MODEL REDUCTION FOR LINEAR SYSTEMS WITH LOW-RANK SWITCHING*

PHILIPP SCHULZE† AND BENJAMIN UNGER†

Abstract. We introduce a novel model order reduction (MOR) method for large-scale linear switched systems (LSS) where the coefficient matrices are affected by a low-rank switching. The key idea is to replace the LSS by a nonswitched system with extended input and output vectors—called the envelope system—which is able to reproduce the dynamical behavior of the original LSS by applying a certain feedback law. The envelope system can be reduced using standard MOR schemes and then transformed back into an LSS. Furthermore, we present an upper bound for the output error of the reduced-order LSS and show how to preserve quadratic Lyapunov stability. The approach is tested by means of a numerical example demonstrating the efficacy of the presented method.

Key words. switched system, model reduction, balanced truncation, rational interpolation, stability, port-Hamiltonian system

AMS subject classifications. 93C30, 93A15, 30E05, 93B52

DOI. 10.1137/18M1167887

1. Introduction.

Many real world applications are a combination of continuous and discrete dynamics, which are mathematically described by *hybrid systems*. Popular examples are manufacturing processes [35], automotive engine control [6], and transportation systems [25]. In the context of control design the discrete dynamics may be used to switch between operating modes of the dynamical system. Thus the focus is on the continuous dynamics only. This results in the concept of *switched systems*.

In this paper we study linear switched systems (LSS) of the form

$$\Sigma = \begin{cases} \dot{\boldsymbol{x}}(t) = A_\sigma \boldsymbol{x}(t) + B_\sigma \boldsymbol{u}(t), \\ \boldsymbol{y}(t) = C_\sigma \boldsymbol{x}(t) + D_\sigma \boldsymbol{u}(t), \\ \boldsymbol{x}(0) = \boldsymbol{x}_0, \end{cases} \qquad t \geq 0, \tag{1.1}$$

with matrices $A_i \in \mathbb{R}^{n\times n}$, $B_i \in \mathbb{R}^{n\times m}$, $C_i \in \mathbb{R}^{p\times n}$, $D_i \in \mathbb{R}^{p\times m}$ for $i = 1,\dots,\ell$, and *switching signal* $\sigma : [0,\infty) \times \mathbb{R}^p \to \{1,\dots,\ell\}$. We refer to $\boldsymbol{x}$, $\boldsymbol{u}$, and $\boldsymbol{y}$ as the *state*, *input*, and *output* of the system, respectively. Without loss of generality, we assume a zero initial condition, i.e., $\boldsymbol{x}_0 = 0$, since otherwise we can simply append the initial condition to the input matrices as in [20] or solve two independent problems as in [10]. We include the case when the switching signal depends on the output of the system (1.1), thus allowing feedback-based switching laws. Although each subsystem in (1.1) is linear, the overall dynamics are nonlinear with a discontinuous vector field, and thus even linear time-invariant (LTI) switched systems exhibit complex phenomena.

Funding: The first author was supported by the DFG Collaborative Research Center 1029 *Substantial efficiency increase in gas turbines through direct use of coupled unsteady combustion and flow dynamics*, project A02. The second author was supported by the DFG Collaborative Research Center 910 *Control of self-organizing nonlinear systems: Theoretical methods and concepts of application*, project A2.

†Institut für Mathematik, Technische Universität Berlin, 10623 Berlin, Germany (pschulze@math.tu-berlin.de, unger@math.tu-berlin.de).

If (1.1) results from semidiscretization of a switched partial differential equation, then n is large, and the controller design, optimization, or system verification of such a large-scale system is computationally demanding or even intractable. For LTI systems without switching, the large computational cost is nowadays often reduced by replacing the system with a cheap-to-evaluate surrogate model. A standard approach to build reliable surrogate models is model order reduction (MOR), which usually constructs a low-dimensional approximation via Petrov–Galerkin projection; see [1, 2, 8, 12] for an overview. Popular reduction methods are balanced truncation [29, 30], rational interpolation (formerly known as moment matching) [18], and proper orthogonal decomposition [21, 48].

A straightforward way to generalize these methods to switched systems is to apply the reduction technique of choice to each subsystem $\Sigma_i = (A_i, B_i, C_i, D_i)$ separately; see, for example, [27, 33, 34]. However, even if during this reduction each input-output mapping is preserved, the approximation error for the switched system may be arbitrarily large (see Example 2.1). Instead, MOR methods for switched systems should be tailored to switching specific phenomena. A generalization of balanced truncation to switched systems is developed in [43] and analyzed in [37]. In these papers the authors assume that all subsystems in (1.1) can be balanced simultaneously via generalized Gramians, which is in general not feasible [26]. To compute the generalized Gramians, 2ℓ coupled Lyapunov inequalities need to be solved, which may become computationally intractable in a large-scale setting. A modified version with similar computational complexity is presented in [44], where the authors claim (without proof or numerical example) that this method performs better in terms of the approximation error. Model reduction of switched systems via interpolation is considered in [4, 5, 38]. In [4, 5] the notion of moments is defined for LSS based on the realization theory derived in [36]. Then a certain selection of moments—called *nice selections*—is interpolated. In contrast, the authors of [38] base their work on the signal generator approach [3]. Let us also mention the recent preprint [17], where model reduction is performed by introducing Gramians for the switched system that are similar to those of bilinear systems.

As is standard in MOR theory, all methods assume that the solution of the dynamical system evolves (approximately) in a low-dimensional subspace. The existence of such a low-dimensional subspace is based on the smoothness of the input-output mapping [31]—a feature that switched systems do not possess in general if arbitrary switching is allowed. In order to expect a good approximation quality with low-dimensional surrogates, we need to assume either that the switching occurs very rarely in the desired time horizon or that the switching affects only a low-dimensional part of the system.

Example 1.1. Similarly to [33, 34] we consider the heating of two adjacent rooms that are connected via a door that is either open or closed. For simplicity we restrict ourselves to a one-dimensional model. The situation is depicted in Figure 1. Opening or closing the door results in an abrupt change in the thermal conductivity and the heat capacity in the domain Ω_2. Since the domain of the door is small compared to the overall domain $\Omega := \Omega_1 \cup \Omega_2 \cup \Omega_3$, the switching influences only a small part of the dynamical system. Thus after a semidiscretization (cf. Example 5.1), only a low-dimensional structure of the overall system is affected.

In this paper we allow arbitrary switching sequences and thus assume that the switching affects only a low-dimensional part of the system. We exploit this to compute a multiple-input multiple-output (MIMO) LTI system without switching—called

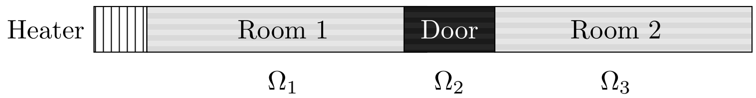

Fig. 1. *Heating of two adjacent rooms.*

the *envelope system*—that covers all the dynamics of the original switched system. More precisely, using a certain feedback control law for the envelope system, the two systems are equivalent. Model reduction is achieved by reducing the envelope system by a standard LTI MOR method. The resulting reduced-order model (ROM) can be transformed back into a switched system of the same form as (1.1). The advantage of this approach is that a comparison of models, and hence an evaluation of approximation properties, can be made for the envelope system using standard tools. Using this abstraction, the difficulties associated with switched systems and different switching signals can be circumvented. Of course, one cannot expect approximation qualities similar to those for a standard LTI system—this is reflected by a slower decay of the Hankel singular values of the envelope system.

The main contributions of this paper are the following:

1. We show (cf. Example 2.1) that independent reduction of the subsystems in (1.1) may result in arbitrarily large errors and thus should be avoided.
2. Example 2.2 demonstrates that if the switching signal depends on the state or the output of the system itself, then any approximation of the state or the output that affects the switching law may result in arbitrarily large errors, and thus no error bounds for such switching signals can be derived.
3. We propose an abstraction of the switched system (1.1) in the form of an LTI system without switching (cf. Proposition 3.4). This abstraction allows us to use standard MOR methods for switched systems and to derive an error bound for the approximation error in the time domain (Theorem 3.8) provided that the switching signal depends only on time.
4. Proposition 4.1 shows that quadratic Lyapunov stability of a switched system implies that all subsystems are dissipative Hamiltonian systems, and hence MOR techniques for port-Hamiltonian (pH) systems can be used to preserve quadratic stability in the reduced model (Theorem 4.2).

Notation. The symbol A^T denotes the transpose of the real matrix A. If A is square and symmetric (i.e., $A^T = A$), we write $A > 0$ ($A \geq 0$) to indicate that the matrix is positive definite (semidefinite). The matrix I_n is the $n \times n$ identity matrix, and $0_{n,m}$ describes the $n \times m$ zero matrix. We use $\mathrm{blkdiag}(A_1, \dots, A_k)$ to denote the block diagonal matrix with matrices $A_1, \dots, A_k$ on the diagonal. For a square-integrable function $\boldsymbol{f} : [0, T] \to \mathbb{R}^n$ we denote the $\mathcal{L}_2$ norm by $\|\boldsymbol{f}\|_{\mathcal{L}_2} := (\int_0^T \|\boldsymbol{f}(t)\|_2^2 \, \mathrm{d}t)^{1/2}$ with Euclidean vector norm $\|\cdot\|_2$. The switched system (1.1) is written as $\Sigma = (A_i, B_i, C_i, D_i \mid i = 1, \dots, \ell)$, and each subsystem

$$\dot{\boldsymbol{x}}(t) = A_j \boldsymbol{x}(t) + B_j \boldsymbol{u}(t), \qquad \boldsymbol{y}(t) = C_j \boldsymbol{x}(t) + D_j \boldsymbol{u}(t)$$

is denoted by $\Sigma_j = (A_j, B_j, C_j, D_j)$ for $j = 1, \dots, \ell$ with transfer function $H_j(s) = C_j(sI_n - A_j)^{-1} B_j + D_j$. We say that Σ_j is stable if all eigenvalues of A_j have nonpositive real part and say that Σ_j is asymptotically stable if all eigenvalues of A_j have negative real part. In the latter case, there exist matrices $\mathcal{P}_j = \mathcal{P}_j^T \geq 0$ and

$\mathcal{Q}_j = \mathcal{Q}_j^T \geq 0$ that satisfy the Lyapunov equations

$$A_j\mathcal{P}_j + \mathcal{P}_j A_j^T + B_j B_j^T = 0 \qquad \text{and} \qquad A_j^T\mathcal{Q}_j + \mathcal{Q}_j A_j + C_j^T C_j = 0.$$

The matrix $\mathcal{P}_j$ is called the controllability Gramian of Σ_j, while $\mathcal{Q}_j$ is called the observability Gramian of Σ_j. Moreover, the square roots of the eigenvalues of $\mathcal{P}_j\mathcal{Q}_j$ are the Hankel singular values of Σ_j. We say that Σ_j is controllable or observable if the condition $\operatorname{rank}(\begin{bmatrix}\lambda I_n - A_j & B_j\end{bmatrix}) = n$ or $\operatorname{rank}(\begin{bmatrix}\lambda I_n - A_j^T & C_j^T\end{bmatrix}) = n$ holds for all $\lambda \in \mathbb{C}$, respectively. The system Σ_j is called minimal if it is controllable and observable. For corresponding definitions for the switched system (1.1), we refer the reader to [36, 45]. The $\mathcal{H}_\infty$ norm of Σ_j is the $\mathcal{L}_2$–$\mathcal{L}_2$ induced norm and denoted by $\|\Sigma_j\|_{\mathcal{H}_\infty}$.

