



**HAL**  
open science

# Data-Driven Model Order Reduction of Linear Switched Systems in the Loewner Framework

I. Gosea, Mihaly Petreczky, A. Antoulas

► **To cite this version:**

I. Gosea, Mihaly Petreczky, A. Antoulas. Data-Driven Model Order Reduction of Linear Switched Systems in the Loewner Framework. *SIAM Journal on Scientific Computing*, 2018, 40 (2), pp.B572 - B610. 10.1137/17M1120233 . hal-01931178

**HAL Id: hal-01931178**

**<https://hal.science/hal-01931178>**

Submitted on 2 Sep 2021

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Data-driven model order reduction of linear switched systems in the Loewner framework

I.V. Gosea<sup>†</sup>, M. Petreczky<sup>‡</sup>, A.C. Antoulas<sup>§</sup> <sup>†</sup>.

January 26, 2018

## Abstract

The Loewner framework for model reduction is extended to the class of linear switched systems. One advantage of this framework is that it introduces a trade-off between accuracy and complexity. Moreover, through this procedure, one can derive state-space models directly from data which is related to the input-output behavior of the original system. Hence, another advantage of the framework is that it does not require the initial system matrices. More exactly, the data used in this framework consists in frequency domain samples of input-output mappings of the original system. The definition of generalized transfer functions for linear switched systems resembles the one for bilinear systems. A key role is played by the coupling matrices, which ensure the transition from one active mode to another.

## 1 Introduction

Model order reduction (MOR) seeks to transform large, complicated models of time dependent processes into smaller, simpler models that are nonetheless capable of representing accurately the behavior of the original process under a variety of operating conditions. The goal is an efficient, methodical strategy that yields a dynamical system evolving in a substantially lower dimension space (hence requiring far less computational resources for realization), yet retaining response characteristics close to the original system. Such reduced order models could be used as efficient surrogates for the original model, replacing it as a component in larger simulations.

Hybrid systems are a class of nonlinear systems which result from the interaction of continuous time dynamical sub-systems with discrete events. More precisely, a hybrid system is a collection of continuous time dynamical systems. The internal variable of each dynamical system is governed by a set of differential equations. Each of the separate continuous time systems are labeled as a discrete mode. The transitions between the discrete states may result in a jump in the continuous internal variable. Linear switched systems (in short LSS) constitute a subclass of hybrid systems;

---

<sup>†</sup>Data-Driven System Reduction and Identification Group, Max Planck Institute for Dynamics of Complex Technical Systems, Sandtorstrasse 1, 39106, Magdeburg, Germany ([gosea@mpi-magdeburg.mpg.de](mailto:gosea@mpi-magdeburg.mpg.de))

<sup>‡</sup>Centre de Recherche en Informatique, Signal et Automatique de Lille (CRISTAL), UMR CNRS 9189, CNRS, Ecole Centrale de Lille, France ([mihaly.petreczky@ec-lille.fr](mailto:mihaly.petreczky@ec-lille.fr))

<sup>§</sup>Department of Electrical and Computer Engineering, Rice University, 6100 Main St, MS-366, Houston, TX 77005, USA ([aca@rice.edu](mailto:aca@rice.edu))

the main property is that these systems switch among a finite number of linear subsystems. Also, the discrete events interacting with the sub-systems are governed by a piecewise continuous function called the switching signal.

Hybrid and switched systems are powerful models for distributed embedded systems design where discrete controls are routinely applied to continuous processes. However, the complexity of verifying and assessing general properties of these systems is very high so that the use of these models is limited in applications where the size of the state-space is not too large. To cope with complexity, abstraction and reduction are useful techniques. In this paper we mainly analyze the reduction part.

In the past years, hybrid and switched systems have received increasing attention in the scientific community. For a detailed characterization of this relatively new class of dynamical systems, we refer the readers to the books [25], [38], [39] and [19]. Such systems are used in modeling, analysis and design of supervisory control systems, mechanical systems with impact, circuits with relays or ideal diodes.

The study of the properties of hybrid systems in general and switched systems in particular is still the subject of intense research, including the problems of stability (see [15, 38]), realization including observability/controllability (see [31, 32]), analysis of switched differential-algebraic equations, or in short DAE's (see [27, 40]) and of its numerical solutions (see [20]). Recently, considerable research has been dedicated to the problem of MOR for linear switched systems. The most prolific method that has been applied is balanced truncation (or a Gramian-based derivation of it). Techniques that are based on balancing have been considered in the following: [18, 13, 10, 37, 28, 34, 30]. Also, another class of methods involve matching of generalized Markov parameters (known also as time domain Krylov methods) such as the ones in [8, 7];  $\mathcal{H}_\infty$  type of reduction methods were developed in [42, 11, 43]. Finally, we mention some publications that are focused on the reduction of discrete LSS, such as [41, 12].

A linear switched system involves switching between a number of linear systems (the modes of the LSS). Hence, to apply balanced truncation techniques to a switched linear system, one may seek a basis of the state space such that the corresponding modes are all in balanced form. It may happen that some state components of the LSS are difficult to reach and observe in some of the modes yet easy to reach and observe in others. In that case, deciding how to truncate the state variables and obtain a reduced order model is not trivial. A solution to this problem is proposed in [28] where it is shown that the average Gramian can be used to obtain a reduced order model. This method will be used as a comparison tool for our new MOR method.

In this paper, we focus on extending the Loewner framework (see [6]) for reducing linear switched systems. This method can be viewed as a special subclass of rational Krylov methods, also referred to as moment matching or interpolatory methods. Roughly speaking, in the linear case, interpolatory methods seek reduced models whose transfer function (and possibly some of its derivatives) matches the transfer function (and possibly some of its derivatives) of the original system at selected frequencies. For the nonlinear case, these methods require appropriate definitions of transfer functions.

The paper is organized as follows. In Section 2, the formal definition of continuous-time linear switched systems is provided. Furthermore, we introduce the generalized transfer functions for LSS as input-output mappings in frequency domain. Section 3 includes a brief introduction of the Loewner framework for the class of linear systems with no switching. In Section 4, we present the extension of the Loewner framework for LSS with 2 modes. This is done in order to familiarize the reader with basic ideas without having to use heavy notation. Then, in Section 5, we propose extensions of the results introduced in the previous section to the general case of LSS that switch

amongst  $D \geq 2$  modes. Afterwards, in Section 6, we discuss the practical applicability of the new introduced method by means of three numerical examples (one of which is large scale). In those examples, we compare our algorithms with the balanced truncation algorithm of [30] for illustration purposes. Finally, in Section 7, we present a summary of the findings and the conclusions.

## 2 Linear switched systems

**Definition 2.1** *A continuous time linear switched system (LSS) is a dynamical system described by the following equations*

$$\Sigma : \begin{cases} \mathbf{E}_{\sigma(t)} \dot{\mathbf{x}}(t) = \mathbf{A}_{\sigma(t)} \mathbf{x}(t) + \mathbf{B}_{\sigma(t)} \mathbf{u}(t), & \mathbf{x}(0) = \mathbf{x}_0, \\ \mathbf{y}(t) = \mathbf{C}_{\sigma(t)} \mathbf{x}(t), \end{cases} \quad (1)$$

where  $\mathbf{Q} = \{1, 2, \dots, D\}$ ,  $D > 1$ , is a set of discrete modes,  $\sigma(t)$  is the switching signal,  $\mathbf{u}$  is the input,  $\mathbf{x}$  is the state, and  $\mathbf{y}$  is the output.

The system matrices  $\mathbf{E}_q, \mathbf{A}_q \in \mathbb{R}^{n_q \times n_q}$ ,  $\mathbf{B}_q \in \mathbb{R}^{n_q \times m}$ ,  $\mathbf{C}_q \in \mathbb{R}^{p \times n_q}$ , where  $q \in \mathbf{Q}$ , correspond to the linear system active in mode  $q \in \mathbf{Q}$ , and  $\mathbf{x}_0 \in \mathbb{R}^{n_{q_s}}$  is the initial state. Here,  $n_1, n_2, \dots, n_D, m$  and  $p$  are positive integers and  $q_s \in \mathbf{Q}$  is the mode in which the system is initialized. We consider the  $\mathbf{E}_q$  matrices to be invertible. Furthermore, the transition from one mode to another is made via the so called switching or coupling matrices  $\mathbf{K}_{q, \tilde{q}} \in \mathbb{R}^{n_{\tilde{q}} \times n_q}$  where  $q, \tilde{q} \in \mathbf{Q}$ .

**Remark 2.1** The case for which the coupling is made between identical modes is excluded, Hence, when  $q = \tilde{q}$ , consider that the coupling matrices are identity matrices, i.e.  $\mathbf{K}_{q,q} = \mathbf{I}_{n_q}$ .

The notation  $\Sigma = (n_1, n_2, \dots, n_D, \{(\mathbf{E}_q, \mathbf{A}_q, \mathbf{B}_q, \mathbf{C}_q) | q \in \mathbf{Q}\}, \{\mathbf{K}_{q, \tilde{q}} | q, \tilde{q} \in \mathbf{Q}\}, \mathbf{x}_0)$  is used as a short-hand representation for LSS's described by the equations in (1). The vector  $\mathbf{n} \in \mathbb{N}^D$ , where  $\mathbf{n} = [n_1 \ n_2 \ \dots \ n_D]$  is the dimension (order) of  $\Sigma$ . The linear system which is active in the  $q^{\text{th}}$  mode of  $\Sigma$  is denoted with  $\Sigma_q$  and it is described by (where  $1 \leq q \leq D$ )

$$\Sigma_k : \begin{cases} \mathbf{E}_q \dot{\mathbf{x}}_q(t) = \mathbf{A}_q \mathbf{x}_q(t) + \mathbf{B}_q \mathbf{u}(t), & \mathbf{x}(t_k) = \mathbf{x}_k, \\ \mathbf{y}(t) = \mathbf{C}_q \mathbf{x}_q(t). \end{cases} \quad (2)$$

The restriction of the switching signal  $\sigma(t)$  to a finite interval of time  $[0, T]$  can be interpreted as a finite sequence of elements of  $\mathbf{Q} \times \mathbb{R}_+$  of the form:

$$\nu(\sigma) = (q_1, t_1)(q_2, t_2) \dots (q_k, t_k),$$

where  $q_1, \dots, q_k \in \mathbf{Q}$  and  $0 < t_1 < t_2 < \dots < t_k \in \mathbb{R}_+$ ,  $t_1 + \dots + t_k = T$ , such that for all  $t \in [0, T]$  we have:

$$\sigma(t) = \begin{cases} q_1 & \text{if } t \in [0, t_1), \\ q_2 & \text{if } t \in [t_1, t_1 + t_2), \\ \vdots & \\ q_i & \text{if } t \in [t_1 + \dots + t_{i-1}, t_1 + \dots + t_{i-1} + t_i), \\ \vdots & \\ q_k & \text{if } t \in [t_1 + \dots + t_{k-1}, t_1 + \dots + t_{k-1} + t_k). \end{cases}$$

In short, by denoting  $T_i := t_1 + \dots + t_{i-1} + t_i$ ,  $T_0 := 0$ ,  $T_k := T$ , write:

$$\sigma(t) = \begin{cases} q_1 & \text{if } t \in [0, T_1), \\ q_i & \text{if } t \in [T_{i-1}, T_i), \quad i > 2. \end{cases}$$

Denote by  $\mathcal{P}_C(\mathbb{R}_+, \mathbb{R}^n)$ ,  $\mathcal{P}_c(\mathbb{R}_+, \mathbb{R}^n)$ , the set of all piecewise-continuous, and piecewise-constant functions, respectively.

**Definition 2.2** A tuple  $(\mathbf{x}, \mathbf{u}, \sigma, \mathbf{y})$ , where  $\mathbf{x} : \mathbb{R}_+ \rightarrow \bigcup_{i=1}^D \mathbb{R}^{n_i}$ ,  $\mathbf{u} \in \mathcal{P}_C(\mathbb{R}_+, \mathbb{R}^m)$ ,  $\sigma \in \mathcal{P}_c(\mathbb{R}_+, \mathbf{Q})$ ,  $\mathbf{y} \in \mathcal{P}_C(\mathbb{R}_+, \mathbb{R}^p)$  is called a solution, if the following conditions simultaneously hold:

1. The restriction of  $\mathbf{x}(t)$  to  $[T_{i-1}, T_i)$  is differentiable, and satisfies  $\mathbf{E}_{q_i} \dot{\mathbf{x}}(t) = \mathbf{A}_{q_i} \mathbf{x}(t) + \mathbf{B} \mathbf{u}(t)$ .
2. Furthermore, when switching from mode  $q_i$  to mode  $q_{i+1}$  at time  $T_i$ , the following holds

$$\mathbf{E}_{q_{i+1}} \mathbf{x}(T_i) = \mathbf{K}_{q_i, q_{i+1}} \lim_{t \nearrow T_i} \mathbf{x}(t).$$

3. Moreover, for all  $t \in \mathbb{R}$ ,  $\mathbf{y}(t) = \mathbf{C}_{\sigma(t)} \mathbf{x}(t)$  holds.

The switching matrices  $\mathbf{K}_{q_i, q_{i+1}}$  allow having different dimensions for the subsystems active in different modes. For instance, the pencil  $(\mathbf{A}_{q_i}, \mathbf{E}_{q_i}) \in \mathbb{R}^{n_{q_i} \times n_{q_i}}$ , while the pencil  $(\mathbf{A}_{q_{i+1}}, \mathbf{E}_{q_{i+1}}) \in \mathbb{R}^{n_{q_{i+1}} \times n_{q_{i+1}}}$  where the values  $n_{q_i}$  and  $n_{q_{i+1}}$  need not be the same. If the  $\mathbf{K}_{q_i, q_{i+1}}$  matrices are not explicitly given, it is considered that they are identity matrices.

The input-output behavior of an LSS can be formalized in time domain as a map  $\mathbf{f}(\mathbf{u}, \sigma)(t)$ . This particular map can be written in *generalized kernel representation* (as suggested in [33]) using the unique family of analytic functions:  $\mathbf{g}_{q_1, \dots, q_k} : \mathbb{R}_+^k \rightarrow \mathbb{R}^p$  and  $\mathbf{h}_{q_1, \dots, q_k} : \mathbb{R}_+^k \rightarrow \mathbb{R}^{p \times m}$  with  $q_1, \dots, q_k \in \mathbf{Q}$ ,  $k \geq 1$  such that for all pairs  $(\mathbf{u}, \sigma)$  and for  $T = t_1 + t_2 + \dots + t_k$  we can write:

$$\mathbf{f}(u, \sigma)(t) = \mathbf{g}_{q_1, q_2, \dots, q_k}(t_1, t_2, \dots, t_k) + \sum_{i=1}^k \int_0^{t_i} \mathbf{h}_{q_i, q_{i+1}, \dots, q_k}(t_i - \tau, t_{i+1}, \dots, t_k) u(\tau + T_{i-1}) d\tau,$$

where the functions  $\mathbf{g}$ ,  $\mathbf{h}$  are defined for  $k \geq 1$ , as follows,

$$\mathbf{g}_{q_1, q_2, \dots, q_k}(t_1, t_2, \dots, t_k) = \mathbf{C}_{q_k} e^{\tilde{\mathbf{A}}_{q_k} t_k} \tilde{\mathbf{K}}_{q_{k-1}, q_k} e^{\tilde{\mathbf{A}}_{q_{k-1}} t_{k-1}} \tilde{\mathbf{K}}_{q_{k-2}, q_{k-1}} \dots \tilde{\mathbf{K}}_{q_1, q_2} e^{\tilde{\mathbf{A}}_{q_1} t_1} \mathbf{x}_0, \quad (3)$$

$$\mathbf{h}_{q_1, q_2, \dots, q_k}(t_1, t_2, \dots, t_k) = \mathbf{C}_{q_k} e^{\tilde{\mathbf{A}}_{q_k} t_k} \tilde{\mathbf{K}}_{q_{k-1}, q_k} e^{\tilde{\mathbf{A}}_{q_{k-1}} t_{k-1}} \tilde{\mathbf{K}}_{q_{k-2}, q_{k-1}} \dots \tilde{\mathbf{K}}_{q_1, q_2} e^{\tilde{\mathbf{A}}_{q_1} t_1} \tilde{\mathbf{B}}_1. \quad (4)$$

Note that, for the functions defined in (3) and (4) we consider the  $\mathbf{E}_{q_i}$  matrices to be incorporated into the  $\mathbf{A}_{q_i}$  and  $\mathbf{B}_{q_i}$  matrices (i.e.  $\tilde{\mathbf{A}}_{q_i} = \mathbf{E}_{q_i}^{-1} \mathbf{A}_{q_i}$ ,  $\tilde{\mathbf{B}}_{q_i} = \mathbf{E}_{q_i}^{-1} \mathbf{B}_{q_i}$ ). Moreover, the transformed coupling matrices are written accordingly  $\tilde{\mathbf{K}}_{q_i, q_{i+1}} = \mathbf{E}_{q_{i+1}}^{-1} \mathbf{K}_{q_i, q_{i+1}}$ .

In the rest of the paper, the LSS we are studying are assumed to have *zero initial conditions*, i.e.  $\mathbf{x}_0 = \mathbf{0}$ . Hence, only the  $\mathbf{h}$  functions in (4) are relevant for characterizing the input-output mapping  $\mathbf{f}$ .

The behavior of the input-output mappings in frequency domain is in turn characterized by a series of multivariate rational functions obtained by taking the multivariable Laplace transform of the regular kernels in (4), as for

$$\mathbf{H}_{q_1}(s_1) = \mathbf{C}_{q_1} \Phi_{q_1}(s_1) \mathbf{B}_{q_1}, \quad \mathbf{H}_{q_1, q_2}(s_1, s_2) = \mathbf{C}_{q_1} \Phi_{q_1}(s_1) \mathbf{K}_{q_2, q_1} \Phi_{q_2}(s_2) \mathbf{B}_{q_2},$$

$$\mathbf{H}_{q_1, q_2, q_3}(s_1, s_2, s_3) = \mathbf{C}_{q_1} \Phi_{q_1}(s_1) \mathbf{K}_{q_2, q_1} \Phi_{q_2}(s_2) \mathbf{K}_{q_3, q_2} \Phi_{q_3}(s_3) \mathbf{B}_{q_3}, \quad \dots$$

In general, for  $k \geq 3$ , write the *level  $k$  generalized transfer function* associated to the switching sequence  $(q_1, q_2, \dots, q_k)$ , and evaluated at the points  $(s_1, s_2, \dots, s_k)$  as,

$$\mathbf{H}_{q_1, q_2, \dots, q_k}(s_1, s_2, \dots, s_k) = \mathbf{C}_{q_1} \Phi_{q_1}(s_1) \mathbf{K}_{q_2, q_1} \Phi_{q_2}(s_2) \cdots \mathbf{K}_{q_k, q_{k-1}} \Phi_{q_k}(s_k) \mathbf{B}_{q_k}, \quad (5)$$

where  $\Phi_q(s) = (s\mathbf{E}_q - \mathbf{A}_q)^{-1}$ ,  $q_j \in \{1, 2, \dots, D\}$ ,  $1 \leq j \leq k$ . These functions are the *generalized transfer functions* of the linear switched system  $\Sigma$ .

By using their samples, we are able to directly construct (reduced) switched models that interpolate the original model - generalization of the Loewner framework to LSS.

We construct LSS reduced models by means of matching samples of input-output mappings corresponding to the original LSS and evaluated at finite sampling points (as opposed to other approaches - see [7, 8], where the behavior at infinity is studied instead, i.e. by matching Markov parameters). For the explicit derivation of these types of transfer functions, which is based on the so-called Volterra series representation, we refer the readers to [35].

**Remark 2.2** Conceptually, the initial state of an LSS is part of its definition, since it is impossible to define the input-output map of an LSS without fixing such a starting state. In this paper we consider only the case when the initial state is zero. Note that the classical Loewner framework (as introduced in [26]) uses only the input-output behavior from the zero initial state.

The proposed extension of the Loewner framework aims at finding a reduced system whose generalized transfer functions match those of the original LSS. That is, the proposed method takes into account only the input-output behavior from the zero initial state.

For the case of LSS, one could include non-zero initial states by including them as an additional column into one of the matrices  $\mathbf{B}_i$ , but the system theoretic interpretation of the thus obtained generalized transfer functions and their moments remains unclear.

**Remark 2.3** Certain issues might arise when the matrices  $\mathbf{E}_q$  are allowed to be singular such as the existence of a solution of the LSS (Definition 2.2) or the formal definition of time-domain kernels in (4). The case of descriptor LSS (with singular  $\mathbf{E}_q$  matrices) was treated in detail in [40]. In this paper we assume that the matrices  $\mathbf{E}_q$  are invertible to avoid further complications.

**Remark 2.4** The structure of the transfer functions in (5) is similar to the one of the functions recently introduced in [4], for the class of bilinear systems. This is because one can formulate an LSS as a bilinear system, by introducing additional input signals. Nevertheless, there are some restrictions, i.e. this works only for LSSs with no coupling matrices and that have the same state dimension in each mode. For the case  $D = 2$ , introduce the signal  $\hat{\mathbf{u}}(t) = q - 1$ , if the system operates in mode  $q \in \{1, 2\}$ . Then, write the dynamics of the LSS by merging the individual dynamics of the two modes, as

$$\dot{\mathbf{x}}(t) = \mathbf{A}_1 \mathbf{x}(t) + (\mathbf{A}_2 - \mathbf{A}_1) \mathbf{x}(t) \hat{\mathbf{u}}(t) + \mathbf{B}_1 \mathbf{u}(t) (1 - \hat{\mathbf{u}}(t)) + \mathbf{B}_2 \mathbf{u}(t) \hat{\mathbf{u}}(t),$$

or equivalently, to emphasize the bilinear multiple input format, as

$$\dot{\mathbf{x}}(t) = \mathbf{A}^{\text{bil}} \mathbf{x}(t) + \mathbf{N}_1^{\text{bil}} \mathbf{x}(t) \mathbf{u}(t) + \mathbf{N}_2^{\text{bil}} \mathbf{x}(t) \hat{\mathbf{u}}(t) + \mathbf{N}_3^{\text{bil}} \mathbf{x}(t) \mathbf{u}(t) \hat{\mathbf{u}}(t) + \mathbf{B}_1^{\text{bil}} \mathbf{u}(t) + \mathbf{B}_2^{\text{bil}} \hat{\mathbf{u}}(t) + \mathbf{B}_3^{\text{bil}} \mathbf{u}(t) \hat{\mathbf{u}}(t),$$

where  $\mathbf{A}^{\text{bil}} = \mathbf{A}_1$ ,  $\mathbf{N}_2^{\text{bil}} = \mathbf{A}_2 - \mathbf{A}_1$ ,  $\mathbf{N}_1^{\text{bil}} = \mathbf{N}_3^{\text{bil}} = \mathbf{0}$ ,  $\mathbf{B}_1^{\text{bil}} = \mathbf{B}_1$ ,  $\mathbf{B}_2 = \mathbf{0}$ ,  $\mathbf{B}_3^{\text{bil}} = \mathbf{B}_2 - \mathbf{B}_1$ . Hence, rewrite the above differential equation equivalently as,

$$\dot{\mathbf{x}}(t) = \mathbf{A}^{\text{bil}} \mathbf{x}(t) + \sum_{i=1}^3 \mathbf{N}_i^{\text{bil}} \mathbf{x}(t) \mathbf{u}_i(t) + \sum_{i=1}^3 \mathbf{B}_i^{\text{bil}} \mathbf{u}_i(t),$$

where the three control inputs are  $\mathbf{u}_1 = \mathbf{u}$ ,  $\mathbf{u}_2 = \hat{\mathbf{u}}$  and  $\mathbf{u}_3 = \mathbf{u}\hat{\mathbf{u}}$ . That is, solutions of LSSs are solutions of a bilinear system with a very specific structure and with specially chosen inputs. In particular, the continuous input  $\mathbf{u}$  and the switching signal have to be merged into a new artificial input  $\mathbf{u}_3$ . Note that not all solutions of the bilinear system correspond to solutions of an LSS: for the correspondence to hold,  $\mathbf{u}_2$  should take values 0, 1 and  $\mathbf{u}_3$  should satisfy  $\mathbf{u}_3 = \mathbf{u}_2\mathbf{u}_1$ .

