in a modular and computationally attractive manner, with accuracy guarantees.

4. Conclusions and outlook

4.1. Conclusions

This paper has presented model reduction techniques for two classes of interconnected dynamical systems: firstly, Lur'e-type systems comprising the feedback interconnection between a large-scale linear system and static nonlinearities and, secondly, interconnections between (many) linear dynamical systems. For the first class, we have provided a unified description of both balancing-based and moment matching approaches that (1) are applicable to large-scale Lur'e-type systems, (2) provide a reduction error bound and (3) warrant (global) stability preservation. Guaranteeing these properties crucially relies on incremental stability properties such as convergence or a bounded incremental $\mathcal{L}_2$ gain. For the second class of systems, we have presented an approach to link reduction accuracy specifications on the level of the sub-systems to related specifications on the level of the interconnected system and vice versa. This allows for a modular approach which preserves the interconnection structure of the high-order system, hence

Table 3

Orders of the full-order model (FOM), the reduced models (and the environments), resulting from sub-system reduction with MGBR (Unneland et al., 2007), structure-preserving reduction with PIBR (as presented here), and PIBR in combination with the abstracted reduction approach of Algorithm 1 (aPIBR). The relative, approximate $\mathcal{L}_2$ -norm $\bar{\epsilon}_A$ is also included.

System	FOM	MGBR	PIBR	aPIBR
G^1	336	24	9	14
G^2	704	34	33	34
G^3	936	38	30	29
E^1	1645	–	1645	40
E^2	1277	–	1277	78
E^3	1045	–	1045	40
G_A	1976	101	77	82
$\bar{\epsilon}_A$	–	7.0×10^{-3}	8.6×10^{-3}	3.0×10^{-3}

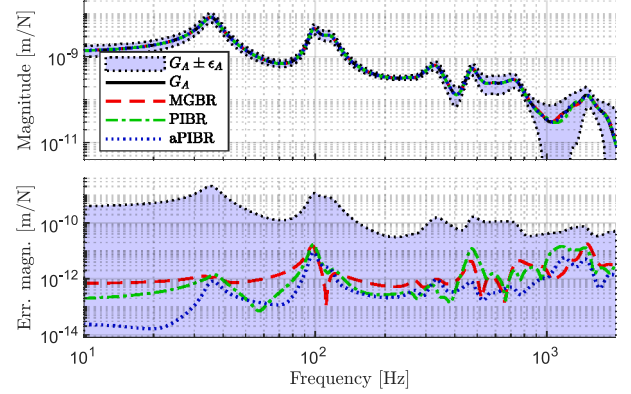


Fig. 17. Top plot: magnitudes of the frequency responses of the interconnected model G_A and the three low-order approximations $\hat{G}_A$. Bottom plot: the magnitudes of the error systems $\hat{G}_A - G_A$ for the three approximate models. In both plots, the system accuracy requirement $\mathcal{E}_A$ is indicated by the blue area.

promoting interpretability of the reduced-order system. In addition, we discussed the concept of abstracted reduction for interconnected systems, which allows performing modular reduction of the sub-systems, while taking into account the dynamics of the rest of the interconnected system in a computationally tractable way. Finally, these methods also provide an error bound, stability preservation, and, in addition, under certain conditions, passivity preservation. The different approaches for the reduction of linear interconnected systems have been demonstrated and compared by means of a structural dynamics case study in the scope of lithography machines.

4.2. Extensions and outlook

Many important challenges remain open in the field of the reduction of interconnected dynamical systems. We will highlight a concise selection of these challenges, in addition to extensions of the work in this paper, below.

The results in this paper considered feedback interconnections involving static nonlinearities. Extensions to feedback interconnections involving nonlinear dynamical systems in the feedback loop exist as well, see, e.g., Besselink et al. (2012), Besselink, van de Wouw, and Nijmeijer (2013), Shakib (2022). The results on nonlinear model reduction in this paper are based on assumptions on global incremental system properties. An extension alleviating this assumption can be found in Padoan, Forni, and Sepulchre (2021), in which dominance theory is used to address a broader class of Lur'e-type systems. Alternatively, (Moreschini & Astolfi, 2024) proposes a closed-loop time-domain moment matching approach which replaces global incremental system properties with related properties on the closed loop. Other extensions include (Cheng, Scherpen, & Zhang, 2019; Dou, Cheng, & Scherpen,

2024) on the complexity reduction of networks of Lur'e-type systems. Finally, (Shakib, Scariotti, & Astolfi, 2024) proposes a data-driven model reduction by moment matching method technique for so-called recurrent equilibrium networks, a model class that also captures Lur'e-type models. In the latter work, the system matrices may be unknown, while the reduction is based on the input–output data directly.