2. Problem setting and motivation. Starting from the LSS (1.1), we want to construct a reduced-order model

$$\widetilde{\Sigma} = \begin{cases} \dot{\widetilde{\boldsymbol{x}}}(t) = \widetilde{A}_\sigma \widetilde{\boldsymbol{x}}(t) + \widetilde{B}_\sigma \boldsymbol{u}(t), \\ \widetilde{\boldsymbol{y}}(t) = \widetilde{C}_\sigma \widetilde{\boldsymbol{x}}(t) + \widetilde{D}_\sigma \boldsymbol{u}(t), \end{cases} \qquad t \geq 0, \tag{2.1}$$

of the same form as (1.1) with $\widetilde{A}_i \in \mathbb{R}^{r\times r}$, $\widetilde{B}_i \in \mathbb{R}^{r\times m}$, $\widetilde{C}_i \in \mathbb{R}^{p\times r}$, and $\widetilde{D}_i \in \mathbb{R}^{p\times m}$ for $i = 1,\ldots,\ell$ and $r \ll n$, such that the approximation error $\boldsymbol{y} - \widetilde{\boldsymbol{y}}$ is small (in some appropriate norm). If we want to emphasize the dependency of the state or the output on the input variable, we write $\boldsymbol{x}(\boldsymbol{u})$ or $\boldsymbol{y}(\boldsymbol{u})$.

As is common for many model reduction techniques [1], the reduced matrices $\widetilde{A}_i$, $\widetilde{B}_i$, $\widetilde{C}_i$, and $\widetilde{D}_i$ are constructed via Petrov–Galerkin projection. More precisely, we determine matrices $W_i, V_i \in \mathbb{R}^{n\times r}$ with $W_i^T V_i$ nonsingular and compute the reduced model $\widetilde{\Sigma}$ via

$$\widetilde{A}_i := \left(W_i^T V_i\right)^{-1} W_i^T A_i V_i, \quad \widetilde{B}_i := \left(W_i^T V_i\right)^{-1} W_i^T B_i, \quad \widetilde{C}_i := C_i V_i, \quad \widetilde{D}_i := D_i \tag{2.2}$$

for $i = 1,\ldots,\ell$. Note that if we use different W_i and V_i for each subsystem, then we assume $V_i\widetilde{\boldsymbol{x}} \approx \boldsymbol{x}$ for each subsystem, and thus the reduced state $\widetilde{\boldsymbol{x}} \approx (W_i^T V_i)^{-1} W_i^T \boldsymbol{x}$ depends on the subsystem. Accordingly we need a state transition transformation. If we switch from subsystem i to subsystem j, we need to transform the state via $\widetilde{\boldsymbol{x}} \mapsto \widetilde{K}_{ij}\widetilde{\boldsymbol{x}}$ with

$$\widetilde{K}_{ij} = \left(W_j^T V_j\right)^{-1} W_j^T V_i. \tag{2.3}$$

A standard approach in the literature [27, 33, 34] is to reduce the subsystems independently. The following example shows that this may result in large approximation errors.

Example 2.1. We consider the RLC circuit given in Figure 2, where the input is given by the entering current i, while the output of interest is i_L, the current of the branch containing the inductance. The corresponding control system is given in [42], but here we modify the system by allowing a switch in the inductance. Precisely, we set $R = 1$, $C = 1$, $L_1 = 1/2$, and $L_2 = 1$. The resulting switched system has two modes, i.e., $\ell = 2$, given by

$$A_1 = \begin{bmatrix} 0 & -1 \\ 2 & -4 \end{bmatrix}, \quad A_2 = \begin{bmatrix} 0 & -1 \\ 1 & -2 \end{bmatrix}, \quad B_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \quad B_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad C_1^T = C_2^T = \begin{bmatrix} 0 \\ 1 \end{bmatrix};$$

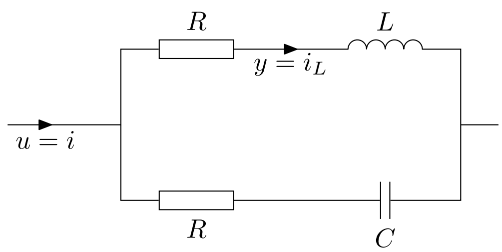

FIG. 2. *A simple RLC circuit* [42].

i.e., the switching occurs only in the matrices A_i and B_i but not in the output matrix. The transfer functions of the two subsystems are given by

$$H_1(s) = \frac{2s+2}{s^2+4s+2} \qquad \text{and} \qquad H_2(s) = \frac{1}{s+1}.$$

Notice that the first subsystem is minimal but the second is not controllable. We only reduce the dimension of the second subsystem by truncating the states that are not controllable (e.g., via Kalman decomposition); i.e., we make no error with respect to the transfer function. This can, for instance, be achieved using the projection matrices

$$V_2^T = \begin{bmatrix} 1 & 1 \end{bmatrix}, \quad W_2^T = \begin{bmatrix} 2 & -1 \end{bmatrix},$$

and via (2.2) the reduced model and its transfer functions are given by

$$\widetilde{A}_2 = -1, \qquad \widetilde{B}_2 = 1, \qquad \widetilde{C}_2 = 1, \qquad \widetilde{H}_2(s) = \frac{1}{s+1}.$$

Now let us suppose that we have the switching function

$$\sigma : [0,2] \to \{1,2\}, \qquad \sigma(t) = \begin{cases} 1 & \text{if } t \in [0,1], \\ 2 & \text{if } t \in (1,2], \end{cases}$$

and we are interested in the output $\boldsymbol{y}$ of the switched system at the final time $t = 2$. We fix $\xi \in \mathbb{R}$ and choose $\boldsymbol{u}\big|_{[0,1]}$ such that $\boldsymbol{x}(1) = \begin{bmatrix} \xi & 2\xi \end{bmatrix}^T$. This is possible since the first system is controllable. Afterward we no longer control the system, i.e., $\boldsymbol{u}\big|_{(1,2]} \equiv 0$. This yields $\boldsymbol{x}(2) = \begin{bmatrix} 0 & \xi \exp(-1) \end{bmatrix}^T$ and hence $\boldsymbol{y}(2) = \xi \exp(-1)$. However, if we use the reduced model for the computation, we obtain $\widetilde{\boldsymbol{y}}(2) = 0$, and thus the error can be arbitrarily large, although no approximation error in the input-output mapping of both subsystems was made.

The reason for the behavior in Example 2.1 is the fact that in a minimal realization of a switched system the subsystems might not be minimal [36]. This can be seen from the solution formula for switched systems. More precisely, assume that the switching signal σ depends only on time, and the switching time points are $0 = t_0 < t_1 < \cdots < t_s$ with switching index sequence $i_k := \sigma(t_k)$. Then for $t \geq t_s$ the solution [45] of the state equation in (1.1) is given by (recall that we assume $\boldsymbol{x}_0 = 0$)

$$\begin{aligned} \boldsymbol{x}(t;\boldsymbol{u}) &= \mathrm{e}^{A_{i_s}(t-t_s)} \cdots \mathrm{e}^{A_{i_1}(t_2-t_1)} \int_{t_0}^{t_1} \mathrm{e}^{A_{i_0}(t_1-\tau)} B_{i_0} \boldsymbol{u}(\tau)\mathrm{d}\tau + \cdots \\ &\quad + \mathrm{e}^{A_{i_s}(t-t_s)} \int_{t_{s-1}}^{t_s} \mathrm{e}^{A_{i_{s-1}}(t_s-\tau)} B_{i_{s-1}} \boldsymbol{u}(\tau)\mathrm{d}\tau + \int_{t_s}^{t} \mathrm{e}^{A_{i_s}(t-\tau)} B_{i_s} \boldsymbol{u}(\tau)\mathrm{d}\tau. \end{aligned} \tag{2.4}$$

If we reduce each subsystem independently, we expect a good approximation of

$$\int_{t_k}^{t} C_{i_k} \mathrm{e}^{A_{i_k}(t-\tau)} B_{i_k} \boldsymbol{u}(\tau) \mathrm{d}\tau + D_{i_k} \boldsymbol{u}(t) \qquad \text{for each } 0 \leq k \leq s.$$

However, this does not necessarily imply a good approximation of the interaction of the subsystems, which is described by the terms

$$\mathrm{e}^{A_{i_k}(t-t_k)} \cdots \mathrm{e}^{A_{i_1}(t_2-t_1)} \int_{t_0}^{t_1} \mathrm{e}^{A_{i_0}(t_1-\tau)} B_{i_0} \boldsymbol{u}(\tau) \mathrm{d}\tau + \cdots \\ + \mathrm{e}^{A_{i_k}(t-t_k)} \int_{t_{k-1}}^{t_k} \mathrm{e}^{A_{i_{k-1}}(t_k-\tau)} B_{i_{k-1}} \boldsymbol{u}(\tau) \mathrm{d}\tau.$$

The interaction terms can be interpreted as (nonzero) initial conditions at the switching time points. Note that also the techniques offered in [10, 20] cannot be used, since it is a priori unclear which initial conditions are important.

A standard question in approximation theory is about the approximation quality in the form of error bounds or error estimators. The following toy example illustrates what we can expect to happen if the switching signal depends on the state or the output of the system.

Example 2.2. Consider the switched system $\Sigma = (A_i, B_i, C_i, D_i \mid i = 1, \ldots, \ell)$, where the first subsystem $\Sigma_1 = (A_1, B_1, C_1, D_1)$ is given by

$$A_1 = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}, \qquad B_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \qquad C_1^T = \begin{bmatrix} 1 \\ p \end{bmatrix}, \qquad D_1 = 0 \tag{2.5}$$

with parameter $p > 0$. The approximation $\widetilde{\Sigma}_1 = (\widetilde{A}_1, \widetilde{B}_1, \widetilde{C}_1, \widetilde{D}_1)$ with

$$\widetilde{A}_1 = -1, \qquad \widetilde{B}_1 = 1, \qquad \widetilde{C}_1 = 1, \qquad \widetilde{D}_1 = 0 \tag{2.6}$$

satisfies (without switching) $\widetilde{\boldsymbol{y}} = (1+p)^{-1}\boldsymbol{y}$, and thus the approximation error becomes arbitrarily small for $p \to 0$. Suppose the switching condition and the input are given by

$$\sigma(t, \boldsymbol{y}) = \begin{cases} 1 & \text{if } \boldsymbol{y} \geq \left(1 + \frac{p}{2}\right)(1 - \exp(-1)), \\ 2 & \text{otherwise,} \end{cases} \qquad \boldsymbol{u}(t) = \begin{cases} 1 & \text{if } t \leq 1, \\ 0 & \text{otherwise.} \end{cases}$$

Then for $p > 0$, we have $\widetilde{\boldsymbol{y}}(t) < \left(1 + \frac{p}{2}\right)(1 - \exp(-1))$ for all $t > 0$, whereas there exists a $t_1 \in (0, 1)$ with $\boldsymbol{y}(t_1) = \left(1 + \frac{p}{2}\right)(1 - \exp(-1))$. The situation is depicted in Figure 3. Thus the approximation (2.6) driven with the above input and switching condition never switches the system, while the original model (2.5) switches to the second (not further specified) subsystem at t_1.