### 3 Interpolatory MOR methods and the Loewner framework

Consider a linear system defined by matrices  $\mathbf{E} \in \mathbb{R}^{n \times n}$ ,  $\mathbf{A} \in \mathbb{R}^{n \times n}$ ,  $\mathbf{B} \in \mathbb{R}^{n \times m}$ ,  $\mathbf{C} \in \mathbb{R}^{p \times n}$ , with transfer function  $\mathbf{H}(s) = \mathbf{C}(s\mathbf{E} - \mathbf{A})^{-1}\mathbf{B}$ . Given left interpolation points:  $\{\mu_j | 1 \leq j \leq q\} \subset \mathbb{C}$ , with *left tangential directions*:  $\{\boldsymbol{\ell}_j | 1 \leq j \leq q\} \subset \mathbb{C}^p$ , and right interpolation points:  $\{\lambda_i | 1 \leq i \leq k\} \subset \mathbb{C}$ , with right tangential directions:  $\{\mathbf{r}_i | 1 \leq i \leq k\} \subset \mathbb{C}^m$ , find a reduced-order system  $\hat{\mathbf{E}}, \hat{\mathbf{A}}, \hat{\mathbf{B}}, \hat{\mathbf{C}}$ , such that the resulting transfer function,  $\hat{\mathbf{H}}(s)$  is a *tangential interpolant* to  $\mathbf{H}(s)$ :

$$\boldsymbol{\ell}_j^T \hat{\mathbf{H}}(\mu_j) = \boldsymbol{\ell}_j^T \mathbf{H}(\mu_j), \quad j = 1, \dots, q, \quad \text{and} \quad \hat{\mathbf{H}}(\lambda_i) \mathbf{r}_i = \mathbf{H}(\lambda_i) \mathbf{r}_i, \quad i = 1, \dots, k. \quad (6)$$

Interpolation points and tangent directions are selected to realize appropriate MOR goals. If instead of state space data, we are given input/output data, the resulting problem is hence modified. Given a set of input-output response measurements specified by left driving frequencies:  $\{\mu_j | 1 \leq j \leq q\} \subset \mathbb{C}$ , using left input directions:  $\{\boldsymbol{\ell}_j | 1 \leq j \leq q\} \subset \mathbb{C}^p$ , producing left responses:  $\{\mathbf{v}_j | 1 \leq j \leq q\} \subset \mathbb{C}^m$ , and right driving frequencies:  $\{\lambda_i | 1 \leq i \leq k\} \subset \mathbb{C}$ , using right input directions:  $\{\mathbf{r}_i | 1 \leq i \leq k\} \subset \mathbb{C}^m$ , producing right responses:  $\{\mathbf{w}_i | 1 \leq i \leq k\} \subset \mathbb{C}^p$ , find (low order) system matrices  $\hat{\mathbf{E}}, \hat{\mathbf{A}}, \hat{\mathbf{B}}, \hat{\mathbf{C}}$ , such that the resulting transfer function,  $\hat{\mathbf{H}}(s)$ , is an (approximate) tangential interpolant to the data:

$$\boldsymbol{\ell}_j^T \hat{\mathbf{H}}(\mu_j) = \mathbf{v}_j^T, \quad j = 1, \dots, q, \quad \text{and} \quad \hat{\mathbf{H}}(\lambda_i) \mathbf{r}_i = \mathbf{w}_i, \quad i = 1, \dots, k. \quad (7)$$

For details on interpolatory or moment matching MOR methods, we refer the reader to [17, 3].

#### 3.1 Overview of the Loewner framework for linear systems

The approach we discuss in this section is data driven. After collecting input/output (e.g. frequency response) measurements for some appropriate range of frequencies, we construct models which fit (or approximately fit) the data and have reduced dimension. The key is that, larger amounts of data than necessary are collected and the essential underlying system structure is extracted appropriately. Thus an advantage of this approach is that it can provide the user with a trade-off between accuracy of fit and complexity of the model.

The Loewner framework was developed in a series of papers; for details we refer the reader to [1], as well as [26, 24, 23, 5, 21]. For a recent overview see [6].

#### 3.2 The Loewner pencil

We will formulate the results for the more general tangential interpolation problem. We are given the right data:  $(\lambda_i; \mathbf{r}_i, \mathbf{w}_i)$ ,  $i = 1, \dots, k$ , and the left data:  $(\mu_j; \boldsymbol{\ell}_j^T, \mathbf{v}_j^T)$ ,  $j = 1, \dots, q$ ; it is assumed for simplicity that all points are distinct. The dimensions are as in (6), (7). The data can be organized as follows: the *right data*:

$$\boldsymbol{\Lambda} = \text{diag}[\lambda_1, \dots, \lambda_k] \in \mathbb{C}^{k \times k}, \quad \mathbf{R} = [\mathbf{r}_1, \dots, \mathbf{r}_k] \in \mathbb{C}^{m \times k}, \quad \mathbf{W} = [\mathbf{w}_1, \dots, \mathbf{w}_k] \in \mathbb{C}^{p \times k},$$

and the left data:

$$\mathbf{M} = \text{diag} [\mu_1, \dots, \mu_q] \in \mathbb{C}^{q \times q}, \quad \mathbf{L}^T = [\boldsymbol{\ell}_1, \dots, \boldsymbol{\ell}_q] \in \mathbb{C}^{q \times p}, \quad \mathbf{V}^T = [\mathbf{v}_1, \dots, \mathbf{v}_q] \in \mathbb{C}^{q \times m}.$$

Then, the associated Loewner pencil, consists of the Loewner and shifted Loewner matrices. The **Loewner matrix**  $\mathbb{L} \in \mathbb{C}^{q \times k}$ , is defined as:

$$\mathbb{L} = \begin{bmatrix} \frac{\mathbf{v}_1^T \mathbf{r}_1 - \boldsymbol{\ell}_1^T \mathbf{w}_1}{\mu_1 - \lambda_1} & \dots & \frac{\mathbf{v}_1^T \mathbf{r}_k - \boldsymbol{\ell}_1^T \mathbf{w}_k}{\mu_1 - \lambda_k} \\ \vdots & \ddots & \vdots \\ \frac{\mathbf{v}_q^T \mathbf{r}_1 - \boldsymbol{\ell}_q^T \mathbf{w}_1}{\mu_q - \lambda_1} & \dots & \frac{\mathbf{v}_q^T \mathbf{r}_k - \boldsymbol{\ell}_q^T \mathbf{w}_k}{\mu_q - \lambda_k} \end{bmatrix}.$$

Note that the matrix  $\mathbb{L}$  satisfies the Sylvester equation  $\mathbf{M}\mathbb{L} - \mathbb{L}\mathbf{A} = \mathbf{V}\mathbf{R} - \mathbf{L}\mathbf{W}$ . Suppose that the underlying transfer function is  $\mathbf{H}(s) = \mathbf{C}(s\mathbf{E} - \mathbf{A})^{-1}\mathbf{B}$ , and define the generalized observability/controllability matrices:

$$\mathcal{O} = \begin{bmatrix} \mathbf{C}(\mu_1\mathbf{E} - \mathbf{A})^{-1} \\ \vdots \\ \mathbf{C}(\mu_q\mathbf{E} - \mathbf{A})^{-1} \end{bmatrix}, \quad \mathcal{R} = [(\lambda_1\mathbf{E} - \mathbf{A})^{-1}\mathbf{B} \quad \dots \quad (\lambda_k\mathbf{E} - \mathbf{A})^{-1}\mathbf{B}]. \quad (8)$$

It readily follows that the Loewner matrix can be factored as  $\mathbb{L} = -\mathcal{O}\mathbf{E}\mathcal{R}$ . The **shifted Loewner matrix**  $\mathbb{L}_s \in \mathbb{C}^{q \times k}$ , is defined as:

$$\mathbb{L}_s = \begin{bmatrix} \frac{\mu_1 \mathbf{v}_1^T \mathbf{r}_1 - \boldsymbol{\ell}_1^T \mathbf{w}_1 \lambda_1}{\mu_1 - \lambda_1} & \dots & \frac{\mu_1 \mathbf{v}_1^T \mathbf{r}_k - \boldsymbol{\ell}_1^T \mathbf{w}_k \lambda_k}{\mu_1 - \lambda_k} \\ \vdots & \ddots & \vdots \\ \frac{\mu_q \mathbf{v}_q^T \mathbf{r}_1 - \boldsymbol{\ell}_q^T \mathbf{w}_1 \lambda_1}{\mu_q - \lambda_1} & \dots & \frac{\mu_q \mathbf{v}_q^T \mathbf{r}_k - \boldsymbol{\ell}_q^T \mathbf{w}_k \lambda_k}{\mu_q - \lambda_k} \end{bmatrix}.$$

Note that the matrix  $\mathbb{L}_s$  satisfies the Sylvester equation  $\mathbf{M}\mathbb{L}_s - \mathbb{L}_s\mathbf{A} = \mathbf{M}\mathbf{V}\mathbf{R} - \mathbf{L}\mathbf{W}\mathbf{A}$ , and can be factored in terms of the generalized controllability/observability matrices as  $\mathbb{L}_s = -\mathcal{O}\mathcal{A}\mathcal{R}$ . Finally notice that the following relations hold:  $\mathbf{V} = \mathbf{C}\mathcal{R}$ ,  $\mathbf{W} = \mathcal{O}\mathbf{B}$ .

### 3.3 Construction of reduced order models

We will distinguish two cases namely, the right amount of data and the more realistic redundant amount of data cases. The following lemma covers the first case.

**Lemma 3.1** *Assume that  $k = q$ , and let  $(\mathbb{L}_s, \mathbb{L})$ , be a regular pencil, such that none of the interpolation points  $\lambda_i, \mu_j$  are its eigenvalues. Then  $\mathbf{E} = -\mathbb{L}$ ,  $\mathbf{A} = -\mathbb{L}_s$ ,  $\mathbf{B} = \mathbf{V}$ ,  $\mathbf{C} = \mathbf{W}$ , is a minimal realization of an interpolant of the data, i.e. the rational function  $\mathbf{H}(s) = \mathbf{W}(\mathbb{L}_s - s\mathbb{L})^{-1}\mathbf{V}$ , interpolates the data (the conditions in (7) are hence matched).*

If the pencil  $(\mathbb{L}_s, \mathbb{L})$  is singular we are dealing with the case of *redundant data*. In this case if the following assumption is satisfied:

$$\text{rank}(\gamma\mathbb{L} - \mathbb{L}_s) = \text{rank} \begin{bmatrix} \mathbb{L} \\ \mathbb{L}_s \end{bmatrix} = \text{rank} [\mathbb{L} \quad \mathbb{L}_s] = r \leq k, \quad (9)$$

for all  $\gamma \in \{\lambda_i | 1 \leq i \leq k\} \cup \{\mu_j | 1 \leq j \leq q\}$ , we consider the following SVD factorizations:



$$\begin{bmatrix} \mathbb{L} & \mathbb{L}_s \end{bmatrix} = \mathbf{Y}^{(1)} \mathbf{S}^{(1)} (\mathbf{X}^{(1)})^T, \quad \begin{bmatrix} \mathbb{L} \\ \mathbb{L}_s \end{bmatrix} = \mathbf{Y}^{(2)} \mathbf{S}^{(2)} (\mathbf{X}^{(2)})^T, \quad (10)$$

where  $\mathbf{Y}^{(1)}, \mathbf{X}^{(2)} \in \mathbb{C}^{k \times k}$ . The projection matrices  $\mathbf{Y} \in \mathbb{C}^{k \times r}$  and  $\mathbf{X} \in \mathbb{C}^{k \times r}$  are obtained by selecting the first  $r$  columns of the matrices  $\mathbf{Y}^{(1)}$  and, respectively  $\mathbf{X}^{(2)}$ . The following result is used in practical applications.

**Lemma 3.2** *A realization  $(\mathbf{E}, \mathbf{A}, \mathbf{B}, \mathbf{C})$  of an approximate interpolant is given by the system matrices  $\mathbf{E} = -\mathbf{Y}^T \mathbb{L} \mathbf{X}$ ,  $\mathbf{A} = -\mathbf{Y}^T \mathbb{L}_s \mathbf{X}$ ,  $\mathbf{B} = \mathbf{Y}^T \mathbf{V}$ ,  $\mathbf{C} = \mathbf{W} \mathbf{X}$ . Hence, the rational function  $\mathbf{H}(s) = \mathbf{W} \mathbf{X} (\mathbf{Y}^T \mathbb{L}_s \mathbf{X} - s \mathbf{Y}^T \mathbb{L} \mathbf{X})^{-1} \mathbf{Y}^T \mathbf{V}$  approximately matches the data (the conditions in (7) are approximately fulfilled, i.e.  $\mathbf{H}(\lambda_i) \mathbf{r}_i = \mathbf{w}_i + \epsilon_i^r$  and  $\ell_j^T \mathbf{H}(\mu_j) = \mathbf{v}_j^T + (\epsilon_j^\ell)^T$ , where the residual errors are collected in the vectors  $\epsilon_i^r$  and  $\epsilon_j^\ell$ ).*

Thus, if we have more data than necessary, we can consider  $(\mathbb{L}_s, \mathbb{L}, \mathbf{V}, \mathbf{W})$ , as a *singular* model of the data. An appropriate projection then yields a reduced system of order  $k$  (see [26, 2]).

A direct consequence is that the Loewner framework offers a trade-off between accuracy and complexity of the reduced order system, by means of the singular values of  $\mathbb{L}$ .

**Remark 3.1** For an error bound that links the quality of approximation to the singular values of the Loewner pencil (which is valid only at the interpolation points  $\mu_j$  and  $\lambda_i$ ), we refer the readers to [6].

**Remark 3.2** For the classical Loewner framework applied to linear systems (see [26]), it is not mandatory that the samples used in the modeling step, come from systems with an invertible  $\mathbf{E}$  matrix. We believe that similarly to the LTI case, the Loewner framework can be extended to LSSs for which the matrices  $\mathbf{E}_k$  are not invertible. Indeed, since one needs to evaluate the multi-variable transfer functions on certain frequency grids, the only condition that is mandatory is that the pencils  $(\mathbf{A}_i, \mathbf{E}_i)$  are regular for all  $i \in \mathbf{Q}$ . This is because the resolvent of such pencils enters the transfer functions for various interpolation points  $s$ , as described in (5).

## 4 The Loewner framework for LSS - the case $D=2$

The characterization of linear switched systems by means of rational functions suggests that reduction of such systems can be performed by means of interpolatory methods. In the following we propose a way to generalize the Loewner framework to LSS by interpolating appropriately defined transfer functions on a chosen grid of frequencies (interpolation points).

As for the linear case, the given set of sampling (interpolation) points is first partitioned into the two following categories: left interpolation points:  $\{\mu_j | 1 \leq j \leq \ell\} \subset \mathbb{C}$  and right interpolation points:  $\{\lambda_i | 1 \leq i \leq k\} \subset \mathbb{C}$ .

In this paper we consider the case of SISO linear switched systems and thus, the left and right tangential directions can be considered to be scalar (i.e. taking the value 1). The transfer functions which will be matched are not single variable functions anymore (they depend on multiple variables as described in (5)). Hence, the structure of the interpolation points used in the new framework will be modified. Instead of having singleton values as in Section 3, we will use instead n-tuples that include multiple singleton values.

For simplicity of the exposition, we first consider the simplified case  $D = 2$  (the system switches between two modes only). This situation is encountered in most of the numerical examples in the literature that we encountered. Nevertheless, all the results presented in this section can be

generalized for higher number of modes in a direct way (as presented in Section 5). Depending on the switching signal  $\sigma(t)$ , we either have,

$$\Sigma_1 : \begin{cases} \mathbf{E}_1 \dot{\mathbf{x}}_1(t) = \mathbf{A}_1 \mathbf{x}_1(t) + \mathbf{B}_1 u(t), \\ y(t) = \mathbf{C}_1 \mathbf{x}_1(t), \end{cases} \quad \text{or} \quad \Sigma_2 : \begin{cases} \mathbf{E}_2 \dot{\mathbf{x}}_2(t) = \mathbf{A}_2 \mathbf{x}_2(t) + \mathbf{B}_2 u(t), \\ y(t) = \mathbf{C}_2 \mathbf{x}_2(t), \end{cases}$$

where  $\dim(\Sigma_1) = n_1$  (i.e.  $\mathbf{x}_1 \in \mathbb{R}^{n_1}$  and  $\mathbf{E}_1, \mathbf{A}_1 \in \mathbb{R}^{n_1 \times n_1}, \mathbf{B}_1, \mathbf{C}_1^T \in \mathbb{R}^{n_1}$ ) and also  $\dim(\Sigma_2) = n_2$  (i.e.  $\mathbf{x}_2 \in \mathbb{R}^{n_2}$  and  $\mathbf{E}_2, \mathbf{A}_2 \in \mathbb{R}^{n_2 \times n_2}, \mathbf{B}_2, \mathbf{C}_2^T \in \mathbb{R}^{n_2}$ ).

Denote, for simplicity, with  $\mathbf{K}_1$  the coupling matrix when switching from mode 1 to mode 2 (instead of  $\mathbf{K}_{1,2}$ ) and, with  $\mathbf{K}_2$ , the coupling matrix when switching from mode 2 to mode 1 (instead of  $\mathbf{K}_{2,1}$ ) with  $\mathbf{K}_1 \in \mathbb{R}^{n_2 \times n_1}$  and  $\mathbf{K}_2 \in \mathbb{R}^{n_1 \times n_2}$ .

Let  $\Phi_q(s) = (s\mathbf{E}_q - \mathbf{A}_q)^{-1}$ , for  $q \in \{1, 2\}$ ,  $s \in \mathbb{C}$  so that  $\det(s\mathbf{E}_q - \mathbf{A}_q) \neq 0$ . The generalized transfer functions corresponding to the first three levels, are written as:

### Level 1

$$\mathbf{H}_1(s_1) = \mathbf{C}_1 \Phi_1(s_1) \mathbf{B}_1 \quad \mathbf{H}_2(s_2) = \mathbf{C}_2 \Phi_2(s_2) \mathbf{B}_2,$$

### Level 2

$$\mathbf{H}_{1,2}(s_1, s_2) = \mathbf{C}_1 \Phi_1(s_1) \mathbf{K}_2 \Phi_2(s_2) \mathbf{B}_2 \quad \mathbf{H}_{2,1}(s_2, s_1) = \mathbf{C}_2 \Phi_2(s_2) \mathbf{K}_1 \Phi_1(s_1) \mathbf{B}_1,$$

### Level 3

$$\begin{cases} \mathbf{H}_{1,2,1}(s_1, s_2, s_3) = \mathbf{C}_1 \Phi_1(s_1) \mathbf{K}_2 \Phi_2(s_2) \mathbf{K}_1 \Phi_1(s_3) \mathbf{B}_1, \\ \mathbf{H}_{2,1,2}(s_1, s_2, s_3) = \mathbf{C}_2 \Phi_2(s_1) \mathbf{K}_1 \Phi_1(s_2) \mathbf{K}_2 \Phi_2(s_3) \mathbf{B}_2. \end{cases}$$

**Definition 4.1** Consider two LSS  $\hat{\Sigma} = (n_1, n_2, \{(\hat{\mathbf{E}}_i, \hat{\mathbf{A}}_i, \hat{\mathbf{B}}_i, \hat{\mathbf{C}}_i) | i \in \mathbf{Q}\}, \{\hat{\mathbf{K}}_{i,j} | i, j \in \mathbf{Q}\}, \mathbf{0})$  and  $\bar{\Sigma} = (n_1, n_2, \{(\bar{\mathbf{E}}_i, \bar{\mathbf{A}}_i, \bar{\mathbf{B}}_i, \bar{\mathbf{C}}_i) | i \in \mathbf{Q}\}, \{\bar{\mathbf{K}}_{i,j} | i, j \in \mathbf{Q}\}, \mathbf{0})$  with  $\mathbf{Q} = \{1, 2\}$ . These systems are said to be equivalent if there exist non-singular matrices  $\mathbf{Z}_j^L, \mathbf{Z}_j^R \in \mathbb{R}^{n_j \times n_j}$  so that

$$\bar{\mathbf{E}}_j = \mathbf{Z}_j^L \hat{\mathbf{E}}_j \mathbf{Z}_j^R, \quad \bar{\mathbf{A}}_j = \mathbf{Z}_j^L \hat{\mathbf{A}}_j \mathbf{Z}_j^R, \quad \bar{\mathbf{B}}_j = \mathbf{Z}_j^L \hat{\mathbf{B}}_j, \quad \bar{\mathbf{C}}_j = \hat{\mathbf{C}}_j \mathbf{Z}_j^R, \quad j \in \{1, 2\},$$

and also  $\bar{\mathbf{K}}_1 = \mathbf{Z}_2^L \hat{\mathbf{K}}_1 \mathbf{Z}_1^R$ ,  $\bar{\mathbf{K}}_2 = \mathbf{Z}_1^L \hat{\mathbf{K}}_2 \mathbf{Z}_2^R$ . In this configuration, one can easily show that the transfer functions defined above are the same for each LSS and for all sampling points  $s_k$ .

## 4.1 The generalized controllability and observability matrices

Let  $\Sigma$  be an LSS as described in (1) with  $\dim(\Sigma_k) = n_k$  for  $k = 1, 2$  and let  $\mathbf{K}_1 \in \mathbb{R}^{n_2 \times n_1}$  and  $\mathbf{K}_2 \in \mathbb{R}^{n_1 \times n_2}$  be the coupling matrices. Before stating the general definitions, we first clarify how the newly introduced matrices are constructed through a simple self-explanatory example.

**Example 4.1** Let  $\mathcal{U}$  be a set with  $\ell = 12$  elements, interpreted as left interpolation points, where  $\mathcal{U} = \{\mu_1, \mu_2, \dots, \mu_{12}\}$ . Partition  $\mathcal{U}$  as  $\mathcal{U} = \mathcal{U}_1 \cup \mathcal{U}_2$ , where  $\mathcal{U}_1 = \{\mu_1^{(1)}, \mu_3^{(1)}, \mu_1^{(2)}, \mu_3^{(2)}, \mu_5^{(2)}, \mu_1^{(3)}\}$  and  $\mathcal{U}_2 = \{\mu_2^{(1)}, \mu_4^{(1)}, \mu_2^{(2)}, \mu_4^{(2)}, \mu_6^{(2)}, \mu_2^{(3)}\}$ . Here,  $\mathcal{U}_i$  contains points associated to mode  $i$ . Introduce the nested multi-tuples corresponding to each mode of the LSS, as:

$$\underline{\text{Mode 1}} : \quad \boldsymbol{\mu}_1^{(1)} = \begin{cases} (\mu_1^{(1)}) \\ (\mu_2^{(1)}, \mu_3^{(1)}) \end{cases}, \quad \boldsymbol{\mu}_1^{(2)} = \begin{cases} (\mu_1^{(2)}) \\ (\mu_2^{(2)}, \mu_3^{(2)}) \\ (\mu_1^{(2)}, \mu_4^{(2)}, \mu_5^{(2)}) \end{cases}, \quad \boldsymbol{\mu}_1^{(3)} = \{(\mu_1^{(3)})\},$$

$$\text{Mode 2 : } \boldsymbol{\mu}_2^{(1)} = \begin{cases} (\mu_2^{(1)}) \\ (\mu_1^{(1)}, \mu_4^{(1)}) \end{cases}, \quad \boldsymbol{\mu}_2^{(2)} = \begin{cases} (\mu_2^{(2)}) \\ (\mu_1^{(2)}, \mu_4^{(2)}) \\ (\mu_2^{(2)}, \mu_3^{(2)}, \mu_6^{(2)}) \end{cases}, \quad \boldsymbol{\mu}_2^{(3)} = \{(\mu_2^{(3)})\}.$$

We explicitly write the generalized observability matrices  $\mathcal{O}_1$  and  $\mathcal{O}_2$  as follows:

$$\mathcal{O}_1 = \begin{bmatrix} \mathbf{C}_1 \boldsymbol{\Phi}_1(\mu_1^{(1)}) \\ \mathbf{C}_2 \boldsymbol{\Phi}_2(\mu_2^{(1)}) \mathbf{K}_1 \boldsymbol{\Phi}_1(\mu_3^{(1)}) \\ \mathbf{C}_1 \boldsymbol{\Phi}_1(\mu_1^{(2)}) \\ \mathbf{C}_2 \boldsymbol{\Phi}_2(\mu_2^{(2)}) \mathbf{K}_1 \boldsymbol{\Phi}_1(\mu_3^{(2)}) \\ \mathbf{C}_1 \boldsymbol{\Phi}_1(\mu_1^{(2)}) \mathbf{K}_2 \boldsymbol{\Phi}_2(\mu_4^{(2)}) \mathbf{K}_1 \boldsymbol{\Phi}_1(\mu_5^{(2)}) \\ \mathbf{C}_1 \boldsymbol{\Phi}_1(\mu_1^{(3)}) \end{bmatrix}, \quad \mathcal{O}_2 = \begin{bmatrix} \mathbf{C}_2 \boldsymbol{\Phi}_2(\mu_2^{(1)}) \\ \mathbf{C}_1 \boldsymbol{\Phi}_1(\mu_1^{(1)}) \mathbf{K}_2 \boldsymbol{\Phi}_2(\mu_4^{(1)}) \\ \mathbf{C}_2 \boldsymbol{\Phi}_2(\mu_2^{(2)}) \\ \mathbf{C}_1 \boldsymbol{\Phi}_1(\mu_1^{(2)}) \mathbf{K}_2 \boldsymbol{\Phi}_2(\mu_4^{(2)}) \\ \mathbf{C}_2 \boldsymbol{\Phi}_2(\mu_2^{(2)}) \mathbf{K}_1 \boldsymbol{\Phi}_1(\mu_3^{(2)}) \mathbf{K}_2 \boldsymbol{\Phi}_2(\mu_6^{(2)}) \\ \mathbf{C}_2 \boldsymbol{\Phi}_2(\mu_2^{(3)}) \end{bmatrix}.$$

For an interpolation point  $\mu_j^{(i)}$ , the subscript  $j$  is related to the mode with which the point is associated to. This mode is given by the residue  $\eta(j)$ , where  $\eta(j) = 1$ , if  $j$  is odd, and  $\eta(j) = 2$ , if  $j$  is even. The superscript  $i$  stands for the block index. In this particular example, we considered three such blocks for each of the two modes with the following dimensions:  $p_1 = 2, p_2 = 3$  and  $p_3 = 1$  (here  $p_i$  represents the dimension of the block index  $i$  for  $i \in \{1, 2, 3\}$ ).

