Technische Universität Berlin

Institut für Mathematik

Systematic Discretization of Input/Output Maps and Control of Partial Differential Equations

J. Heiland, V. Mehrmann and M. Schmidt

Preprint 2010/23

Preprint-Reihe des Instituts für Mathematik
Technische Universität Berlin

Report 2010/23 October 2010

Systematic Discretization of Input/Output Maps and Control of Partial Differential Equations

Jan Heiland* Volker Mehrmann* Michael Schmidt †

October 13, 2010

Abstract

We present a framework for the direct discretization of the input/output map of dynamical systems governed by linear partial differential equations with distributed inputs and outputs. The approximation consists of two steps. First, the input and output signals are discretized in space and time, resulting in finite dimensional spaces for the input and output signals. These are then used to approximate the dynamics of the system. The approximation errors in both steps are balanced and a matrix representation of an approximate input/output map is constructed which can be further reduced using singular value decompositions. We present the discretization framework, corresponding error estimates, and the SVD-based system reduction method. The theoretical results are illustrated with some applications in the optimal control of partial differential equations.

Keywords input/output maps, discretization, control of partial differential equations
AMS subject classification. 39C20, 35B37.

1 Introduction

The real-time control of complex physical systems is a major challenge in many engineering applications as well as in mathematical research. Typically, these control systems are modeled by infinite-dimensional state space

*Institut für Mathematik, TU Berlin,
10623 Berlin, Germany
Email: {heiland,mehrmann}@math.tu-berlin.de

†GE Global Research,
85748 Garching bei München, Germany,
Email: mail@schmidt-michael.de

systems on the basis of (instationary and nonlinear) partial differential equations (PDEs). The challenge arises from the fact that on the one hand, space-discretizations resolving most of the state information typically lead to very large semi-discrete systems, on the other hand, popular design techniques for real-time controllers like optimal and robust control techniques require models of very moderate size.

Numerous approaches to bridge this gap are proposed in the literature [1, 2]. In many applications it is sufficient to approximate the high-order model by a low-order model that captures the essential state dynamics. To determine such low-order models one can use physical insight [3, 4, 5] and/or mathematical methods like proper orthogonal decomposition [6] or balanced truncation [1, 7]. In this paper we focus on the situation, where for the design of appropriate controllers it is sufficient to approximate the *input/output (I/O) map* of the system, schematically illustrated in Figure 1.

For such configurations, empirical or simulation-based black-box system identification [8, 9], and mathematical model reduction techniques like balanced truncation [10], moment matching [11] or recent variants of proper orthogonal decomposition [12] are common tools to extract appropriate low-order models. Typically, the bottleneck in these methods is the computational effort to compute the reduced order model from the semi-discretized model which often is of very high order.

In contrast to this, we present a new approach to construct low-order I/O maps (with error estimates) directly from the I/O map

$$\mathbb{G} : \mathcal{U} \to \mathcal{Y}, \qquad u = u(t, \theta) \mapsto y = y(t, \xi)$$

of *original* infinite-dimensional system. We suggest a new framework for the direct discretization of $\mathbb{G}$ for a general class of *infinite dimensional linear*

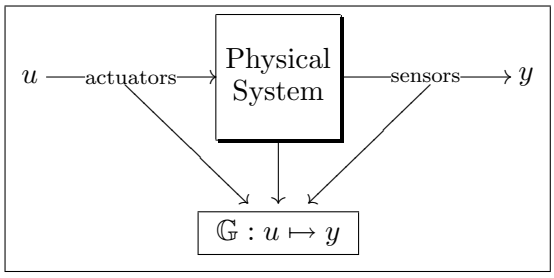


Figure 1: Schematic illustration of the I/O map corresponding to a physical system.

time-invariant state space systems (introduced in Section 2). Here u and y are input and output signals from Hilbert spaces $\mathcal{U}$ and $\mathcal{Y}$, respectively, which may vary in time t and space $\theta \in \Theta$ and $\xi \in \Xi$, with appropriate spatial domains Θ and Ξ. The framework consists of two steps.

1. *Approximation of signals (cf. Section 3).* We choose finite-dimensional subspaces $\bar{\mathcal{U}} \subset \mathcal{U}$ and $\bar{\mathcal{Y}} \subset \mathcal{Y}$ with bases $\{u_1, \dots, u_{\bar{p}}\} \subset \bar{\mathcal{U}}$ and $\{y_1, \dots, y_{\bar{q}}\} \subset \bar{\mathcal{Y}}$, and denote the corresponding orthogonal projections onto these subspaces by $\mathbb{P}_{\bar{\mathcal{U}}}$ and $\mathbb{P}_{\bar{\mathcal{Y}}}$, respectively. Then, the approximation
$$\mathbb{G}_S = \mathbb{P}_{\bar{\mathcal{Y}}} \mathbb{G} \mathbb{P}_{\bar{\mathcal{U}}}$$
has a matrix representation $\mathbf{G} \in \mathbb{R}^{\bar{q} \times \bar{p}}$.

2. *Approximation of the system dynamics (cf. Section 4).* Since $\mathbb{G}$ arises from a linear state space model, the components $\mathbf{G}_{ij} = (y_i, \mathbb{G} u_j)_{\mathcal{Y}}$ can be approximated by *numerically simulating* the state space model successively for inputs u_j, $j = 1, \dots, \bar{p}$ and by testing the resulting outputs against all $y_1, \dots, y_{\bar{q}}$.

We discuss several features of this framework.