The failure of the approximation in the previous example is caused by the fact that for small p the output approaches the switching surface approximately tangentially, which makes the system ill-posed for this input, and one may need a regularization [23]. Since we do not know the input a priori, we cannot exclude such cases. Anyhow, the previous example shows that, independent from the approximation quality of the subsystem, the error may become arbitrarily large if the switching depends on the state or the output. Thus we cannot expect to have error estimators for output-dependent switching laws, which implies that we need the following assumption to derive an error bound.

Assumption 2.3. The switching signal $\sigma : [0, \infty) \to \{1, \ldots, \ell\}$ depends only on time.

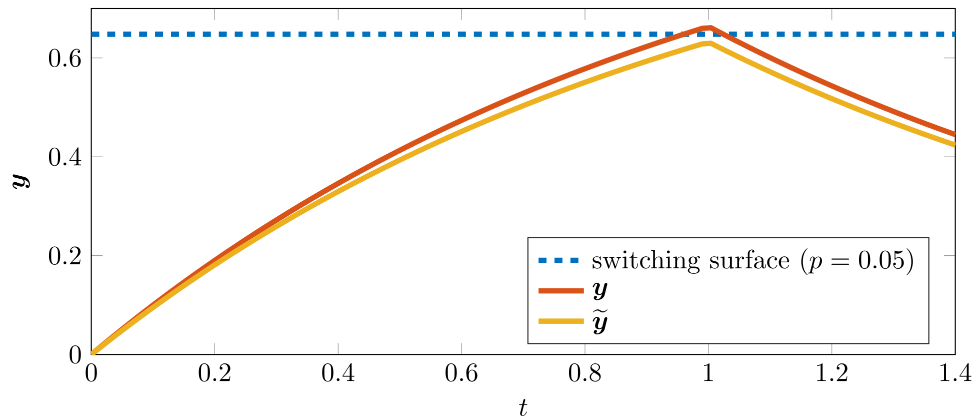

FIG. 3. *Output-dependent switching: full model and approximation.*

3. Envelope system and model reduction. Consider the switched system (1.1) with switching time points $0 = t_0 < t_1 < \cdots < t_s$ and switching sequence $i_k := \sigma(t_k)$. In order to circumvent the problems that come with switching between different subsystems, we propose the following view. For a given input $\boldsymbol{u}$, is it possible to modify $\boldsymbol{u}$ on the interval $(t_1, t_2]$ such that the unswitched system driven with the modified input produces the same output as the switched system with the original input in the interval $[0, t_2]$? If this is the case, then we could run the first subsystem without needing to switch and still observe the same behavior.

However, in general it is not possible to excite the first subsystem such that it behaves as the second subsystem; for an example, we refer the reader to [45]. Instead we construct an *envelope system* of the first subsystem with extended inputs and outputs that is able to mimic the behavior of all other modes of the switched system. Note that a similar viewpoint was considered for parametric MOR in [7] and for MOR of systems with time delay in [46].

Let us describe for $i = 2, \ldots, \ell$ the difference between the ith subsystem and the first subsystem via the matrices

$$\begin{aligned} \Delta_{A,i} &:= A_1 - A_i \in \mathbb{R}^{n\times n}, \quad \Delta_{B,i} := B_1 - B_i \in \mathbb{R}^{n\times m}, \\ \Delta_{C,i} &:= C_1 - C_i \in \mathbb{R}^{p\times n}, \quad \Delta_{D,i} := D_1 - D_i \in \mathbb{R}^{p\times m}, \end{aligned} \tag{3.1}$$

and we additionally assume a decomposition

$$\Delta_{A,i} = S_i M_i T_i^T \qquad \text{with} \qquad S_i, T_i \in \mathbb{R}^{n\times\beta_i}, \quad M_i \in \mathbb{R}^{\beta_i\times\beta_i} \tag{3.2}$$

with $\beta_i = \operatorname{rank}(\Delta_{A,i})$. Notice that we can always obtain such a decomposition via singular value decomposition (SVD) or Schur decomposition. Using the decomposition (3.2), we introduce for $i = 2, \ldots, \ell$ the new variables

$$\boldsymbol{z}_i(t) := T_i^T \boldsymbol{x}(t) \in \mathbb{R}^{\beta_i}.$$

Thus using the Kronecker delta $\delta_{i,j}$, (1.1) can be rewritten as the descriptor system

$$\left\{\begin{aligned} \dot{\boldsymbol{x}}(t) &= A_1\boldsymbol{x}(t) - \sum_{i=2}^{\ell} S_i\delta_{i,\sigma}M_i\boldsymbol{z}_i(t) + \left(B_1 - \sum_{i=2}^{\ell}\delta_{i,\sigma}\Delta_{B,i}\right)\boldsymbol{u}(t),\\ \boldsymbol{z}_i(t) &= T_i^T\boldsymbol{x}(t),\\ \boldsymbol{y}(t) &= \left(C_1 - \sum_{i=2}^{\ell}\delta_{i,\sigma}\Delta_{C,i}\right)\boldsymbol{x}(t) + \left(D_1 - \sum_{i=2}^{\ell}\delta_{i,\sigma}\Delta_{D,i}\right)\boldsymbol{u}(t). \end{aligned}\right. \tag{3.3}$$

Viewing $\delta_{i,\sigma}M_i\boldsymbol{z}_i$ as a feedback control law and defining the variables

$$\boldsymbol{x}_{\mathrm{E}} = \boldsymbol{x}, \qquad \boldsymbol{u}_{\mathrm{E}} = \begin{bmatrix} \boldsymbol{u}\\ \delta_{2,\sigma}\boldsymbol{u}\\ \vdots\\ \delta_{\ell,\sigma}\boldsymbol{u}\\ \delta_{2,\sigma}M_2\boldsymbol{z}_2\\ \vdots\\ \delta_{\ell,\sigma}M_\ell\boldsymbol{z}_\ell \end{bmatrix}, \qquad \boldsymbol{y}_{\mathrm{E}} = \begin{bmatrix} C_1\boldsymbol{x} + D_1\boldsymbol{u}\\ \Delta_{C,2}\boldsymbol{x} + \Delta_{D,2}\delta_{2,\sigma}\boldsymbol{u}\\ \vdots\\ \Delta_{C,\ell}\boldsymbol{x} + \Delta_{D,\ell}\delta_{\ell,\sigma}\boldsymbol{u}\\ \boldsymbol{z}_2\\ \vdots\\ \boldsymbol{z}_\ell \end{bmatrix}$$

motivates the following abstraction of the switched system (1.1).

Definition 3.1 (envelope system). *Consider the switched system* (1.1) *and matrices* $\Delta_{B,i}$, $\Delta_{C,i}$, $\Delta_{D,i}$, S_i, *and* T_i *as in* (3.1) *and* (3.2). *Define*

$$\begin{aligned} A_{\mathrm{E}} &:= A_1 \in \mathbb{R}^{n_{\mathrm{E}}\times n_{\mathrm{E}}},\\ B_{\mathrm{E}} &:= \begin{bmatrix} B_1 & \Delta_{B,2} & \cdots & \Delta_{B,\ell} & S_2 & \cdots & S_\ell \end{bmatrix} \in \mathbb{R}^{n_{\mathrm{E}}\times m_{\mathrm{E}}},\\ C_{\mathrm{E}}^T &:= \begin{bmatrix} C_1^T & \Delta_{C,2}^T & \cdots & \Delta_{C,\ell}^T & T_2 & \cdots & T_\ell \end{bmatrix} \in \mathbb{R}^{n_{\mathrm{E}}\times p_{\mathrm{E}}},\\ D_{\mathrm{E}} &:= \operatorname{blkdiag}(D_1, \Delta_{D,2}, \ldots, \Delta_{D,\ell}, 0_{\mathfrak{L},\mathfrak{L}}) \in \mathbb{R}^{p_{\mathrm{E}}\times m_{\mathrm{E}}}, \end{aligned}$$

with dimensions $n_{\mathrm{E}} := n$, $m_{\mathrm{E}} := \ell m + \mathfrak{L}$, $p_{\mathrm{E}} := \ell p + \mathfrak{L}$, *and* $\mathfrak{L} := \sum_{i=2}^{\ell}\beta_i$. *Then we call*

$$\Sigma_{\mathrm{E}} = \begin{cases} \dot{\boldsymbol{x}}_{\mathrm{E}}(t) = A_{\mathrm{E}}\boldsymbol{x}_{\mathrm{E}}(t) + B_{\mathrm{E}}\boldsymbol{u}_{\mathrm{E}}(t),\\ \boldsymbol{y}_{\mathrm{E}}(t) = C_{\mathrm{E}}\boldsymbol{x}_{\mathrm{E}}(t) + D_{\mathrm{E}}\boldsymbol{u}_{\mathrm{E}}(t), \end{cases} \qquad t \geq 0,$$

an envelope system *for the switched system* (1.1).

Notice that we have included the T_i matrices in the output matrix C_{E}, since in terms of model reduction, it is not sufficient to have a good approximation of the output variable $\boldsymbol{y}$, but we need in addition a good approximation of $\boldsymbol{z}_i$ to compute the feedback control law.

Remark 3.2. In a large-scale setting the computation of (3.2) via a full SVD is computationally expensive. However, the decomposition (3.2) requires only a skinny SVD, which can be computed efficiently; see [13].

Remark 3.3. Notice that the envelope system is not unique. In addition to different low-rank factorizations (3.2), also the ordering of the subsystems in (1.1) is an arbitrary choice that affects the construction of the envelope system. Moreover, it may be computationally advantageous to compute the difference of the subsystems to a hypothesized system $\Sigma = (A, B, C, D)$. This may be particularly useful if neither of the subsystems is asymptotically stable but all subsystems are low-rank perturbations of an asymptotically stable system. Moreover, it is a priori not clear that B_{E} and C_{E} have full rank, and thus a rank-revealing decomposition of B_{E} and C_{E} may be used to compress the number of inputs and outputs.