**Definition 4.2** Given a non-empty set  $\mathbf{X}$ , denote with  $\mathbf{X}^i$  the set of all  $i^{\text{th}}$  tuples with elements from  $\mathbf{X}$ . Introduce the concatenation of two tuples composed of elements (symbols)  $\alpha_1, \dots, \alpha_i$ , and  $\beta_1, \dots, \beta_j$  from  $\mathbf{X}$  as the mapping  $\odot : \mathbf{X}^i \times \mathbf{X}^j \rightarrow \mathbf{X}^{i+j}$  with the following property:

$$(\alpha_1, \alpha_2, \dots, \alpha_i) \odot (\beta_1, \beta_2, \dots, \beta_j) = (\alpha_1, \alpha_2, \dots, \alpha_i, \beta_1, \beta_2, \dots, \beta_j).$$

In the following we denote the  $\ell^{\text{th}}$  element of the ordered set  $\boldsymbol{\mu}_j^{(i)}$  with  $\boldsymbol{\mu}_j^{(i)}(\ell)$  (where  $j \in \mathbf{Q}$  and  $i \geq 1$ ). For instance,  $\boldsymbol{\mu}_1^{(2)}(3) := (\mu_1^{(2)}, \mu_4^{(2)}, \mu_5^{(2)})$ . For convenience, use the notation  $\mathbf{H}_{1,2,1}(\mu_1^{(2)}, \mu_4^{(2)}, \mu_5^{(2)})$  instead of  $\mathbf{H}_{(1,2,1)}((\mu_1^{(2)}, \mu_4^{(2)}, \mu_5^{(2)}))$ , when referring to the function evaluation:

$$\mathbf{H}_{1,2,1}(\mu_1^{(2)}, \mu_4^{(2)}, \mu_5^{(2)}) = \mathbf{C}_1 \boldsymbol{\Phi}_1(\mu_1^{(2)}) \mathbf{K}_2 \boldsymbol{\Phi}_2(\mu_4^{(2)}) \mathbf{K}_1 \boldsymbol{\Phi}_1(\mu_5^{(2)}) \mathbf{B}_1.$$

**Definition 4.3** Let  $\mathcal{V} = \{\lambda_1, \lambda_2, \dots, \lambda_k\} \subset \mathbb{C}$  be a set composed of  $k$  right interpolation points. Partition  $\mathcal{V}$  in two sets  $\mathcal{V}_1$  and  $\mathcal{V}_2$ , as  $\mathcal{V} = \mathcal{V}_1 \cup \mathcal{V}_2$ , where:

$$\mathcal{V}_1 = \{\lambda_{2g-1}^{(i)} | 1 \leq g \leq m_i, 1 \leq i \leq \mathcal{K}\}, \quad \mathcal{V}_2 = \{\lambda_{2g}^{(i)} | 1 \leq g \leq m_i, 1 \leq i \leq \mathcal{K}\}. \quad (11)$$

Here,  $\mathcal{V}_1$  and  $\mathcal{V}_2$  correspond to interpolation points associated to the  $I^{\text{st}}$  mode and, respectively, associated to the  $II^{\text{nd}}$  mode. Let  $\mathcal{K} \geq 1$  be a positive integer. For each  $i = 1, \dots, \mathcal{K}$ , introduce the blocks of right  $i^{\text{th}}$  tuples in terms of the points from (11), as:

$$\boldsymbol{\lambda}_1^{(i)} = \begin{cases} (\lambda_1^{(i)}), \\ (\lambda_3^{(i)}, \lambda_2^{(i)}), \\ (\lambda_5^{(i)}, \lambda_4^{(i)}, \lambda_1^{(i)}), \\ \vdots \\ (\lambda_{2m_i-3}^{(i)}, \dots, \lambda_4^{(i)}, \lambda_1^{(i)}), \\ (\lambda_{2m_i-1}^{(i)}, \lambda_{2m_i-2}^{(i)}, \dots, \lambda_3^{(i)}, \lambda_2^{(i)}). \end{cases}, \quad \boldsymbol{\lambda}_2^{(i)} = \begin{cases} (\lambda_2^{(i)}), \\ (\lambda_4^{(i)}, \lambda_1^{(i)}), \\ (\lambda_6^{(i)}, \lambda_3^{(i)}, \lambda_2^{(i)}), \\ \vdots \\ (\lambda_{2m_i-2}^{(i)}, \dots, \lambda_3^{(i)}, \lambda_2^{(i)}), \\ (\lambda_{2m_i}^{(i)}, \lambda_{2m_i-3}^{(i)}, \dots, \lambda_4^{(i)}, \lambda_1^{(i)}). \end{cases}, \quad (12)$$

where  $\mathcal{K}$  is the number of blocks, and  $m_i$  is the dimension of the index  $i$  block so that the equality  $m_1 + \dots + m_{\mathcal{K}} = k$  holds. Finally define the nested right multi-tuples, as:

$$\boldsymbol{\lambda}_1 = \left\{ \boldsymbol{\lambda}_1^{(1)}, \boldsymbol{\lambda}_1^{(2)}, \dots, \boldsymbol{\lambda}_1^{(\mathcal{K})} \right\}, \quad \boldsymbol{\lambda}_2 = \left\{ \boldsymbol{\lambda}_2^{(1)}, \boldsymbol{\lambda}_2^{(2)}, \dots, \boldsymbol{\lambda}_2^{(\mathcal{K})} \right\}. \quad (13)$$

Note that the right tuples in (12) are constructed based on the following recurrence relations (where  $\boldsymbol{\lambda}_1^{(i)}(1) = (\lambda_1^{(i)})$  and  $\boldsymbol{\lambda}_2^{(i)}(1) = (\lambda_2^{(i)})$ )

$$\boldsymbol{\lambda}_1^{(i)}(g) = (\lambda_{2g-1}^{(i)}) \odot \boldsymbol{\lambda}_2^{(i)}(g-1), \quad \boldsymbol{\lambda}_2^{(i)}(g) = (\lambda_{2g}^{(i)}) \odot \boldsymbol{\lambda}_1^{(i)}(g-1). \quad (14)$$

**Definition 4.4** Let  $\mathcal{U} = \{\mu_1, \mu_2, \dots, \mu_k\} \subset \mathbb{C}$  be a set composed of  $k$  left interpolation points. Partition  $\mathcal{U}$  in two sets  $\mathcal{U}_1$  and  $\mathcal{U}_2$ , as  $\mathcal{U} = \mathcal{U}_1 \cup \mathcal{U}_2$ , where:

$$\mathcal{U}_1 = \{\mu_{2h-1}^{(j)} | 1 \leq h \leq p_j, 1 \leq j \leq \mathcal{L}\}, \quad \mathcal{U}_2 = \{\mu_{2h}^{(j)} | 1 \leq h \leq p_j, 1 \leq j \leq \mathcal{L}\}. \quad (15)$$

Here,  $\mathcal{U}_1$  and  $\mathcal{U}_2$  correspond to interpolation points associated to the  $I^{\text{st}}$  mode and, respectively, associated to the  $II^{\text{nd}}$  mode. Let  $\mathcal{L} \geq 1$  be a positive integer. For each  $j = 1, \dots, \mathcal{L}$ , introduce the blocks of left  $j^{\text{th}}$  tuples in terms of the points from (15), as:

$$\boldsymbol{\mu}_1^{(j)} = \begin{cases} (\mu_1^{(j)}), \\ (\mu_2^{(j)}, \mu_3^{(j)}), \\ (\mu_1^{(j)}, \mu_4^{(j)}, \mu_5^{(j)}), \\ \vdots \\ (\mu_1^{(j)}, \mu_4^{(j)}, \dots, \mu_{2p_j-3}^{(j)}), \\ (\mu_2^{(j)}, \mu_3^{(j)}, \dots, \mu_{2p_j-2}^{(j)}, \mu_{2p_j-1}^{(j)}). \end{cases}, \quad \boldsymbol{\mu}_2^{(j)} = \begin{cases} (\mu_2^{(j)}), \\ (\mu_1^{(j)}, \mu_4^{(j)}), \\ (\mu_2^{(j)}, \mu_3^{(j)}, \mu_6^{(j)}), \\ \vdots \\ (\mu_2^{(j)}, \mu_3^{(j)}, \dots, \mu_{2p_j-2}^{(j)}), \\ (\mu_1^{(j)}, \mu_4^{(j)}, \dots, \mu_{2p_j-3}^{(j)}, \mu_{2p_j}^{(j)}). \end{cases}, \quad (16)$$

where  $\mathcal{L} \geq 1$  is the number of blocks, and  $p_j$  is the dimension of the index  $j$  block. Additionally, note that the equality  $p_1 + \dots + p_{\mathcal{L}} = \ell$ , holds. Finally define the nested left multi-tuples, as:

$$\boldsymbol{\mu}_1 = \left\{ \boldsymbol{\mu}_1^{(1)}, \boldsymbol{\mu}_1^{(2)}, \dots, \boldsymbol{\mu}_1^{(\mathcal{L})} \right\}, \quad \boldsymbol{\mu}_2 = \left\{ \boldsymbol{\mu}_2^{(1)}, \boldsymbol{\mu}_2^{(2)}, \dots, \boldsymbol{\mu}_2^{(\mathcal{L})} \right\}. \quad (17)$$

Note that the left tuples are constructed based on the following recurrence relations (where  $\boldsymbol{\mu}_1^{(j)}(1) = (\mu_1^{(j)})$  and  $\boldsymbol{\mu}_2^{(j)}(1) = (\mu_2^{(j)})$ )

$$\boldsymbol{\mu}_1^{(j)}(h) = \boldsymbol{\mu}_2^{(j)}(h-1) \odot (\mu_{2h-1}^{(j)}), \quad \boldsymbol{\mu}_2^{(j)}(h) = \boldsymbol{\mu}_1^{(j)}(h-1) \odot (\mu_{2h}^{(j)}). \quad (18)$$

**Condition 4.1** The right interpolation points  $\lambda_u$ ,  $u \in \{1, 2, \dots, k\}$  are chosen in such a way so that they do not coincide with the poles of any of the subsystems  $\boldsymbol{\Sigma}_1$  or  $\boldsymbol{\Sigma}_2$ . More exactly, the following conditions are imposed for all  $i = 1, \dots, \mathcal{K}$  and  $g = 1, \dots, m_i$ ,

$$\det(\lambda_{2g-1}^{(i)} \mathbf{E}_1 - \mathbf{A}_1) \neq 0, \quad \det(\lambda_{2g}^{(i)} \mathbf{E}_2 - \mathbf{A}_2) \neq 0. \quad (19)$$

We associate the following matrices to the set of right tuples in (12), as

$$\mathcal{R}_1^{(i)} = \left[ \Phi_1(\lambda_1^{(i)}) \mathbf{B}_1, \Phi_1(\lambda_3^{(i)}) \mathbf{K}_2 \Phi_2(\lambda_2^{(i)}) \mathbf{B}_2, \dots, \Phi_1(\lambda_{2m_i-1}^{(i)}) \mathbf{K}_2 \dots \mathbf{K}_1 \Phi_1(\lambda_3^{(i)}) \mathbf{K}_2 \Phi_2(\lambda_2^{(i)}) \mathbf{B}_2 \right],$$

$$\mathcal{R}_2^{(i)} = \left[ \Phi_2(\lambda_2^{(i)}) \mathbf{B}_2, \Phi_2(\lambda_4^{(i)}) \mathbf{K}_1 \Phi_1(\lambda_1^{(i)}) \mathbf{B}_1, \dots, \Phi_2(\lambda_{2m_i}^{(i)}) \mathbf{K}_1 \dots \mathbf{K}_2 \Phi_2(\lambda_2^{(i)}) \mathbf{K}_1 \Phi_1(\lambda_1^{(i)}) \mathbf{B}_1 \right],$$

where  $i = 1, \dots, \mathcal{K}$  and  $\mathcal{R}_q^{(i)} \in \mathbb{C}^{n_q \times m_i}$  is attached to  $\boldsymbol{\Lambda}_q^{(i)}$  for  $q \in \{1, 2\}$ .

**Definition 4.5** For the LSS  $\Sigma$  in (1), introduce the generalized controllability matrices  $\mathcal{R}_1$  and  $\mathcal{R}_2$  associated to the right multi-tuples  $\lambda_1$ , and, respectively  $\lambda_2$ , as follows:

$$\mathcal{R}_1 = [\mathcal{R}_1^{(1)}, \mathcal{R}_1^{(2)}, \dots, \mathcal{R}_1^{(\mathcal{L})}] \in \mathbb{C}^{n_1 \times k}, \quad \mathcal{R}_2 = [\mathcal{R}_2^{(1)}, \mathcal{R}_2^{(2)}, \dots, \mathcal{R}_2^{(\mathcal{L})}] \in \mathbb{C}^{n_2 \times k}. \quad (20)$$

**Condition 4.2** The left interpolation points  $\mu_v$ ,  $v \in \{1, 2, \dots, \ell\}$  are chosen in such a way so that they do not coincide with the poles of any of the subsystems  $\Sigma_1$  or  $\Sigma_2$ . More exactly, the following conditions are imposed for all  $j = 1, \dots, \mathcal{L}$  and  $h = 1, \dots, p_j$ ,

$$\det(\mu_{2h-1}^{(j)} \mathbf{E}_1 - \mathbf{A}_1) \neq 0, \quad \det(\mu_{2h}^{(j)} \mathbf{E}_2 - \mathbf{A}_2) \neq 0. \quad (21)$$

Next, associate the following matrices to the set of right tuples in (16), as

$$\mathcal{O}_1^{(j)} = \begin{bmatrix} \mathbf{C}_1 \Phi_1(\mu_1^{(j)}) \\ \mathbf{C}_2 \Phi_2(\mu_2^{(j)}) \mathbf{K}_1 \Phi_1(\mu_3^{(j)}) \\ \vdots \\ \mathbf{C}_2 \Phi_2(\mu_2^{(j)}) \mathbf{K}_1 \Phi_1(\mu_3^{(j)}) \mathbf{K}_2 \cdots \mathbf{K}_1 \Phi_1(\mu_{2p_j-1}^{(j)}) \end{bmatrix} \in \mathbb{C}^{p_j \times n_1}, \quad j = 1, \dots, \mathcal{L},$$

$$\mathcal{O}_2^{(j)} = \begin{bmatrix} \mathbf{C}_2 \Phi_2(\mu_1^{(j)}) \\ \mathbf{C}_1 \Phi_1(\mu_1^{(j)}) \mathbf{K}_2 \Phi_2(\mu_4^{(j)}) \\ \vdots \\ \mathbf{C}_1 \Phi_1(\mu_1^{(j)}) \mathbf{K}_2 \Phi_2(\mu_2^{(j)}) \mathbf{K}_1 \cdots \mathbf{K}_2 \Phi_2(\mu_{2p_j}^{(j)}) \end{bmatrix} \in \mathbb{C}^{p_j \times n_2}, \quad j = 1, \dots, \mathcal{L}.$$

**Definition 4.6** For the LSS  $\Sigma$  in (1), introduce the generalized observability matrices  $\mathcal{O}_1$  and  $\mathcal{O}_2$  associated to the right multi-tuples  $\mu_1$ , and, respectively  $\mu_2$ , as follows:

$$\mathcal{O}_1 = \begin{bmatrix} \mathcal{O}_1^{(1)} \\ \vdots \\ \mathcal{O}_1^{(\mathcal{L})} \end{bmatrix} \in \mathbb{C}^{\ell \times n_1}, \quad \mathcal{O}_2 = \begin{bmatrix} \mathcal{O}_2^{(1)} \\ \vdots \\ \mathcal{O}_2^{(\mathcal{L})} \end{bmatrix} \in \mathbb{C}^{\ell \times n_2}. \quad (22)$$

**Definition 4.7** For  $\nu \in \{1, 2\}$ , let  $\mathbf{Q}^{\nu,+}$  and  $\mathbf{Q}^{+, \nu}$  be the ordered sets containing all tuples that can be constructed with symbols from the  $\mathbf{Q} = \{1, 2\}$  and that **start** (and respectively **end**) with the symbol  $\nu$ . Also, no two consecutive characters are allowed to be the same. Hence, explicitly write the new introduced sets as follows:

$$\mathbf{Q}^{1,+} = \{(1), (1, 2), (1, 2, 1), \dots\}, \quad \mathbf{Q}^{2,+} = \{(2), (2, 1), (2, 1, 2), \dots\}, \quad (23)$$

$$\mathbf{Q}^{+,1} = \{(1), (2, 1), (1, 2, 1), \dots\}, \quad \mathbf{Q}^{+,2} = \{(2), (1, 2), (2, 1, 2), \dots\}. \quad (24)$$

**Remark 4.1** In the following we denote the  $\ell^{\text{th}}$  element of the ordered set  $\mathbf{Q}^{\nu,+}$  with  $\mathbf{Q}^{\nu,+}(\ell)$ . For example, one writes  $\mathbf{Q}^{1,+}(4) := (1, 2, 1, 2)$ . Moreover, we have  $\mathbf{Q}^{+,2}(3) \odot \mathbf{Q}^{1,+}(2) = (2, 1, 2, 1, 2)$ .

The compact notation  $\mathbf{H}_{\mathbf{Q}^{+,1}}(\mu_1(2))$  is used instead of  $\mathbf{H}_{2,1}(\mu_2, \mu_3)$ , where  $\mu_1(2) := (\mu_2, \mu_3)$

**Definition 4.8** Let the  $i^{\text{th}}$  unit vector of length  $k$  be denoted with  $\mathbf{e}_{i,k} = [0 \dots, 1, \dots, 0]^T \in \mathbb{R}^k$ . Additionally, let  $\mathbf{0}_{k,\ell} \in \mathbb{R}^{k \times \ell}$  be an all zero matrix.

In the following, use the notation  $\hat{\mathbf{H}}$  to emphasize that we are referring to the *generalized transfer functions* corresponding to the LSS  $\hat{\Sigma}$ .

**Definition 4.9** We say that an LSS  $\hat{\Sigma} = (n_1, n_2, \{(\hat{\mathbf{E}}_i, \hat{\mathbf{A}}_i, \hat{\mathbf{B}}_i, \hat{\mathbf{C}}_i) | i \in \mathbf{Q}\}, \{\hat{\mathbf{K}}_{i,j} | i, j \in \mathbf{Q}\}, \mathbf{0})$  matches the data associated with the right tuples  $\{\lambda_a^{(1)}, \dots, \lambda_a^{(\mathcal{K})}\}$  as well as with the left tuples  $\{\mu_b^{(1)}, \dots, \mu_b^{(\mathcal{L})}\}$ ,  $a, b \in \mathbf{Q}$  and corresponding to the original LSS  $\Sigma = (n_1, n_2, \{(\mathbf{E}_i, \mathbf{A}_i, \mathbf{B}_i, \mathbf{C}_i) | i \in \mathbf{Q}\}, \{\mathbf{K}_{i,j} | i, j \in \mathbf{Q}\}, \mathbf{0})$ , if the following  $2(k^2 + 2k)$  relations

$$\begin{cases} \mathbf{H}_{\mathbf{Q}^{+1}(h)}(\mu_1^{(j)}(h)) = \hat{\mathbf{H}}_{\mathbf{Q}^{+1}(h)}(\mu_1^{(j)}(h)), & \mathbf{H}_{\mathbf{Q}^{+2}(h)}(\mu_2^{(j)}(h)) = \hat{\mathbf{H}}_{\mathbf{Q}^{+2}(h)}(\mu_2^{(j)}(h)), \\ \mathbf{H}_{\mathbf{Q}^{1+(g)}}(\lambda_1^{(i)}(g)) = \hat{\mathbf{H}}_{\mathbf{Q}^{1+(g)}}(\lambda_1^{(i)}(g)), & \mathbf{H}_{\mathbf{Q}^{2+(g)}}(\lambda_2^{(i)}(g)) = \hat{\mathbf{H}}_{\mathbf{Q}^{2+(g)}}(\lambda_2^{(i)}(g)), \\ \mathbf{H}_{\mathbf{Q}^{+1}(h) \odot \mathbf{Q}^{2+(g)}}(\mu_1^{(j)}(h) \odot \lambda_2^{(i)}(g)) = \hat{\mathbf{H}}_{\mathbf{Q}^{+1}(h) \odot \mathbf{Q}^{2+(g)}}(\mu_1^{(j)}(h) \odot \lambda_2^{(i)}(g)), \\ \mathbf{H}_{\mathbf{Q}^{+2}(h) \odot \mathbf{Q}^{1+(g)}}(\mu_2^{(j)}(h) \odot \lambda_1^{(i)}(g)) = \hat{\mathbf{H}}_{\mathbf{Q}^{+2}(h) \odot \mathbf{Q}^{1+(g)}}(\mu_2^{(j)}(h) \odot \lambda_1^{(i)}(g)), \end{cases} \quad (25)$$

hold for  $j = 1, \dots, \mathcal{K}$ ,  $h = 1, \dots, p_j$  and  $i = 1, \dots, \mathcal{L}$ ,  $g = 1, \dots, m_i$ , where

$$p_1 + p_2 + \dots + p_{\mathcal{K}} = m_1 + m_2 + \dots + m_{\mathcal{L}} = k.$$

The following lemma extends the rational interpolation idea for linear systems approximation to the linear switched system case.

**Lemma 4.1 Interpolation of LSS.** Let  $\Sigma = (n_1, n_2, \{(\mathbf{E}_i, \mathbf{A}_i, \mathbf{B}_i, \mathbf{C}_i) | i \in \mathbf{Q}\}, \{\mathbf{K}_{i,j} | i, j \in \mathbf{Q}\}, \mathbf{0})$  be an LSS of order  $(n_1, n_2)$ . Consider that the number of left and right interpolation points is the same for each mode, i.e.  $\ell = k$ . Additionally, assume the matrices in (20) and (22) have full rank, i.e.  $\text{rank}(\mathcal{R}_i) = \text{rank}(\mathcal{O}_i) = k$ ,  $i \in \{1, 2\}$ . An order  $k$  reduced LSS  $\hat{\Sigma} = (n_1, n_2, \{(\hat{\mathbf{E}}_i, \hat{\mathbf{A}}_i, \hat{\mathbf{B}}_i, \hat{\mathbf{C}}_i) | i \in \mathbf{Q}\}, \{\hat{\mathbf{K}}_{i,j} | i, j \in \mathbf{Q}\}, \mathbf{0})$  is constructed by projection, i.e. by using right and left projectors chosen as

$$\mathbf{X}_1 = \mathcal{R}_1, \quad \mathbf{X}_2 = \mathcal{R}_2 \quad \text{and} \quad \mathbf{Y}_1^T = \mathcal{O}_1, \quad \mathbf{Y}_2^T = \mathcal{O}_2.$$

The projected matrices corresponding to the  $I^{\text{st}}$  subsystem  $\hat{\Sigma}_1$  are computed as,

$$\hat{\mathbf{E}}_1 = \mathbf{Y}_1^T \mathbf{E}_1 \mathbf{X}_1, \quad \hat{\mathbf{A}}_1 = \mathbf{Y}_1^T \mathbf{A}_1 \mathbf{X}_1, \quad \hat{\mathbf{B}}_1 = \mathbf{Y}_1^T \mathbf{B}_1, \quad \hat{\mathbf{C}}_1 = \mathbf{C}_1 \mathbf{X}_1, \quad \hat{\mathbf{K}}_1 = \mathbf{Y}_2^T \mathbf{K}_1 \mathbf{X}_1, \quad (26)$$

while the projected matrices corresponding to the  $II^{\text{nd}}$  subsystem  $\hat{\Sigma}_2$  can also be computed as,

$$\hat{\mathbf{E}}_2 = \mathbf{Y}_2^T \mathbf{E}_2 \mathbf{X}_2, \quad \hat{\mathbf{A}}_2 = \mathbf{Y}_2^T \mathbf{A}_2 \mathbf{X}_2, \quad \hat{\mathbf{B}}_2 = \mathbf{Y}_2^T \mathbf{B}_2, \quad \hat{\mathbf{C}}_2 = \mathbf{C}_2 \mathbf{X}_2, \quad \hat{\mathbf{K}}_2 = \mathbf{Y}_1^T \mathbf{K}_2 \mathbf{X}_2. \quad (27)$$

It follows that the reduced-order system  $\hat{\Sigma}$  matches the data of the system  $\Sigma$  (as it was previously introduced in Definition 4.9).