The philosophy of representing nonlinear system dynamics by linear dynamics with a nonlinear operator in the feedback loop has been extended to delay differential equations (van de Wouw, Michiels, & Besselink, 2015a, 2015b) and to sampled-data systems with static feedback controllers (Abbasi & van de Wouw, 2022). Herein, the operators representing either the delay of the sampling-and-hold operator are exactly retained in the reduced-order model and the linear dynamics is reduced in a delay-aware of sampled-data-aware fashion. It remains an open question on how to use these ideas to pursue complexity reduction for sampled-data systems with dynamic feedback controllers.

The reduction methods in this article preserve the static nonlinearities, which provides interpretability. In some scenarios, one might trade such interpretability for a reduced dimension of the nonlinearities. It is, however, an open question how to perform such dimension reduction of the static nonlinearities in Lur'e-type systems while still preserving incremental stability properties and providing error bounds. A related problem arises when the same nonlinearity is repeated many times. For example, it has been shown that dynamics whose vector fields are described through neural network mappings can be represented by Lur'e-type models, see, e.g., El Ghaoui, Gu, Travacca, Askari, and Tsai (2021), Martinelli, Galimberti, Manchester, Furieri, and Ferrari-Trecate (2023), Revay, Wang, and Manchester (2023), Suykens, Vandewalle, and De Moor (1999), Yin, Seiler, and Arcaç (2021). The static nonlinearities (typically many) in this Lur'e-type model are nothing more than the activation functions of the neural networks. In such problems, it is interesting to perform dimensionality reduction of the nonlinearities, i.e., the number of activation functions, as was, e.g., proposed in Drummond, Turner, and Duncan (2022) for small-sized networks.

In Section 3, we presented an analysis framework for linking accuracy specifications of reduced-order linear sub-systems to those for the linear system as a whole and vice versa. The latter, i.e., the distribution of the accuracy specification of the overall, interconnected system over those for the sub-systems can, however, be done in many ways. This freedom allows imposing additional optimality criteria or constraints (e.g., on the accuracy or order of the reduced sub-systems) and, therewith, to tailor the reduction process to the requirements on the use of the reduced-order model. In addition, the modular framework for the distribution of accuracy specifications is not only applicable to guide balancing-based model reduction, but can also be used for smart eigenmode selection in component mode synthesis techniques for the complexity reduction of structural dynamics models (Janssen, Fey, Besselink, & van de Wouw, 2024c). It would also be interesting to investigate how the presented approaches for structure-preserving model reduction based on balancing could be extended to moment-matching methods. Although moment-matching methods have been employed to retain internal model structure, such as second-order structure (Bai, Meerbergen, & Su, 2005) or Port-Hamiltonian structure (Beattie, Gugercin, & Mehrmann, 2022; Polyuga & van der Schaft, 2010), combining these methods with preservation of the interconnection structure is an open challenge.

We also envision such an analysis framework to be highly beneficial not only for model complexity reduction, but also for the modular design of dynamical systems, see Janssen et al. (2023), Janssen, Fey, Besselink, and van de Wouw (2024b) for preliminary work in this direction.

Another perspective on the complexity reduction of networks of dynamical systems is taken in Besselink, Sandberg, and Johansson (2016), Cheng, Kawano, and Scherpen (2019), Cheng, Scherpen, and Besselink (2019), Jongsma, Mlinarić, Grundel, Benner, and Trentelman (2018), Monshizadeh, Trentelman, and Camlibel (2014), where not

only reduction of the sub-system dynamics is pursued, but also the reduction of the network topology (e.g., using clustering techniques). How to jointly pursue such complexity reduction on both network topology and system dynamics remains, in the general sense, an open challenge.

Finally, it is an open problem to extend the framework for modular model reduction from linear interconnected systems to the nonlinear realm, especially in a non-conservative manner. A first attempt has been pursued in Egelmeers et al. (2024) for a class of mechanical systems with position-dependent dynamics, which invites to pursue extensions to linear parameter-varying systems.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

No data was used for the research described in the article.