The connection of the envelope system (3.1) to the original switched system (1.1) is as follows. For a given input $\boldsymbol{u}$, consider the feedback control $\boldsymbol{u}_{\mathrm{E}} = K(\sigma)\boldsymbol{y}_{\mathrm{E}} + K_0(\sigma)\boldsymbol{u}$ with

$$
\begin{aligned}
K(\sigma) &:= \operatorname{blkdiag}(0_{\ell m,\ell p}, -\delta_{2,\sigma}M_2, \ldots, -\delta_{\ell,\sigma}M_\ell) \in \mathbb{R}^{m_{\mathrm{E}}\times p_{\mathrm{E}}},\\
K_0(\sigma) &:= \begin{bmatrix} I_m & -\delta_{2,\sigma}I_m & \cdots & -\delta_{\ell,\sigma}I_m & 0_{m,\mathfrak{L}}\end{bmatrix}^T \in \mathbb{R}^{m_{\mathrm{E}}\times m},
\end{aligned} \tag{3.4}
$$

and the matrix function

$$
C_0(\sigma) := \begin{bmatrix} I_p & -\delta_{2,\sigma}I_p & \cdots & -\delta_{\ell,\sigma}I_p & 0_{p,\mathfrak{L}}\end{bmatrix} \in \mathbb{R}^{p\times p_{\mathrm{E}}}. \tag{3.5}
$$

Then by construction, the envelope system (3.1) with feedback law (3.4) reproduces the output of the LSS (1.1). More precisely, we have the following result.

PROPOSITION 3.4. *Consider the LSS* (1.1) *and its envelope system* (3.1)*. For a given input $\boldsymbol{u}$ let $\boldsymbol{y}(t;\boldsymbol{u})$ denote the output of* (1.1) *under this input, and accordingly, let $\boldsymbol{y}_{\mathrm{E}}(t;\boldsymbol{u}_{\mathrm{E}})$ denote the output of the envelope system with input $\boldsymbol{u}_{\mathrm{E}}$. Then*

$$
\boldsymbol{y}(t;\boldsymbol{u}) = C_0(\sigma)\boldsymbol{y}_{\mathrm{E}}\left(t; K(\sigma)\boldsymbol{y}_{\mathrm{E}} + K_0(\sigma)\boldsymbol{u}\right),
$$

where K, K_0, and C_0 are defined as in (3.4) *and* (3.5)*.*

Remark 3.5. If the envelope system is not minimal, then neither is the switched system, and we can remove the uncontrollable and unobservable parts of the envelope system. Unfortunately, the converse is not true. Consider, for example, the switched system $\Sigma = (A_i, B_i, C_i, 0 \mid i = 1, 2)$ with

$$
A_1 = \begin{bmatrix} -1 & 0 \\ 0 & -2 \end{bmatrix}, \qquad \Delta_{A,2} = \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}\begin{bmatrix} 1 & 0 \end{bmatrix}, \qquad B_1 = B_2 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.
$$

The reachable set of this system from the zero initial condition under arbitrary switching is $\{0\}$, and thus, following [45], the switched system is not controllable. However, the envelope system

$$
A_{\mathrm{E}} = \begin{bmatrix} -1 & 0 \\ 0 & -2 \end{bmatrix}, \qquad B_{\mathrm{E}} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}
$$

is controllable.

The advantage of the envelope system (3.1) over the switched system (1.1) is that we can apply all standard MOR methods for MIMO LTI systems, for example, balanced truncation [29] or the iterative rational Krylov algorithm (IRKA) [18]. Since the construction of the envelope system is computationally tractable (provided that the matrices $\Delta_{A,i}$ have low rank), MOR of switched systems within our framework is possible whenever MOR for one of the subsystems is computationally feasible.

Assume that the MOR method of choice has determined projection matrices $W, V \in \mathbb{R}^{n,r}$ for the envelope system (3.1). Then the reduced envelope system is given via

$$
\widetilde{A}_{\mathrm{E}} = (W^TV)^{-1}W^TA_{\mathrm{E}}V, \quad \widetilde{B}_{\mathrm{E}} = (W^TV)^{-1}W^TB_{\mathrm{E}}, \quad \widetilde{C}_{\mathrm{E}} = C_{\mathrm{E}}V, \quad \widetilde{D}_{\mathrm{E}} = D_{\mathrm{E}}. \tag{3.6}
$$

On the other hand, we can reduce the switched system (1.1) with W and V according to (2.2) via

$$\widetilde{A}_i = (W^T V)^{-1} W^T A_i V, \quad \widetilde{B}_i = (W^T V)^{-1} W^T B_i, \quad \widetilde{C}_i = C_i V, \quad \widetilde{D}_i = D_i \tag{3.7}$$

for $i = 1, \ldots, \ell$. We infer from

$$\Delta_{\widetilde{A},i} := \widetilde{A}_1 - \widetilde{A}_i = (W^T V)^{-1} W^T \Delta_{A,i} V = (W^T V)^{-1} W^T S_i M_i T_i^T V$$

that the reduced envelope system (3.6) is an envelope system for the reduced system (3.7), and thus Proposition 3.4 applies.

Note that we can drive the envelope system with input signals other than the feedback law from Proposition 3.4 and thus can generate dynamics that the switched system cannot produce. Similarly, standard LTI MOR methods, such as balanced truncation and IRKA, aim for a good approximation for all possible input signals and are not tailored to a restricted input space, which in our case is parameterized due to the feedback law.

Remark 3.6. The larger the part of the system that is affected by the switching, i.e., the larger $\sum_{i=2}^{\ell} \beta_i$, the more we have to extend the input and the output dimensions of the envelope system, which thus makes the system harder to reduce (provided $\operatorname{rank}(B_{\mathrm{E}}) = m_{\mathrm{E}}$ and $\operatorname{rank}(C_{\mathrm{E}}) = p_{\mathrm{E}}$; see Remark 3.3). Notably, the controllability Gramian $\mathcal{P}_{\mathrm{E}}$ of the envelope system Σ_{E} satisfies

$$0 = A_{\mathrm{E}} \mathcal{P}_{\mathrm{E}} + \mathcal{P}_{\mathrm{E}} A_{\mathrm{E}}^T + B_{\mathrm{E}} B_{\mathrm{E}}^T = A_1 \mathcal{P}_{\mathrm{E}} + \mathcal{P}_{\mathrm{E}} A_1^T + B_1 B_1^T + \mathcal{Q}$$

for some $\mathcal{Q} = \mathcal{Q}^T \geq 0$ and is thus a generalized Gramian for the first subsystem; cf. [43]. In particular, the Hankel singular values of the envelope system are greater than or equal to the Hankel singular values of the first subsystem.

Remark 3.7. As outlined in Remark 3.3, the construction of the envelope system is not unique. In particular, the decomposition (3.2) allows us to weight the subsystems differently (for example, via scaling of S_i, M_i, and T_i) for the model reduction procedure. This may be exploited if the switching sequence is known beforehand. A theoretical discussion and optimal scaling for the modes is, however, beyond the scope of this paper. It is worth mentioning that scaling issues are not unique to our methodology but appear in other MOR methods, e.g., balanced truncation. Heuristics for the scaling are, for instance, proposed in [20].

As outlined in Example 2.2, in general it is not possible to derive an error bound or error estimator for LSS with state- or output-dependent switching. Thus, we need to invoke Assumption 2.3, which guarantees that the full-order model and the reduced model switch at the same time to the same subsystem. This in turn allows us to use error bounds that we obtain based on the model reduction of the envelope system. For instance, if we assume that the envelope system is minimal and asymptotically stable, then the error bound for balanced truncation [1] for square-integrable inputs $\boldsymbol{u}_{\mathrm{E}}$ is given by

$$\|\boldsymbol{y}_{\mathrm{E}}(\boldsymbol{u}_{\mathrm{E}}) - \widetilde{\boldsymbol{y}}_{\mathrm{E}}(\boldsymbol{u}_{\mathrm{E}})\|_{\mathcal{L}_2} \leq \left(2 \sum_{i=k}^{q} \sigma_i \right) \|\boldsymbol{u}_{\mathrm{E}}\|_{\mathcal{L}_2},$$

where $\sigma_1 > \sigma_2 > \cdots > \sigma_q > 0$ denote the distinct Hankel singular values of Σ_{E} with multiplicities $\nu_1, \ldots, \nu_q \in \mathbb{N}$ and σ_k denotes the largest neglected Hankel singular value. Recall that $\boldsymbol{y}_{\mathrm{E}}(\boldsymbol{u}_{\mathrm{E}})$ denotes the solution of the envelope system (3.1) with input $\boldsymbol{u}_{\mathrm{E}}$, and $\widetilde{\boldsymbol{y}}_{\mathrm{E}}(\boldsymbol{u}_{\mathrm{E}})$ denotes the solution of the reduced envelope system (3.6) with the same input $\boldsymbol{u}_{\mathrm{E}}$. Unfortunately, this error bound is not carried over to the error of the switched system, since the feedback in Proposition 3.4 is different for the envelope system and for the reduced envelope system. Thus we need to compare $\boldsymbol{y}_{\mathrm{E}}(\boldsymbol{u}_{\mathrm{E}})$ and $\widetilde{\boldsymbol{y}}_{\mathrm{E}}(\widetilde{\boldsymbol{u}}_{\mathrm{E}})$. Moreover, any error bound that is obtained in this way is of a posteriori type, since the input for the envelope system depends on the actual solution of the system.

THEOREM 3.8. *Let σ satisfy Assumption 2.3, and for any square-integrable input $\boldsymbol{u}$ let $\widetilde{\boldsymbol{u}}_{\mathrm{E}} := K(\sigma)\widetilde{\boldsymbol{y}}_{\mathrm{E}} + K_0(\sigma)\boldsymbol{u}$ denote the feedback law from Proposition 3.4 for the reduced envelope system. Moreover, assume that*

$$\max_{i=2,\ldots,\ell} \|M_i\|_2 \|\Sigma_{\mathrm{E}}\|_{\mathcal{H}_\infty} < 1. \tag{3.8}$$

Then there exists a constant $\eta = \eta(\Sigma_{\mathrm{E}}) > 0$ independent of $\boldsymbol{u}$ and $\widetilde{\Sigma}_{\mathrm{E}}$ such that

$$\|\boldsymbol{y}(\boldsymbol{u}) - \widetilde{\boldsymbol{y}}(\boldsymbol{u})\|_{\mathcal{L}_2} \leq \eta \|\Sigma_{\mathrm{E}} - \widetilde{\Sigma}_{\mathrm{E}}\|_{\mathcal{H}_\infty} \|\widetilde{\boldsymbol{u}}_{\mathrm{E}}\|_{\mathcal{L}_2} \tag{3.9}$$

holds.