**Proof of Lemma 4.1** For simplicity, assume that we have one set of right multi-tuples, and one set of left multi-tuples with  $k$  interpolation points for each mode. This corresponds to the case  $\mathcal{L} = \mathcal{K} = 1$  and  $m_1 = p_1 = k$  (using the notations introduced in Definitions 4.3 and 4.4). For the first mode, write down the interpolation points as follows:

$$\begin{cases} \lambda_1 = \{(\lambda_1), (\lambda_3, \lambda_2), \dots, (\lambda_{2k-1}, \dots, \lambda_3, \lambda_2)\}, \\ \mu_1 = \{(\mu_1), (\mu_2, \mu_3), \dots, (\mu_2, \mu_3, \dots, \mu_{2k-1})\}. \end{cases} \quad (28)$$

For the second mode, write down the interpolation points as follows:

$$\begin{cases} \boldsymbol{\lambda}_2 = \{(\lambda_2), (\lambda_4, \lambda_1), \dots, (\lambda_{2k}, \dots, \lambda_2, \lambda_1)\}, \\ \boldsymbol{\mu}_2 = \{(\mu_2), (\mu_1, \mu_4), \dots, (\mu_1, \mu_4, \dots, \mu_{2k})\}. \end{cases} \quad (29)$$

It follows that the interpolation conditions stated in Definition 4.9, can be rewritten by taking into account the aforementioned simplification as,

$$\text{2k conditions: } \begin{cases} \mathbf{H}_{\mathbf{Q}^{+,1}(j)}(\boldsymbol{\mu}_1(j)) = \hat{\mathbf{H}}_{\mathbf{Q}^{+,1}(j)}(\boldsymbol{\mu}_1(j)) \\ \mathbf{H}_{\mathbf{Q}^{+,2}(j)}(\boldsymbol{\mu}_2(j)) = \hat{\mathbf{H}}_{\mathbf{Q}^{+,2}(j)}(\boldsymbol{\mu}_2(j)) \end{cases}, \quad j \in \{1, \dots, k\}, \quad (30)$$

$$\text{2k conditions: } \begin{cases} \mathbf{H}_{\mathbf{Q}^{1,+}(i)}(\boldsymbol{\lambda}_1(i)) = \hat{\mathbf{H}}_{\mathbf{Q}^{1,+}(i)}(\boldsymbol{\lambda}_1(i)) \\ \mathbf{H}_{\mathbf{Q}^{2,+}(i)}(\boldsymbol{\lambda}_2(i)) = \hat{\mathbf{H}}_{\mathbf{Q}^{2,+}(i)}(\boldsymbol{\lambda}_2(i)) \end{cases}, \quad i \in \{1, \dots, k\}, \quad (31)$$

$$\text{k}^2 \text{ conditions: } \left\{ \mathbf{H}_{\mathbf{Q}^{+,1}(j) \odot \mathbf{Q}^{2,+}(i)}(\boldsymbol{\mu}_1(j) \odot \boldsymbol{\lambda}_2(i)) = \hat{\mathbf{H}}_{\mathbf{Q}^{+,1}(j) \odot \mathbf{Q}^{2,+}(i)}(\boldsymbol{\mu}_1(j) \odot \boldsymbol{\lambda}_2(i)), \right. \quad (32)$$

$$\left. \text{k}^2 \text{ conditions: } \left\{ \mathbf{H}_{\mathbf{Q}^{+,2}(j) \odot \mathbf{Q}^{1,+}(i)}(\boldsymbol{\mu}_2(j) \odot \boldsymbol{\lambda}_1(i)) = \hat{\mathbf{H}}_{\mathbf{Q}^{+,2}(j) \odot \mathbf{Q}^{1,+}(i)}(\boldsymbol{\mu}_2(j) \odot \boldsymbol{\lambda}_1(i)). \right. \right. \quad (33)$$

With the assumptions in (28) and (29), it follows that the associated generalized controllability and observability matrices defined previously in (20) and (22), are rewritten as:

$$\mathcal{R}_1 = [ \boldsymbol{\Phi}_1(\lambda_1)\mathbf{B}_1, \boldsymbol{\Phi}_1(\lambda_3)\mathbf{K}_2\boldsymbol{\Phi}_2(\lambda_2)\mathbf{B}_2, \dots, \boldsymbol{\Phi}_1(\lambda_{2k-1})\mathbf{K}_2 \cdots \mathbf{K}_2\boldsymbol{\Phi}_2(\lambda_2)\mathbf{B}_2 ] \in \mathbb{C}^{n \times k},$$

$$\mathcal{R}_2 = [ \boldsymbol{\Phi}_2(\lambda_2)\mathbf{B}_2, \boldsymbol{\Phi}_2(\lambda_4)\mathbf{K}_1\boldsymbol{\Phi}_1(\lambda_1)\mathbf{B}_1, \dots, \boldsymbol{\Phi}_2(\lambda_{2k})\mathbf{K}_1 \cdots \mathbf{K}_1\boldsymbol{\Phi}_1(\lambda_1)\mathbf{B}_1 ] \in \mathbb{C}^{n \times k},$$

$$\mathcal{O}_1 = \begin{bmatrix} \mathbf{C}_1\boldsymbol{\Phi}_1(\mu_1) \\ \mathbf{C}_2\boldsymbol{\Phi}_2(\mu_2)\mathbf{K}_1\boldsymbol{\Phi}_1(\mu_3) \\ \vdots \\ \mathbf{C}_2\boldsymbol{\Phi}_2(\mu_2)\mathbf{K}_1\boldsymbol{\Phi}_1(\mu_3) \cdots \mathbf{K}_1\boldsymbol{\Phi}_1(\mu_{2k-1}) \end{bmatrix}, \quad \mathcal{O}_2 = \begin{bmatrix} \mathbf{C}_2\boldsymbol{\Phi}_2(\mu_2) \\ \mathbf{C}_1\boldsymbol{\Phi}_1(\mu_1)\mathbf{K}_2\boldsymbol{\Phi}_2(\mu_4) \\ \vdots \\ \mathbf{C}_1\boldsymbol{\Phi}_1(\mu_1)\mathbf{K}_2\boldsymbol{\Phi}_2(\mu_4) \cdots \mathbf{K}_2\boldsymbol{\Phi}_2(\mu_{2k}) \end{bmatrix},$$

with  $\mathcal{O}_1, \mathcal{O}_2 \in \mathbb{C}^{k \times n}$ . Additionally, introduce the notation  $\hat{\boldsymbol{\Phi}}_i(s) = (s\hat{\mathbf{E}} - \hat{\mathbf{A}})^{-1}$ .

From (26) and (27), using that  $\mathbf{X}_i = \mathcal{R}_i$  for  $i = 1, 2$ , it readily follows that:

$$\text{(a) } \hat{\boldsymbol{\Phi}}_1(\lambda_1)\hat{\mathbf{B}}_1 = \mathbf{e}_{1,k} \quad \text{and} \quad \text{(b) } \hat{\boldsymbol{\Phi}}_1(\lambda_{2i-1})\hat{\mathbf{K}}_2\mathbf{e}_{i-1,k} = \mathbf{e}_{i,k}, \quad i = 2, \dots, k,$$

$$\text{(c) } \hat{\boldsymbol{\Phi}}_2(\lambda_2)\hat{\mathbf{B}}_2 = \mathbf{e}_1 \quad \text{and} \quad \text{(d) } \hat{\boldsymbol{\Phi}}_2(\lambda_{2i})\hat{\mathbf{K}}_1\mathbf{e}_{i-1,k} = \mathbf{e}_{i,k}, \quad i = 2, \dots, k.$$

These equalities imply the right-hand conditions in (31). Similarly, from (26) and (27), using that  $\mathbf{Y}_j^T = \mathcal{O}_j$  for  $j = 1, 2$ , it follows that:

$$\text{(e) } \mathbf{C}_1\hat{\boldsymbol{\Phi}}_1(\mu_1) = \mathbf{e}_{1,k}^T \quad \text{and} \quad \text{(f) } \mathbf{e}_{j-1,k}^T\mathbf{K}_2\hat{\boldsymbol{\Phi}}_2(\mu_{2j}) = \mathbf{e}_{j,k}^T, \quad j = 2, \dots, k,$$

$$\text{(g) } \mathbf{C}_2\hat{\boldsymbol{\Phi}}_2(\mu_2) = \mathbf{e}_{1,k}^T \quad \text{and} \quad \text{(h) } \mathbf{e}_{j-1}^T\mathbf{K}_1\hat{\boldsymbol{\Phi}}_1(\mu_{2j-1}) = \mathbf{e}_{j,k}^T, \quad j = 2, \dots, k,$$

which imply left-hand conditions in (30). Finally, with  $\mathbf{X} = \mathcal{R}$ ,  $\mathbf{Y}^T = \mathcal{O}$ , and combining (a)-(h), all interpolation conditions in (32) and (33) are hence satisfied.  $\blacksquare$

**Remark 4.2** For instance, in Example 4.2, the conditions stated in (48) are satisfied.

#### 4.1.1 Sylvester equations for $\mathcal{O}$ and $\mathcal{R}$

The motivation behind this subsection is closely related to building parametrized reduced order models. The idea is that, one can use only one sided interpolation conditions, either left as in (30) or right as in (31), to reduce the original LSS. Then, one can choose the free parameters to impose additional conditions (not necessarily interpolatory).

Further development of this strategy was studied in [4] (in Section 4.4), for the case of generalized Sylvester equations for bilinear systems.

The generalized controllability and observability matrices satisfy Sylvester equations. To state the corresponding result we need to define particular quantities. First introduce the vectors

$$\mathbf{R} = [\mathbf{e}_{1,m_1}^T \cdots \mathbf{e}_{1,m_{\mathcal{X}}}^T] \in \mathbb{R}^{1 \times k}, \quad \mathbf{L}^T = [\mathbf{e}_{1,p_1}^T \cdots \mathbf{e}_{1,p_{\mathcal{L}}}^T] \in \mathbb{R}^{1 \times \ell}, \quad (34)$$

where  $m_1 + \dots + m_{\mathcal{X}} = k$  and  $p_1 + \dots + p_{\mathcal{L}} = \ell$ . Next, introduce the block-shift matrices

$$\begin{cases} \mathbf{S}_{\mathbf{R}} = \text{blkdiag} [\mathbf{J}_{m_1}, \dots, \mathbf{J}_{m_{\mathcal{X}}}], \\ \mathbf{S}_{\mathbf{L}} = \text{blkdiag} [\mathbf{J}_{p_1}^T, \dots, \mathbf{J}_{p_{\mathcal{L}}}^T]. \end{cases} \quad \text{where } \mathbf{J}_u = \begin{bmatrix} 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \\ 0 & 0 & \cdots & 0 \end{bmatrix} \in \mathbb{R}^{u \times u}. \quad (35)$$

Finally we arrange the left interpolation points in the diagonal matrices  $\mathbf{M}_1, \mathbf{M}_2 \in \mathbb{R}^{\ell \times \ell}$  as,

$$\mathbf{M}_1 = \text{blkdiag} [\mathbf{M}_1^{(1)}, \mathbf{M}_1^{(2)}, \dots, \mathbf{M}_1^{(\mathcal{L})}], \quad \mathbf{M}_2 = \text{blkdiag} [\mathbf{M}_2^{(1)}, \mathbf{M}_2^{(2)}, \dots, \mathbf{M}_2^{(\mathcal{L})}], \quad (36)$$

where  $\mathbf{M}_1^{(j)} = \text{diag} [\mu_1^{(j)}, \mu_3^{(j)}, \dots, \mu_{2p_j-1}^{(j)}]$  and  $\mathbf{M}_2^{(j)} = \text{diag} [\mu_2^{(j)}, \mu_4^{(j)}, \dots, \mu_{2p_j}^{(j)}]$ ; we used the MATLAB notation 'blkdiag' which outputs a block diagonal matrix with each input entry as a block. Also arrange the right interpolation points in the diagonal matrices  $\mathbf{\Lambda}_1, \mathbf{\Lambda}_2 \in \mathbb{R}^{k \times k}$  as,

$$\mathbf{\Lambda}_1 = \text{blkdiag} [\mathbf{\Lambda}_1^{(1)}, \mathbf{\Lambda}_1^{(2)}, \dots, \mathbf{\Lambda}_1^{(\mathcal{X})}], \quad \mathbf{\Lambda}_2 = \text{blkdiag} [\mathbf{\Lambda}_2^{(1)}, \mathbf{\Lambda}_2^{(2)}, \dots, \mathbf{\Lambda}_2^{(\mathcal{X})}], \quad (37)$$

where  $\mathbf{\Lambda}_1^{(i)} = \text{diag} [\lambda_1^{(i)}, \lambda_3^{(i)}, \dots, \lambda_{2m_i-1}^{(i)}]$  and  $\mathbf{\Lambda}_2^{(i)} = \text{diag} [\lambda_2^{(i)}, \lambda_4^{(i)}, \dots, \lambda_{2m_i}^{(i)}]$ . The next results represent extensions of the linear case and hence follow naturally.

**Lemma 4.2** *Consider that the assumption in Condition 4.1 holds, i.e. (19) is valid. Then, the generalized controllability matrices  $\mathcal{R}_1, \mathcal{R}_2$  defined in (20) are the unique solutions of the following coupled Sylvester equations:*

$$\begin{cases} \mathbf{A}_1 \mathcal{R}_1 + \mathbf{K}_2 \mathcal{R}_2 \mathbf{S}_{\mathbf{R}} + \mathbf{B}_1 \mathbf{R} = \mathbf{E}_1 \mathcal{R}_1 \mathbf{\Lambda}_1, \\ \mathbf{A}_2 \mathcal{R}_2 + \mathbf{K}_1 \mathcal{R}_1 \mathbf{S}_{\mathbf{R}} + \mathbf{B}_2 \mathbf{R} = \mathbf{E}_2 \mathcal{R}_2 \mathbf{\Lambda}_2. \end{cases} \quad (38)$$

**Proof of Lemma 4.2** Assume again, for simplicity of the proof, that the assumptions made in (28)-(29) are valid. Hence, we have one set of right multi-tuples for each of the two modes with same number of interpolation points  $k$  (with  $k$  even). Multiplying the first equation in (38) on the right with the unit vector  $\mathbf{e}_{1,k}$  we obtain:

$$\mathbf{A}_1 \mathcal{R}_1^{(1)} + \mathbf{B}_1 = \lambda_1 \mathbf{E}_1 \mathcal{R}_1^{(1)} \Leftrightarrow \mathcal{R}_1^{(1)} = (\lambda_1 \mathbf{E}_1 - \mathbf{A}_1)^{-1} \mathbf{B}_1 = \mathbf{\Phi}_1(\lambda_1) \mathbf{B}_1. \quad (39)$$

where  $\mathcal{R}_i^{(j)}$  is the  $j^{\text{th}}$  column of  $\mathcal{R}_i$  (with  $j \leq k$  and  $i \in \{1, 2\}$ ). Thus the first column of the matrix which is the solution of the first equation in (38) is indeed equal to the first column of



the generalized controllability matrix  $\mathcal{R}_1$ . By multiplying the second equation in (38) on the right with the unit vector  $\mathbf{e}_{1,k}$ , we obtain:

$$\mathbf{A}_2 \mathcal{R}_2^{(1)} + \mathbf{B}_2 = \lambda_2 \mathbf{E}_2 \mathcal{R}_2^{(1)} \Leftrightarrow \mathcal{R}_2^{(1)} = (\lambda_2 \mathbf{E}_2 - \mathbf{A}_2)^{-1} \mathbf{B}_2 = \Phi_2(\lambda_2) \mathbf{B}_2. \quad (40)$$

Thus the first column of the matrix which is the solution of the second equation in (38) is indeed equal to the first column of the generalized controllability matrix  $\mathcal{R}_2$ . By multiplying first equation in (38) on the right with the  $j^{\text{th}}$  unit vector  $\mathbf{e}_{j,k}$ , we obtain:

$$\mathbf{A}_1 \mathcal{R}_1^{(j)} + \mathbf{K}_2 \mathcal{R}_2^{(j-1)} = \lambda_{2j-1} \mathbf{E}_1 \mathcal{R}_1^{(j)} \Leftrightarrow \mathcal{R}_1^{(j)} = (\lambda_{2j-1} \mathbf{E}_1 - \mathbf{A}_1)^{-1} \mathbf{K}_2 \mathcal{R}_2^{(j-1)}. \quad (41)$$

By multiplying the second equation in (38) on the right with the  $j^{\text{th}}$  unit vector  $\mathbf{e}_{j,k}$ , write:

$$\mathbf{A}_2 \mathcal{R}_2^{(j)} + \mathbf{K}_1 \mathcal{R}_1^{(j-1)} = \lambda_{2j} \mathbf{E}_2 \mathcal{R}_2^{(j)} \Leftrightarrow \mathcal{R}_2^{(j)} = (\lambda_{2j} \mathbf{E}_2 - \mathbf{A}_2)^{-1} \mathbf{K}_1 \mathcal{R}_1^{(j-1)}. \quad (42)$$

From (41) and (42) derive the following linear recursive system of equations:

$$\begin{cases} \mathcal{R}_1^{(j)} = \Phi_1(\lambda_{2j-1}) \mathbf{K}_2 \mathcal{R}_2^{(j-1)}, \\ \mathcal{R}_2^{(j)} = \Phi_2(\lambda_{2j}) \mathbf{K}_1 \mathcal{R}_1^{(j-1)}. \end{cases} \quad (43)$$

with initial conditions (39) and (40). Hence, by solving the recursive system of equations, we conclude that any solution of (38) is given by pairs of generalized controllability matrices defined as in (20). Conversely, it automatically follows that the matrices defined in (20) satisfy the relations in (38). In this case, the assumption made in Condition 4.1 insures that the equations are solvable.

This proof can be straightforward adapted from the simplified case in (28)-(29) to the more general case of interpolation tuples considered in (12)-(16).  $\blacksquare$

**Lemma 4.3** *Consider that the assumption in Condition 4.2 holds, i.e. (21) is valid. Then, the generalized observability matrices  $\mathcal{O}_1$  and  $\mathcal{O}_2$  defined by (22) satisfy the following coupled generalized Sylvester equations:*

$$\begin{cases} \mathcal{O}_1 \mathbf{A}_1 + \mathbf{S}_L \mathcal{O}_2 \mathbf{K}_1 + \mathbf{L} \mathbf{C}_1 = \mathbf{M}_1 \mathcal{O}_1 \mathbf{E}_1, \\ \mathcal{O}_2 \mathbf{A}_2 + \mathbf{S}_L \mathcal{O}_1 \mathbf{K}_2 + \mathbf{L} \mathbf{C}_2 = \mathbf{M}_2 \mathcal{O}_2 \mathbf{E}_2. \end{cases} \quad (44)$$

**Proof of Lemma 4.3** Similar to the proof of Lemma 4.2.  $\blacksquare$

## 4.2 The generalized Loewner pencil

**Definition 4.10** *Given a linear switched system  $\Sigma$  as defined in (1), let  $\{\mathcal{R}_1, \mathcal{R}_2\}$  and  $\{\mathcal{O}_1, \mathcal{O}_2\}$  be the controllability and observability matrices defined in (20), (22) respectively, and associated with the multi-tuples in (13), (17) respectively. The **Loewner matrices**  $\mathbb{L}_1$  and  $\mathbb{L}_2$  are defined as*

$$\mathbb{L}_1 = -\mathcal{O}_1 \mathbf{E}_1 \mathcal{R}_1, \quad \mathbb{L}_2 = -\mathcal{O}_2 \mathbf{E}_2 \mathcal{R}_2. \quad (45)$$

*Additionally, the **shifted Loewner matrices**  $\mathbb{L}_{s_1}$  and  $\mathbb{L}_{s_2}$  are defined as*

$$\mathbb{L}_{s_1} = -\mathcal{O}_1 \mathbf{A}_1 \mathcal{R}_1, \quad \mathbb{L}_{s_2} = -\mathcal{O}_2 \mathbf{A}_2 \mathcal{R}_2. \quad (46)$$

*Also define the quantities*

$$\begin{cases} \mathbf{W}_1 = \mathbf{C}_1 \mathcal{R}_1 \\ \mathbf{W}_2 = \mathbf{C}_2 \mathcal{R}_2 \end{cases}, \quad \begin{cases} \mathbf{V}_1 = \mathcal{O}_1 \mathbf{B}_1 \\ \mathbf{V}_2 = \mathcal{O}_2 \mathbf{B}_2 \end{cases} \quad \text{and} \quad \begin{cases} \mathbf{\Xi}_1 = \mathcal{O}_2 \mathbf{K}_1 \mathcal{R}_1 \\ \mathbf{\Xi}_2 = \mathcal{O}_1 \mathbf{K}_2 \mathcal{R}_2 \end{cases}. \quad (47)$$

**Remark 4.3** In general, the Loewner matrices defined above need not have only real entries. For instance, it may happen that the samples points are purely imaginary values (on the  $j\omega$  axis). In this case, we refer the readers to Section 4.3.1 in [4]. We propose a similar method to enforce all system matrices have only real entries. In short, the sampling points have to be chosen as complex conjugate pairs; after the data is arranged into matrix format, use projection matrices as in equation (4.26) in [4] to multiply the matrices in (45), (46) and (47) to the left and to the right. In this way, the LSS does not change as pointed out in Definition 4.1.

**Remark 4.4** Note that  $\mathbb{L}_k$  and  $\mathbb{L}_{sk}$  (where  $k \in \{1, 2\}$ ), as defined above, are indeed Loewner matrices, that is, they can be expressed as divided differences of appropriate transfer function values of the underlying LSS (see the following example).

**Example 4.2** Given the LSS described by  $(\mathbf{C}_j, \mathbf{E}_j, \mathbf{A}_j, \mathbf{B}_j)$  ( $D = 2$  and  $j \in \{1, 2\}$ ), consider the ordered tuples of left interpolation points:  $\{(\mu_1), (\mu_2, \mu_3)\}$ ,  $\{(\mu_2), (\mu_1, \mu_4)\}$  and right interpolation points  $\{(\lambda_1), (\lambda_3, \lambda_2)\}$ ,  $\{(\lambda_2), (\lambda_4, \lambda_1)\}$ . The associated *generalized observability and controllability* matrices are computed as follows

$$\begin{aligned} \mathcal{O}_1 &= \begin{bmatrix} \mathbf{C}_1 \Phi_1(\mu_1) \\ \mathbf{C}_2 \Phi_2(\mu_2) \mathbf{K}_1 \Phi_1(\mu_3) \end{bmatrix}, \quad \mathcal{O}_2 = \begin{bmatrix} \mathbf{C}_2 \Phi_2(\mu_2) \\ \mathbf{C}_1 \Phi_3(\mu_1) \mathbf{K}_2 \Phi_2(\mu_4) \end{bmatrix}, \\ \mathcal{R}_1 &= \begin{bmatrix} \Phi_1(\lambda_1) \mathbf{B}_1 & \Phi_1(\lambda_3) \mathbf{K}_2 \Phi_2(\lambda_2) \mathbf{B}_2 \end{bmatrix}, \quad \mathcal{R}_2 = \begin{bmatrix} \Phi_2(\lambda_2) \mathbf{B}_2 & \Phi_2(\lambda_4) \mathbf{K}_1 \Phi_1(\lambda_1) \mathbf{B}_1 \end{bmatrix}. \end{aligned}$$

The projected Loewner matrices can be written in terms of the samples in the following way:

$$\begin{aligned} \mathbb{L}_1 &= \begin{bmatrix} \frac{\mathbf{H}_1(\mu_1) - \mathbf{H}_1(\lambda_1)}{\mu_1 - \lambda_1} & \frac{\mathbf{H}_{1,2}(\mu_1, \lambda_2) - \mathbf{H}_{1,2}(\lambda_3, \lambda_2)}{\mu_1 - \lambda_3} \\ \frac{\mathbf{H}_{2,1}(\mu_2, \mu_3) - \mathbf{H}_{2,1}(\mu_2, \lambda_1)}{\mu_3 - \lambda_1} & \frac{\mathbf{H}_{2,1,2}(\mu_2, \mu_3, \lambda_2) - \mathbf{H}_{2,1,2}(\mu_2, \lambda_3, \lambda_2)}{\mu_3 - \lambda_3} \end{bmatrix} = -\mathcal{O}_1 \mathbf{E}_1 \mathcal{R}_1, \\ \mathbb{L}_2 &= \begin{bmatrix} \frac{\mathbf{H}_2(\mu_2) - \mathbf{H}_2(\lambda_2)}{\mu_2 - \lambda_2} & \frac{\mathbf{H}_{2,1}(\mu_2, \lambda_1) - \mathbf{H}_{2,1}(\lambda_4, \lambda_1)}{\mu_2 - \lambda_4} \\ \frac{\mathbf{H}_{1,2}(\mu_1, \mu_4) - \mathbf{H}_{1,2}(\mu_1, \lambda_2)}{\mu_4 - \lambda_2} & \frac{\mathbf{H}_{1,2,1}(\mu_1, \mu_4, \lambda_4) - \mathbf{H}_{1,2,1}(\mu_1, \lambda_4, \lambda_1)}{\mu_4 - \lambda_4} \end{bmatrix} = -\mathcal{O}_2 \mathbf{E}_2 \mathcal{R}_2. \end{aligned}$$