Error estimation (cf. Section 5). The total error ϵ_{DS} of the approximation can be estimated by combining the *signal* approximation error ϵ_S and the *dynamical* approximation error ϵ_D, i.e.,
$$\underbrace{\|\mathbb{G} - \mathbb{G}_{DS}\|}_{=:\epsilon_{DS}} \leq \underbrace{\|\mathbb{G} - \mathbb{G}_S\|}_{=:\epsilon_S} + \underbrace{\|\mathbb{G}_S - \mathbb{G}_{DS}\|}_{=:\epsilon_D},$$
where the norms still have to be specified. Here $\mathbb{G}_{DS}$ denotes the numerically estimated approximation of $\mathbb{G}_S$. Theorem 13 shows how to choose $\bar{\mathcal{U}}$ and $\bar{\mathcal{Y}}$ and the accuracy tolerances for the numerical solutions of the underlying PDEs such that ϵ_S and ϵ_D are balanced and that $\epsilon_S + \epsilon_D < \texttt{tol}$ for a given tolerance `tol`. Choosing hierarchical bases in $\bar{\mathcal{U}}$ and $\bar{\mathcal{Y}}$, the error ϵ_S can be progressively reduced by adding further basis functions $u_{\bar{p}+1}, u_{\bar{p}+2}, \dots$ and $y_{\bar{q}+1}, y_{\bar{q}+2}, \dots$ resulting in additional columns and rows of the matrix representation.

Applications and examples in control design (cf. Section 6). We explicitly construct the error estimates for the control problem associated with a 2D heat equation. Furthermore, we show how the matrix representation $\mathbf{G} = [\mathbf{G}_{ij}]$ may directly be used in control design, or a state realization of the I/O model $\mathbb{G}_{DS}$ can be used as basis for many classical control design algorithms.

Notation

For $\Omega \subset \mathbb{R}^d$, $d \in \mathbb{N}$, $L^2(\Omega)$ denotes the usual Lebesgue space of square-integrable functions, and $H^\alpha(\Omega)$, $\alpha \in \mathbb{N}_0$ denotes the corresponding Sobolev spaces of α-times weakly differentiable functions. We interpret functions v, which vary in space and time, optionally as classical functions $v : [0,T] \times \Omega \to \mathbb{R}$ with values $v(t;x) \in \mathbb{R}$, or as *abstract* functions $v : [0,T] \to X$ with values in a function space X such as $X = H^\alpha(\Omega)$. Correspondingly, $H^\alpha(0,T;H^\beta(\Omega))$, with $\alpha, \beta \in \mathbb{N}_0$, denotes the space of equivalence classes of functions $v : [0,T] \to H^\beta(\Omega)$ with $t \mapsto \|v\|_{H^\beta(\Omega)}$ being α-times weakly differentiable [13]. We introduce Hilbert spaces [14]

$$\begin{aligned} H^{\alpha,\beta}((0,T)\times\Omega) &:= H^\alpha(0,T;L^2(\Omega)) \cap L^2(0,T;H^\beta(\Omega)), \\ \|v\|_{H^{\alpha,\beta}((0,T)\times\Omega)} &:= \|v\|_{H^\alpha(0,T;L^2(\Omega))} + \|v\|_{L^2(0,T;H^\beta(\Omega))}. \end{aligned}$$

By $C([0,T];X)$ and $C^\alpha([0,T];X)$ we denote the space of functions $v : [0,T] \to X$ which are continuous or α-times continuously differentiable. For two normed spaces X and Y, $\mathscr{L}(X,Y)$ denotes the set of bounded linear operators $X \to Y$, and we abbreviate $\mathscr{L}(X) := \mathscr{L}(X,X)$. For $\alpha \in \mathbb{N}$, $L^\alpha(0,T;\mathscr{L}(X,Y))$ denotes the space of operator-valued functions $K : [0,T] \to \mathscr{L}(X,Y)$ with $t \mapsto \|K(t)\|_{\mathscr{L}(X,Y)} = \sup_{x\neq 0}\|K(t)x\|_Y/\|x\|_X$ lying in $L^\alpha(0,T)$. Vectors, often representing a discretization of a function v, are written in corresponding small bold letters $\mathbf{v}$, whereas matrices, often representing a discrete version of an operator like $\mathbb{G}$ or G, are written in bold capital letters $\mathbf{G}$. By $\mathbb{R}^{\alpha\times\beta}$ we denote the set of real $\alpha \times \beta$ matrices, and $\mathbf{A} \otimes \mathbf{B}$ denotes the Kronecker product of matrices $\mathbf{A}$ and $\mathbf{B}$.

2 I/O maps of ∞-dimensional LTI state space systems

We consider infinite-dimensional, linear, time-invariant systems of first order

$$\partial_t z(t) = Az(t) + Bu(t), \quad t \in (0,T], \tag{1a}$$
$$z(0) = z^0, \tag{1b}$$
$$y(t) = Cz(t), \quad t \in [0,T]. \tag{1c}$$

Here for every time $t \in [0,T]$, the state $z(t)$ is supposed to belong to a Hilbert space Z like $Z = L^2(\Omega)$, where Ω is a subset of $\mathbb{R}^{d_\Omega}$ with $d_\Omega \in \mathbb{N}$. A is a

densely defined unbounded operator $A : Z \supset D(A) \to Z$, generating a C^0-semigroup $(S(t))_{t\geq 0}$ on Z. The control operator B belongs to $\mathscr{L}(U, Z)$ and the observation operator C to $\mathscr{L}(Z, Y)$, where $U = L^2(\Theta)$ and $Y = L^2(\Xi)$ with subsets $\Theta \subset \mathbb{R}^{d_1}$ and $\Xi \subset \mathbb{R}^{d_2}$, $d_1, d_2 \in \mathbb{N}$.