Proof. First, note that due to Assumption 2.3, the matrix functions C_0, K, and K_0 are identical for the envelope system and the reduced envelope system. Let $\boldsymbol{u}_{\mathrm{E}} = K(\sigma)\boldsymbol{y}_{\mathrm{E}} + K_0(\sigma)\boldsymbol{u}$ denote the feedback law from Proposition 3.4 for the envelope system. From Proposition 3.4 we infer $\boldsymbol{y}(\boldsymbol{u}) = C_0(\sigma)\boldsymbol{y}_{\mathrm{E}}(\boldsymbol{u}_{\mathrm{E}})$ and $\widetilde{\boldsymbol{y}}(\boldsymbol{u}) = C_0(\sigma)\widetilde{\boldsymbol{y}}_{\mathrm{E}}(\widetilde{\boldsymbol{u}}_{\mathrm{E}})$. The triangle inequality implies

$$\begin{aligned} \|\boldsymbol{y}_{\mathrm{E}}(\boldsymbol{u}_{\mathrm{E}}) - \widetilde{\boldsymbol{y}}_{\mathrm{E}}(\widetilde{\boldsymbol{u}}_{\mathrm{E}})\|_{\mathcal{L}_2} &\leq \|\boldsymbol{y}_{\mathrm{E}}(\boldsymbol{u}_{\mathrm{E}}) - \boldsymbol{y}_{\mathrm{E}}(\widetilde{\boldsymbol{u}}_{\mathrm{E}})\|_{\mathcal{L}_2} + \|\boldsymbol{y}_{\mathrm{E}}(\widetilde{\boldsymbol{u}}_{\mathrm{E}}) - \widetilde{\boldsymbol{y}}_{\mathrm{E}}(\widetilde{\boldsymbol{u}}_{\mathrm{E}})\|_{\mathcal{L}_2} \\ &\leq \|\boldsymbol{y}_{\mathrm{E}}(\boldsymbol{u}_{\mathrm{E}}) - \boldsymbol{y}_{\mathrm{E}}(\widetilde{\boldsymbol{u}}_{\mathrm{E}})\|_{\mathcal{L}_2} + \|\Sigma_{\mathrm{E}} - \widetilde{\Sigma}_{\mathrm{E}}\|_{\mathcal{H}_\infty} \|\widetilde{\boldsymbol{u}}_{\mathrm{E}}\|_{\mathcal{L}_2}. \end{aligned} \tag{3.10}$$

Since for any $\sigma \in \{1, \ldots, \ell\}$ the largest singular value of C_0 is $\sqrt{2}$, we obtain

$$\begin{aligned} \|\boldsymbol{y}(\boldsymbol{u}) - \widetilde{\boldsymbol{y}}(\boldsymbol{u})\|_{\mathcal{L}_2} &= \|C_0(\sigma)\left(\boldsymbol{y}_{\mathrm{E}}(\boldsymbol{u}_{\mathrm{E}}) - \widetilde{\boldsymbol{y}}_{\mathrm{E}}(\widetilde{\boldsymbol{u}}_{\mathrm{E}})\right)\|_{\mathcal{L}_2} \leq \sqrt{2}\|\boldsymbol{y}_{\mathrm{E}}(\boldsymbol{u}_{\mathrm{E}}) - \widetilde{\boldsymbol{y}}_{\mathrm{E}}(\widetilde{\boldsymbol{u}}_{\mathrm{E}})\|_{\mathcal{L}_2} \\ &\leq \sqrt{2}\left(\|\boldsymbol{y}_{\mathrm{E}}(\boldsymbol{u}_{\mathrm{E}}) - \boldsymbol{y}_{\mathrm{E}}(\widetilde{\boldsymbol{u}}_{\mathrm{E}})\|_{\mathcal{L}_2} + \|\Sigma_{\mathrm{E}} - \widetilde{\Sigma}_{\mathrm{E}}\|_{\mathcal{H}_\infty} \|\widetilde{\boldsymbol{u}}_{\mathrm{E}}\|_{\mathcal{L}_2}\right). \end{aligned}$$

Due to the definitions of $\boldsymbol{u}_{\mathrm{E}}$ and $\widetilde{\boldsymbol{u}}_{\mathrm{E}}$ and Assumption 2.3 we get

$$\begin{aligned} \|\boldsymbol{y}_{\mathrm{E}}(\boldsymbol{u}_{\mathrm{E}}) - \boldsymbol{y}_{\mathrm{E}}(\widetilde{\boldsymbol{u}}_{\mathrm{E}})\|_{\mathcal{L}_2} &\leq \|\Sigma_{\mathrm{E}}\|_{\mathcal{H}_\infty} \|\boldsymbol{u}_{\mathrm{E}} - \widetilde{\boldsymbol{u}}_{\mathrm{E}}\|_{\mathcal{L}_2} \\ &= \|\Sigma_{\mathrm{E}}\|_{\mathcal{H}_\infty} \|K(\sigma)\left(\boldsymbol{y}_{\mathrm{E}}(\boldsymbol{u}_{\mathrm{E}}) - \widetilde{\boldsymbol{y}}_{\mathrm{E}}(\widetilde{\boldsymbol{u}}_{\mathrm{E}})\right)\|_{\mathcal{L}_2} \\ &\leq \|\Sigma_{\mathrm{E}}\|_{\mathcal{H}_\infty} \max_{i=2,\ldots,\ell} \|M_i\|_2 \|\boldsymbol{y}_{\mathrm{E}}(\boldsymbol{u}_{\mathrm{E}}) - \widetilde{\boldsymbol{y}}_{\mathrm{E}}(\widetilde{\boldsymbol{u}}_{\mathrm{E}})\|_{\mathcal{L}_2}. \end{aligned}$$

Equation (3.10) implies

$$\begin{aligned} &\|\boldsymbol{y}_{\mathrm{E}}(\boldsymbol{u}_{\mathrm{E}}) - \boldsymbol{y}_{\mathrm{E}}(\widetilde{\boldsymbol{u}}_{\mathrm{E}})\|_{\mathcal{L}_2} \\ &\qquad \leq \max_{i=2,\ldots,\ell} \|M_i\|_2 \|\Sigma_{\mathrm{E}}\|_{\mathcal{H}_\infty} \left(\|\boldsymbol{y}_{\mathrm{E}}(\boldsymbol{u}_{\mathrm{E}}) - \boldsymbol{y}_{\mathrm{E}}(\widetilde{\boldsymbol{u}}_{\mathrm{E}})\|_{\mathcal{L}_2} + \|\Sigma_{\mathrm{E}} - \widetilde{\Sigma}_{\mathrm{E}}\|_{\mathcal{H}_\infty} \|\widetilde{\boldsymbol{u}}_{\mathrm{E}}\|_{\mathcal{L}_2}\right), \end{aligned}$$

yielding

$$\|\boldsymbol{y}_{\mathrm{E}}(\boldsymbol{u}_{\mathrm{E}}) - \boldsymbol{y}_{\mathrm{E}}(\widetilde{\boldsymbol{u}}_{\mathrm{E}})\|_{\mathcal{L}_2} \leq \frac{\max_{i=2,\ldots,\ell} \|M_i\|_2 \|\Sigma_{\mathrm{E}}\|_{\mathcal{H}_\infty}}{1 - \max_{i=2,\ldots,\ell} \|M_i\|_2 \|\Sigma_{\mathrm{E}}\|_{\mathcal{H}_\infty}} \|\Sigma_{\mathrm{E}} - \widetilde{\Sigma}_{\mathrm{E}}\|_{\mathcal{H}_\infty} \|\widetilde{\boldsymbol{u}}_{\mathrm{E}}\|_{\mathcal{L}_2},$$

where we have used (3.8). Summarizing all calculations yields

$$\begin{aligned}\|\boldsymbol{y}(\boldsymbol{u})-\widetilde{\boldsymbol{y}}(\boldsymbol{u})\|_{\mathcal{L}_2} &\le \sqrt{2}\left(\|\boldsymbol{y}_{\mathrm{E}}(\boldsymbol{u}_{\mathrm{E}})-\boldsymbol{y}_{\mathrm{E}}(\widetilde{\boldsymbol{u}}_{\mathrm{E}})\|_{\mathcal{L}_2}+\|\Sigma_{\mathrm{E}}-\widetilde{\Sigma}_{\mathrm{E}}\|_{\mathcal{H}_\infty}\|\widetilde{\boldsymbol{u}}_{\mathrm{E}}\|_{\mathcal{L}_2}\right)\\ &\le \sqrt{2}\left(\frac{\max\limits_{i=2,\dots,\ell}\|M_i\|_2\|\Sigma_{\mathrm{E}}\|_{\mathcal{H}_\infty}}{1-\max\limits_{i=2,\dots,\ell}\|M_i\|_2\|\Sigma_{\mathrm{E}}\|_{\mathcal{H}_\infty}}+1\right)\|\Sigma_{\mathrm{E}}-\widetilde{\Sigma}_{\mathrm{E}}\|_{\mathcal{H}_\infty}\|\widetilde{\boldsymbol{u}}_{\mathrm{E}}\|_{\mathcal{L}_2}\\ &\le \frac{\sqrt{2}}{1-\max\limits_{i=2,\dots,\ell}\|M_i\|_2\|\Sigma_{\mathrm{E}}\|_{\mathcal{H}_\infty}}\|\Sigma_{\mathrm{E}}-\widetilde{\Sigma}_{\mathrm{E}}\|_{\mathcal{H}_\infty}\|\widetilde{\boldsymbol{u}}_{\mathrm{E}}\|_{\mathcal{L}_2},\end{aligned}$$

which completes the proof. □

Note that the switching information is encoded in $\|\widetilde{\boldsymbol{u}}_{\mathrm{E}}\|_{\mathcal{L}_2}$ and thus explicitly taken into account in the error bound. Nevertheless, the bound (3.9) seems rather restrictive, since in the proof we bound the output error $\boldsymbol{y}-\widetilde{\boldsymbol{y}}$ by adding up all output errors of the envelope system $\boldsymbol{y}_{\mathrm{E}}-\widetilde{\boldsymbol{y}}_{\mathrm{E}}$.

Remark 3.9. Although the theory presented above is tailored to systems in standard state-space form, it is easily generalized to generalized state-space form, i.e., a switched system of the form

$$E_\sigma\dot{\boldsymbol{x}}(t)=A_\sigma\boldsymbol{x}(t)+B_\sigma\boldsymbol{u}(t),\qquad \boldsymbol{y}(t)=C_\sigma\boldsymbol{x}(t)+D_\sigma\boldsymbol{u}(t)$$

with nonsingular matrices $E_i\in\mathbb{R}^{n\times n}$. Setting $\Delta_{E,i}:=E_1-E_i$, we have

$$E_i^{-1}=(E_1-\Delta_{E,i})^{-1}=E_1^{-1}+E_1^{-1}\Delta_{E,i}E_i^{-1}$$

for $i=2,\dots,\ell$. Transformation to standard state-space form thus transforms the perturbations in (3.2) into

$$\Delta_{A,i}\mapsto E_1^{-1}\Delta_{A,i}-E_1^{-1}\Delta_{E,i}E_i^{-1}A_i,\qquad \Delta_{B,i}\mapsto E_1^{-1}\Delta_{B,i}-E_1^{-1}\Delta_{E,i}E_i^{-1}B_i.$$

In particular, if the perturbations have a low-rank structure, then the transformed perturbations are also low-rank, and thus the framework presented above can be applied.

4. Structure preservation. In the last decades, a lot of research in model reduction was devoted to structure preservation [9, 15, 19, 22, 24, 41] and preservation of system properties, such as stability or passivity. The latter properties are directly encoded in pH systems [47], which can be written in the form

$$\text{(4.1)}\qquad \begin{cases}\dot{\boldsymbol{x}}(t)=(J-R)\,Q\boldsymbol{x}(t)+B\boldsymbol{u}(t),\\ \boldsymbol{y}(t)=B^TQ\boldsymbol{x}(t),\end{cases}$$

with $J=-J^T$, $R=R^T\ge 0$, and $Q=Q^T>0$. In this context any square matrix A of the form $A=(J-R)Q$ is called a *dissipative Hamiltonian* matrix [16, 28]. The special structure in (4.1) allows us to use the Hamiltonian $\mathcal{H}(\boldsymbol{x})=1/2\boldsymbol{x}^TQ\boldsymbol{x}$, which often represents the total energy of the system as a Lyapunov function, and thus shows that pH systems are stable. Let us mention the recent work [11, 28] for a more general form of (4.1).