References

- Aarsnes, U. J., & van de Wouw, N. (2019). Axial and torsional self-excited vibrations of a distributed drill-string. *Journal of Sound and Vibration*, 444, 127–151.
- Abbasi, M. H., & van de Wouw, N. (2022). Model order reduction of linear sampled-data control systems. In *2022 European control conference* (pp. 711–716).
- Abbasi, M., Iapichino, L., Schilders, W. H. A., & van de Wouw, N. (2022). A data-based stability-preserving model order reduction method for hyperbolic partial differential equations. *Nonlinear Dynamics*, 107, 3729–3748.
- Amiri, Anahita, Shakib, Mohammad Fahim, Lopez Arteaga, Ines, & van de Wouw, Nathan (2025). Data-driven modeling for electro-active liquid crystal polymer networks. *Discover Applied Sciences*, 7(62).
- Anderson, B. D. O., & Vongpanitlerd, S. (2006). *Network analysis and synthesis: a modern systems theory approach*. Mineola, USA: Dover Publications, Originally published by Prentice-Hall, Englewood Cliffs, USA, 1973.
- Angeli, David (2002). A Lyapunov approach to incremental stability properties. *IEEE Transactions on Automatic Control*, 47(3), 410–421.
- Antoulas, A. C. (2005a). *Approximation of large-scale dynamical systems*. Philadelphia, USA: SIAM.
- Antoulas, A. C. (2005b). An overview of approximation methods for large-scale dynamical systems. *Annual Reviews in Control*, 29(2), 181–190.
- Arcaç, M., Meissen, C., & Packard, A. (2016). Networks of dissipative systems: Compositional certification of stability, performance, and safety. In *SpringerBriefs in control, automation and robotics*, Springer International Publishing, Switzerland.
- Astolfi, A. (2010). Model reduction by moment matching for linear and nonlinear systems. *IEEE Transactions on Automatic Control*, 55(10), 2321–2336.
- Bai, Z. (2002). Krylov subspace techniques for reduced-order modeling of large-scale dynamical systems. *Applied Numerical Mathematics*, 43(1–2), 9–44.
- Bai, Z., Meerbergen, K., & Su, Y. (2005). Arnoldi Methods for Structure-Preserving Dimension Reduction of Second-Order Dynamical Systems. In *Dimension reduction of large-scale systems* (pp. 173–189). Springer-Verlag.
- Bao, Jie, & Lee, Peter L. (2007). *Process Control: The Passive Systems Approach*. In *Advances in industrial control*, London: Springer.
- Beattie, Christopher, Gugercin, Serkan, & Mehrmann, Volker (2022). Structure-preserving interpolatory model reduction for port-Hamiltonian differential-algebraic systems. In Christopher Beattie, Peter Benner, Mark Embree, Serkan Gugercin, & Sanda Lefteri (Eds.), *Realization and model reduction of dynamical systems: a festschrift in honor of the 70th birthday of thanos antoulas* (pp. 235–254). Cham: Springer International Publishing.
- Benner, P., Hinze, M., & ter Maten, E. J. W. (2011). *Model reduction for circuit simulation*. Springer.
- Berkooz, G., Holmes, P., & Lumley, J. L. (1993). The proper orthogonal decomposition in the analysis of turbulent flows. *Annual Review of Fluid Mechanics*, 25, 539–575.
- Besselink, B., van de Wouw, N., & Nijmeijer, H. (2009). An error bound for model reduction of lur'e-type systems. In *Proceedings of the joint 48th IEEE conference on decision and control and 28th Chinese control conference, shanghai, China* (pp. 3264–3269).
- Besselink, B., van de Wouw, N., & Nijmeijer, H. (2012). Model reduction for a class of convergent nonlinear systems. *IEEE Transactions on Automatic Control*, 57(4), 1071–1076.
- Besselink, B., van de Wouw, N., & Nijmeijer, H. (2013). Model reduction for nonlinear systems with incremental gain or passivity properties. *Automatica*, 49(4), 861–872.