The projected shifted Loewner matrices can also be written in terms of the samples as:

$$\begin{aligned} \mathbb{L}_{s1} &= \begin{bmatrix} \frac{\mu_1 \mathbf{H}_1(\mu_1) - \lambda_1 \mathbf{H}_1(\lambda_1)}{\mu_1 - \lambda_1} & \frac{\mu_1 \mathbf{H}_{1,2}(\mu_1, \lambda_2) - \lambda_3 \mathbf{H}_{1,2}(\lambda_3, \lambda_2)}{\mu_1 - \lambda_3} \\ \frac{\mu_3 \mathbf{H}_{2,1}(\mu_2, \mu_3) - \lambda_1 \mathbf{H}_{2,1}(\mu_2, \lambda_1)}{\mu_3 - \lambda_1} & \frac{\mu_3 \mathbf{H}_{2,1,2}(\mu_2, \mu_3, \lambda_2) - \lambda_3 \mathbf{H}_{2,1,2}(\mu_2, \lambda_3, \lambda_2)}{\mu_3 - \lambda_3} \end{bmatrix} = -\mathcal{O}_1 \mathbf{A}_1 \mathcal{R}_1, \\ \mathbb{L}_{s2} &= \begin{bmatrix} \frac{\mu_2 \mathbf{H}_2(\mu_2) - \lambda_2 \mathbf{H}_2(\lambda_2)}{\mu_2 - \lambda_2} & \frac{\mu_2 \mathbf{H}_{2,1}(\mu_2, \lambda_1) - \lambda_4 \mathbf{H}_{2,1}(\lambda_4, \lambda_1)}{\mu_2 - \lambda_4} \\ \frac{\mu_4 \mathbf{H}_{1,2}(\mu_1, \mu_4) - \lambda_2 \mathbf{H}_{1,2}(\mu_1, \lambda_2)}{\mu_4 - \lambda_2} & \frac{\mu_4 \mathbf{H}_{1,2,1}(\mu_1, \mu_4, \lambda_4) - \lambda_4 \mathbf{H}_{1,2,1}(\mu_1, \lambda_4, \lambda_1)}{\mu_4 - \lambda_4} \end{bmatrix} = -\mathcal{O}_2 \mathbf{A}_2 \mathcal{R}_2. \end{aligned}$$

The same property applies for the  $\mathbf{V}_i$  and  $\mathbf{W}_j$  vectors and  $\mathbf{E}_j$  matrices:

$$\begin{aligned} \mathbf{V}_1 &= \begin{bmatrix} \mathbf{H}_1(\mu_1) \\ \mathbf{H}_{2,1}(\mu_2, \mu_3) \end{bmatrix} = \mathcal{O}_1 \mathbf{B}_1, \quad \mathbf{V}_2 = \begin{bmatrix} \mathbf{H}_2(\mu_2) \\ \mathbf{H}_{1,2}(\mu_1, \mu_4) \end{bmatrix} = \mathcal{O}_2 \mathbf{B}_2, \\ \mathbf{W}_1 &= \begin{bmatrix} \mathbf{H}_1(\lambda_1) & \mathbf{H}_{1,2}(\lambda_3, \lambda_2) \end{bmatrix} = \mathbf{C}_1 \mathcal{R}_1, \quad \mathbf{W}_2 = \begin{bmatrix} \mathbf{H}_2(\lambda_2) & \mathbf{H}_{2,1}(\lambda_4, \lambda_1) \end{bmatrix} = \mathbf{C}_2 \mathcal{R}_2, \\ \mathbf{E}_1 &= \begin{bmatrix} \mathbf{H}_{2,1}(\mu_2, \lambda_1) & \mathbf{H}_{2,1,2}(\mu_2, \lambda_3, \lambda_2) \\ \mathbf{H}_{1,2,1}(\mu_1, \mu_4, \lambda_1) & \mathbf{H}_{1,2,1,2}(\mu_1, \mu_4, \lambda_3, \lambda_2) \end{bmatrix} = \mathcal{O}_2 \mathbf{K}_1 \mathcal{R}_1, \\ \mathbf{E}_2 &= \begin{bmatrix} \mathbf{H}_{1,2}(\mu_1, \lambda_2) & \mathbf{H}_{1,2,1}(\mu_1, \lambda_4, \lambda_1) \\ \mathbf{H}_{2,1,2}(\mu_2, \mu_3, \lambda_2) & \mathbf{H}_{2,1,2,1}(\mu_2, \mu_3, \lambda_4, \lambda_1) \end{bmatrix} = \mathcal{O}_1 \mathbf{K}_2 \mathcal{R}_2. \end{aligned}$$

It readily follows that, given the original system  $\Sigma$ , a reduced LSS of order two can be obtained without computation (matrix factorizations or solves) as:

$$\hat{\mathbf{E}}_k = \mathcal{O}\mathbf{E}\mathcal{R}, \quad \hat{\mathbf{A}} = \mathcal{O}\mathbf{A}\mathcal{R}, \quad \hat{\mathbf{N}} = \mathcal{O}\mathbf{N}\mathcal{R}, \quad \hat{\mathbf{B}} = \mathcal{O}\mathbf{B}, \quad \hat{\mathbf{C}} = \mathcal{C}\mathcal{R}.$$

This reduced system matches sixteen moments of the original system, namely:

$$\begin{aligned} \text{four of } \mathbf{H}_1/\mathbf{H}_2 &: \mathbf{H}_1(\mu_1), \mathbf{H}_2(\mu_2), \mathbf{H}_1(\lambda_1), \mathbf{H}_2(\lambda_2), \\ \text{three of } \mathbf{H}_{1,2} &: \mathbf{H}_{1,2}(\mu_1, \mu_4), \mathbf{H}_{1,2}(\mu_1, \lambda_2), \mathbf{H}_{1,2}(\lambda_3, \lambda_2), \\ \text{three of } \mathbf{H}_{2,1} &: \mathbf{H}_{2,1}(\mu_2, \mu_3), \mathbf{H}_{2,1}(\mu_2, \lambda_1), \mathbf{H}_{2,1}(\lambda_4, \lambda_1), \\ &\vdots \\ \text{one of } \mathbf{H}_{1,2,1,2} &: \mathbf{H}_{1,2,1,2}(\mu_1, \mu_4, \lambda_3, \lambda_2), \\ \text{one of } \mathbf{H}_{2,1,2,1} &: \mathbf{H}_{2,1,2,1}(\mu_2, \mu_3, \lambda_4, \lambda_1). \end{aligned} \tag{48}$$

i.e. in total  $2(2k + k^2) = 16$  moments are matched using this procedure. ■

#### 4.2.1 Properties of the Loewner pencil

We will now show that the quantities defined earlier satisfy various equations which generalize the ones in the linear or bilinear case.

The equations that are presented in this section are used to automatically find the Loewner and shifted Loewner matrices by means of solving Sylvester equations (instead of building the divided difference matrices from the computed samples at the sampling points).

**Proposition 4.1** *The Loewner matrix  $\mathbb{L}_1$  and the shifted Loewner matrix  $\mathbb{L}_{s1}$  (corresponding to mode 1) satisfy the following relations (where  $\mathbf{L}, \mathbf{R}, \mathbf{\Lambda}_k, \mathbf{M}_k, \mathbf{S}_L, \mathbf{S}_R$  are given in (34), (35) and (36)):*

$$\mathbb{L}_{s1} = \mathbb{L}_1\mathbf{\Lambda}_1 + \mathbf{V}_1\mathbf{R} + \mathbf{\Xi}_2\mathbf{S}_R, \tag{49}$$

$$\mathbb{L}_{s1} = \mathbf{M}_1\mathbb{L}_1 + \mathbf{L}\mathbf{W}_1 + \mathbf{S}_L\mathbf{\Xi}_1. \tag{50}$$

*The Loewner matrix  $\mathbb{L}_2$  and the shifted Loewner matrix  $\mathbb{L}_{s2}$  (corresponding to mode 2) satisfy the following relations:*

$$\mathbb{L}_{s2} = \mathbb{L}_2\mathbf{\Lambda}_2 + \mathbf{V}_2\mathbf{R} + \mathbf{\Xi}_1\mathbf{S}_R, \tag{51}$$

$$\mathbb{L}_{s2} = \mathbf{M}_2\mathbb{L}_2 + \mathbf{L}\mathbf{W}_2 + \mathbf{S}_L\mathbf{\Xi}_2. \tag{52}$$

**Proof of Proposition 4.1** By multiplying the first equation in (38) with  $\mathcal{O}_1$  to the left we obtain:

$$\mathcal{O}_1\mathbf{A}_1\mathcal{R}_1 + \mathcal{O}_1\mathbf{K}_2\mathcal{R}_2\mathbf{S}_R + \mathcal{O}_1\mathbf{B}_1\mathbf{R} = \mathcal{O}_1\mathbf{E}_1\mathcal{R}_1\mathbf{\Lambda}_1 \Rightarrow -\mathbb{L}_{s1} + \mathbf{\Xi}_2\mathbf{S}_R + \mathbf{V}_1\mathbf{R} = -\mathbb{L}_1\mathbf{\Lambda}_1,$$

and hence relation (49) is proven. Similarly we prove (51). By multiplying the first equation in (44) with  $\mathcal{R}_1$  to the right we obtain:

$$\mathcal{O}_1\mathbf{A}_1\mathcal{R}_1 + \mathbf{S}_L\mathcal{O}_2\mathbf{K}_1\mathcal{R}_1 + \mathbf{L}\mathcal{C}_1\mathcal{R}_1 = \mathbf{M}_1\mathcal{O}_1\mathbf{E}_1\mathcal{R}_1 \Rightarrow -\mathbb{L}_{s1} + \mathbf{S}_L\mathbf{\Xi}_1 + \mathbf{L}\mathbf{W}_1 = -\mathbf{M}_1\mathbb{L}_1,$$

and hence relation (50) is proven. Similarly we prove (52).

**Proposition 4.2** *The Loewner matrices  $\mathbb{L}_1$  and  $\mathbb{L}_2$  satisfy the following Sylvester equations:*

$$\mathbf{M}_1\mathbb{L}_1 - \mathbb{L}_1\mathbf{\Lambda}_1 = (\mathbf{V}_1\mathbf{R} - \mathbf{L}\mathbf{W}_1) + (\mathbf{\Xi}_2\mathbf{S}_R - \mathbf{S}_L\mathbf{\Xi}_1), \tag{53}$$

$$\mathbf{M}_2\mathbb{L}_2 - \mathbb{L}_2\mathbf{\Lambda}_2 = (\mathbf{V}_2\mathbf{R} - \mathbf{L}\mathbf{W}_2) + (\mathbf{\Xi}_1\mathbf{S}_R - \mathbf{S}_L\mathbf{\Xi}_2). \tag{54}$$

**Proof of Proposition 4.2** By subtracting equation (49) from (50) we directly obtain (53) and also, by subtracting equation (51) from (52) we directly obtain (54).

**Proposition 4.3** *The shifted Loewner matrices  $\mathbb{L}_{s_1}$  and  $\mathbb{L}_{s_2}$  satisfy the following Sylvester equations:*

$$\mathbf{M}_1 \mathbb{L}_{s_1} - \mathbb{L}_{s_1} \mathbf{A}_1 = (\mathbf{M}_1 \mathbf{V}_1 \mathbf{R} - \mathbf{L} \mathbf{W}_1 \mathbf{A}_1) + (\mathbf{M}_1 \mathbf{\Xi}_2 \mathbf{S}_R - \mathbf{S}_L \mathbf{\Xi}_1 \mathbf{A}_1), \quad (55)$$

$$\mathbf{M}_2 \mathbb{L}_{s_2} - \mathbb{L}_{s_2} \mathbf{A}_2 = (\mathbf{M}_2 \mathbf{V}_2 \mathbf{R} - \mathbf{L} \mathbf{W}_2 \mathbf{A}_2) + (\mathbf{M}_2 \mathbf{\Xi}_1 \mathbf{S}_R - \mathbf{S}_L \mathbf{\Xi}_2 \mathbf{A}_2). \quad (56)$$

**Proof of Proposition 4.3** By subtracting equation (49) after being multiplied with  $\mathbf{M}_1$  to the left from equation (50) after being multiplied with  $\mathbf{A}_1$  to the right, we directly obtain (55). Similar procedure is applied to prove (56).

**Remark 4.5** The right hand side of the equations (53) - (56) contains constant  $\{0, 1\}$  matrices (i.e.  $\mathbf{R}, \mathbf{L}, \mathbf{S}_R, \mathbf{S}_L$ ) as well as matrices (i.e.  $\mathbf{V}_j, \mathbf{W}_j, \mathbf{\Xi}_j, j \in \{1, 2\}$ ) which are directly constructed by putting together the given samples values as pointed out in Example 4.2.

### 4.3 Construction of reduced order models

As we already noted, the interpolation data for the LSS has different format than the one used for the linear case without switching, as higher order transfer function values are matched as shown in the previous sections. However, the procedure itself is similar to the one previously presented in Section 3, i.e. in Lemma 3.1.

**Lemma 4.4** *Assume that  $k = \ell$  and that the interpolation points are chosen to satisfy the conditions in (19) and (21). Moreover, assume that the Loewner matrices  $\mathbb{L}_1$  and  $\mathbb{L}_2$  to be invertible. Then, a realization of a reduced order LSS  $\hat{\Sigma}$  that matches the data of the original LSS  $\Sigma$  (as introduced in Definition 4.9) is given by the following matrices,*

$$\begin{cases} \hat{\mathbf{E}}_1 = -\mathbb{L}_1, & \hat{\mathbf{A}}_1 = -\mathbb{L}_{s_1}, & \hat{\mathbf{B}}_1 = \mathbf{V}_1, & \hat{\mathbf{C}}_1 = \mathbf{W}_1, \\ \hat{\mathbf{E}}_2 = -\mathbb{L}_2, & \hat{\mathbf{A}}_2 = -\mathbb{L}_{s_2}, & \hat{\mathbf{B}}_2 = \mathbf{V}_2, & \hat{\mathbf{C}}_2 = \mathbf{W}_2 \end{cases} \quad \text{and} \quad \hat{\mathbf{K}}_1 = \mathbf{\Xi}_1, \quad \hat{\mathbf{K}}_2 = \mathbf{\Xi}_2.$$

*If  $k = n$ , then the proposed realization is equivalent to the original one (as in Definition 4.1).*

**Proof of Lemma 4.4** This result directly follows from Lemma 4.1 by taking into consideration the notations introduced in (45-47).

In the case of redundant data, at least one of the pencils  $(\mathbb{L}_{s_j}, \mathbb{L}_j)$  is singular (for  $j \in \{1, 2\}$ ), and hence construct pairs of projectors  $(\mathbf{X}_j, \mathbf{Y}_j)$  (corresponding to mode  $j$ ) similar to (10). The MOR procedure for approximate data matching is presented as follows.

**Procedure 1** Consider the rank revealing singular value factorization of the following matrices composed of the Loewner and shifted Loewner matrices corresponding to mode  $j \in \{1, 2\}$ , as

$$\begin{aligned} \begin{bmatrix} \mathbb{L}_j & \mathbb{L}_{s_j} \end{bmatrix} &= \begin{bmatrix} \mathbf{Y}_j^{(1)} & \tilde{\mathbf{Y}}_j^{(1)} \end{bmatrix} \begin{bmatrix} \mathbf{S}_j^{(1)} & \mathbf{O} \\ \mathbf{O} & \tilde{\mathbf{S}}_j^{(1)} \end{bmatrix} \begin{bmatrix} \mathbf{X}_j^{(1)} & \tilde{\mathbf{X}}_j^{(1)} \end{bmatrix}^T = \mathbf{Y}_j^{(1)} \mathbf{S}_j^{(1)} (\mathbf{X}_j^{(1)})^T + \tilde{\mathbf{Y}}_j^{(1)} \tilde{\mathbf{S}}_j^{(1)} (\tilde{\mathbf{X}}_j^{(1)})^T. \\ \begin{bmatrix} \mathbb{L}_j \\ \mathbb{L}_{s_j} \end{bmatrix} &= \begin{bmatrix} \mathbf{Y}_j^{(2)} & \tilde{\mathbf{Y}}_j^{(2)} \end{bmatrix} \begin{bmatrix} \mathbf{S}_j^{(2)} & \mathbf{O} \\ \mathbf{O} & \tilde{\mathbf{S}}_j^{(2)} \end{bmatrix} \begin{bmatrix} \mathbf{X}_j^{(2)} & \tilde{\mathbf{X}}_j^{(2)} \end{bmatrix}^T = \mathbf{Y}_j^{(2)} \mathbf{S}_j^{(2)} (\mathbf{X}_j^{(2)})^T + \tilde{\mathbf{Y}}_j^{(2)} \tilde{\mathbf{S}}_j^{(2)} (\tilde{\mathbf{X}}_j^{(2)})^T, \end{aligned} \quad (57)$$

where  $\mathbf{Y}_j^{(i)}, \mathbf{X}_j^{(i)} \in \mathbb{R}^{k \times r_j}$  and  $\mathbf{S}_j^{(i)} \in \mathbb{R}^{r_j \times r_j}$  for  $i \in \{1, 2\}$ . The projected system matrices corresponding to subsystem  $\hat{\Sigma}_j$  are computed as,

$$\hat{\mathbf{E}}_j = -(\mathbf{Y}_j^{(1)})^T \mathbb{L}_j \mathbf{X}_j^{(2)}, \hat{\mathbf{A}}_j = -(\mathbf{Y}_j^{(1)})^T \mathbb{L}_{s_j} \mathbf{X}_j^{(2)}, \hat{\mathbf{B}}_j = (\mathbf{Y}_j^{(1)})^T \mathbf{V}_j, \hat{\mathbf{C}}_j = \mathbf{W}_j \mathbf{X}_j^{(2)}, \text{ for } j \in \{1, 2\}$$

Moreover, the projected coupling matrices are computed in the following way

$$\hat{\mathbf{K}}_1 = (\mathbf{Y}_2^{(1)})^T \mathbf{\Xi}_1 \mathbf{X}_1^{(2)}, \hat{\mathbf{K}}_2 = (\mathbf{Y}_1^{(1)})^T \mathbf{\Xi}_2 \mathbf{X}_2^{(2)}.$$

By choosing  $r_j$  as the numerical rank of the Loewner matrix  $\mathbb{L}_j$  (i.e. the largest neglected singular value corresponding to index  $r_j + 1$  is less than machine precision  $\epsilon$ ), ensure that the  $\hat{\mathbf{E}}_j$  matrices are not singular. Hence, construct a reduced order LSS denoted with  $\hat{\Sigma}$ , that approximately matches the data of the original LSS  $\Sigma$ . If the truncated singular values are all 0 (the ones on the main diagonal of the matrices  $\tilde{\mathbf{S}}_j^{(i)}$ ), then the matching is exact.

We provide a qualitative rather than quantitative result for the projected Loewner case. The quality of approximation is directly linked to the singular values of the Loewner pencils which represent an indicator of the desired accuracy. For linear systems with no switching, an error bound is provided in [6] as a quantitative measure.

The dimensions of the subsystems  $\hat{\Sigma}_1$  and  $\hat{\Sigma}_2$ , corresponding to the reduced order LSS, need not be the same (i.e.  $r_1 \neq r_2$ ). In this case the coupling matrices are not square anymore.

The projectors are computed via singular value factorization of the Loewner matrices. The use of the *Drazin* or *Moore-Penrose* pseudo inverses also holds (as shown in [2]).

## 5 The Loewner framework for linear switched systems - the general case

In this section we are mainly concerned with generalizing some of the results presented in Section 4. Most of the findings can be smoothly extended to the cases with more complex switching patterns (more modes). By enforcing a prefix/suffix closure structure in the proposed framework, we can show that all interpolation conditions can be written in matrix equation format.

**Definition 5.1** Let  $\Gamma$  and  $\Theta$  be finite sets of tuples so that  $\Gamma, \Theta \subseteq \bigcup_{k=1}^{\infty} \mathbf{Q}^k \times \mathbb{C}^k$  so that  $\Gamma$  has the prefix closure property, i.e.

$$(q_1, q_2, \dots, q_i, \lambda_1, \dots, \lambda_i) \in \Gamma \Rightarrow (q_2, \dots, q_i, \lambda_2, \dots, \lambda_i) \in \Gamma \quad \forall i \geq 2,$$

and  $\Theta$  has the suffix closure property, i.e.

$$(q_1, q_2, \dots, q_j, \mu_1, \dots, \mu_j) \in \Theta \Rightarrow (q_1, \dots, q_{j-1}, \mu_1, \dots, \mu_{j-1}) \in \Theta \quad \forall j \geq 2.$$

Now consider the specific subset  $\Gamma_q$  (for any  $q \in \mathbf{Q}$ ) of the set  $\Gamma$ , defined in the following way:

$$\Gamma_q = \{(q_1, q_2, \dots, q_i, \lambda_1, \dots, \lambda_i) \in \Gamma \mid q_1 = q, i \leq \delta_\Gamma\}, \quad \delta_\Gamma = \max_{w \in \Gamma} (|w|)/2.$$

Denote the cardinality of  $\Gamma_q$  with  $k_q = \text{card}(\Gamma_q)$  and explicitly enumerate the elements of this set as follows:  $\Gamma_q = \{w_q^{(1)}, w_q^{(2)}, \dots, w_q^{(k_q)}\}$ . Consider the following function (mapping)  $\mathbf{r} : \Gamma_q \rightarrow \mathbb{C}^{n_q \times 1}$  that maps a word form  $\Gamma_q$  into a column vector of size  $n_q$ :

$$\mathbf{r}((q, q_2, \dots, q_i, \lambda_1, \dots, \lambda_i)) = \mathbf{\Phi}_q(\lambda_1) \mathbf{K}_{q_2, q} \mathbf{\Phi}_{q_2}(\lambda_2) \cdots \mathbf{K}_{q_i, q_{i-1}} \mathbf{\Phi}_{q_i}(\lambda_i) \mathbf{B}_{q_i}.$$

Next, construct the controllability matrix  $\mathcal{R}_q$  corresponding to the mode  $q$  of the system  $\Sigma$  as follows:

$$\mathcal{R}_q = \begin{bmatrix} \mathbf{r}(w_q^{(1)}) & \mathbf{r}(w_q^{(2)}) & \cdots & \mathbf{r}(w_q^{(k_q)}) \end{bmatrix} \in \mathbb{C}^{n_q \times k_q}. \quad (58)$$

Similarly, define the specific subset  $\Theta_q$  (for any  $q \in \mathbf{Q}$ ) of the set  $\Theta$  in the following way:

$$\Theta_q = \{(q_1, q_2, \dots, q_j, \mu_1, \dots, \mu_j) \in \Gamma \mid q_j = q, j \leq \delta_\Theta\}, \quad \delta_\Theta = \max_{w \in \Theta} (|w|)/2.$$

Consider the cardinality of  $\Theta_q$  to be the same as the one of  $\Gamma_q$ , i.e.  $k_q = \text{card}(\Theta_q)$ . Although this additional constraint is not necessarily needed, we would like to enforce the construction of reduced systems with square matrices  $\mathbf{A}_k$  and  $\mathbf{E}_k$ . Next we explicitly enumerate the elements of this set as follows:  $\Theta_q = \{v_q^{(1)}, v_q^{(2)}, \dots, v_q^{(k_q)}\}$ . Consider the following mapping  $\mathbf{o} : \Theta_q \rightarrow \mathbb{C}^{1 \times n_q}$  that maps a word form  $\Theta_q$  into a row vector of size  $n_q$ :

$$\mathbf{o}((q_1, q_2, \dots, q_{j-1}, q, \mu_1, \dots, \mu_j)) = \mathbf{C}_{q_1} \Phi_{q_1}(\mu_1) \mathbf{K}_{q_2, q_1} \Phi_{q_2}(\mu_2) \cdots \mathbf{K}_{q, q_{j-1}} \Phi_q(\mu_j).$$

Next, construct the observability matrix  $\mathcal{O}_q \in \mathbb{C}^{k_q \times n_q}$  corresponding to the mode  $q$  of the system  $\Sigma$  as follows

$$\mathcal{O}_q = \begin{bmatrix} \mathbf{o}(v_q^{(1)})^T & \mathbf{o}(v_q^{(2)})^T & \cdots & \mathbf{o}(v_q^{(k_q)})^T \end{bmatrix}^T \in \mathbb{C}^{k_q \times n_q}. \quad (59)$$

Consider the following example to show how the general procedure is extended from the linear case (no switching) to the case when switching occurs.