In terms of switched systems, stability is usually not defined for each subsystem separately but in terms of a joint Lyapunov function. More precisely, an LSS (1.1) is called *quadratic Lyapunov stable* [26] if there exists a symmetric positive definite matrix $Q \in \mathbb{R}^{n\times n}$ that satisfies the linear matrix inequalities

$$A_i^T Q + QA_i < 0 \qquad \text{for all } i \in \{1, \dots, \ell\}. \tag{4.2}$$

Following [16, Lemma 2], we immediately have the following connection between quadratic Lyapunov stable LSS and dissipative Hamiltonian systems.

PROPOSITION 4.1. *The LSS* (1.1) *is quadratic Lyapunov stable if and only if there exist some $Q \in \mathbb{R}^{n\times n}$ with $Q = Q^T > 0$ and matrices $J_i, R_i \in \mathbb{R}^{n\times n}$ with $J_i^T = -J_i$ and $R_i = R_i^T > 0$ such that $A_i = (J_i - R_i)Q$ for $i = 1, \dots, \ell$.*

Proof. Suppose that the switched system (1.1) is quadratic Lyapunov stable; i.e., there exists $Q \in \mathbb{R}^{n\times n}$ with $Q = Q^T > 0$ such that (4.2) holds. For $i = 1, \dots, \ell$ define

$$J_i := \frac{1}{2}\left(A_i Q^{-1} - Q^{-1}A_i^T\right) \qquad \text{and} \qquad R_i := -\frac{1}{2}\left(A_i Q^{-1} + Q^{-1}A_i^T\right).$$

Then $J_i^T = -J_i$, $R_i = R_i^T$, and $A_i = (J_i - R_i)Q$. The R_i's are positive definite if and only if (4.2) holds, since

$$\boldsymbol{x}^T R_i \boldsymbol{x} = -\frac{1}{2}\boldsymbol{x}^T\left(A_i Q^{-1} + Q^{-1}A_i^T\right)\boldsymbol{x} = -\frac{1}{2}(Q^{-1}\boldsymbol{x})^T\left(A_i^T Q + QA_i\right)Q^{-1}\boldsymbol{x}$$

for each $\boldsymbol{x} \in \mathbb{R}^n$, which also shows the converse direction. ☐

Using Proposition 4.1, we see that for a quadratic Lyapunov stable LSS the envelope system (3.1) is also a dissipative Hamiltonian system [16], and hence we can use any of the techniques derived for structure-preserving model reduction for pH systems [14, 19, 49] to ensure quadratic stability in the reduced switched system as well. Following [19] we obtain, for instance, the following result.

THEOREM 4.2. *Suppose that the LSS* (1.1) *is quadratic Lyapunov stable, i.e.,* (1.1) *satisfies* (4.2) *for some $Q \in \mathbb{R}^{n\times n}$ with $Q = Q^T > 0$. For any $V \in \mathbb{R}^{n\times r}$ with full column rank, set $W = QV(V^TQV)^{-1}$. Then the reduced switched system $\widetilde{\Sigma} = (\widetilde{A}_i, \widetilde{B}_i, \widetilde{C}_i, \widetilde{D}_i \mid i = 1, \dots, \ell)$ with $\widetilde{A}_i, \widetilde{B}_i, \widetilde{C}_i, \widetilde{D}_i$ obtained as in* (3.7) *is quadratic Lyapunov stable.*

Proof. For $i = 1, \dots, \ell$ we have

$$\begin{aligned}\widetilde{A}_i^T V^T QV + V^T QV\widetilde{A}_i &= V^T A_i^T W V^T QV + V^T QV W^T A_i V \\ &= V^T(A_i^T Q + QA_i)V < 0.\end{aligned}$$ ☐

From a modeling perspective, the matrix Q is usually not known beforehand, and hence we may only assume that each mode of the switched system exhibits a pH structure, i.e.,

$$\begin{cases}\dot{\boldsymbol{x}}(t) = (J_\sigma - R_\sigma)\,Q_\sigma \boldsymbol{x}(t) + B_\sigma \boldsymbol{u}(t), \\ \boldsymbol{y}(t) = B_\sigma^T Q_\sigma \boldsymbol{x}(t),\end{cases} \tag{4.3}$$

with $J_i = -J_i^T$, $R_i = R_i^T \geq 0$, and $Q_i = Q_i^T > 0$. Instead of asking to preserve potential quadratic stability of the switched system, we can also aim for preserving the pH structure of each subsystem, which in turn implies that each subsystem is stable and passive. Note that stability of the subsystems does not imply quadratic stability of the switched system; for a counterexample we refer the reader to [26].

Remark 4.3. If for $i = 1, \dots, \ell$ the matrices R_i in (4.3) are positive definite and $Q_i = \alpha_i Q_0$ for some $\alpha_i > 0$ and a symmetric positive definite matrix Q_0, then the previous discussion immediately implies that the LSS (4.3) is quadratic Lyapunov stable.

A structure-preserving model reduction scheme for the LSS (4.3) can be obtained by a straightforward extension of Theorem 4.2 and [19]. This is summarized in the following result.

THEOREM 4.4. *Consider the LSS* (4.3). *For any* $V_i \in \mathbb{R}^{n\times r}$ *with full column rank, set* $W_i = Q_i V_i (V_i^T Q_i V_i)^{-1}$ *for* $i = 1, \dots, \ell$. *Then each mode of the reduced switched system in* (2.1) *exhibits a pH structure.*

Theorems 4.2 and 4.4 both do not further specify the projection matrices V and V_i. Following our framework from section 3, we can compute the envelope system and apply any MOR method of choice to determine projection matrices $W, V \in \mathbb{R}^{n\times r}$. In general, the left projection matrix W does not satisfy the condition in Theorems 4.2 and 4.4 and thus needs to be modified accordingly. Let us mention that one can modify the construction of V within the IRKA framework such that some necessary $\mathcal{H}_2$ optimality conditions are satisfied, while at the same time the pH structure is preserved [19].

So far, we have considered cases when either a common quadratic Lyapunov function or a pH structure of the subsystems of (1.1) is known. However, if neither of them is available, we can still preserve potential stability of the subsystems, for example, by formulating different envelope systems for the different modes with $A_{\mathrm{E}} = A_i$, etc., for $i = 2, \dots, \ell$ and reducing each of them separately with a stability-preserving model reduction scheme. In this case the transformation matrices in (2.3) are, in general, not identity matrices. However, we postpone a detailed analysis of this situation to future work. An alternative way of preserving stability is to write each stable A_i as $A_i = (J_i - R_i)Q_i$ (cf. [16, Lemma 2]) and afterward apply a projection as in Theorem 4.4.

5. Numerical examples. We illustrate the proposed method by means of numerical examples and compare the ROMs with the respective full-order model. Further examples are presented in the preprint [40] of this article. In Figures 5 and 6, we use the notation FOM for the full-order model and $\widetilde{\Sigma}_{\mathrm{E}}$ for the ROMs obtained by the method presented in section 3. More precisely, $\widetilde{\Sigma}_{\mathrm{E}}$ refers to the usage of IRKA [2] for reducing the envelope system, and $\widetilde{\Sigma}_{\mathrm{E}}$ BT refers to balanced truncation [29].

For constructing the envelope system, we use an SVD $\Delta_{A,i} = U_i \Sigma_i V_i^T$ for the decomposition in (3.2) and assign $S_i = U_i$, $M_i = I_{\beta_i}$, and $T_i = V_i \Sigma_i$. Furthermore, all time simulations are performed with the MATLAB solver `ode45` with default settings, except for Example 5.2 where we used a backward Euler scheme. The source code for the numerical examples is available from [39].

Example 5.1. Continuing with Example 1.1, we ignore the effects of convection and model the time evolution of the temperature distribution by the heat equation

$$\zeta(\xi,\sigma)\partial_t z(t,\xi) - \partial_\xi \left(k(\xi,\sigma)\partial_\xi z(t,\xi)\right) = 0 \quad \text{for } (t,\xi) \in [0,T] \times \Omega, \tag{5.1}$$

where k is the *thermal conductivity*, ζ is the *volumetric heat capacity*, and z is the relative temperature difference with respect to the ambient temperature θ_∞, i.e., $z(t,\xi) = \theta(t,\xi) - \theta_\infty$ with absolute temperature θ. The thermal conductivity and the volumetric heat capacity are assumed to be homogeneous over the total domain apart from the fact that the properties in the second subdomain Ω_2 are affected by a switch, i.e.,

$$k(\xi,\sigma) = \begin{cases} k_{\Omega_2}(\sigma) & \text{if } \xi \in \Omega_2, \\ k_2 & \text{otherwise,} \end{cases} \qquad \zeta(\xi,\sigma) = \begin{cases} \zeta_{\Omega_2}(\sigma) & \text{if } \xi \in \Omega_2, \\ \zeta_2 & \text{otherwise,} \end{cases}$$

$$k_{\Omega_2}(\sigma) = \begin{cases} k_1 & \text{if } \sigma = 1, \\ k_2 & \text{if } \sigma = 2, \end{cases} \qquad \zeta_{\Omega_2}(\sigma) = \begin{cases} \zeta_1 & \text{if } \sigma = 1, \\ \zeta_2 & \text{if } \sigma = 2. \end{cases}$$

The heater on the left side acts as a boundary control, and we assume that on the right side of room 2 there is a heat transfer to the environment; i.e., we assume homogeneous Robin boundary conditions at this point; cf. [32]. More precisely, we have the boundary conditions

$$-k_2 \partial_\xi z(t,0) = \boldsymbol{u}(t) \quad \text{for } t \in [0,T], \tag{5.2}$$

$$k_2 \partial_\xi z(t,L) + h z(t,L) = 0 \quad \text{for } t \in [0,T], \tag{5.3}$$

where L is the total length of the computational domain Ω and h denotes the heat transfer coefficient at the right boundary. Furthermore, we assume that the initial temperature distribution in both rooms and in the door equals the ambient temperature yielding the initial condition $z(0,\xi) = 0$ for $\xi \in \Omega$. As output we use the average temperature in the second room, which is

$$\boldsymbol{y}(t) = \frac{1}{|\Omega_3|} \int_{\Omega_3} z(t,\xi)\,\mathrm{d}\xi \quad \text{for } t \in [0,T].$$

Discretization with finite volumes leads to a switched system of the form (1.1), where the first mode corresponds to the closed door and the second to the open door. Details about the discretization are given in the preprint of this article [40]. The used physical, geometry, and discretization parameters are listed in Tables 1 and 2. Here we assume that the (closed) door acts like an isolator, i.e., with low thermal conductivity k_1 and high heat capacity ζ_1.