- Besselink, B., van de Wouw, N., Scherpen, J. M. A., & Nijmeijer, H. (2013). Generalized incremental balanced truncation for nonlinear systems. In *Proceedings of the 52nd IEEE conference on decision and control, Florence, Italy* (pp. 5552–5557).
- Besselink, B., van de Wouw, N., Scherpen, J. M. A., & Nijmeijer, H. (2014). Model reduction for nonlinear systems by incremental balanced truncation. *IEEE Transactions on Automatic Control*, 59(10), 2739–2753.
- Besselink, B., Sandberg, H., & Johansson, K. H. (2016). Clustering-based model reduction of networked passive systems. *IEEE Transactions on Automatic Control*, 61(10), 2958–2973.
- Besselink, B., Tabak, U., Lutowska, A., van de Wouw, N., Nijmeijer, H., Rixen, D. J., et al. (2013). A comparison of model reduction techniques from structural dynamics, numerical mathematics and systems and control. *Journal of Sound and Vibration*, 332(19), 4403–4422.
- Boyd, S., Ghaoui, L. El, Feron, E., & Balakrishnan, V. (1994). Linear matrix inequalities in system and control theory. *Studies in applied mathematics: Vol. 15*, Philadelphia, PA: SIAM.
- Breiten, T., & Unger, B. (2022). Passivity preserving model reduction via spectral factorization. *Automatica*, 142(110368).
- Brogliato, B., Lozano, R., Maschke, B., & Egeland, O. (2020). Dissipative systems analysis and control: Theory and applications, *Communications and control engineering series*, (3rd ed.).
- Brogliato, B., & Tanwani, A. (2020). Dynamical systems coupled with monotone set-valued operators: Formalisms, applications, well-posedness, and stability. *SIAM Review*, 62.
- Bruinsma, N. A., & Steinbuch, M. (1990). A fast algorithm to compute the H_∞ -norm of a transfer function matrix. *Systems & Control Letters*, 14(4), 287–293.
- Brunton, S. L., Proctor, J. L., & Kutz, J. N. (2016). Discovering governing equations from data by sparse identification of nonlinear dynamical systems. *Proceedings of the National Academy of Sciences*, 113(15), 3932–3937.
- Burohman, A. M., Besselink, B., Scherpen, J. M. A., & Camlibel, M. K. (2023). From data to reduced-order models via generalized balanced truncation. *IEEE Transactions on Automatic Control*, 68(10), 6160–6175.
- Burohman, A. M., Besselink, B., Scherpen, J. M. A., & Camlibel, M. K. (2024). From data to reduced-order models via moment matching. *Systems & Control Letters*, 194, Article 105965.
- Carr, J. (1981). *Applications of centre manifold theory*. New York, USA: Springer Science & Business Media.
- Ceton, C., Wortelboer, Pepijn M. R., & Bosgra, Okko H. (1993). Frequency weighted closed loop balanced reduction. *Proceedings of the 2nd European Control Conference*, 697–701.
- Chaturantabut, S., & Sorensen, D. C. (2010). Reduction via discrete empirical interpolation. *SIAM Journal on Scientific Computing*, 32(5), 2737–2764.
- Cheng, X., Kawano, Y., & Scherpen, J. M. A. (2019). Model reduction of multiagent systems using dissimilarity-based clustering. *IEEE Transactions on Automatic Control*, 64(4), 1663–1670.
- Cheng, X., Scherpen, J. M. A., & Besselink, B. (2019). Balanced truncation of networked linear passive systems. *Automatica*, 104, 17–25.
- Cheng, X., Scherpen, J. M. A., & Zhang, F. (2019). Model reduction of synchronized homogeneous lur'e networks with incrementally sector-bounded nonlinearities. *European Journal of Control*, 50, 11–19.
- Craig Jr., R. R. (2000). Coupling of substructures for dynamic analysis: an overview. In *Proceedings of the 41st AIAA/ASME/ASCE/AHS/aSC structures, structural dynamics, and materials conference*. Atlanta, USA.
- de Klerk, D., Rixen, D. J., & Voormeeren, S. N. (2003). Model reduction methods based on Krylov subspaces. *Acta Numerica*, 12, 267–319.
- de Klerk, D., Rixen, D. J., & Voormeeren, S. N. (2008). General framework for dynamics substructuring: history, review and classification of techniques. *AIAA Journal*, 46(5), 1169–1181.
- Deenen, D. A., Sharif, B., van den Eijnden, S., Nijmeijer, H., Heemels, W. P. M. H., & Heertjes, M. F. (2021). Projection-based integrators for improved motion control: Formalization, well-posedness and stability of hybrid integrator-gain systems. *Automatica*, 133, Article 109830.
- Desai, U. B., & Pal, D. (1984). A Transformation Approach to Stochastic Model Reduction. *IEEE Transactions on Automatic Control*, 29(12), 1097–1100.
- Dou, Y., Cheng, X., & Scherpen, J. M. A. (2024). Stability-preserving model reduction of networked lur'e systems. In *2024 European control conference* (pp. 2059–2064). IEEE.
- Drummond, R., Turner, M. C., & Duncan, S. R. (2022). Reduced-order neural network synthesis with robustness guarantees. *IEEE Transactions on Neural Networks and Learning Systems*, 35(1), 1182–1191.
- Egelsmeers, R. A., Janssen, L. A. L., Fey, R. H. B., Gerritsen, J. W., & van de Wouw, N. (2024). Reduced-order modeling of modular, position-dependent systems with translating interfaces. *Mechatronics*, 102, Article 103224.
- El Ghaoui, L., Gu, F., Travacca, B., Askari, A., & Tsai, A. (2021). Implicit deep learning. *SIAM Journal on Mathematics of Data Science*, 3(3), 930–958.
- Enns, D. F. (1984). Model reduction with balanced realizations: an error bound and a frequency weighted generalization. In *Proceedings of the 23rd IEEE conference on decision and control, las vegas, USA* (pp. 127–132).
- Faassen, R. P. H., van de Wouw, N., Oosterling, J. A. J., & Nijmeijer, H. (2003). Prediction of regenerative chatter by modelling and analysis of high-speed milling. *International Journal of Machine Tools and Manufacture*, 43(14), 1437–1446.