Let us recall how a linear bounded I/O-map $\mathbb{G} \in \mathscr{L}(\mathcal{U}, \mathcal{Y})$ with

$$\mathcal{U} = L^2(0, T; U) \quad \text{and} \quad \mathcal{Y} = L^2(0, T; Y)$$

can be associated[15] to (1). It is well-known that for initial values $z_0 \in D(A)$ and controls $u \in C^1([0, T]; Z)$, a unique *classical solution* $z \in C([0, T]; Z) \cap C^1((0, T); Z)$ of (1) exists. For $z_0 \in Z$ and $u \in \mathcal{U}$, the well-defined function

$$z(t) = S(t)z_0 + \int_0^t S(t - s)Bu(s)\, ds, \quad t \in [0, T], \tag{2}$$

is called a *mild solution* of (1). A mild solution of (1) is unique, belongs to $C([0, T]; Z)$, and is the uniform limit of classical solutions [15]. Hence, the output signal $y(t) = Cz(t)$ is well-defined and belongs to $\mathcal{Y} \cap C([0, T]; Y)$. In particular, the output signals $y(u) \in \mathcal{Y}$ arising from input signals $u \in \mathcal{U}$ and zero initial conditions $z_0 \equiv 0$ allow to define the linear I/O-map $\mathbb{G} : \mathcal{U} \to \mathcal{Y}$ of the system (1) by $u \mapsto y(u)$. It is possible to represent $\mathbb{G}$ as a convolution with the kernel function $K \in L^2(-T, T; \mathscr{L}(U, Y))$,

$$K(t) = \begin{cases} CS(t)B, & t \geq 0 \\ 0, & t < 0 \end{cases}.$$

Lemma 1. *The I/O-map $\mathbb{G}$ of (1) has the representation*

$$(\mathbb{G}u)(t) = \int_0^T K(t - s)u(s)\, ds, \quad t \in [0, T], \tag{3}$$

belongs to $\mathscr{L}(\mathcal{U}, \mathcal{Y}) \cap \mathscr{L}(\mathcal{U}, C([0, T], \mathcal{Y}))$, and satisfies

$$\|\mathbb{G}\|_{\mathscr{L}(\mathcal{U},\mathcal{Y})} \leq \sqrt{T}\|K\|_{L^2(0,T;\mathscr{L}(U,Y))}. \tag{4}$$

Proof. Since C is bounded, the representation of $y = Cz$ based on (2) can be reformulated as in (3), calling on the theory of Bochner integrals [13]. For general $K \in L^2(-T, T; \mathscr{L}(U, Y))$, using a generalized Hölder inequality implies that for fixed $t \in [0, T]$ the function $s \to K(t - s)u(s)$ belongs to $L^1(0, T; \mathscr{L}(U, Y))$ with

$$\|(\mathbb{G}u)(t)\|_Y \leq \|u\|_{\mathcal{U}}\|K(t - \cdot)\|_{L^2(0,T;\mathscr{L}(U,Y)},$$

and by integrating over $[0, T]$ we obtain (4). □

Remark 2. *The I/O-map $\mathbb{G}$ is* causal *in the sense that $y(t)$ only depends on $u_{|[0,t)}$ for all $t \in [0,T]$, and $\mathbb{G}$ is* time-invariant *in the sense that if $y = \mathbb{G}u$ then $\sigma_\tau y = \mathbb{G}(\sigma_\tau u)$ for all $\tau \in [0,T]$. Here σ_τ is a time-shift operator with $(\sigma_\tau u)(t) = u(t-\tau)$ for $t \in [\tau, T]$ and $(\sigma_\tau u)(t) = 0$ for $t \in [0,\tau)$.*

Example 3. *As prototypical system, we consider the heat equation with homogeneous Dirichlet boundary conditions and assume that Ω has a C^2-boundary. In this case, $Z = L^2(\Omega)$ and the operator A in (1) coincides with the Laplace operator*

$$A = \triangle : D(A) = H^2(\Omega) \cap H_0^1(\Omega) \subset Z \to Z.$$

Since A is the infinitesimal generator of an analytic C^0-semigroup of contractions $(S(t))_{t\geq 0}$, the mild solution z of (1) exhibits the following stability and regularity properties [15, 16].

(i) If $z_0 = 0$ and $u \in \mathcal{U}$, then $z \in H^{1,2}((0,T)\times\Omega)$ with

$$\|z\|_{H^{1,2}((0,T)\times\Omega)} \leq c\|u\|_{\mathcal{U}}. \tag{5}$$

(ii) Assume that $u \equiv 0$. For $z_0 \in D(A)$ we have $z \in C^1([0,T]; D(A))$, but for $z_0 \in Z$ we only have $z \in C^1((0,T]; D(A))$.

We will consider concrete choices of Ω, B and C in Section 6. We note that if the observation preserves the inherent state regularity in the sense that

$$C_{|H^2(\Omega)} \in \mathscr{L}(H^2(\Omega), H^2(\Xi)),$$

then $\mathbb{G} \in \mathscr{L}(\mathcal{U}, \mathcal{Y}_s)$ and also

$$\mathbb{G}_{|\mathcal{U}_s} \in \mathscr{L}(\mathcal{U}_s, \mathcal{Y}_s), \quad \text{with } \mathcal{U}_s = H^{1,2}((0,T)\times\Theta), \quad \mathcal{Y}_s = H^{1,2}((0,T)\times\Xi). \tag{6}$$

In fact, for $u \in \mathcal{U}_s$, we have $\|u\|_{\mathcal{U}} \leq \|u\|_{\mathcal{U}_s}$, and for $u \in \mathcal{U}$ we have

$$\|\mathbb{G}u\|_{\mathcal{Y}_s} \leq c'\|z\|_{H^{1,2}((0,T)\times\Omega)} \leq c\,c'\|u\|_{\mathcal{U}},$$

where $c' = \max\{\|C\|_{\mathscr{L}(L^2(\Omega),L^2(\Xi))}, \|C\|_{\mathscr{L}(H^2(\Omega),H^2(\Xi))}\}$ and c is the constant in (5).