**Example 5.1** Take  $D = 3$  (3 active modes) and hence  $\mathbf{Q} = \{1, 2, 3\}$ . The following interpolation points are given:  $\{s_1, s_2, \dots, s_{18}\} \subset \mathbb{C}$ . The first step is to partition this set into two disjoint subsets (each having 9 points):

$$\underline{\text{left interpolation points}} : \{\mu_1, \mu_2, \dots, \mu_9\}, \quad \underline{\text{right interpolation points}} : \{\lambda_1, \lambda_2, \dots, \lambda_9\}.$$

The set  $\Gamma$  is composed of three subsets  $\Gamma = \Gamma_1 \cup \Gamma_2 \cup \Gamma_3$  which are defined by imposing the previously defined suffix closure property, as

$$\begin{cases} \Gamma_1 = \{(1, \lambda_1), (1, 3, \lambda_4, \lambda_3), (1, 3, 2, \lambda_7, \lambda_6, \lambda_2)\}, \\ \Gamma_2 = \{(2, \lambda_2), (2, 1, \lambda_5, \lambda_1), (2, 1, 3, \lambda_8, \lambda_4, \lambda_3)\}, \\ \Gamma_3 = \{(3, \lambda_3), (3, 2, \lambda_6, \lambda_2), (3, 2, 1, \lambda_9, \lambda_5, \lambda_1)\}. \end{cases}$$

To the sets  $\Gamma_j$ , we attach the following controllability matrices  $\mathcal{R}_j$

$$\begin{aligned} \mathcal{R}_1 &= \begin{bmatrix} \Phi_1(\lambda_1) & \Phi_1(\lambda_4) \mathbf{K}_{3,1} \Phi_3(\lambda_3) \mathbf{B}_3 & \Phi_1(\lambda_7) \mathbf{K}_{3,1} \Phi_3(\lambda_6) \mathbf{K}_{2,3} \Phi_2(\lambda_2) \mathbf{B}_2 \end{bmatrix}, \\ \mathcal{R}_2 &= \begin{bmatrix} \Phi_2(\lambda_2) & \Phi_2(\lambda_5) \mathbf{K}_{1,2} \Phi_1(\lambda_1) \mathbf{B}_1 & \Phi_2(\lambda_8) \mathbf{K}_{1,2} \Phi_1(\lambda_4) \mathbf{K}_{3,1} \Phi_3(\lambda_3) \mathbf{B}_3 \end{bmatrix}, \\ \mathcal{R}_3 &= \begin{bmatrix} \Phi_3(\lambda_3) & \Phi_3(\lambda_6) \mathbf{K}_{2,3} \Phi_2(\lambda_2) \mathbf{B}_2 & \Phi_3(\lambda_9) \mathbf{K}_{2,3} \Phi_2(\lambda_5) \mathbf{K}_{1,2} \Phi_1(\lambda_1) \mathbf{B}_1 \end{bmatrix}. \end{aligned}$$

In the same manner, the set  $\Theta$  is composed of three subsets  $\Theta = \Theta_1 \cup \Theta_2 \cup \Theta_3$ . These are defined by means of imposing the previously defined prefix closure property, as follows

$$\begin{cases} \Theta_1 = \{(1, \mu_1), (3, 1, \mu_3, \mu_4), (1, 2, 1, \mu_1, \mu_5, \mu_7)\}, \\ \Theta_2 = \{(2, \mu_2), (1, 2, \mu_1, \mu_5), (2, 3, 2, \mu_2, \mu_6, \mu_8)\}, \\ \Theta_3 = \{(3, \mu_3), (2, 3, \mu_2, \mu_6), (3, 1, 3, \mu_3, \mu_4, \mu_9)\}. \end{cases}$$

To the sets  $\Theta_i$ , we attach the observability matrices  $\mathcal{O}_i$  defined as follows:

$$\mathcal{O}_1 = \begin{bmatrix} \mathbf{C}_1 \Phi_1(\mu_1) \\ \mathbf{C}_3 \Phi_3(\mu_3) \mathbf{K}_{1,3} \Phi_1(\mu_4) \\ \mathbf{C}_1 \Phi_1(\mu_1) \mathbf{K}_{2,1} \Phi_2(\mu_5) \mathbf{K}_{1,2} \Phi_1(\mu_7) \end{bmatrix}, \quad \mathcal{O}_2 = \begin{bmatrix} \mathbf{C}_2 \Phi_2(\mu_2) \\ \mathbf{C}_1 \Phi_1(\mu_1) \mathbf{K}_{2,1} \Phi_2(\mu_5) \\ \mathbf{C}_2 \Phi_2(\mu_2) \mathbf{K}_{3,2} \Phi_3(\mu_6) \mathbf{K}_{2,3} \Phi_2(\mu_8) \end{bmatrix},$$

$$\mathcal{O}_3 = \begin{bmatrix} \mathbf{C}_3 \Phi_3(\mu_3) \\ \mathbf{C}_2 \Phi_2(\mu_2) \mathbf{K}_{3,2} \Phi_3(\mu_6) \\ \mathbf{C}_3 \Phi_3(\mu_3) \mathbf{K}_{1,3} \Phi_1(\mu_4) \mathbf{K}_{3,1} \Phi_3(\mu_9) \end{bmatrix}.$$

## 5.1 Sylvester equations for $\mathcal{R}_q$ and $\mathcal{O}_q$

In this section we would like to generalize the results presented in Lemma 4.2 and Lemma 4.3, and hence extend the framework to a general number of operational modes denoted with  $D$ .

**Definition 5.2** Introduce the special concatenation of tuples composed of mixed elements (symbols) from the sets  $\mathbf{Q}$  and  $\mathbb{C}$ , as the mapping with the following property:

$$(\alpha_1 \odot \beta_1) \odot (\alpha_2 \odot \beta_2) = \left( (\alpha_1 \odot \alpha_2) \odot (\beta_1 \odot \beta_2) \right),$$

where  $\alpha_k \in \mathbf{Q}^{i_k}$  and  $\beta_k \in \mathbb{C}^{j_k}$  for  $i_k, j_k \geq 1$  and  $k = 1, 2$ .

**Definition 5.3** For  $g, i = 1, \dots, D$ , let  $\mathbf{S}_i^{(g)} = \left[ \mathbf{S}_i^{(g)}(1) \ \dots \ \mathbf{S}_i^{(g)}(k_g) \right] \in \mathbb{R}^{k_i \times k_g}$  be constant matrices that contain only 0/1 entries constructed so that  $\mathbf{S}_i^{(g)}(1) = \mathbf{0}_{k_i,1}$  and for  $u = 2, \dots, k_g$ , we write:

$$\mathbf{S}_i^{(g)}(u) = \begin{cases} \mathbf{e}_{u-1, k_i}, & \text{if } \exists \tilde{\lambda} \in \mathbb{C}, \text{ s.t. } \mathbf{w}_g^{(u)} = (g, \tilde{\lambda}) \odot \mathbf{w}_i^{(u-1)}, \\ \mathbf{0}_{k_i,1}, & \text{else} \end{cases}. \quad (60)$$

Also, introduce the matrices  $\mathbf{R}^{(i)}$  and  $\Lambda_i$  that are defined similarly as in (34) and (37),

$$\mathbf{R}^{(i)} = [\mathbf{e}_{1, m_1}^T \ \dots \ \mathbf{e}_{1, m_{\mathcal{K}}}^T] \in \mathbb{R}^{1 \times k_i}, \quad \Lambda_i = \text{blkdiag}[\Lambda_i^{(1)}, \Lambda_i^{(2)}, \dots, \Lambda_i^{(\mathcal{L})}] \in \mathbb{R}^{k_i \times k_i}, \quad (61)$$

where the diagonal matrices  $\Lambda_i^{(a)}$ ,  $a = 1, \dots, \mathcal{K}$  contain the right interpolation points associated to mode  $i$ . For a set  $\Gamma$  with general structure (as in Definition 5.1), it follows that the controllability matrices  $\mathcal{R}_i \in \mathbb{R}^{n_i \times k_i}$ ,  $1 \leq i \leq D$  satisfy the following system of generalized Sylvester equations:

$$\begin{cases} \mathbf{A}_1 \mathcal{R}_1 + \sum_{i=1}^D \mathbf{K}_{i,1} \mathcal{R}_i \mathbf{S}_i^{(1)} + \mathbf{B}_1 \mathbf{R}^{(1)} = \mathbf{E}_1 \mathcal{R}_1 \Lambda_1, \\ \mathbf{A}_2 \mathcal{R}_2 + \sum_{i=1}^D \mathbf{K}_{i,2} \mathcal{R}_i \mathbf{S}_i^{(2)} + \mathbf{B}_2 \mathbf{R}^{(2)} = \mathbf{E}_2 \mathcal{R}_2 \Lambda_2, \\ \vdots \\ \mathbf{A}_D \mathcal{R}_D + \sum_{i=1}^D \mathbf{K}_{i,D} \mathcal{R}_i \mathbf{S}_i^{(D)} + \mathbf{B}_D \mathbf{R}^{(D)} = \mathbf{E}_D \mathcal{R}_D \Lambda_D. \end{cases} \quad (62)$$

Note that  $\mathbf{S}_i^{(i)} = \mathbf{0}_{k_i, k_i}$ , and if  $k_1 = k_2 = \dots = k_D = k$ , the above defined matrices  $\mathbf{S}_i^{(g)}$  satisfy the following equality  $\forall g \in \mathbf{Q}$  :

$$\sum_{i=1}^D \mathbf{S}_i^{(g)} = \text{blkdiag} [\mathbf{J}_{m_1}, \dots, \mathbf{J}_{m_{\mathcal{K}}}] . \quad (63)$$

where  $\mathbf{J}_l$  is the Jordan block of size  $l$  defined in (35).

To directly find  $\mathcal{R}_g$ ,  $g = 1, 2, 3$  for the case presented in Example 5.1, we have to solve the following system of coupled generalized Sylvester equations

$$\begin{cases} \mathbf{A}_1 \mathcal{R}_1 + \mathbf{K}_{3,1} \mathcal{R}_3 \mathbf{S}_3^{(1)} + \mathbf{B}_1 \mathbf{R} = \mathbf{E}_1 \mathcal{R}_1 \mathbf{\Lambda}_1, \\ \mathbf{A}_2 \mathcal{R}_2 + \mathbf{K}_{1,2} \mathcal{R}_1 \mathbf{S}_1^{(2)} + \mathbf{B}_2 \mathbf{R} = \mathbf{E}_2 \mathcal{R}_2 \mathbf{\Lambda}_2, \\ \mathbf{A}_3 \mathcal{R}_3 + \mathbf{K}_{2,3} \mathcal{R}_2 \mathbf{S}_2^{(3)} + \mathbf{B}_3 \mathbf{R} = \mathbf{E}_3 \mathcal{R}_3 \mathbf{\Lambda}_3. \end{cases}$$

where:

$$\mathbf{\Lambda}_1 = \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_4 & 0 \\ 0 & 0 & \lambda_7 \end{bmatrix}, \quad \mathbf{\Lambda}_2 = \begin{bmatrix} \lambda_2 & 0 & 0 \\ 0 & \lambda_5 & 0 \\ 0 & 0 & \lambda_8 \end{bmatrix}, \quad \mathbf{\Lambda}_3 = \begin{bmatrix} \lambda_3 & 0 & 0 \\ 0 & \lambda_6 & 0 \\ 0 & 0 & \lambda_9 \end{bmatrix},$$

$$\mathbf{R} = [1 \ 0 \ 0], \quad \mathbf{S}_3^{(1)} = \mathbf{S}_1^{(2)} = \mathbf{S}_2^{(3)} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}.$$

This corresponds to the case  $k_1 = k_2 = k_3 = 3$ ,  $\mathcal{K} = 1$  and  $m_1 = 3$ .

**Definition 5.4** For  $h, j = 1, \dots, D$ , let  $\mathbf{T}_j^{(h)} = \left[ (\mathbf{T}_j^{(h)})^T(1) \ \dots \ (\mathbf{T}_j^{(h)})^T(k_h) \right]^T \in \mathbb{R}^{\ell_h \times \ell_j}$  be constant matrices that contain only 0/1 entries constructed so that  $(\mathbf{T}_j^{(h)})^T(1) = \mathbf{0}_{\ell_j, 1}$  and for  $v = 2, \dots, k_g$ , we write:

$$(\mathbf{T}_j^{(h)})^T(v) = \begin{cases} \mathbf{e}_{v-1, k_j}, & \text{if } \exists \tilde{\mu} \in \mathbb{C}, \text{ s.t. } \mathbf{w}_h^{(v)} = \mathbf{w}_j^{(v-1)} \odot (h, \tilde{\mu}), \\ \mathbf{0}_{\ell_j, 1}, & \text{else} \end{cases} . \quad (64)$$

Also, introduce the following matrices

$$(\mathbf{L}^{(j)})^T = [\mathbf{e}_{1, p_1}^T \ \dots \ \mathbf{e}_{1, p_{\mathcal{L}}}^T] \in \mathbb{R}^{1 \times \ell_j}, \quad \mathbf{M}_j = \text{blkdiag} [\mathbf{M}_j^{(1)}, \mathbf{M}_j^{(2)}, \dots, \mathbf{M}_j^{(\mathcal{L})}] \in \mathbb{R}^{\ell_j \times \ell_j}, \quad (65)$$

where the diagonal matrices  $\mathbf{M}_j^{(v)}$  for  $v = 1, \dots, \ell_j$  contain the left interpolation points associated to mode  $j$ . For a set  $\Theta$  with general structure (as in Definition 5.1), one can conclude that the observability matrices  $\mathcal{O}_j \in \mathbb{R}^{\ell_j \times n_j}$ ,  $1 \leq j \leq D$  satisfy the following system of generalized Sylvester equations:

$$\begin{cases} \mathcal{O}_1 \mathbf{A}_1 + \sum_{j=1}^D \mathbf{T}_j^{(1)} \mathcal{O}_j \mathbf{K}_{1,j} + \mathbf{L}^{(1)} \mathbf{C}_1 = \mathbf{M}_1 \mathcal{O}_1 \mathbf{E}_1, \\ \mathcal{O}_2 \mathbf{A}_2 + \sum_{j=1}^D \mathbf{T}_j^{(2)} \mathcal{O}_j \mathbf{K}_{2,j} + \mathbf{L}^{(2)} \mathbf{C}_2 = \mathbf{M}_2 \mathcal{O}_2 \mathbf{E}_2, \\ \vdots \\ \mathcal{O}_D \mathbf{A}_D + \sum_{j=1}^D \mathbf{T}_j^{(D)} \mathcal{O}_j \mathbf{K}_{D,j} + \mathbf{L}^{(D)} \mathbf{C}_D = \mathbf{M}_D \mathcal{O}_D \mathbf{E}_D. \end{cases} \quad (66)$$



Note that  $\mathbf{T}_j^{(j)} = \mathbf{0}_{\ell_j, \ell_j}$ , and if  $\ell_1 = \ell_2 = \dots = \ell_D = \ell$ , the square matrices  $\mathbf{T}_j^{(h)} \in \mathbb{R}^{\ell \times \ell}$  satisfy the following equality,  $\forall h \in Q$ :

$$\sum_{j=1}^D \mathbf{T}_j^{(h)} = \text{blkdiag} [\mathbf{J}_{p_1}, \dots, \mathbf{J}_{p_{\mathcal{L}}}]^T. \quad (67)$$

Again to find the matrices  $\mathcal{O}_h$ ,  $h = 1, 2, 3$  in Example 5.1, it is required to solve the following system of coupled generalized Sylvester equations

$$\begin{cases} \mathcal{O}_1 \mathbf{A}_1 + \mathbf{T}_3^{(1)} \mathcal{O}_3 \mathbf{K}_{1,3} + \mathbf{T}_2^{(1)} \mathcal{O}_2 \mathbf{K}_{1,2} + \mathbf{L} \mathbf{C}_1 = \mathbf{M}_1 \mathcal{O}_1 \mathbf{E}_1, \\ \mathcal{O}_2 \mathbf{A}_2 + \mathbf{T}_1^{(2)} \mathcal{O}_1 \mathbf{K}_{2,1} + \mathbf{T}_3^{(2)} \mathcal{O}_3 \mathbf{K}_{2,3} + \mathbf{L} \mathbf{C}_2 = \mathbf{M}_2 \mathcal{O}_2 \mathbf{E}_2, \\ \mathcal{O}_3 \mathbf{A}_3 + \mathbf{T}_2^{(3)} \mathcal{O}_2 \mathbf{K}_{3,2} + \mathbf{T}_1^{(3)} \mathcal{O}_1 \mathbf{K}_{3,1} + \mathbf{L} \mathbf{C}_3 = \mathbf{M}_3 \mathcal{O}_3 \mathbf{E}_3. \end{cases}$$

where:

$$\mathbf{M}_1 = \begin{bmatrix} \mu_1 & 0 & 0 \\ 0 & \mu_4 & 0 \\ 0 & 0 & \mu_7 \end{bmatrix}, \quad \mathbf{M}_2 = \begin{bmatrix} \mu_2 & 0 & 0 \\ 0 & \mu_5 & 0 \\ 0 & 0 & \mu_8 \end{bmatrix}, \quad \mathbf{M}_3 = \begin{bmatrix} \mu_3 & 0 & 0 \\ 0 & \mu_6 & 0 \\ 0 & 0 & \mu_9 \end{bmatrix},$$

$$\mathbf{T}_3^{(1)} = \mathbf{T}_1^{(2)} = \mathbf{T}_2^{(3)} = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad \mathbf{T}_2^{(1)} = \mathbf{T}_3^{(2)} = \mathbf{T}_1^{(3)} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}, \quad \mathbf{L} = \mathbf{e}_{1,3}.$$

This corresponds to the case  $\ell_1 = \ell_2 = \ell_3 = 3$ ,  $\mathcal{L} = 1$  and  $p_1 = 3$ . Note that the relation in (67) hold, i.e.  $\mathbf{T}_2^{(1)} + \mathbf{T}_3^{(1)} = \mathbf{T}_1^{(2)} + \mathbf{T}_3^{(2)} = \mathbf{T}_1^{(3)} + \mathbf{T}_2^{(3)} = \mathbf{J}_3^T$ .

Throughout this section, we chose cycling switching to make the exposition more comprehensible and to directly relate it to the case  $D = 2$ . The cyclic property of the switching scenarios comes naturally in many applications (for example in the boost power converter model in [9], Section 5.2), but does not necessarily need to be enforced in our framework (Definition 5.1 is not constrained to only such type of switching). Consequently, in Example 5.1, the construction of the sets  $\Gamma_k$  is indeed cyclic, while the sets  $\Theta_k$  do not possess this property.

## 5.2 The Loewner matrices

For the case of linear switched systems with  $D$  active modes, the generalization of the Loewner framework includes one important feature. Instead of only one pair of Loewner matrices (as in the linear case without switching which is covered in Section 3), we define a pair of Loewner matrices for each individual active mode; hence in total  $D$  pairs of Loewner matrices.

**Definition 5.5** *Given a linear switched system  $\Sigma$ , let  $\{\mathcal{R}_i | i \in \mathbf{Q}\}$  and  $\{\mathcal{O}_j | j \in \mathbf{Q}\}$  be the controllability and observability matrices associated with the multi-tuples  $\Gamma_i$  and  $\Theta_j$ . The **Loewner matrices**  $\{\mathbb{L}_i | i \in \mathbf{Q}\}$  are defined as*

$$\mathbb{L}_1 = -\mathcal{O}_1 \mathbf{E}_1 \mathcal{R}_1, \quad \mathbb{L}_2 = -\mathcal{O}_2 \mathbf{E}_2 \mathcal{R}_2, \quad \dots, \quad \mathbb{L}_D = -\mathcal{O}_D \mathbf{E}_D \mathcal{R}_D, \quad (68)$$

*Additionally, the **shifted Loewner matrices**  $\{\mathbb{L}_{s_i} | i \in \mathbf{Q}\}$  are defined as*

$$\mathbb{L}_{s_1} = -\mathcal{O}_1 \mathbf{A}_1 \mathcal{R}_1, \quad \mathbb{L}_{s_2} = -\mathcal{O}_2 \mathbf{A}_2 \mathcal{R}_2, \quad \dots, \quad \mathbb{L}_{s_D} = -\mathcal{O}_D \mathbf{A}_D \mathcal{R}_D. \quad (69)$$

*Also introduce the matrices  $\forall i, j \in Q$*

$$\mathbf{W}_i = \mathbf{C}_i \mathcal{R}_i, \quad \mathbf{V}_j = \mathcal{O}_j \mathbf{B}_j, \quad \text{and} \quad \mathbf{\Xi}_{i,j} = \mathcal{O}_j \mathbf{K}_{i,j} \mathcal{R}_i.$$

**Remark 5.1** The number of Loewner matrices, shifted Loewner matrices,  $W_i$  row vectors and  $V_j$  column vectors is the same as the number of active modes (i.e  $D$ ). On the other hand, the number of matrices  $\Xi_{i,j}$  increases quadratically with  $D$  (i.e in total  $D^2$  matrices).

**Remark 5.2** Note that the matrices  $\mathbb{L}_i$  and  $\mathbb{L}_{s_i}$  as defined in (68) and (69) (for  $i \in \{1, 2, \dots, D\}$ ) are indeed Loewner matrices, that is, they can be expressed as *divided differences* of generalized transfer function values of the underlying LSS.

**Proposition 5.1** *The Loewner matrices  $\mathbb{L}_h$  satisfy the following Sylvester equations:*

$$\mathbf{M}_h \mathbb{L}_h - \mathbb{L}_h \mathbf{A}_h = (\mathbf{V}_h \mathbf{R} - \mathbf{L} \mathbf{W}_h) + \sum_{j=1}^D \left( \Xi_{j,h} \mathbf{S}_j^{(h)} - \mathbf{T}_j^{(h)} \Xi_{h,j} \right), \quad h \in Q. \quad (70)$$

**Proposition 5.2** *The shifted Loewner matrices  $\mathbb{L}_{s_h}$  satisfy the following Sylvester equations:*

$$\mathbf{M}_h \mathbb{L}_{s_h} - \mathbb{L}_{s_h} \mathbf{A}_h = (\mathbf{M}_h \mathbf{V}_h \mathbf{R} - \mathbf{L} \mathbf{W}_h \mathbf{A}_h) + \sum_{j=1}^D \left( \mathbf{M}_h \Xi_{j,h} \mathbf{S}_j^{(h)} - \mathbf{T}_j^{(h)} \Xi_{h,j} \mathbf{A}_h \right), \quad h \in Q. \quad (71)$$

**Remark 5.3** The proof of the results stated in (70)-(71) is performed in a similar manner as for the results obtained for the special case  $D = 2$  in Section 4 (i.e. for (53)-(56)).

### 5.3 Construction of reduced order models

The general procedure for the case with  $D$  switching modes is more or less similar to the one covered in Section 4.3 (where  $D = 2$ ).

**Lemma 5.1** *Let  $\mathbb{L}_j$  be invertible matrices for  $1 \leq j \leq D$ , such that none of the interpolation points  $\lambda_i, \mu_k$  are eigenvalues of any of the Loewner pencils  $(\mathbb{L}_{s_j}, \mathbb{L}_j)$ . Then, the matrices*

$$\{\hat{\mathbf{E}}_j = -\mathbb{L}_j, \quad \hat{\mathbf{A}}_1 = -\mathbb{L}_{s_j}, \quad \hat{\mathbf{B}}_j = \mathbf{V}_j, \quad \hat{\mathbf{C}}_j = \mathbf{W}_j, \quad \hat{\mathbf{K}}_{i,j} = \Xi_{i,j}, \quad i, j \in \{1, \dots, D\},$$

*form a realization of a reduced order LSS  $\hat{\Sigma}$  that matches the data of the original LSS  $\Sigma$ . If  $k_j = n_j$  for  $1 \leq j \leq D$ , the proposed realization is equivalent to the original one.*

The concept of an LSS matching the data of another LSS in the case  $D > 2$  is formulated in a similar manner as to the case  $D = 2$ , which is covered in Definition 4.9. Also, the definition of equivalent LSS for the case  $D > 2$  is formulated similarly as to Definition 4.1.

In the case of redundant data, at least one of the pencils  $(\mathbb{L}_{s_j}, \mathbb{L}_j)$  is singular (for  $j \in \{1, \dots, D\}$ ). The main procedure is presented as follows.

**Procedure 2** Consider the rank revealing singular value factorization of the matrices composed of the Loewner matrices  $\mathbb{L}_j$  and of the shifted Loewner matrices  $\mathbb{L}_{s_j}$  as in (57), this time for  $j \in \{1, \dots, D\}$ . Again,  $\mathbf{X}_j^{(\ell)}, \mathbf{Y}_j^{(\ell)} \in \mathbb{R}^{k_j \times r_j}$ ,  $\mathbf{S}_j^{(\ell)} \in \mathbb{R}^{r_j \times r_j}$ ,  $j = \{1, \dots, D\}$  and  $\ell = \{1, 2\}$ . Here, choose  $r_j$  as the numerical rank of the Loewner matrix  $\mathbb{L}_j$  (i.e. the largest neglected singular value corresponding to index  $r_j + 1$  is less than machine precision  $\epsilon$ ). The projected system matrices computed as

$$\hat{\mathbf{E}}_j = -(\mathbf{Y}_j^{(1)})^T \mathbb{L}_j \mathbf{X}_j^{(2)}, \quad \hat{\mathbf{A}}_j = -(\mathbf{Y}_j^{(1)})^T \mathbb{L}_{s_j} \mathbf{X}_j^{(2)}, \quad \hat{\mathbf{B}}_j = (\mathbf{Y}_j^{(1)})^T \mathbf{V}_j, \quad \hat{\mathbf{C}}_j = \mathbf{W}_j \mathbf{X}_j^{(2)}, \quad \text{for } j \in \{1, \dots, D\},$$

and the projected coupling matrices computed in the following way

$$\hat{\mathbf{K}}_{i,j} = (\mathbf{Y}_j^{(1)})^T \tilde{\boldsymbol{\Xi}}_{i,j} \mathbf{X}_i^{(2)}, \quad \forall i, j \in \{1, \dots, D\},$$

form a realization of a reduced order LSS denoted with  $\hat{\boldsymbol{\Sigma}}$  that approximately matches the data of the original LSS  $\boldsymbol{\Sigma}$ . Each reduced subsystem  $\hat{\boldsymbol{\Sigma}}_j$  has dimension  $r_j$ ,  $j \in \{1, \dots, D\}$ .

**Remark 5.4** If the truncated singular values are all 0 (the ones on the main diagonal of the matrices  $\tilde{\mathbf{S}}_j$ ), then the interpolation is exact.