- Filanova, Y., Pontes Duff, I., Goyal, P., & Benner, P. (2023). An operator inference oriented approach for linear mechanical systems. *Mechanical Systems and Signal Processing*, 200, Article 110620.
- Fujimoto, K., & Scherpen, J. M. A. (2010). Balanced realization and model order reduction for nonlinear systems based on singular value analysis. *SIAM Journal on Control and Optimization*, 48(7), 4591–4623.
- Gallivan, K. A., Grimme, E., & Van Dooren, P. M. (1999). Model reduction of large-scale systems: rational krylov versus balancing techniques. In H. Bulgak, & C. Zenger (Eds.), *Error control and adaptivity in scientific computing* (pp. 177–190). Kluwer Academic.
- Gallivan, K. A., Vandendorpe, A., & Van Dooren, P. M. (2004). Model reduction of MIMO systems via tangential interpolation. In *SIAM journal on matrix analysis and applications SIAM Journal on Matrix Analysis and Applications*, 26(2), 328–349.
- Germay, C., Denoël, V., & Detournay, E. (2009). Multiple mode analysis of the self-excited vibrations of rotary drilling systems. *Journal of Sound and Vibration*, 325, 362–381.
- Glover, K. (1984). All optimal Hankel-norm approximations of linear multivariable systems and their L^∞ -error bounds. *International Journal of Control*, 39(6), 1115–1193.
- Gugercin, S., & Antoulas, A. C. (2004). A survey of model reduction by balanced truncation and some new results. *International Journal of Control*, 77(8), 748–766.
- Harshavardhana, P., Jonckheere, E. A., & Silverman, L. M. (1984). Stochastic balancing and approximation - stability and minimality. *IEEE Transactions on Automatic Control*, AC-29(8), 744–746.
- Hazeleger, L., Haring, M. A. M., & van de Wouw, N. (2020). Extremum-seeking control for optimization of time-varying steady-state responses of nonlinear systems. *Automatica*, 119, Article 109068.
- Herting, D. N. (1985). A general purpose, multi-stage, component modal synthesis method. *Finite Elements in Analysis and Design*, 1(2), 153–164.
- Hesthaven, J. S., Rozza, G., & Stamm, B. (2016). *Certified reduced basis methods for parametrized partial differential equations*. Springer.
- Horn, R. A., & Johnson, C. R. (2013). *Matrix analysis* (2nd ed.). Cambridge, United Kingdom: Cambridge University Press.
- Inspurger, T., & Stépán, G. (2000). Stability of the milling process. *Periodica Polytechnica Mechanical Engineering*, 44(1), 47–57.
- Ishizaki, T., Kashima, K., Imura, J., & Aihara, K. (2014). Model reduction and clusterization of large-scale bidirectional networks. *IEEE Transactions on Automatic Control*, 59(1), 48–63.
- Janssen, L. A. L., Besselink, B., Fey, R. H. B., & van de Wouw, N. (2023). Modular model reduction of interconnected systems: A top-down approach. *IFAC-PapersOnLine*, 56(2), 4246–4251.
- Janssen, L. A. L., Besselink, B., Fey, R. H. B., & van de Wouw, N. (2024a). Modular model reduction of interconnected systems: A robust performance analysis perspective. *Automatica*, 160, Article 111423.
- Janssen, L. A. L., Fey, R. H. B., Besselink, B., Gerritsen, J., Kostić, D., & van de Wouw, N. (2023). Managing complexity in interconnected system models. *DSPE Mikroniek*, 63(6), 5–9.
- Janssen, L. A. L., Fey, R. H. B., Besselink, B., & van de Wouw, N. (2024b). Modular redesign of mechatronic systems: Formulation of module specifications guaranteeing system dynamics specifications. *Mechatronics*, 103, Article 103236.
- Janssen, L. A. L., Fey, R. H. B., Besselink, B., & van de Wouw, N. (2024c). Mode selection for component mode synthesis with guaranteed assembly accuracy. *Journal of Sound and Vibration*, 589, Article 118596.
- Jonckheere, E. A., & Silverman, L. M. (1983). A New Set of Invariants for Linear Systems—Application to Reduced Order Compensator Design. *IEEE Transactions on Automatic Control*, 28(10), 953–964.
- Jongsma, H.-J., Mlinarić, P., Grundel, S., Benner, P., & Trentelman, H. L. (2018). Model reduction of linear multi-agent systems by clustering with H_2 and H_∞ error bounds. *Mathematics of Control, Signals, and Systems*, 30(6).
- Kawano, Y. (2022). Controller reduction for nonlinear systems by generalized differential balancing. *IEEE Transactions on Automatic Control*, 67(11), 5856–5871.
- Kawano, Y., & Scherpen, J. M. A. (2015). Model reduction by generalized differential balancing. In M. K. Camlibel, A. A. Julius, R. Pasumath, & J. M. A. Scherpen (Eds.), *Lecture notes in control and information sciences: Vol. 461, Mathematical control theory i* (pp. 349–362). Springer International Publishing, Cham, Switzerland.
- Kawano, Y., & Scherpen, J. M. A. (2017). Model reduction by differential balancing based on nonlinear Hankel operators. *IEEE Transactions on Automatic Control*, 62(7), 3293–3308.
- Khalil, H. K. (1996). *Nonlinear systems*. New Jersey, USA: Prentice-Hall.
- Khalin, A., Efimov, D., & Ushirobira, R. (2023). On observer design for a class of Persidskii systems based on steady-state estimation. *IFAC-PapersOnLine*, 56(2), 10805–10810.
- Knockaert, L., Dhaene, T., Ferranti, F., & de Zutter, D. (2011). Model order reduction with preservation of passivity, non-expansivity and Markov moments. *Systems & Control Letters*, 60(1), 53–61.
- Lohmiller, W., & Slotine, J.-J. E. (1998). On contraction analysis for non-linear systems. *Automatica*, 34(6), 683–696.
- Lutowska, A. (2012). *Model order reduction for coupled systems using low-rank approximations* (Ph.D. thesis), Eindhoven University of Technology, Eindhoven, the Netherlands.