Remark 4. *Many other linear time-invariant systems with distributed controls and observations admit a representation of the I/O map via (3) and exhibit properties similar to (6). This is, for instance, the case for the heat equation with homogeneous Neumann boundary conditions, and also*

for more general parabolic equations [14, 17]. Wave equations with second order time derivatives can be represented in the form of (1) *and* (3) *by means of an order reduction. Though hyperbolic systems do not have the smoothing properties of parabolic systems, they preserve the regularity of the data and results similar to* (6) *can be obtained by restricting the input signals to be of higher regularity in time [14].*

The presented framework can also be used for linearized flow systems. For the Stokes equation, results similar to (3) *and* (6) *are obtained by working with appropriate subspaces of divergence-free functions [18] and for the spatially discretized Oseen equations, which arise as linearizations of the Navier-Stokes equations, it has been shown in [19, 20] how the framework can be extended to linear time invariant descriptor systems.*

Note, however, that systems with boundary control or pointwise observations do not fit directly into the setting (1)*.*

3 Discretization of Signals

In order to discretize the input signals $u \in \mathcal{U}$ and $y \in \mathcal{Y}$ in space and time, we choose four families $\{U_{h_1}\}_{h_1>0}$, $\{Y_{h_2}\}_{h_2>0}$, $\{\mathcal{R}_{\tau_1}\}_{\tau_1>0}$ and $\{\mathcal{S}_{\tau_2}\}_{\tau_2>0}$ of subspaces $U_{h_1} \subset U$, $Y_{h_2} \subset Y$, $\mathcal{R}_{\tau_1} \subset L^2(0,T)$, and $\mathcal{S}_{\tau_2} \subset L^2(0,T)$ of finite dimensions $p(h_1) = \dim(U_{h_1})$, $q(h_2) = \dim(Y_{h_2})$, $r(\tau_1) = \dim(\mathcal{R}_{\tau_1})$ and $s(\tau_2) = \dim(\mathcal{S}_{\tau_2})$. We then define

$$\mathcal{U}_{h_1,\tau_1} = \{u \in \mathcal{U} : u(t;\cdot) \in U_{h_1},\, u(\cdot;\theta) \in \mathcal{R}_{\tau_1}, \quad t \in [0,T] \text{ a.e. }, \theta \in \Theta\},$$
$$\mathcal{Y}_{h_2,\tau_2} = \{y \in \mathcal{Y} : y(t;\cdot) \in Y_{h_2},\, y(\cdot;\xi) \in \mathcal{S}_{\tau_2}, \quad t \in [0,T] \text{ a.e. }, \xi \in \Xi\}.$$

We denote the orthogonal projections onto these subspaces by $P_{\mathcal{S},\tau_2} \in \mathscr{L}(L^2(0,T))$, $\mathbb{P}_{\mathcal{U},h_1,\tau_1} \in \mathscr{L}(\mathcal{U})$, and $\mathbb{P}_{\mathcal{Y},h_2,\tau_2} \in \mathscr{L}(\mathcal{Y})$. As first step of the approximation of $\mathbb{G}$, we define

$$\mathbb{G}_S = \mathbb{G}_S(h_1,\tau_1,h_2,\tau_2) = \mathbb{P}_{\mathcal{Y},h_2,\tau_2}\mathbb{G}\mathbb{P}_{\mathcal{U},h_1,\tau_1} \in \mathscr{L}(\mathcal{U},\mathcal{Y}).$$

In order to obtain a matrix representation of $\mathbb{G}_S$, we introduce families of bases $\{\mu_1,\dots,\mu_p\}$ of U_{h_1}, $\{\nu_1,\dots,\nu_q\}$ of Y_{h_2}, $\{\phi_1,\dots,\phi_r\}$ of $\mathcal{R}_{\tau_1}$, and $\{\psi_1,\dots,\psi_s\}$ of $\mathcal{S}_{\tau_2}$ and corresponding mass matrices $\mathbf{M}_{U,h_1} \in \mathbb{R}^{p\times p}$, $\mathbf{M}_{Y,h_2} \in \mathbb{R}^{q\times q}$, $\mathbf{M}_{\mathcal{R},\tau_1} \in \mathbb{R}^{r\times r}$ and $\mathbf{M}_{\mathcal{S},\tau_2} \in \mathbb{R}^{s\times s}$, for instance via

$$[\mathbf{M}_{U,h_1}]_{ij} = (\mu_j,\mu_i)_U, \qquad i,j = 1,\dots,p.$$

These mass matrices induce, via

$$(\mathbf{v},\mathbf{w})_{\mathbb{R}^p;w} = \mathbf{v}^T\mathbf{M}_{U,h_1}\mathbf{w} \quad \text{for all } \mathbf{v},\mathbf{w} \in \mathbb{R}^p,$$

weighted scalar products and corresponding norms in the respective spaces, which we indicate by a subscript w, like $\mathbb{R}^p_w$ with $(\cdot,\cdot)_{\mathbb{R}^p;w}$ and $\|\cdot\|_{\mathbb{R}^p;w}$, in contrast to the canonical spaces like $\mathbb{R}^p$, with $(\cdot,\cdot)_{\mathbb{R}^p}$ and $\|\cdot\|_{\mathbb{R}^p}$. We represent signals $u \in \mathcal{U}_{h_1,\tau_1}$ and $y \in \mathcal{Y}_{h_2,\tau_2}$ as

$$u(t;\theta) = \sum_{k=1}^{p}\sum_{i=1}^{r} \mathbf{u}_i^k \phi_i(t)\mu_k(\theta), \qquad y(t;\xi) = \sum_{l=1}^{q}\sum_{j=1}^{s} \mathbf{y}_j^l \psi_j(t)\nu_k(\xi),$$

where $\mathbf{u}_i^k$ are the elements of a block-structured vector $\mathbf{u} \in \mathbb{R}^{pr}$ with p blocks $\mathbf{u}^k \in \mathbb{R}^r$, and the vector $\mathbf{y} \in \mathbb{R}^{qs}$ is defined similarly. Then

$$\|u\|_{\mathcal{U}} = \|\mathbf{u}\|_{\mathbb{R}^{pr};w}, \quad \text{and} \quad \|y\|_{\mathcal{Y}} = \|\mathbf{y}\|_{\mathbb{R}^{qs};w},$$

where $\|\cdot\|_{\mathbb{R}^{pr};w}$ and $\|\cdot\|_{\mathbb{R}^{qs};w}$ denote the weighted norms with respect to the mass matrices

$$\mathbf{M}_{\mathcal{U},h_1,\tau_1} = \mathbf{M}_{U,h_1} \otimes \mathbf{M}_{\mathcal{R},\tau_1} \in \mathbb{R}^{pr\times pr}, \quad \mathbf{M}_{\mathcal{Y},h_2,\tau_2} = \mathbf{M}_{Y,h_2} \otimes \mathbf{M}_{\mathcal{S},\tau_2} \in \mathbb{R}^{qs\times qs},$$

i.e., the corresponding coordinate isomorphisms $\kappa_{\mathcal{U},h_1,\tau_1} \in \mathscr{L}(\mathcal{U}_{h_1,\tau_1}, \mathbb{R}^{pr}_w)$ and $\kappa_{\mathcal{Y},h_2,\tau_2} \in \mathscr{L}(\mathcal{Y}_{h_2,\tau_2}, \mathbb{R}^{qs}_w)$ are unitary.