## 6 Numerical experiments

In this section we illustrate the new method by means of three numerical examples. We use a certain generalization of the balanced truncation (BT) method for LSS (as presented in [28]) to compare the performance of our new introduced method. The main ingredient of the BT method is to compute the controllability and observability Gramians  $\mathbf{P}_i$  and  $\mathbf{Q}_i$  (where  $i \in \{1, 2, \dots, D\}$ ) as the solutions of the following Lyapunov equations:

$$\mathbf{A}_i \mathbf{P}_i \mathbf{E}_i^T + \mathbf{E}_i \mathbf{P}_i \mathbf{A}_i^T + \mathbf{B}_i \mathbf{B}_i^T = \mathbf{0}, \quad (72)$$

$$\mathbf{A}_i^T \mathbf{Q}_i \mathbf{E}_i + \mathbf{E}_i^T \mathbf{P}_i \mathbf{A}_i + \mathbf{C}_i^T \mathbf{C}_i = \mathbf{0}. \quad (73)$$

### 6.1 Balanced Truncation

In [28] it was shown that, if certain conditions are satisfied, the technique of simultaneous balanced truncation can be applied to switched linear systems. In some special cases, the existence of a global transformation matrix  $\mathbf{V}_{bal}$  is guaranteed, provided that (*Corollary IV.3* in [28]):

1. The matrices  $\mathcal{P}_i \mathbf{Q}_i$  and  $\mathcal{P}_j \mathbf{Q}_j$  commute for all  $i, j \in \{1, 2, \dots, D\}$ ;
2. The conditions  $\mathcal{P}_i \mathbf{Q}_j = \mathcal{P}_j \mathbf{Q}_i$  are satisfied for all  $i, j \in \{1, 2, \dots, D\}$ .

Hence, it follows that:

$$\mathbf{V}_{bal} \mathcal{P}_i \mathbf{V}_{bal}^T = \mathbf{V}_{bal}^{-T} \mathbf{Q}_i \mathbf{V}_{bal}^{-1} = \mathbf{U}_i, \quad (74)$$

where  $\mathbf{U}_i$  are diagonal matrices. Although conceptually attractive as a MOR method, in general the conditions are rather restrictive in practice. This motivates the search for a more general MOR approach for the case where simultaneous balancing cannot be achieved.

The problem of finding a balancing transformation for a single linear system can be formulated as finding a nonsingular matrix such that the following cost function is minimized (see [1]):

$$f(\mathbf{V}) = \text{trace}[\mathbf{V} \mathcal{P} \mathbf{V}^T + \mathbf{V}^{-T} \mathbf{Q} \mathbf{V}^{-1}]. \quad (75)$$

For the class of LSS with distinct operational modes, we hence have to minimize not one but a number of  $D$  cost functions:

$$f_i(\mathbf{V}) = \text{trace}[\mathbf{V} \mathcal{P}_i \mathbf{V}^T + \mathbf{V}^{-T} \mathbf{Q}_i \mathbf{V}^{-1}], \quad i \in \{1, 2, \dots, D\}. \quad (76)$$

If the conditions of Corollary IV.3 from [28] hold, simultaneous balancing is possible, and there exists a transformation  $\mathbf{V}$  which simultaneously minimizes  $f_i$  for all  $i = 1, 2, \dots, D$ . Instead of

having  $D$  functions as in (76), one can introduce a single overall cost function (i.e the average of the cost functions of the individual modes). Define the function  $f_{av}$  as in [28]:

$$f_{av}(\mathbf{V}) = \frac{1}{D} \sum_{i=1}^D \text{trace}[\mathbf{V}\mathcal{P}_i\mathbf{V}^T + \mathbf{V}^{-T}\mathcal{Q}_i\mathbf{V}^{-1}] = \text{trace}[\mathbf{V}\mathcal{P}_{av}\mathbf{V}^T + \mathbf{V}^{-T}\mathcal{Q}_{av}\mathbf{V}^{-1}], \quad (77)$$

where

$$\mathcal{P}_{av} = \frac{1}{D} \sum_{i=1}^D \mathcal{P}_i, \quad \mathcal{Q}_{av} = \frac{1}{D} \sum_{i=1}^D \mathcal{Q}_i. \quad (78)$$

In the case of LSS, the BT method computes a basis where the sum of the eigenvalues of  $\mathcal{P}_i$  and  $\mathcal{Q}_i$  over all modes is minimal. Hence, minimizing the proposed overall cost function provides a natural extension of classical BT to the case of LSS.

It follows that the transformation  $\tilde{\mathbf{V}}$  that minimizes the cost function in (77) is precisely the one which balances the pair  $(\mathcal{P}_{av}, \mathcal{Q}_{av})$  of average Gramians.

By applying  $\tilde{\mathbf{V}}$  to the individual modes and truncating, a reduced order model is obtained. After applying the transformation  $\tilde{\mathbf{V}}$ , the new state space representations of the individual modes need not be balanced. Nevertheless, as stated in [28], it is expected to be relatively close to being balanced. Moreover, a downside of this method is given by the fact that it does not allow different state-space dimensionality for different modes.

## 6.2 First example

As first example we consider the simple model of an evaporator vessel from [29]. There is a constant inflow of liquid into a tank as well as an outflow that depends on the pressure in the tank and the Bernoulli resistance  $R_b$ . To keep the level of fluid in the evaporator vessel at or below a pre-specified maximum, an overflow mechanism is activated when the level of fluid  $L$  in the evaporator exceeds the threshold value  $L_{th}$ . This causes a flow through a narrow pipe with resistance  $R_p$  and inertia  $I$  that builds up flow momentum  $p$ . The system is modeled in two distinct operation modes: mode 1, where there is no overflow (the fluid level is below the overflow level), and mode 2, where the overflow mechanism is active. The ordinary differential equations describing the system in the two operation modes are given by

$$\begin{aligned} \begin{bmatrix} I & 0 \\ 0 & C \end{bmatrix} \begin{bmatrix} \dot{p} \\ \dot{L} \end{bmatrix} &= \begin{bmatrix} -R_p & 0 \\ 0 & -1/R_b \end{bmatrix} \begin{bmatrix} p \\ L \end{bmatrix} + \begin{bmatrix} 0 \\ f_{in} \end{bmatrix} \quad (\text{mode 1}), \\ \begin{bmatrix} I & 0 \\ 0 & C \end{bmatrix} \begin{bmatrix} \dot{p} \\ \dot{L} \end{bmatrix} &= \begin{bmatrix} -R_p & 1 \\ -1 & -1/R_b \end{bmatrix} \begin{bmatrix} p \\ L \end{bmatrix} + \begin{bmatrix} 0 \\ f_{in} \end{bmatrix} \quad (\text{mode 2}). \end{aligned}$$

Additionally, note that the observed output  $y$  is chosen to be the average of the two system variables  $p$  and  $L$ , i.e.  $y = (p + L)/2$  (for both modes). Assuming the system is initially in mode 1, the inflow then causes the tank to start filling, which causes an outflow through resistance  $R_b$ . In this mode the outflow through the narrow pipe is zero. If  $L$  exceeds the level  $L_{th}$ , a switch from mode 1 to mode 2 occurs at the point in time when  $L = L_{th}$ .

In the following, use the parameters  $R_b = 1, R_p = 0.5, I = 1, C = 1, f_{in} = 1, L_{th} = 0.08$  and compute the following system matrices:

$$\underline{\text{Mode 1}} : \mathbf{E}_1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \mathbf{A}_1 = \begin{bmatrix} -\frac{1}{2} & 0 \\ 0 & -1 \end{bmatrix}, \quad \mathbf{B}_1 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad \mathbf{C}_1 = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \end{bmatrix},$$

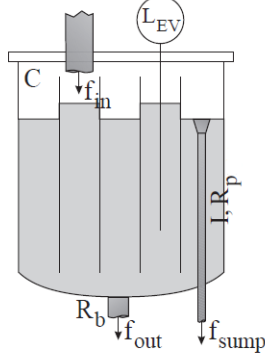


Figure 1: Schematic of the evaporator vessel [29].

$$\text{Mode 2 : } \mathbf{E}_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \mathbf{A}_2 = \begin{bmatrix} -\frac{1}{2} & 1 \\ -1 & -1 \end{bmatrix}, \quad \mathbf{B}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad \mathbf{C}_2 = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \end{bmatrix}.$$

The coupling matrices are chosen to be identity matrices, i.e.  $\mathbf{K}_1 = \mathbf{K}_2 = \mathbf{I}_2$ . Next, consider the following tuples of left and right interpolation points corresponding to each mode, as

$$\left\{ \begin{array}{l} \lambda_1 = \{(-1.5), (-2, 1)\}, \\ \mu_1 = \{(2), (0, 0.5)\}. \end{array} \right\}, \quad \left\{ \begin{array}{l} \lambda_2 = \{(1), (1.5, -1.5)\}, \\ \mu_2 = \{(0), (2, -0.5)\}. \end{array} \right\}.$$

Hence, following the procedure described in Section 4, we recover the following system matrices:

$$\text{Mode 1 : } \hat{\mathbf{E}}_1 = \begin{bmatrix} -\frac{1}{3} & -\frac{23}{240} \\ -\frac{2}{3} & -\frac{1}{8} \end{bmatrix}, \quad \hat{\mathbf{A}}_1 = \begin{bmatrix} \frac{1}{3} & \frac{19}{240} \\ \frac{2}{3} & \frac{1}{8} \end{bmatrix}, \quad \hat{\mathbf{B}}_1 = \begin{bmatrix} \frac{1}{6} \\ \frac{1}{3} \end{bmatrix}, \quad \hat{\mathbf{C}}_1 = \begin{bmatrix} -1 & -\frac{13}{48} \end{bmatrix},$$

$$\text{Mode 2 : } \hat{\mathbf{E}}_2 = \begin{bmatrix} \frac{3}{16} & -\frac{1}{3} \\ \frac{7}{120} & -\frac{1}{9} \end{bmatrix}, \quad \hat{\mathbf{A}}_2 = \begin{bmatrix} -\frac{5}{16} & \frac{1}{2} \\ -\frac{17}{120} & \frac{7}{30} \end{bmatrix}, \quad \hat{\mathbf{B}}_2 = \begin{bmatrix} \frac{1}{2} \\ \frac{1}{5} \end{bmatrix}, \quad \hat{\mathbf{C}}_2 = \begin{bmatrix} \frac{5}{16} & -\frac{1}{2} \end{bmatrix}.$$

Note that the recovered realization is equivalent to the original one (no reduction has been enforced since the task was to recover the initial system only). The coupling matrices are also computed:

$$\hat{\mathbf{K}}_1 = \begin{bmatrix} -1 & -\frac{3}{16} \\ -\frac{2}{5} & -\frac{23}{360} \end{bmatrix}, \quad \hat{\mathbf{K}}_2 = \begin{bmatrix} \frac{9}{80} & -\frac{8}{45} \\ \frac{1}{8} & -\frac{2}{9} \end{bmatrix}.$$

### 6.3 Second example

For the next experiment, consider the CD player system from the SLICOT benchmark examples for MOR (see [14]). This linear system of order 120 has two inputs and two outputs. We consider that, at any given instance of time, only one input and one output are active (the others are not functional due to mechanical failure). More precisely, consider mode  $j$  to be activated whenever the  $j^{\text{th}}$  input and the  $j^{\text{th}}$  output are simultaneously failing (where  $j \in \{1, 2\}$ ).

In this way, we construct an LSS with two operational modes (hence  $D = 2$ ). Both subsystems are stable SISO linear systems of order 120. This initial linear switched system  $\Sigma$  will be reduced by means of the Loewner framework to obtain  $\hat{\Sigma}_L$  and by means of the balanced truncation method proposed in [28] to obtain  $\hat{\Sigma}_B$ . Denote with  $\hat{\Sigma}_{L_j}$  and  $\hat{\Sigma}_{B_j}$  the  $j^{\text{th}}$  linear subsystem corresponding to  $\hat{\Sigma}_L$ , and respectively, to  $\hat{\Sigma}_B$ .

The frequency response of each original subsystem is depicted in Fig. 2. Note that, an amplitude scaling of the original system was performed, in order to enforce a similar gain for each of the two subsystems.

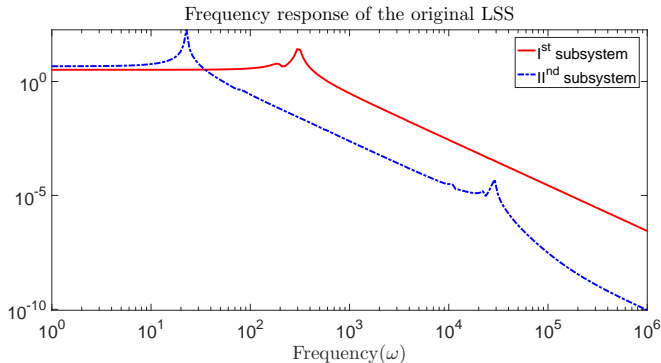
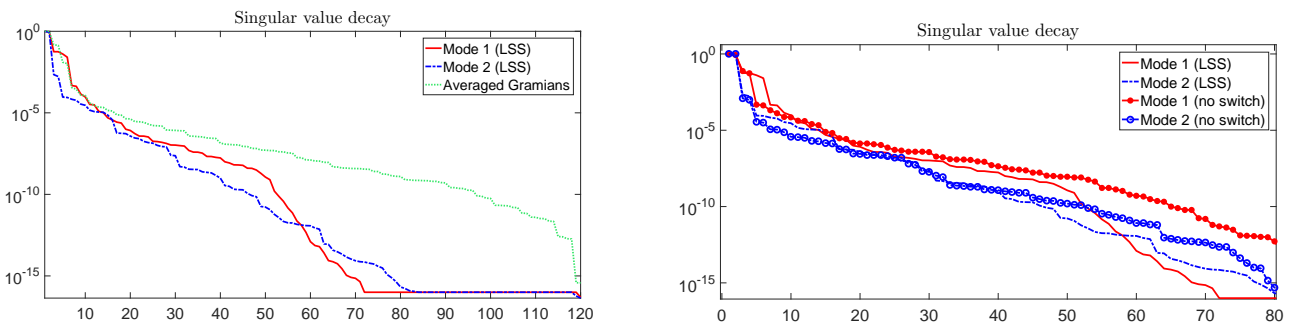


Figure 2: Frequency response of the original subsystems.

For the Loewner method, we choose 120 logarithmically distributed interpolation points in the interval  $[10^1, 10^5]j$ . In Fig. 3 (a), we depict the singular value decay of the appended Loewner matrices, i.e.  $[\mathbb{L}_j \ \mathbb{L}_{sj}]$ ,  $j \in \{1, 2\}$ . These matrices are used in Procedure 1, i.e. in (57). Note that the Loewner matrices  $\mathbb{L}_j$  and the shifted Loewner matrices  $\mathbb{L}_{sj}$  are defined as in Section 4.2. Additionally, Fig. 3 (a) depicts the decay of the Hankel singular values of the averaged Gramians corresponding to  $\Sigma$ , as defined in (78). We observe that the 80<sup>th</sup> value attains machine precision ( $\epsilon \approx 10^{-16}$ ) in the case of the Loewner matrices, while in the case of the balancing procedure, the same truncation order provides a  $10^{-9}$  decay. The decay presented in Fig. 3 is a good indicator for choosing the desired truncation order.



(a) Appended Loewner matrices + averaged Gramians.

(b) Appended Loewner matrices.

Figure 3: Decay of the singular values for different matrices.

For both reduced order systems,  $\hat{\Sigma}_L$  and  $\hat{\Sigma}_B$ , we decide to truncate at order  $k_1 = k_2 = 27$  for the two subsystems. This choice was made so that the neglected singular values corresponding to all three curves in Fig. 3 (a) are less than the chosen tolerance value, i.e.  $\epsilon_1 = 10^{-6}$ . More specifically, this corresponds to the following values for the 27<sup>th</sup> singular value:  $\sigma_{27}^{(1)} = 1.5498 \cdot 10^{-7}$  for the first Loewner subsystem and  $\sigma_{27}^{(2)} = 8.5741 \cdot 10^{-8}$  for the second Loewner subsystem. The last kept singular value corresponding to the averaged balanced model is higher, i.e.  $\sigma_{27}^{(3)} = 1.5516 \cdot 10^{-6}$ . Hence, we would like to emphasize that the singular value decay is faster for Loewner compared to BT and thus certain errors can be achieved with lower order models.

Additionally, in Fig. 3 (b) we present the singular value decay of appended Loewner matrices  $[\mathbb{L}_j \ \mathbb{L}_{sj}]$ ,  $j \in \{1, 2\}$ , defined as in Section 4.2 (the LSS case for which samples of higher order transfer functions are used in the process). The first and second curves correspond to these quantities. In the same figure, we present the singular value decay of appended Loewner matrices defined as in Section 3.2 (the classical linear case with no switching and in which only samples of

first order linear transfer functions are used in the process). The third and fourth curves correspond to these quantities. Note that, the singular value decay of the latter matrices is slightly slower than that of the first matrices.

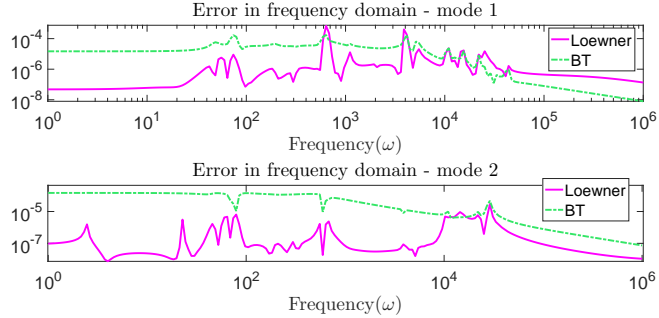


Figure 4: Frequency domain approximation error.

We assess the performance of the two MOR methods mentioned above in the following ways: depicting the frequency domain simulation error, computing the  $\mathcal{H}_2$  and  $\mathcal{H}_\infty$  norms of the error subsystems and by depicting the time domain simulation error.

First, start by comparing the quality of approximation of the frequency response. In Fig. 4 the frequency domain error is depicted for both MOR methods (Loewner and BT).

Observe that the error curve corresponding to the Loewner method is lower than that of the BT method for most of the frequency points considered in this experiment. This behavior is especially noticeable in the low frequency range. If one compares the maximum of the error, i.e. the  $\mathcal{H}_\infty$  norm, then one can notice that for mode 1 the Loewner method produces a higher peak error, while for mode 2, the Loewner method produces a lower peak error.

Next, we explicitly compute the relative approximation errors for each mode individually, for both reduction methods, with respect to the  $\mathcal{H}_2$  norm. More specifically,  $\|\hat{\Sigma}_{L_j} - \Sigma_j\|_{\mathcal{H}_2} / \|\Sigma_j\|_{\mathcal{H}_2}$  for the Loewner method, and  $\|\Sigma_{B_j} - \Sigma_j\|_{\mathcal{H}_2} / \|\Sigma_j\|_{\mathcal{H}_2}$  where  $j \in \{1, 2\}$  for the BT method. Additionally, we compute the  $\mathcal{H}_\infty$  norm relative errors, i.e.  $\|\hat{\Sigma}_{L_j} - \Sigma_j\|_{\mathcal{H}_\infty} / \|\Sigma_j\|_{\mathcal{H}_\infty}$  for the Loewner method, and  $\|\Sigma_{B_j} - \Sigma_j\|_{\mathcal{H}_\infty} / \|\Sigma_j\|_{\mathcal{H}_\infty}$  where  $j \in \{1, 2\}$  for the BT method. The results are presented in Tab. 1. Note that, for mode 1, the balanced truncation method produces slightly lower errors, while for mode 2, the errors corresponding to the Loewner method are lower.

$\mathcal{H}_2$	Loewner	Bal Trunc	$\mathcal{H}_\infty$	Loewner	Bal Trunc
Mode 1	$3.9210 \cdot 10^{-5}$	$2.0303 \cdot 10^{-5}$	Mode 1	$3.3888 \cdot 10^{-5}$	$5.5266 \cdot 10^{-6}$
Mode 2	$9.9040 \cdot 10^{-6}$	$2.2922 \cdot 10^{-5}$	Mode 2	$2.6036 \cdot 10^{-7}$	$6.3798 \cdot 10^{-7}$

Table 1: Relative approximation error for the two modes in the  $\mathcal{H}_2$  and  $\mathcal{H}_\infty$  norms

Finally, we compare the time domain response of the original linear switched system against the ones corresponding to the two reduced models. We first use a simple sinusoidal signal, i.e.  $\mathbf{u}(t) = \cos(t)/10$  as the control input. The piecewise-constant signal in the upper part of Fig. 5 (a) represents the switching signal. The switching times are chosen within the interval  $[0, 10]$  seconds so that fast switching is also exhibited. More precisely, note that in the range  $[5.5, 7]$ s, this signal switches between modes 1 and 2 much more frequently than in the rest of the time axis.

As it can be seen in the lower part of Fig. 5 (a), the output of the LSS is well approximated for both MOR methods (all three curves are indistinguishable from one another). The blue circles

located on the observed output curve in the lower part of Fig. 5 (a) are used to mark the switching times.

Finally, by inspecting the time domain error between the original response and the responses coming from the two reduced models (depicted in Fig. 5 (b)), we notice that the error curve corresponding to the Loewner method is two orders of magnitude below the error curve corresponding to the BT method for most of the points on the time axis.

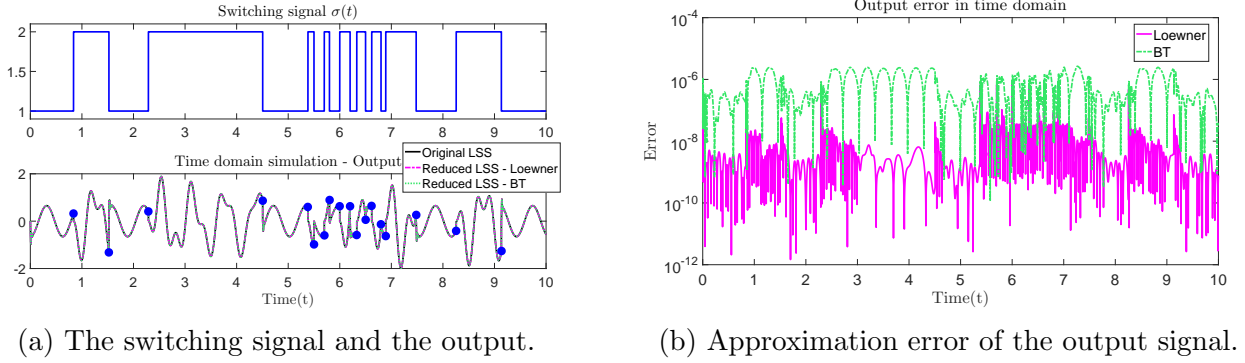


Figure 5: Time domain simulation - first choice of input.

Additionally, we repeat the time domain simulation experiment depicted in Fig. 5, by considering another control input signal. Let  $\mathbf{u}(t) = (\sum_{k=1}^{10} \sin(\omega_k t))/10$  be a richer frequency spectrum signal, where the frequency points  $\omega_k$  are logarithmically spaced in the interval  $[1, 50]$ . We also choose another switching signal  $\sigma(t)$ , with random switching times (and with no particular imposed conditions to the frequency of switching). In Fig. 6 we depict the switching signal, the observed output and the approximation error. Similar conclusions to the ones mentioned above for the results in Fig. 5 can be drawn by inspecting Fig. 6.

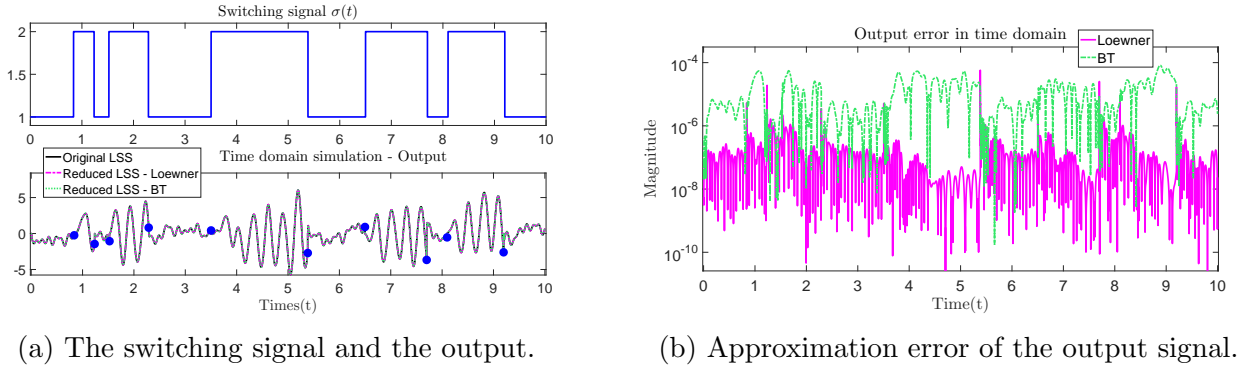


Figure 6: Time domain simulation - second choice of input.