TABLE 1
Example 5.1. *Physical parameters.*

$\zeta_1\ \left[\frac{\mathrm{J}}{\mathrm{m}^3\mathrm{K}}\right]$	$\zeta_2\ \left[\frac{\mathrm{J}}{\mathrm{m}^3\mathrm{K}}\right]$	$k_1\ \left[\frac{\mathrm{W}}{\mathrm{mK}}\right]$	$k_2\ \left[\frac{\mathrm{W}}{\mathrm{mK}}\right]$	$h\ \left[\frac{\mathrm{W}}{\mathrm{m}^2\mathrm{K}}\right]$
2×10^6	700	0.015	3	100

TABLE 2
Example 5.1. *Geometry and discretization parameters.*

	$\lvert\Omega_i\rvert\ [m]$	$\Delta\xi_i\ [m]$	n_i
$i \in \{1,3\}$	5	0.1	50
$i = 2$	0.3	0.1	3

The Hankel singular values of the two subsystems and of the envelope system are depicted in Figure 4. Since the Hankel singular values decay is fast for all three systems, we expect a good approximation quality with a small dimension.

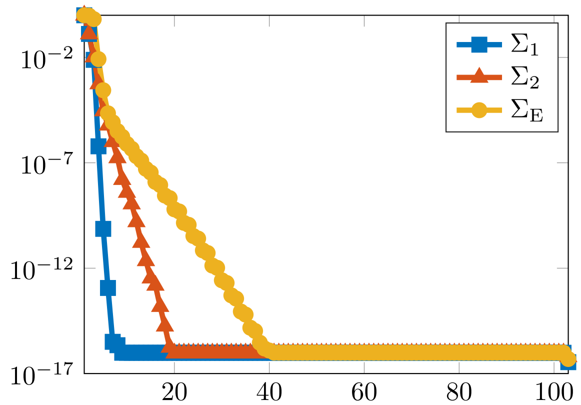

FIG. 4. *Example* 5.1. *Normalized Hankel singular values of the two subsystems (Σ_1 and Σ_2) and of the envelope system Σ_{E}.*

We construct ROMs based on IRKA and on BT of the envelope system with reduced dimensions $r = 6$ and $r = 10$. As input we use a constant heat flux density of $\boldsymbol{u}(t) \equiv 1\mathrm{W}/(\mathrm{m}^2)$, and the switching sequence is given by

$$\sigma(t) = \begin{cases} 2, & t \in [0, 1.1] \cup (1.6, 1.7], \\ 1, & t \in (1.1, 1.6] \cup (1.7, 6]. \end{cases}$$

Figure 5 shows the time evolution of the output of the FOM and of the ROMs. At the beginning the door is open, and the temperature in the second room increases until the system is switched the first time. Afterward, the output temperature decreases since the isolating door separates the second room from the heated (first) room. After each switching the output dynamics change abruptly depending on whether the door is open or not. This behavior of the FOM ($n = 103$) is captured by the ROMs of dimension 10, whose graphs lie on top of that of the FOM (the $\mathcal{L}_\infty$ errors are less than 10^{-2}). On the other hand, the ROMs of dimension 6 show clear quantitative deviations from the FOM, while the qualitative behavior is still captured reasonably well. The significant difference between the reduced models of dimension 6 and those of dimension 10 can be explained by the strong decay between the sixth and the tenth Hankel singular values of the envelope system; cf. Figure 4.

In Example 2.2 we demonstrated with a toy example that output-dependent switching may pose a severe challenge for MOR. Indeed, the next example illustrates that using an output-dependent switching law in the previous example may result in inaccurate approximations.

Example 5.2. We revisit Example 5.1 and consider an output-dependent switching signal that closes the door whenever the average temperature in the second room becomes too high and that opens the door whenever the average temperature falls below a certain threshold. The corresponding switching law reads

$$\begin{cases} \text{if } (\sigma = 1 \wedge \boldsymbol{y}(t) < \vartheta_1), & \text{switch to } \sigma = 2, \\ \text{if } (\sigma = 2 \wedge \boldsymbol{y}(t) > \vartheta_2), & \text{switch to } \sigma = 1. \end{cases}$$

To avoid a permanent switching between the modes, the threshold temperatures ϑ_1 and ϑ_2 should be chosen such that $\vartheta_1 < \vartheta_2$. Here, we choose $\vartheta_1 = 0.2\mathrm{K}$ and $\vartheta_2 = 0.5\mathrm{K}$.

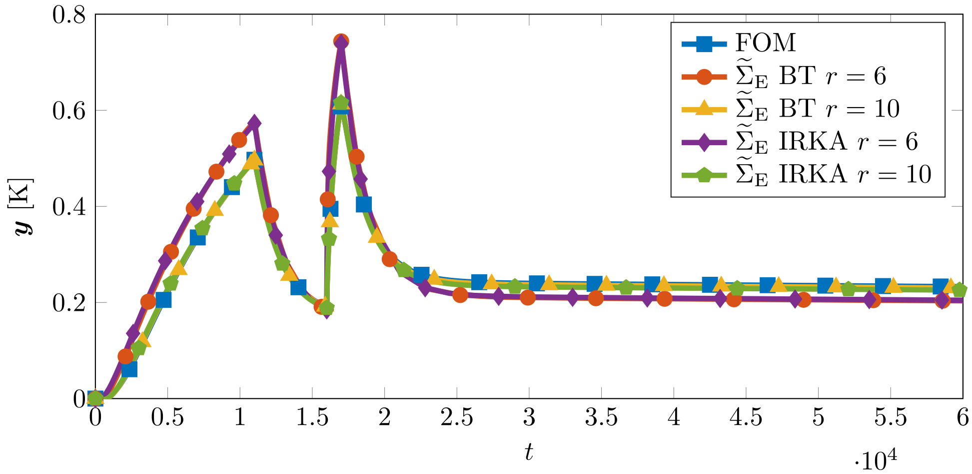

FIG. 5. *Example* 5.1. *Time domain simulation of the FOM and the ROMs.*

The other parameters and the input signal are chosen as before; see Example 5.1. Moreover, the initial value for σ is set to 2 again, which refers to an open door.

Figure 6 depicts the simulated time evolution of the output of the FOM and of the ROMs from Example 5.2. The dynamics of the FOM look similar to those in the case of the time-dependent switching. The ROM of dimension 10 is again sufficient to approximate the dynamics very well, although the approximation error is larger compared to Example 5.1. While in Example 5.1 the ROM of dimension 6 was at least able to capture the qualitative behavior, in the output-dependent switching setting, the output of the ROM of order 6 reveals more switches than the full-order model, and consequently the error is large; see Figure 6.

6. Summary. For switched LTI systems, we introduced in Definition 3.1 a so-called envelope system, which is a standard LTI system without switching and which allows us to reproduce the output of the original switched system by using a suitable feedback law (Proposition 3.4). The envelope system can be reduced by classical MOR techniques, and the reduced envelope system can be transformed back into a linear switched system by the same feedback law as that for the full-order envelope system. Our theoretical findings are confirmed by several numerical examples. The strengths of this new MOR framework for switched systems are summarized in the following:

- The model reduction is applied to the nonswitched envelope system, which allows the usage of standard model reduction techniques and, hence, existing implementations.
- The approach allows for efficient treatment of large-scale systems since the state dimension of the envelope system equals the state dimension of the modes of the switched system. The main cost for constructing the envelope system is due to the matrix decompositions in (3.2), which can be computed efficiently; cf. Remark 3.2.
- The LTI envelope system is more easily accessible for analysis than the original switched system. Specifically, an error bound for the output error of the reduced-order switched system has been presented.
- For the case when the original switched system is quadratic Lyapunov stable

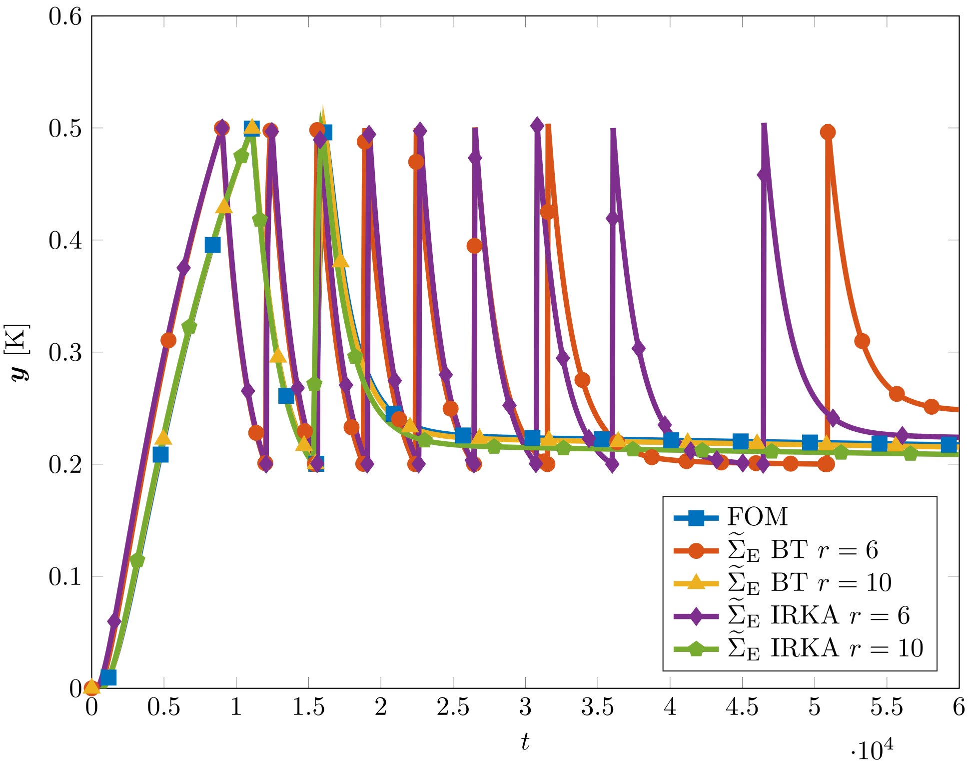

FIG. 6. *Example* 5.2. *Time domain simulation of the FOM and the ROMs.*

or consists of pH subsystems, we can adapt the model reduction scheme such that these properties are preserved in the reduced switched system.

In contrast to the strengths are the following weaknesses:

- While the envelope system behaves like the original switched system when applying a certain feedback law, other inputs are also possible and allow for different dynamics. Thus the envelope reduction framework may be improved if one modifies existing MOR methods to handle restricted input spaces.
- The derived error bound seems to be quite conservative since it is based on a sum of the output errors of the envelope system.

An interesting direction for future work is to analyze the approximation quality of the ROM with respect to the freedoms in the construction of the envelope system. In particular, it is an open problem to determine optimal scaling strategies for the low-rank decomposition in (3.2) and the effect of the different orderings of the subsystems.

REFERENCES

[1] A. C. Antoulas, *Approximation of Large-Scale Dynamical Systems*, Adv. Des. Control 6, SIAM, Philadelphia, 2005, https://doi.org/10.1137/1.9780898718713.

[2] A. C. Antoulas, C. A. Beattie, and S. Gugercin, *Interpolatory Model Reduction of Large-Scale Dynamical Systems*, in Efficient Modeling and Control of Large-Scale Systems, J. Mohammadpour and K. M. Grigoriadis, eds., Springer, New York, 2010, pp. 3–58.