Finally, we obtain a matrix representation $\mathbf{G}$ of $\mathbb{G}_S$ by setting

$$\mathbf{G} = \mathbf{G}(h_1,\tau_1,h_2,\tau_2) = \kappa_{\mathcal{Y}}\mathbb{P}_{\mathcal{Y}}\mathbb{G}\mathbb{P}_{\mathcal{U}}\kappa_{\mathcal{U}}^{-1} \in \mathbb{R}^{qs\times pr}, \tag{7}$$

where the dependencies on h_1,τ_1,h_2,τ_2 have been partially omitted. Considering

$$\mathbf{H} = \mathbf{H}(h_1,\tau_1,h_2,\tau_2) := \mathbf{M}_{\mathcal{Y},h_2,\tau_2}\mathbf{G} \in \mathbb{R}^{qs\times pr}$$

as a block-structured matrix with $q\times p$ blocks $\mathbf{H}^{kl} \in \mathbb{R}^{s\times r}$ and block elements $\mathbf{H}^{kl}_{ij} \in \mathbb{R}$, we obtain the representation

$$\mathbf{H}^{kl}_{ij} = [\mathbf{M}_{\mathcal{Y}}\kappa_{\mathcal{Y}}\mathbb{P}_{\mathcal{Y}}\mathbb{G}(\mu_l\phi_j)]^k_i = (\nu_k\psi_i, \mathbb{G}(\mu_l\phi_j))_{\mathcal{Y}}. \tag{8}$$

To have a discrete analogon of the $\mathscr{L}(\mathcal{U},\mathcal{Y})$-norm, for given $\mathcal{U}_{h_1,\tau_1}$ and $\mathcal{Y}_{h_2,\tau_2}$, we introduce the weighted matrix norm

$$\begin{aligned}\|\mathbf{G}(h_1,\tau_1,h_2,\tau_2)\|_{\mathbb{R}^{qs}\times\mathbb{R}^{pr};w} &:= \sup_{\mathbf{u}\in\mathbb{R}^{pr}} \frac{\|\mathbf{G}\mathbf{u}\|_{\mathbb{R}^{qs};w}}{\|\mathbf{u}\|_{\mathbb{R}^{pr};w}} \\ &= \|\mathbf{M}^{1/2}_{\mathcal{Y},h_2,\tau_2}\mathbf{G}\mathbf{M}^{-1/2}_{\mathcal{U},h_1,\tau_1}\|_{\mathbb{R}^{qs}\times\mathbb{R}^{pr}},\end{aligned}$$

and we write $(h_1',\tau_1',h_2',\tau_2') \le (h_1,\tau_1,h_2,\tau_2)$ if the inequality holds componentwise.

Lemma 5. *For all* $(h_1, \tau_1, h_2, \tau_2) \in \mathbb{R}^4_+$, *we have*

$$\|\mathbf{G}(h_1, \tau_1, h_2, \tau_2)\|_{\mathbb{R}^{qs}\times\mathbb{R}^{pr};w} = \|\mathbb{G}_S(h_1, \tau_1, h_2, \tau_2)\|_{\mathscr{L}(\mathcal{U},\mathcal{Y})} \le \|\mathbb{G}\|_{\mathscr{L}(\mathcal{U},\mathcal{Y})}. \quad (9)$$

If the subspaces $\{\mathcal{U}_{h_1,\tau_1}\}_{h_1,\tau_1>0}$ *and* $\{\mathcal{Y}_{h_2,\tau_2}\}_{h_2,\tau_2>0}$ *are nested, in the sense that*

$$\mathcal{U}_{h_1,\tau_1} \subset \mathcal{U}_{h_1',\tau_1'}, \quad \mathcal{Y}_{h_2,\tau_2} \subset \mathcal{Y}_{h_2',\tau_2'} \quad \text{for } (h_1', \tau_1', h_2', \tau_2') \le (h_1, \tau_1, h_2, \tau_2), \quad (10)$$

then $\|\mathbf{G}(h_1, \tau_1, h_2, \tau_2)\|_{\mathbb{R}^{qs}\times\mathbb{R}^{pr};w}$ *is monotonically increasing for decreasing* $(h_1, \tau_1, h_2, \tau_2) \in \mathbb{R}^4_+$, *and* $\|\mathbf{G}(h_1, \tau_1, h_2, \tau_2)\|_{\mathbb{R}^{qs}\times\mathbb{R}^{pr};w}$ *is convergent for* $(h_1, \tau_1, h_2, \tau_2) \searrow 0$.

Proof. In order to show (9), we calculate

$$\|\mathbb{G}_S\|\mathscr{L}(\mathcal{U},\mathcal{Y}) = \sup_{u\in\mathcal{U}_{h_1,\tau_1}} \frac{\|\mathbb{P}_{\mathcal{Y},h_2,\tau_2}\mathbb{G}u\|_{\mathcal{Y}}}{\|u\|_{\mathcal{U}}} \le \sup_{u\in\mathcal{U}_{h_1,\tau_1}} \frac{\|\mathbb{G}u\|_{\mathcal{Y}}}{\|u\|_{\mathcal{U}}} \le \|\mathbb{G}\|_{\mathscr{L}(\mathcal{U},\mathcal{Y})},$$

and observe that for $u \in \mathcal{U}_{h_1,\tau_1}$ and $\mathbf{u} = \kappa_{\mathcal{U},h_1,\tau_1}u \in \mathbb{R}^{pr}$, we have

$$\|\mathbb{G}_S u\|_{\mathcal{Y}} = \|\mathbf{G}\mathbf{u}\|_{\mathbb{R}^{qs};w} \le \|\mathbf{G}\|_{\mathbb{R}^{qs}\times\mathbb{R}^{pr};w}\|u\|_{\mathcal{U}} \quad \text{and}$$
$$\|\mathbb{G}_S u\|_{\mathcal{Y}} \le \|\mathbb{G}_S\|_{\mathscr{L}(\mathcal{U},\mathcal{Y}}\|\mathbf{u}\|_{\mathbb{R}^{pr};w}.$$