## 6.4 Third example

For the last experiment, consider a large scale LSS constructed as in [22] from the original machine stand example given in [16]. In this example, the system variability is induced by a moving tool slide on the guide rails of the stand (see Fig. 7). The aim is to determine the thermally driven displacement of the machine stand structure. Following the model setting in [16], consider the heat



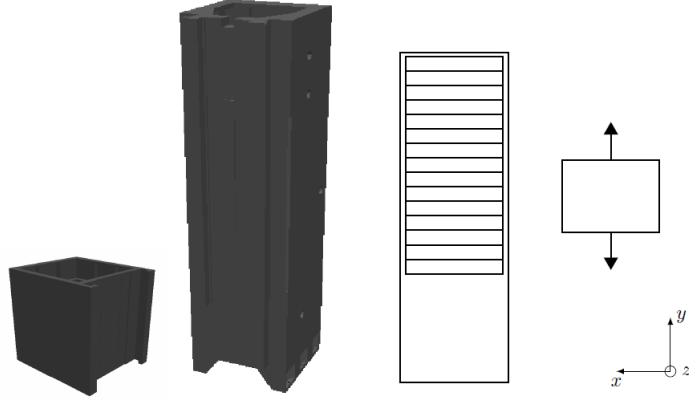


Figure 7: Schematic of the tool slide on the guide rails of the stand [22].

equation with Robin boundary conditions. Using a finite element (FE) discretization and denoting the external influences as the system input  $z$ , we obtain the dynamical heat model

$$\mathbf{E}_{th}\dot{\mathbf{x}}(t) = \mathbf{A}_{th}(t)\mathbf{x}(t) + \mathbf{B}_{th}(t)\mathbf{z}(t), \quad (79)$$

describing the deformation independent evolution of the temperature field  $\mathbf{x}$  with the system matrices  $\mathbf{E}_{th}$ ,  $\mathbf{A}_{th}(t)$  and  $\mathbf{B}_{th}(t)$ . The variability of the model is described by time dependent matrices  $\mathbf{A}_{th}(t)$  and  $\mathbf{B}_{th}(t)$ . This leads directly to the linear time varying system described by (79). Since model reduction for LTV systems is a highly storage consuming procedure, the authors in [22] exploit properties of the spatially semi-discretized model to set up an LSS consisting of LTI subsystems only. As described in [16], the guide rails of the machine stand are modeled as 15 equally distributed horizontal segments (see Fig. 7). Any of these segments is said to be completely covered by the tool slide if its midpoint lies within the height of the slide. On the other hand, each segment whose midpoint is not covered is treated as not in contact and therefore the slide always covers exactly 5 segments at each time. This in fact allows the stand to reach 11 distinct, discrete positions given by the model restrictions. In this way, one can define the subsystems of the LSS as follows:

$$\Sigma_\ell : \begin{cases} \mathbf{E}_{th}\dot{\mathbf{x}} = \mathbf{A}_{th}^\ell\mathbf{x} + \mathbf{B}_{th}^\ell\mathbf{z}^\ell, \\ \mathbf{y} = \mathbf{C}\mathbf{x}, \end{cases} \quad (80)$$

where  $\ell \in \{1, \dots, 11\}$ . Note that the change of the input operator  $\mathbf{B}_{th}(t)$  is hidden in the input itself, since it is sufficient to activate the correct boundary parts by choosing the corresponding columns in  $\mathbf{B}_{th}$  via the input  $\mathbf{z}^\ell$ . Therefore, the input operator  $\mathbf{B}_{th}(t) := \mathbf{B}_{th}$  becomes constant and the input variability is represented by the input  $z^\ell$

$$\mathbf{z}_i^\ell := \begin{cases} \mathbf{z}_i, & \text{segment } i \text{ is in contact,} \\ 0, & \text{otherwise,} \end{cases}, \quad \text{for } i \in \{1, \dots, 15\}. \quad (81)$$

Here,  $\mathbf{z}_i \in \mathbb{R}$  is the thermal input as described in [22]. The only varying part influencing the model reduction process left in the dynamical system is the system matrix  $\mathbf{A}_{th}(t) := \mathbf{A}_{th}^\ell$ .

Since the application of this example is to study the thermally driven deformation at particular points, the output equation  $\mathbf{y} = \mathbf{C}\mathbf{z}$  is used to explicitly select these points (such as the ones located around the tool center point or around the connections to neighboring assembly groups). The rows of the matrix  $\mathbf{C}$  are unit vectors with only 0 and 1 values and hence, the output  $\mathbf{y}$  contains selected entries of the internal variable  $\mathbf{x}$ .

After the finite element discretization was performed, obtain an LSS with 11 active modes (denote it with  $\Sigma$ ). Each subsystem has dimension  $n = 16626$ . The  $\mathbf{E}$  and  $\mathbf{C}$  matrices are the same for all modes of the LSS. The  $\mathbf{B}$  matrices have 6 columns (corresponding to different inputs) and the  $\mathbf{C}$  matrix has 9 rows (corresponding to different outputs). All the aforementioned matrices are saved in sparse format.

The proposed extension of the Loewner method to LSSs (described in Section 4 and 5) can be generalized in a straightforward manner to the MIMO case similarly to the linear case (see [6]), by introducing left and right tangential direction vectors.

In the following experiments, we take into consideration three active modes (the first, the third and the fifth). This corresponds to the particular case of  $D = 3$  (treated in Section 5). Furthermore, consider only the pairs of the first and third inputs as well as the first and third outputs to be activated at any time, for each of the three modes. Hence, the measurements used in the Loewner framework are  $2 \times 2$  matrices.

We analyze a simplified model composed of three modes and certain input/output pairs in order to ensure that the numerical results can be depicted in a clear and distinguishable manner (without overcrowding the figures).

All subsystems are stable linear systems of order 16626 in sparse format. The large scale LSS  $\Sigma$  will be again reduced, as in the second example, by means of the Loewner method and of the balanced truncation method proposed in [28].

We perform a time domain simulation to investigate the approximation quality of the observed output, which in this case has two components. More exactly, the choice of outputs is as follows: the first output represents the 9163<sup>th</sup> entry of the deformation vector  $\mathbf{x}$ , while the second output is the 9814<sup>th</sup> entry of  $\mathbf{x}$ .

The singular values of the frequency response of each original subsystem are depicted in Fig. 8. More exactly, for all three modes and for each frequency point  $j\omega$ , compute the two singular values corresponding to  $\mathbf{H}_k(j\omega) \in \mathbb{C}^{2 \times 2}$ .

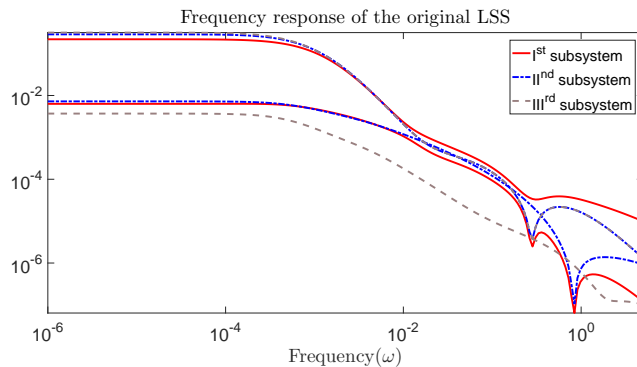


Figure 8: Frequency response of the original subsystems.

For the Loewner method, we choose 200 logarithmically spaced interpolation points in the interval  $[10^{-6}, 5 \cdot 10^0]j$ . The decay of the singular values of the appended Loewner matrices  $[\mathbb{L}_j \ \mathbb{L}_{s_j}]$  corresponding to mode  $j$ , for  $j \in \{1, 2, 3\}$ , is presented in Fig. 9. We notice that all three curves are close to each other and already the 100<sup>th</sup> singular values attain machine precision. Additionally, in the same figure one can observe the decay of the approximate averaged Hankel singular values corresponding to  $\Sigma$ . Of course, one can not compute exact Gramians for the large order system, but only approximate low rank factors by using, for instance, such software tools as in [36].

For the Loewner reduced order LSS (i.e  $\Sigma_1$ ), we decide to truncate at order  $k = 66$  for all three subsystems. This corresponds to eliminating the singular values that are smaller than  $10^{-13}$

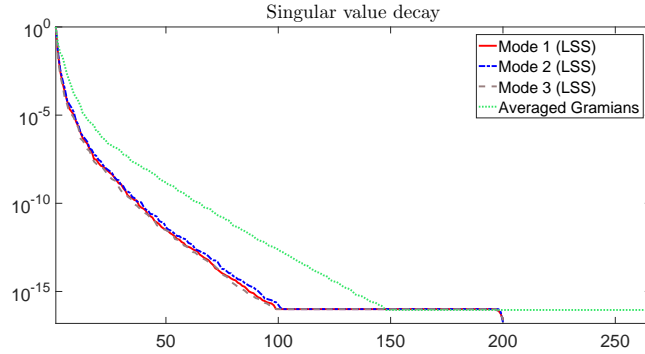


Figure 9: Decay of the singular values of the different matrices.

(for each of the three appended Loewner matrices). The same truncation order is chosen for the reduced order model computed via BT for which the last kept singular value (before the balancing truncation procedure is applied) is  $\sigma_{66}^{\text{avg}} = 8.0693 \cdot 10^{-11}$ .

In the upper part of figure Fig. 10, the control input signals  $u_1$  and  $u_2$  are depicted.

$$u_1(t) = \frac{1}{2} \sin(t/20)e^{-t/500} + \frac{1}{20} e^{-t/500}, \quad u_2(t) = \frac{1}{10}. \quad (82)$$

As it can be observed in the x axis of this figure, the running time of the performed experiment is 1 hour (the control is active form time  $t_s = 0s$  to  $t_e = 3600s$ ). In the lower part of Fig. 10, the switching signal  $\sigma : \mathbb{R} \rightarrow \{1, 2, 3\}$  is presented. Note that the time axis is restricted from 380 to 440 seconds. The switching signal follows a simple periodical rule, by repeating the sequence of modes (1, 2, 3, 2, 1).

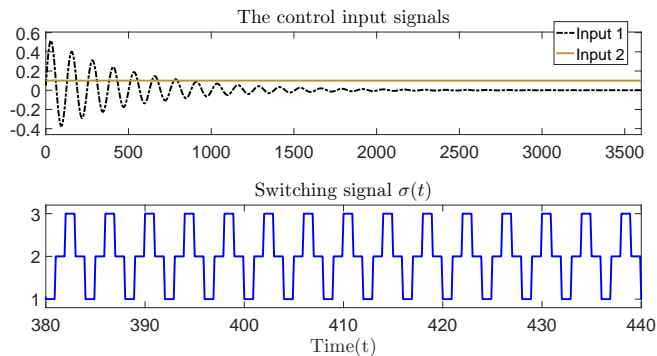


Figure 10: The control and switched input signals.

Next, compare the time domain response of the original LSS against the ones corresponding to the two reduced models. Notice that the two outputs of the LSS are well approximated for both MOR methods, as it can be seen in Fig. 11 (the upper part depicts the first observed output, while in the lower part, the second output is shown). As for Fig. 5 in Section 6.3, the blue circles included on the output curves are used to mark the exact time instances when switching occurs. Again, the time axis is restricted from 380 to 440 seconds.

Finally, we inspect the time domain error between the original response and the responses of the two reduced models (depicted in Fig. 12). Note that the time axis is again restricted to one minute, in between  $[380, 440]$  seconds. We observe that, the error curve corresponding to our proposed method is always below the error curve corresponding to the BT method, in the considered time range.

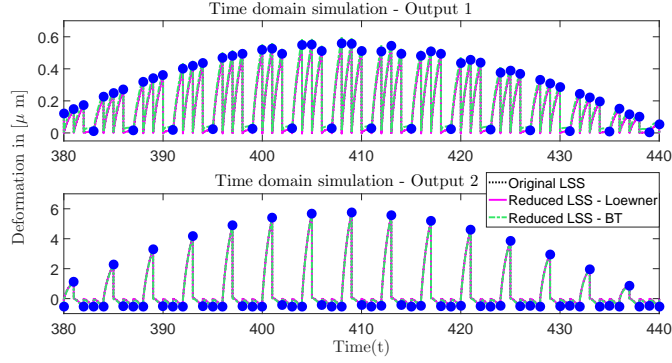


Figure 11: Time domain simulation.

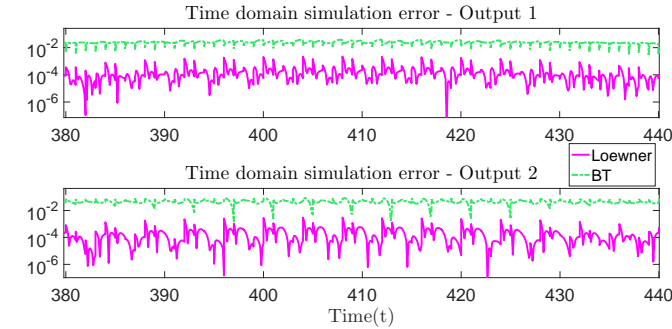


Figure 12: Time domain approximation error.

## 7 Summary and conclusions

In this paper we address the problem of model reduction of linear switched systems from data consisting of values of high order transfer functions. More specifically, we have extended the Loewner framework to the reduction of LSS. The Loewner approach to model reduction, first developed for linear time-invariant systems (see [6] for a survey), was later extended to linear parametrized systems [5, 21] and to bilinear systems [4].

The underlying philosophy of the Loewner framework is to collect data and then extract the desired information. For the case of linear switched systems, the data must be computed a priori, rather than measured (as for linear systems with no switching where one can use Vector Network Analyzers for instance). Having the required data, the next step would be to arrange it into matrix format. We have shown that the Loewner matrices (which represent the recovered  $\mathbf{E}$  and  $\mathbf{A}$  matrices of the underlying LSS) can be automatically calculated as solutions of Sylvester equations. In the proposed framework, the transition/coupling matrices can be recovered from the given computed data as well. Since these matrices need not be square, they allow having different dimensions of the reduced state space in different modes.

Three numerical examples demonstrate the effectiveness of the proposed approach. The quality of approximation for the reduced models was determined by performing both frequency and time domain tests. We have chosen a generalization of the classical balanced truncation method to LSS for comparison purposes. As opposed to most of the balancing methods we encountered in the literature ([13], [10], [37] and [34]), the method we choose (i.e [28]) does not require solving systems of LMI (linear matrix inequalities) which might be difficult for very large systems such as the one in Section 6.4. The results of the new proposed method turned out to be overall better than the ones obtained when using the BT method. More precisely, we would like to emphasize that, the

Loewner method seems to be able to achieve similar approximation accuracy to the BT method, but with lower dimension of the ROMs constructed. This observation is confirmed by the faster decay of the singular values computed for our method as compared to the decay of the (averaged) Hankel singular values corresponding to the BT method (presented in Fig. 3 and Fig. 9).

In Section 6.3, we also investigated switching signals  $\sigma(t)$  that exhibit a fast switching behavior (hence smaller dwell times) in a particular time interval (Fig. 5 (b)). We conclude that there are no reasons to suspect that our method will fail for such cases.

In general, stability preservation of the reduced order model is still an open issue for moment matching MOR methods, even for linear systems. In fact, even for linear systems, the Loewner framework does not guarantee preservation of stability. Since the model reduction method presented in this paper is a direct extension of the Loewner framework for linear systems, in general, it will not preserve stability. However, although it is likely that under suitable assumptions the proposed method could be modified to preserve stability, this remains a topic of future research.

The Loewner framework is conceptually a data driven MOR method that builds reduced order models that interpolate the frequency response of the large scale original system. In principle, this method requires only data and not solving any type of equations. In the linear case, the data can be measured using VNAs (vector network analyzers). In the nonlinear case (as for bilinear, quadratic-bilinear or linear switched), the data is obtained by direct numerical simulation (DNS). In the LSS case, no equation needs to be solved (the data is gathered and put together in the fashion described in the paper). The computational complexity of our proposed method is related to the DNS process. Hence, one needs to compute samples of the generalized transfer functions (in the case for which these values are not provided via real time measurements). This can be performed in a fast way by avoiding the explicit inversion of the matrices  $\omega_j \mathbf{E}_k - \mathbf{A}_k$  and using, for example, Gaussian elimination instead.

## Acknowledgements

The first and third authors would like to thank Mert Bastug and Rafael Wisniewski for suggesting the problem of model reduction of LSS and for useful discussions.

Also, the authors want to thank Norman Lang for providing the LSS large scale system from Section 6.3 (used in [22]).

Last but not least, we wish to thank Jens Saak for helpful discussions and also for recommending the M-M.E.S.S. toolbox (see [36]) for efficiently computing solutions of sparse large scale algebraic equations (which was very helpful for implementing the BT method for the 16626<sup>th</sup> order system).

## References

- [1] A. C. ANTOULAS, Approximation of Large-Scale Dynamical Systems, Series: Advances in Design and Control, <https://doi.org/10.1137/1.9780898718713>, SIAM, Philadelphia, 2005.
- [2] ———, The Loewner framework and transfer functions of singular/rectangular systems, Applied Mathematics Letters, 54 (2016), pp. 36–47.
- [3] A. C. ANTOULAS, C. BEATTIE, AND S. GUGERCIN, Interpolatory Model Reduction of Large-Scale Dynamical Systems, Efficient Modeling and Control of Large-Scale Systems, Springer US, Boston, MA, 2010, pp. 3–58.
- [4] A. C. ANTOULAS, I. V. GOSEA, AND A. C. IONITA, Model reduction of bilinear systems in the Loewner framework, SIAM Journal on Scientific Computing, 38 (2016), pp. B889–B916.
- [5] A. C. ANTOULAS, A. C. IONITA, AND S. LEFTERIU, On two-variable rational interpolation, Linear Algebra and Its Applications, 436 (2012), pp. 2889–2915.

- [6] A. C. ANTOULAS, S. LEFTERIU, AND A. C. IONITA, A tutorial introduction to the Loewner Framework for Model Reduction, Model Reduction and Approximation for Complex Systems, Edited by P. Benner, A. Cohen, M. Ohlberger, and K. Willcox, Series: Computational Science & Engineering, <https://doi.org/10.1137/1.9781611974829>, SIAM, 2017, pp. 335–376.
- [7] M. BASTUG, Model Reduction of Linear Switched Systems and LPV State-Space Models, PhD thesis, Department of Electronic Systems, Automation and Control, Aalborg University, December, 2015.
- [8] M. BASTUG, M. PETRECZKY, R. WISNIEWSKI, AND J. LETH, Model reduction by nice selections for linear switched systems, IEEE Transactions on Automatic Control, 61 (2016), pp. 3422–3437.
- [9] M. BASTUG, M. PETRECZKY, R. WISNIEWSKI, AND J. LETH, Reachability and observability reduction for linear switched systems with constrained switching, Automatica, 74 (2016), pp. 162–170.
- [10] A. BIROUCHE, J. GUILLET, B. MOURILLON, AND M. BASSET, Gramian based approach to model order-reduction for discrete-time switched linear systems, in Proceedings of the 18th Mediterranean Conference on Control and Automation, 2010, pp. 1224–1229.
- [11] A. BIROUCHE, B. MOURILLON, AND M. BASSET, Model reduction for discrete-time switched linear time-delay systems via the  $H_\infty$  stability, Control and Intelligent Systems, 39 (2011), pp. 1–9.
- [12] ———, Model order-reduction for discrete-time switched linear systems, Int. J. Systems Science, 43 (2012), pp. 1753–1763.
- [13] Y. CHAHLAOUI, Model reduction of hybrid switched systems, in Proceeding of the 4th Conference on Trends in Applied Mathematics in Tunisia, Algeria and Morocco, 2009.
- [14] Y. CHAHLAOUI AND P. V. DOOREN, A collection of benchmark examples for model reduction of linear time invariant dynamical systems. <http://slicot.org/20-site/126-benchmark-examples-for-model-reduction>, February 2002.
- [15] J. DAAFOUZ, P. RIEDINGER, AND C. IUNG, Stability analysis and control synthesis for switched systems: a switched Lyapunov function approach, IEEE Transactions on Automatic Control, 47 (2002), pp. 1883–1887.
- [16] A. GALANT, K. GROSSMANN, AND A. MÜHL, Model order reduction (MOR) for thermo-elastic models of frame structural components on machine tools, in ANSYS Conference and 29th CADFEM Users Meeting, Stuttgart, Germany, October 19-21, 2011.
- [17] K. GALLIVAN, A. VANDERDORPE, AND P. VAN DOOREN, Model reduction of MIMO systems via tangential interpolation, SIAM Journal on Matrix Analysis and Applications, 26 (2004), pp. 328–349.
- [18] H. GAO, J. LAM, AND C. WANG, Model simplification for switched hybrid systems, Systems and Control Letters, 55 (2006), pp. 1015–1021.
- [19] R. GOEBEL, R. G. SANFELICE, AND A. R. TEEL, Hybrid Dynamical Systems: Modeling, Stability, and Robustness, Princeton University Press, 2012.
- [20] P. HAMANN AND V. MEHRMANN, Numerical solution of hybrid systems of differential-algebraic equations, Computer Methods in Applied Mechanics and Engineering, 197 (2008), pp. 693–705.
- [21] A. C. IONITA AND A. C. ANTOULAS, Data-driven parametrized model reduction in the Loewner framework, SIAM Journal on Scientific Computing, 36 (2014), pp. A984–A1007.
- [22] N. LANG, J. SAAK, AND P. BENNER, Model order reduction for systems with moving loads, Automatisierungstechnik, 62 (2014), pp. 512–522.
- [23] S. LEFTERIU AND A. C. ANTOULAS, A new approach to modeling multiport systems from frequency-domain data, IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems, 29 (2010), pp. 14–27.
- [24] ———, Topics in model order reduction with applications to circuit simulation, Model Reduction for Circuit Simulation, Springer, 2011, ch. 74, pp. 81–104.
- [25] D. LIBERZON, Switching in Systems and Control, Birkhäuser, 2008.
- [26] A. J. MAYO AND A. C. ANTOULAS, A framework for the generalized realization problem, Linear Algebra and Its Applications, 426 (2007), pp. 634–662.

- [27] V. MEHRMANN AND L. WUNDERLICH, Hybrid systems of differential-algebraic equations - Analysis and numerical solution, *Journal of Process Control*, 19 (2009), pp. 1218–1228.
- [28] N. MONSHIZADEH, H. L. TRENTelman, AND M. K. CAMLIBEL, A simultaneous balanced truncation approach to model reduction of switched linear systems, *IEEE Transactions on Automatic Control*, 57 (2012), pp. 3118–3131.
- [29] P. J. MOSTERMAN, F. ZHAO, AND G. BISWAS, Sliding mode model semantics and simulation for hybrid systems, *Lecture Notes in Computer Science, Hybrid Systems V*, Springer, 1999, pp. 218–237.
- [30] A. V. PAPADOPOULOS AND M. PRANDINI, Model reduction of switched affine systems, *Automatica*, 70 (2016), pp. 57–65.
- [31] M. PETRECKZY, Realization theory for linear and bilinear switched systems: formal power series approach - Part I: realization theory of linear switched systems, *ESAIM Control, Optimization and Calculus of Variations*, 17 (2011), pp. 410–445.
- [32] M. PETRECKZY, A. TANWANI, AND S. TRENN, Observability of switched linear systems, *Hybrid Dynamical System Control and Observation, from Theory to Application*, Springer Verlag, 2013.
- [33] M. PETRECKZY AND J. H. VAN SCHUPPEN, Realization theory for linear hybrid systems, *IEEE Transactions on Automatic Control*, 55 (2010), pp. 2282–2297.
- [34] M. PETRECKZY, R. WISNIEWSKI, AND J. LETH, Balanced truncation for linear switched systems, in *Special Issue related to IFAC Conference on Analysis and Design of Hybrid Systems (ADHS 12)*, 2013, pp. 4–20.
- [35] W. J. RUGH, Nonlinear system theory - The Volterra/Wiener Approach, The Johns Hopkins University Press, 1981.
- [36] J. SAAK, M. KÖHLER, AND P. BENNER, M-M.E.S.S. 1.0.1, The Matrix Equations Sparse Solvers Library. <http://www.mpi-magdeburg.mpg.de/projects/mess>, 2016.
- [37] H. R. SHAKER AND R. WISNIEWSKI, Generalized Gramian framework for model/controller order reduction of switched systems, *International Journal of Systems Science*, 42 (2011), pp. 1277–1291.
- [38] Z. SUN AND S. S. GE, Switched linear systems: control and design, Springer, 2005.
- [39] ———, Stability Theory of Switched Dynamical Systems, Springer, 2011.
- [40] S. TRENN, Switched differential algebraic equations, *Advances in Industrial Control, Dynamics and Control of Switched Electronic Systems*, Springer Verlag, 2012, ch. 6.
- [41] L. ZHANG, E. BOUKAS, AND P. SHI,  $\mu$ -Dependent model reduction for uncertain discrete-time switched linear systems with average dwell time, *International Journal of Control*, 82 (2009), pp. 378–388.
- [42] L. ZHANG, P. SHI, E. BOUKAS, AND C. WANG, H-infinity model reduction for uncertain switched linear discrete-time systems, *Automatica*, 44 (2008), pp. 2944–2949.
- [43] L. ZHENG-FAN, C. CHEN-XIAO, AND D. WEN-YONG, Stability analysis and  $H_\infty$  model reduction for switched discrete-time time-delay systems, *Mathematical Problems in Engineering*, 15 (2014).