[3] A. Astolfi, *Model reduction by moment matching for linear and nonlinear systems*, IEEE Trans. Automat. Control, 55 (2010), pp. 2321–2336, https://doi.org/10.1109/TAC.2010.2046044.

[4] M. Baştuğ, M. Petreczky, R. Wisniewski, and J. Leth, *Model reduction by moment matching for linear switched systems*, in Proceedings of the 2014 American Control Conference, Portland, OR, 2014, pp. 3942–3947, https://doi.org/10.1109/ACC.2014.6858983.

[5] M. Baştuğ, M. Petreczky, R. Wisniewski, and J. Leth, *Model reduction by nice selections for linear switched systems*, IEEE Trans. Automat. Control, 61 (2016), pp. 3422–3437, https://doi.org/10.1109/TAC.2016.2518023.

[6] A. Balluchi, L. Benvenuti, M. D. di Benedetto, C. Pinello, and A. L. Sangiovanni-Vincentelli, *Automotive engine control and hybrid systems: Challenges and opportunities*, Proc. IEEE, 88 (2000), pp. 888–912.

[7] U. Baur, C. Beattie, and P. Benner, *Mapping parameters across system boundaries: Parameterized model reduction with low rank variability in dynamics*, Proc. Appl. Math. Mech., 14 (2014), pp. 19–22, https://doi.org/10.1002/pamm.201410006.

[8] U. Baur, P. Benner, and L. Feng, *Model order reduction for linear and nonlinear systems: A system-theoretic perspective*, Arch. Comput. Methods Eng., 21 (2014), pp. 331–358.

[9] C. Beattie and S. Gugercin, *Interpolatory projection methods for structure-preserving model reduction*, Syst. Control Lett., 58 (2009), pp. 225–232, https://doi.org/10.1016/j.sysconle.2008.10.016.

[10] C. Beattie, S. Gugercin, and V. Mehrmann, *Model reduction for systems with inhomogeneous initial conditions*, Syst. Control Lett., 99 (2017), pp. 99–106, https://doi.org/10.1016/j.sysconle.2016.11.007.

[11] C. Beattie, V. Mehrmann, H. Xu, and H. Zwart, *Port-Hamiltonian Descriptor Systems*, preprint, https://arxiv.org/abs/1705.09081, 2017.

[12] P. Benner, A. Cohen, M. Ohlberger, and K. Willcox, *Model Reduction and Approximation. Theory and Algorithms*, Comput. Sci. Eng. 15, SIAM, Philadelphia, 2017, https://doi.org/10.1137/1.9781611974829.

[13] M. W. Berry, *Large scale sparse singular value computations*, Int. J. High Performance Comput. Appl., 6 (1992), pp. 13–49, https://doi.org/10.1177/109434209200600103.

[14] S. Chaturantabut, C. Beattie, and S. Gugercin, *Structure-preserving model reduction for nonlinear port-Hamiltonian systems*, SIAM J. Sci. Comput., 38 (2016), pp. B837–B865, https://doi.org/10.1137/15M1055085.

[15] R. W. Freund, *Structure-preserving model order reduction of RCL circuit equations*, in Model Order Reduction: Theory, Research Aspects and Applications, W. H. A. Schilders, H. A. van der Vorst, and J. Rommes, eds., Springer, Berlin, 2008, pp. 49–73, https://doi.org/10.1007/978-3-540-78841-6_3.

[16] N. Gillis and P. Sharma, *On computing the distance to stability for matrices using linear dissipative Hamiltonian systems*, Automatica, 85 (2017), pp. 113–121, https://doi.org/10.1016/j.automatica.2017.07.047.

[17] I. V. Gosea, M. Petreczky, A. C. Antoulas, and C. Fiter, *Balanced Truncation for Linear Switched Systems*, preprint, https://arxiv.org/abs/1712.02158, 2017.

[18] S. Gugercin, A. C. Antoulas, and C. Beattie, *$\mathcal{H}_2$ model reduction for large-scale linear dynamical systems*, SIAM J. Matrix Anal. Appl., 30 (2008), pp. 609–638, https://doi.org/10.1137/060666123.

[19] S. Gugercin, R. V. Polyuga, C. Beattie, and A. van der Schaft, *Structure-preserving tangential interpolation for model reduction of port-Hamiltonian systems*, Automatica, 48 (2012), pp. 1963–1974.

[20] M. Heinkenschloss, T. Reis, and A. C. Antoulas, *Balanced truncation model reduction for systems with inhomogeneous initial conditions*, Automatica, 47 (2011), pp. 559–564, https://doi.org/10.1016/j.automatica.2010.12.002.

[21] M. Hinze and S. Volkwein, *Proper orthogonal decomposition surrogate models for nonlinear dynamical systems: Error estimates and suboptimal control*, in Dimension Reduction of Large-Scale Systems, P. Benner, D. C. Sorensen, and V. Mehrmann, eds., Springer-Verlag, Berlin, 2005, pp. 261–306, https://doi.org/10.1007/3-540-27909-1_10.

[22] E. Jarlebring, T. Damm, and W. Michiels, *Model reduction of time-delay systems using position balancing and delay Lyapunov equations*, Math. Control Signals Systems, 25 (2013), pp. 147–166, https://doi.org/10.1007/s00498-012-0096-9.

[23] P. Kunkel and V. Mehrmann, *Regular Solutions of DAE Hybrid Systems and Regularization Techniques*, preprint 05-2017, Institut für Mathematik, TU Berlin, 2017.

[24] S. Lall, P. Krysl, and J. E. Marsden, *Structure-preserving model reduction for mechanical systems*, Phys. D Nonlinear Phenom., 184 (2003), pp. 304–318, https://doi.org/10.1016/S0167-2789(03)00227-6.

[25] J. Lee and S. Song, *Modeling urban transportation systems with hybrid systems: An overview*, in Proceedings of the 2003 IEEE International Conference on Intelligent Transportation Systems, Shanghai, China, 2003, pp. 1269–1272.

[26] D. Liberzon, *Switching in Systems and Control*, Birkhäuser Boston, 2003, https://doi.org/10.1007/978-1-4612-0017-8.

[27] E. Mazzi, A. S. Vincentelli, A. Balluchi, and A. Bicchi, *Hybrid system reduction*, in Proceedings of the 47th IEEE Conference on Decision and Control, Cancun, Mexico, 2008, pp. 227–232, https://doi.org/10.1109/CDC.2008.4739350.

[28] C. Mehl, V. Mehrmann, and M. Wojtylak, *Linear Algebra Properties of Dissipative Hamiltonian Descriptor Systems*, preprint, https://arxiv.org/abs/1801.02214, 2018.

[29] B. Moore, *Principal component analysis in linear systems: Controllability, observability, and model reduction*, IEEE Trans. Automat. Control, 26 (1981), pp. 17–32.

[30] C. Mullis and R. Roberts, *Synthesis of minimum roundoff noise fixed point digital filters*, IEEE Trans. Circuits Syst., 23 (1976), pp. 551–562.

[31] M. Ohlberger and S. Rave, *Reduced basis methods: Success, limitations and future challenges*, in Proceedings of Algoritmy 2016, Conference on Scientific Computing, Podbanské, Slovakia, 2016, pp. 1–12.

[32] M. N. Özisik, *Boundary Value Problems of Heat Conduction*, Dover Publications, Mineola, NY, 1968.

[33] A. V. Papadopoulos and M. Prandini, *Model reduction of switched affine systems: A method based on balanced truncation and randomized optimization*, in Proceedings of the 17th International Conference on Hybrid Systems: Computation and Control, Berlin, 2014, pp. 113–122, https://doi.org/10.1145/2562059.2562131.

[34] A. V. Papadopoulos and M. Prandini, *Model reduction of switched affine systems*, Automatica, 70 (2016), pp. 57–65.

[35] D. L. Pepyne and C. G. Cassandras, *Optimal control of hybrid systems in manufacturing*, Proc. IEEE, 88 (2000), pp. 1108–1123.

[36] M. Petreczky, *Realization theory for linear and bilinear switched systems: Formal power series approach—Part* I*: Realization theory of linear switched systems*, ESAIM Control Optim. Calc. Var., 17 (2011), pp. 410–445, https://doi.org/10.1051/cocv/2010014.

[37] M. Petreczky, R. Wisniewski, and J. Leth, *Balanced truncation for linear switched systems*, Nonlinear Anal. Hybrid Syst., 10 (2013), pp. 4–20, https://doi.org/10.1016/j.nahs.2013.03.007.

[38] G. Scarciotti and A. Astolfi, *Model reduction for hybrid systems with state-dependent jumps*, IFAC-PapersOnLine, 49 (2016), pp. 850–855, https://doi.org/10.1016/j.ifacol.2016.10.272.

[39] P. Schulze and B. Unger, *Code for the Paper "Model Reduction for Linear Systems with Low-Rank Switching,"* preprint, https://doi.org/10.5281/zenodo.1161994, 2018.

[40] P. Schulze and B. Unger, *Model Reduction for Linear Systems with Low-Rank Switching*, preprint, https://arxiv.org/abs/1801.09445, 2018.

[41] P. Schulze, B. Unger, C. Beattie, and S. Gugercin, *Data-driven structured realization*, Linear Algebra Appl., 537 (2018), pp. 250–286, https://doi.org/10.1016/j.laa.2017.09.030.

[42] K. Selvaraj, M. N. Belur, and R. Abdulrazak, *Dissipative Systems: Uncontrollability, Observability and RLC Realizability*, preprint, https://arxiv.org/abs/1110.2102v1, 2011.

[43] H. R. Shaker and R. Wisniewski, *Generalized Gramian framework for model reduction of switched systems*, in Proceedings of the 2009 European Control Conference (ECC), Budapest, Hungary, 2009, pp. 1029–1034.

[44] H. R. Shaker and R. Wisniewski, *Model reduction of switched systems based on switching generalized Gramians*, Int. J. Innov. Comput. Inform. Control, 8 (2012), pp. 5025–5044.

[45] Z. Sun, S. S. Ge, and T. H. Lee, *Controllability and reachability criteria for switched linear systems*, Automatica, 38 (2002), pp. 775–786, https://doi.org/10.1016/S0005-1098(01)00267-9.

[46] N. van de Wouw, W. Michiels, and B. Besselink, *Model reduction for delay differential equations with guaranteed stability and error bound*, Automatica, 55 (2015), pp. 132–139, https://doi.org/10.1016/j.automatica.2015.02.031.

[47] A. van der Schaft and D. Jeltsema, *Port-Hamiltonian Systems Theory: An Introductory Overview*, Now Publishers Inc., Hanover, MA, 2014.

[48] S. Volkwein, *Optimal control of a phase-field model using proper orthogonal decomposition*, ZAMM Z. Angew. Math. Mech., 81 (2001), pp. 83–97.

[49] T. Wolf, B. Lohmann, R. Eid, and P. Kotyczka, *Passivity and structure preserving order reduction of linear port-Hamiltonian systems using Krylov subspaces*, Eur. J. Control, 16 (2010), pp. 401–406.