If (10) holds, then since $\|\mathbb{P}_{\mathcal{Y},h_2,\tau_2}y\|_{\mathcal{Y}} \le \|\mathbb{P}_{\mathcal{Y},h_2',\tau_2'}y\|_{\mathcal{Y}}$ for all $y \in \mathcal{Y}$, we have

$$\begin{aligned}\|\mathbb{G}_S(h_1, \tau_1, h_2, \tau_2)\|_{\mathbb{R}^{qs}\times\mathbb{R}^{pr};w} &\le \sup_{u\in\mathcal{U}_{h_1',\tau_1'}} \frac{\|\mathbb{P}_{\mathcal{Y},h_2',\tau_2'}\mathbb{G}u\|_{\mathcal{Y}}}{\|u\|_{\mathcal{U}}} \\ &= \|\mathbb{G}_S(h_1', \tau_1', h_2', \tau_2')\|_{\mathbb{R}^{q's'}\times\mathbb{R}^{p'r'};w}.\end{aligned}$$

Hence, (9) ensures the convergence of $\|\mathbb{G}_S(\mathbf{h})\|_{\mathbb{R}^{qs}\times\mathbb{R}^{pr};w}$. □

3.1 Signal discretization via finite elements

There are many possibilities to choose the finite dimensional subspaces in U, Y. As an example, consider the case $U = Y = L^2(0, 1)$, choose U_{h_1} and Y_{h_2} as spaces of continuous, piecewise linear functions and $\mathcal{R}_{\tau_1}$ and $\mathcal{S}_{\tau_2}$ as spaces of piecewise constant functions, all with respect to equidistant grids.

For $p \in \mathbb{N}$, $p \ge 2$ and $h_1(p) = 1/(p-1)$, let $\mathscr{T}_{h_1} = \{I_k\}_{1\le k\le p-1}$ be the equidistant partition of $(0, 1]$ into intervals $I_k = ((k-1)h_1, kh_1]$. The corresponding space U_{h_1} is spanned by the nodal basis

$$\{\mu_1^{(h_1)}, \ldots, \mu_{p(h_1)}^{(h_1)}\} \subset U_{h_1}, \quad \text{with } \mu_l^{(h_1)}(kh_1) = \delta_{l-1}(k), \quad k = 0, \ldots, p.$$

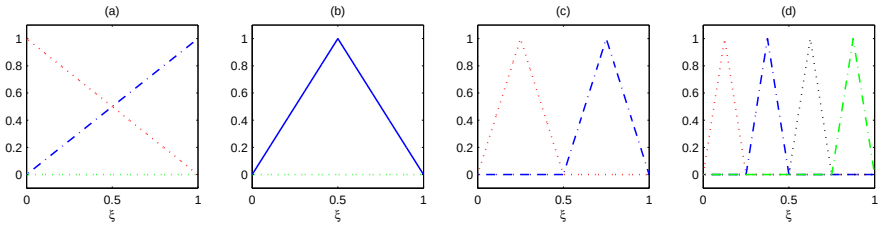


Figure 2: Hierarchical basis for $L^2(0,1)$-subspaces of piecewise linear functions: (a) μ_1 and μ_2 (b) μ_3 (c) μ_4 and μ_5 (d) $\mu_6, \ldots, \mu_9$.

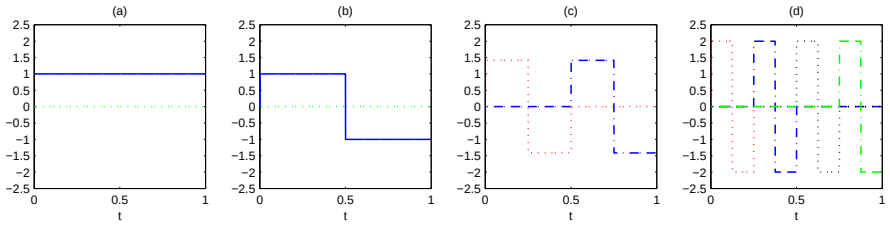


Figure 3: Haar wavelet basis for $L^2(0,1)$-subspaces of piecewise constant functions: (a) ϕ_1 (b) ϕ_2 (c) ϕ_3 and ϕ_4 (d) $\phi_5, \ldots, \phi_8$.

The subspaces $\{U_{h_1}\}$ are nested if the choice is restricted to $h_1 \in \{2^{-n}\}_{n\in\mathbb{N}_0}$ and $p \in \{2^n + 1\}_{n\in\mathbb{N}_0}$. Since the *nodal* bases of U_{h_1} and $U_{h_1'}$ do not have any common element for $h_1 \neq h_1'$, one may prefer to choose a *hierarchical* basis of finite element functions[21, 22] $\hat{\mu}_l$, as in Fig. 2. Then, $U_{h_1} = \text{span}\{\hat{\mu}_1, \ldots, \hat{\mu}_{p(h_1)}\}$ for all $h_1 \in \{2^{-n}\}_{n\in\mathbb{N}_0}$ with basis functions $\hat{\mu}_k$ independent of h_1. For $r \in \mathbb{N}$ and $\tau_1 = T/r$, let $\Gamma_{\tau_1} = \{I_j\}_{1\leq j\leq r}$ be the equidistant partition of $(0,T]$ into intervals $I_j = ((j-1)\tau_1, j\tau_1]$. The corresponding space $\mathcal{R}_{\tau_1}$ of piecewise constant functions is, for instance, spanned by the nodal and orthogonal basis

$$\{\phi_1^{(\tau_1)}(t), \ldots, \phi_r^{(\tau_1)}(t)\}, \quad \text{with } \phi_j^{(\tau_1)}(t) = \chi_{I_j}(t), \qquad j = 1, \ldots, r, \tag{11}$$

with χ_{I_j} denoting the characteristic function on I_j. The spaces are nested by requiring $\tau_1 \in \{2^{-n}T\}_{n\in\mathbb{N}_0}$. An orthonormal hierarchical basis for $\mathcal{R}_{\tau_1}$ is obtained by choosing ϕ_j as Haar-wavelets, cf. Fig. 3 and [23].