- Martinelli, D., Galimberti, C., Lucia, Manchester, I. R., Furieri, L., & Ferrari-Trecate, G. (2023). Unconstrained parametrization of dissipative and contracting neural ordinary differential equations. In *2023 62nd IEEE conference on decision and control* (pp. 3043–3048).
- Monshizadeh, N., Trentelman, H. L., & Camlibel, M. K. (2013). Stability and synchronization preserving model reduction of multi-agent systems. *Systems & Control Letters*, 62(1), 1–10.
- Monshizadeh, N., Trentelman, H. L., & Camlibel, M. K. (2014). Projection-based model reduction of multi-agent systems using graph partitions. *IEEE Transactions on Control of Network Systems*, 1(2), 145–154.
- Moore, B. C. (1981). Principal component analysis in linear systems - controllability, observability, and model reduction. *IEEE Transactions on Automatic Control*, AC-26(1), 17–32.
- Moreschini, A., & Astolfi, A. (2024). Closed-loop interpolation by moment matching for linear and nonlinear systems. *IEEE Transactions on Automatic Control*.
- Mullis, C. T., & Roberts, R. A. (1976). Synthesis of minimum roundoff noise fixed point digital filters. *IEEE Transactions on Circuits and Systems*, CAS-23(9), 551–562.
- Muscato, G., Nunnari, G., & Fortuna, L. (1997). Singular perturbation approximation of bounded real balanced and stochastically balanced transfer matrices. *International Journal of Control*, 66(2), 253–270.
- Nakatsukasa, Y., Sète, O., & Trefethen, L. N. (2018). The AAA algorithm for rational approximation. *SIAM Journal on Scientific Computing*, 40(3), A1494–A1522.
- Ober, R. (1991). Balanced parametrization of classes of linear systems. *SIAM Journal on Control and Optimization*, 29(6), 1251–1287.
- Opdenacker, P. C., & Jonckheere, E. A. (1988). A contraction mapping preserving balanced reduction scheme and its infinity norm error bounds. *IEEE Transactions on Circuits and Systems*, 35(2), 184–189.
- Padoan, A., Forni, F., & Sepulchre, R. (2021). Model reduction of dominant feedback systems. *Automatica*, 130, Article 109695.
- Pavlov, A. V., Steur, E., & van de Wouw, N. (2022). Nonlinear integral coupling for synchronization in networks of nonlinear systems. *Automatica*, 140.
- Pavlov, A., & van de Wouw, Nathan (2008). Fast computation of frequency response functions for a class of nonlinear systems. In *Proceedings of the 47th conference on decision and control* (pp. 1180–1186).
- Pavlov, Alexey, van de Wouw, Nathan, & Nijmeijer, Henk (2006). *Uniform output regulation of nonlinear systems: a convergent dynamics approach*. Boston, USA: Springer Science & Business Media, Birkhäuser.
- Pavlov, A. V., van de Wouw, N., & Nijmeijer, H. (2007). Frequency response functions for nonlinear convergent systems. *IEEE Transactions on Automatic Control*, 52(6).
- Pavlov, A., Hunnekens, B. G. B., van de Wouw, N., & Nijmeijer, H. (2013). Steady-state performance optimization for nonlinear control systems of Lur'e type. *Automatica*, 49(7), 2087–2097.
- Pavlov, A., Pogromsky, A., van de Wouw, N., & Nijmeijer, H. (2004). Convergent dynamics, a tribute to boris pavlovich demidovich. *Systems & Control Letters*, 52(3), 257–261.
- Pernebo, L., & Silverman, L. M. (1982). Model reduction via balanced state space representations. *IEEE Transactions on Automatic Control*, AC-27(2), 382–387.
- Phillips, J. R., Daniel, L., & Silveira, L. M. (2003). Guaranteed passive balancing transformations for model order reduction. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, 22(8), 1027–1041.
- Polyuga, R. V., & van der Schaft, A. (2010). Structure preserving model reduction of port-Hamiltonian systems by moment matching at infinity. *Automatica*, 46(4), 665–672.
- Poort, L., Besselink, B., Fey, R. H. B., & van de Wouw, N. (2024). Balancing-based reduction for interconnected passive systems. *IEEE Transactions on Control Systems Technology*, 32(5), 1817–1826.
- Poort, L., Besselink, B., Fey, R. H. B., & van de Wouw, N. (2025). Efficient reduction of interconnected subsystem models using abstracted environments. *European Journal of Control*, 85, Article 101257.
- Poort, L., A.L., Janssen, B., Besselink, Fey, R. H. B., & van de Wouw, N. (2025). Abstracted model reduction: a general framework for efficient interconnected system reduction. *IEEE Transactions on Control Systems Technology (Digital Preprint)*, doi: 10.1109/TCST.2025.3550027.
- Reis, T., & Heinkenschloss, M. (2009). Model reduction with a-priori error bounds for a class of nonlinear electrical circuits. In *Proceedings of the joint 48th IEEE conference on decision and control and 28th Chinese control conference, shanghai, China* (pp. 5376–5383).
- Reis, T., & Stykel, T. (2008). A survey on model reduction of coupled systems. In W. H. A. Schilders, H. A. van der Vorst, & J. Rommes (Eds.), *Mathematics in industry: Vol. 13, Model order reduction: theory, research aspects and applications* (pp. 133–155). Berlin Heidelberg, Germany: Springer-Verlag.
- Reis, T., & Stykel, T. (2010). Positive real and bounded real balancing for model reduction of descriptor systems. *International Journal of Control*, 83(1), 74–88.
- Revay, M., Wang, R., & Manchester, I. R. (2023). Recurrent equilibrium networks: Flexible dynamic models with guaranteed stability and robustness. *IEEE Transactions on Automatic Control*, 1–16.
- Sandberg, H., & Murray, R. M. (2009). Model reduction of interconnected linear systems. *Optimal Control Applications and Methods*, 30(3), 225–245.
- Sarkar, A., & Scherpen, J. M. A. (2022). Extended differential balancing for nonlinear dynamical systems. *IEEE Control Systems Letters*, 6, 3170–3175.