Denoting the orthogonal projections onto U_{h_1} and $\mathcal{R}_{\tau_1}$ by P_{U,h_1} and $P_{\mathcal{R},\tau_1}$, respectively, the Poincaré-Friedrich inequality shows that there exist constants $c_U = 1/2$ and $c_{\mathcal{R}} = 1/\sqrt{2}$, independent of h_1, τ_1 and T, such

that[24, 25]

$$\|u - P_{U_{h_1}} u\|_{L^2(0,1)} \le c_U h_1^2 \|\partial_\xi^2 u\|_{L^2(0,1)} \quad \text{for } u \in H^2(0,1),$$
$$\|v - P_{\mathcal{R}_{\tau_1}} v\|_{L^2(0,T)} \le c_{\mathcal{R}} \tau_1 \|\partial_t v\|_{L^2(0,T)} \quad \text{for } v \in H^1(0,T).$$

By the Fubini theorem, it follows that the corresponding projection $\mathbb{P}_{\mathcal{U},h_1,\tau_1}$ onto $\mathcal{U}_{h_1,\tau_1} = \{u \in \mathcal{U},\, u_{|I_j} \equiv u^{(j)},\, u^{(j)} \in U_{h_1}, \quad j = 1,\dots,r\}$ satisfies

$$\|u - \mathbb{P}_{\mathcal{U},h_1,\tau_1} u\|_{\mathcal{U}} \le (c_U h_1^2 + c_{\mathcal{R}} \tau_1) \|u\|_{\mathcal{U}_s} \text{ for all } u \in \mathcal{U}_s = H^{1,2}((0,T)\times(0,1)). \tag{13}$$

We define Y_{h_2},$\mathcal{R}_{\tau_2}$ and $\mathcal{Y}_{h_2,\tau_2}$ accordingly and a corresponding estimate as (13) holds for the projection $\mathbb{P}_{\mathcal{Y},h_2,\tau_2} y$ of elements $y \in \mathcal{Y}_s$.

Remark 6. *Estimates similar to* (13) *also exist for domains* $\Theta \subset \mathbb{R}^d$ *with* $d > 1$ *and are classical results from the interpolation theory in Sobolev spaces [24]. Note that the interpolation constants then often have to be estimated numerically. Estimates with higher approximation order can be obtained, if ansatz functions of higher polynomial degree are used and if the input and output signals exhibit corresponding higher regularity in space and time.*

4 Approximation of the system dynamics

Let us now discuss the efficient approximation of $\mathbb{G}_S$ and its matrix representation $\mathbf{G} = \mathbf{M}_{\mathcal{Y}}^{-1}\mathbf{H}$, respectively. For time-invariant systems with distributed control and observation, this task reduces to the approximation of the convolution kernel $K \in L^2(0,T;\mathscr{L}(U,Y))$.

4.1 Kernel function approximation

Inserting (3) in (8), by a change of variables we obtain

$$\mathbf{H}_{ij}^{kl} = \int_0^T \int_0^T \psi_i(t)\phi_j(s)(\nu_k, K(t-s)\mu_l)_Y \, ds\, dt = \int_0^T \mathbf{W}_{ij}(t)\mathbf{K}_{kl}(t)\, dt,$$

with matrix-valued functions $\mathbf{W} : [0,T] \to \mathbb{R}^{s\times r}$ and $\mathbf{K} : [0,T] \to \mathbb{R}^{q\times p}$,

$$\mathbf{W}_{ij}(t) = \int_0^{T-t} \psi_i(t+s)\phi_j(s)\, ds, \qquad \mathbf{K}_{kl}(t) = (\nu_k, K(t)\mu_l)_Y,$$

and thus

$$\mathbf{H} = \mathbf{M}_{\mathcal{Y}}\mathbf{G} = \int_0^T \mathbf{K}(t) \otimes \mathbf{W}(t)\, dt. \tag{15}$$

Remark 7. $\mathbf{W}(t)$ *can be exactly calculated if piecewise polynomial ansatz functions* $\psi_i(t)$ *and* $\phi_j(t)$ *are chosen. For the special choice* (11)*, we see in this way that* $\mathbf{W}(t) \in \mathbb{R}^{r\times r}$ *is a lower triangular Toeplitz matrix for all* $t \in [0,T]$*, and hence the matrices* $\mathbf{H}_{ij} = \int_0^T \mathbf{W}_{ij}(t)\mathbf{K}(t)\,dt \in \mathbb{R}^{q\times p}$ *satisfy* $\mathbf{H}_{ij} = \mathbf{H}_{i-j}$ *for* $1 \le i,j \le r$ *and* $\mathbf{H}_{ij} = 0$ *for* $1 \le i < j \le r$*.*

For systems of the form (1), the matrix-valued function $\mathbf{K}$ is given by

$$\mathbf{K}_{kl}(t) = (\nu_k, CS(t)B\mu_l)_Y = (c_k^*, S(t)b_l)_Z,$$

where $c_k^* = C^*\nu_k \in Z$ and $b_l = B\mu_l$ for $k = 1,\dots,q$ and $l = 1,\dots,p$. Hence, the entries of $\mathbf{K}$ can be calculated by solving the homogeneous systems

$$\dot{z}_l(t) = Az_l(t), \qquad t \in (0,T], \tag{16a}$$

$$z_l(0) = b_l, \qquad l = 1,\dots,p, \tag{16b}$$

since (16) has the mild solution $z_l(t) = S(t)b_l \in C([0,T];L^2(\Omega))$. We obtain an approximation $\tilde{\mathbf{H}}$ of $\mathbf{H}$ by replacing $z_l(t)$ by numerical approximations $z_{l,\mathtt{tol}}(t)$, i.e.,