- Scarciotti, Giordano (2016). Low computational complexity model reduction of power systems with preservation of physical characteristics. *IEEE Transactions on Power Systems*, 32(1), 743–752.
- Scarciotti, G., & Astolfi, A. (2017a). Data-driven model reduction by moment matching for linear and nonlinear systems. *Automatica*, 79, 340–351.
- Scarciotti, Giordano, & Astolfi, Alessandro (2017b). Data-driven model reduction by moment matching for linear and nonlinear systems. *Automatica*, 79, 340–351.
- Scarciotti, Giordano, & Astolfi, Alessandro (2017c). Nonlinear model reduction by moment matching. *Foundations and Trends® in Systems and Control*, 4(3–4), 224–409.
- Scarciotti, Giordano, & Astolfi, Alessandro (2024). Interconnection-based model order reduction - a survey. *European Journal of Control*, 75, Article 100929.
- van der Schaft, A. J. (2017). L_2 -gain and passivity techniques in nonlinear control, In *Communications and control engineering series*, (3rd ed.). Springer International Publishing, Cham, Switzerland.
- Scherer, C. W. (1990). *The riccati inequality and state-space H_∞ -optimal control*. (Ph.D. thesis), Universität Würzburg, Germany.
- Scherpen, J. M. A. (1993). Balancing for nonlinear systems. *Systems & Control Letters*, 21, 143–153.
- Schilders, W. H. A., & Lutowska, A. (2014). A novel approach to model order reduction for coupled multi- physics problems. In *Modeling, simulation and applications book series*, Springer.
- Schilders, W. H. A., van der Vorst, H. A., & Rommes, J. (Eds.), (2008). Model order reduction: Theory, research aspects and applications. In *Mathematics in industry: Vol. 13*, Berlin Heidelberg, Germany: Springer-Verlag.
- Shakib, M. F. (2022). *Data-driven modeling and complexity reduction for nonlinear systems with stability guarantees* (Ph.D. thesis), Eindhoven University of Technology, Eindhoven, the Netherlands.
- Shakib, M. F., Moreschini, A., & Scarciotti, G. (2025). Dissipativity-preserving model reduction for linear systems using moment matching. In *2025 European control conference*. Accepted for publication in the European Journal of Control.
- Shakib, M. F., Pogromsky, A. Y., Pavlov, A., & van de Wouw, N. (2019). Fast identification of continuous-time lur'e-type systems with stability certification. *IFAC-PapersOnLine*, 52(16), 227–232.
- Shakib, M. F., Pogromsky, A. Y., Pavlov, A. V., & van de Wouw, N. (2022). Computationally efficient identification of continuous-time Lur'e-type systems with stability guarantees. *Automatica*, 136, Article 110012.
- Shakib, M. F., Scarciotti, G., & Astolfi, A. (2024). A parameterised family of neuralODEs optimally fitting steady-state data. In *The proceedings of the 2024 American control conference* (pp. 3359–3364). IEEE.
- Shakib, M. F., Scarciotti, G., Jungers, M., Pogromsky, A. Y., Pavlov, A. V., & van de Wouw, N. (2021). Optimal H_∞ LMI-Based Model Reduction by Moment Matching for Linear Time-Invariant Models. In *Proceedings of the 60th conference on decision and control* (pp. 6914–6919).
- Shakib, M. F., Scarciotti, G., Jungers, M., Pogromsky, A. Y., Pavlov, A. V., & van de Wouw, N. (2024). Optimal model reduction by time-domain moment matching for Lur'e-Type models. *IEEE Transactions on Automatic Control*, 69(12), 8820–8827.
- Shakib, M. F., Scarciotti, G., Pogromsky, A. Y., Pavlov, A. V., & van de Wouw, N. (2023a). Model reduction by moment matching with preservation of global stability for a class of nonlinear models. *Automatica*, 157, Article 111227.
- Shakib, M. F., Scarciotti, G., Pogromsky, A. Y., Pavlov, A. V., & van de Wouw, N. (2023b). Time-domain moment matching for multiple-input multiple-output linear time-invariant models. *Automatica*, 152, Article 110935.
- Shanno, D. F. (1970). Conditioning of quasi-Newton methods for function minimization. *Mathematics of Computation*, 24(111), 647–656.
- Sturk, C., Vanfretti, L., Chompoobutgool, Y., & Sandberg, H. (2014). Coherency-independent structured model reduction of power systems. *IEEE Transactions on Power Systems*, 29(5), 2418–2426.
- Suy, H. M. R., Fey, R. H. B., Galanti, F. M. B., & Nijmeijer, H. (2007). Nonlinear dynamic analysis of a structure with a friction-based seismic base isolation system. *Nonlinear Dynamics*, 50, 523–538.
- Suykens, J. A. K., Vandewalle, J., & De Moor, B. (1999). Lur'e systems with multilayer perceptron and recurrent neural networks: absolute stability and dissipativity. *IEEE Transactions on Automatic Control*, 44(4), 770–774.
- Trentelman, H. L. (2009). Bounded real and positive real balanced truncation using Σ -normalised coprime factors. *Systems & Control Letters*, 58(12), 871–879.
- Unneland, K., Van Dooren, P., & Egeland, O. (2007). New schemes for positive real truncation. *Modeling, Identification and Control*, 28(3), 53–67.
- van de Wouw, N., & Leine, R. I. (2010). Stability and Control of Lur'e-type Measure Differential Inclusions. In G. Leonov, H. Nijmeijer, A. Pogromsky, & A. Fradkov (Eds.), *Dynamics and control of hybrid mechanical systems* (pp. 129–151). World-Scientific, Heidelberg.
- van de Wouw, N., Michiels, W., & Besselink, B. (2015a). Model reduction for a class of nonlinear delay differential equations with time-varying delays. In *Proceedings of the 54th IEEE conference on decision and control, Osaka, Japan* (pp. 6422–6428).
- van de Wouw, N., Michiels, W., & Besselink, B. (2015b). Model reduction for delay differential equations with guaranteed stability and error bound. *Automatica*, 55, 132–139.
- van de Wouw, N., Pastink, H. A., Heertjes, M. F., Pavlov, A. V., & Nijmeijer, H. (2008). Performance of convergence-based variable-gain control of optical storage drives. *Automatica*, 44(1), 15–27.

- Vandendorpe, A., & Van Dooren, P. (2008). Model Reduction of Interconnected Systems. In *Model order reduction: theory, research aspects and applications* (pp. 305–321). Springer Berlin, Heidelberg.
- Vaz, A. F., & Davison, E. J. (1990). Modular model reduction for interconnected systems. *Automatica*, 26(2), 251–261.
- Willems, J. C. (1972a). Dissipative Dynamical Systems Part I: General Theory. *Archive for Rational Mechanics and Analysis*, 45(5), 321–351.
- Willems, J. C. (1972b). Dissipative Dynamical Systems Part II: Linear Systems with Quadratic Supply Rates. *Archive for Rational Mechanics and Analysis*, 45(5), 352–393.
- Wortelboer, P. M. R., & Bosgra, O. H. (1994). Frequency weighted closed-loop order reduction in the control design configuration. *Proceedings of the IEEE Conference on Decision and Control*, 3, 2714–2719.
- Yakubovich, V. A. (1964). Matrix inequalities method in stability theory for nonlinear control systems: I. Absolute stability of forced vibrations. *Automation and Remote Control*, 7, 905–917.
- Yakubovich, V. A., Leonov, G. A., & Gelig, A. K. (2004). *Approximation of large-scale dynamical systems*. World Scientific Publishing.
- Yin, H., Seiler, P., & Arcak, M. (2021). Stability analysis using quadratic constraints for systems with neural network controllers. *IEEE Transactions on Automatic Control*, 67(4), 1980–1987.
- Zamani, M., & Arcak, M. (2017). Compositional abstraction for networks of control systems: A dissipativity approach. *IEEE Transactions on Control of Network Systems*, 5(3), 1003–1015.
- Zames, G. (1966). On the input-output stability of time-varying nonlinear feedback systems. Part i: Conditions derived using concepts of loop gain, conicity, and positivity. *IEEE Transactions on Automatic Control*, AC-11(2), 228–238.
- Zhou, Kemin, & Doyle, John C. (1998). *essentials of robust control* (1st ed.). Prentice Hall.
- Zhou, K., Doyle, J. C., & Glover, K. (1996). *Robust and optimal control*. Prentice Hall, Upper Saddle River, USA.