$$\tilde{\mathbf{H}} = \int_0^T \tilde{\mathbf{K}}(t) \otimes \mathbf{W}(t)\,dt, \tag{17}$$

with $\tilde{\mathbf{K}}_{kl}(t) = (\nu_k, Cz_{l,\mathtt{tol}}(t))_Y = (c_k^*, z_{l,\mathtt{tol}}(t))_Z$. Here the subscript `tol` indicates that the error $z_l - z_{l,\mathtt{tol}}$ is assumed to satisfy some tolerance criterion which will be specified later. The corresponding approximation $\mathbb{G}_{DS}$ of $\mathbb{G}_S$ is given by

$$\mathbb{G}_{DS} = \kappa_{\mathcal{Y}}^{-1}\tilde{\mathbf{G}}\kappa_{\mathcal{U}}\mathbb{P}_{\mathcal{U}}, \quad \text{with } \tilde{\mathbf{G}} = \mathbf{M}_{\mathcal{Y}}^{-1}\tilde{\mathbf{H}}, \tag{18}$$

and depends on h_1, h_2, τ_1, τ_2 and `tol`.

Remark 8. *The matrix function* $\mathbf{K}$ *is approximated* columnwise*. The kernel may also be calculated* rowwise *by solving an adjoint autonomous system, which may be preferable if* $q < p$ *or if the output approximation is successively improved by adding further basis functions* $\nu_{q+1}, \nu_{q+2}, \dots$*.*

Remark 9. *The calculation of* $\tilde{\mathbf{H}}$ *can be parallelized in an obvious way by calculating the* p *solutions* $z_{l,\mathtt{tol}}$ *in parallel and we note that no state trajectories have to be stored. In general, the matrix* $\tilde{\mathbf{H}}$ *is not sparse, such that the memory requirements become significant if a high resolution of the signals in space and time is required, and the question of a data-sparse representation arises. Recalling Remark 7, the blocks* $\tilde{\mathbf{H}}^{kl}$ *are lower triangular Toeplitz matrices for the special choice of time basis funtions* (11) *and thus only* $q \cdot p \cdot r$ *elements have to be stored. Another approach to obtain data-sparse representations uses approximate factorizations* $\check{\mathbf{K}}_{kl}(t-s) = \sum_{m,n=1}^{M} \alpha_{mn}L_m(t)L_n(s)$ *for* $s,t \in [0,T]$ *with suitable ansatz functions[26]* $L_n(t)$ *.*

4.2 The approximation error for the dynamics

The following proposition relates the error ϵ_D in the system dynamics to the errors made in solving the PDE (16) for $l = 1, \dots, p$.

Proposition 10. *The error $\epsilon_D := \|\mathbb{G}_S - \mathbb{G}_{DS}\|_{\mathscr{L}(\mathcal{U},\mathcal{Y})}$ in the system dynamics satisfies*

$$\begin{aligned}\epsilon_D \leq & \sqrt{T}\|\mathbf{K} - \tilde{\mathbf{K}}\|_{L^2(0,T;\mathbb{R}_w^{q\times p})} \\ & \leq p\sqrt{T}\sqrt{\frac{\lambda_{\max}(\mathbf{M}_{Y,h_2})}{\lambda_{\min}(\mathbf{M}_{U,h_1})}}\max_{1\leq l\leq p}\|\mathbf{K}_{:,l} - \tilde{\mathbf{K}}_{:,l}\|_{L^2(0,T;\mathbb{R}^q)}. \end{aligned} \tag{19}$$

Here $\mathbf{K}_{:,l}$ and $\tilde{\mathbf{K}}_{:,l}$ denote the l-th column of $\mathbf{K}(t)$ and $\tilde{\mathbf{K}}(t)$, respectively, $\lambda_{\max}(\mathbf{M}_{Y,h_2})$ is the largest eigenvalue of $\mathbf{M}_{Y,h_2}$ and $\lambda_{\min}(\mathbf{M}_{U,h_1})$ the smallest eigenvalue of $\mathbf{M}_{U,h_1}$. Similar as before, $\mathbb{R}_w^{q\times p}$ denotes the space of real $q \times p$ matrices equipped with the weighted matrix norm $\|\mathbf{M}\|_{\mathbb{R}^{q\times p};w} = \sup_{\mathbf{u}\neq 0}\|\mathbf{M}\mathbf{u}\|_{\mathbb{R}^q;w}/\|\mathbf{u}\|_{\mathbb{R}^p;w}$.

Proof. The matrix $\mathbf{K}$ is the representation of the space-projected kernel function $K_m : [-T, T] \to \mathscr{L}(U, Y)$ with $K_m(t) = P_{Y,h_2}K(t)P_{U,h_1}$, where P_{Y,h_2} and P_{U,h_1} are the orthogonal projections onto the subspaces Y_{h_2} and U_{h_1}, respectively. Introducing the corresponding I/O-map $\mathbb{G}_m = \mathbb{G}_m(h_1, h_2)$,

$$(\mathbb{G}_m u)(t) = \int_0^T K_m(t-s)u(s)\,ds, \qquad t \in [0, T]. \tag{20}$$

we note that $\mathbb{G}_S = \mathbb{P}_{\mathcal{Y},h_2,\tau_2}\mathbb{G}_m\mathbb{P}_{\mathcal{U},h_1,\tau_1}$. Similarly, we associate with $\tilde{\mathbf{K}}(t)$ the kernel function $\tilde{K} : [-T, T] \to \mathscr{L}(U, Y)$ with $\tilde{K}(t) = \kappa_{Y,h_2}^{-1}\tilde{\mathbf{K}}(t)\kappa_{U,h_1}P_{U,h_1}$, and with corresponding I/O-map

$$(\mathbb{G}_D u)(t) = \int_0^T \tilde{K}(t-s)u(s)\,ds, \qquad t \in [0, T].$$

We observe that $\mathbb{G}_{DS}$ as defined in (18) satisfies $\mathbb{G}_{DS} = \mathbb{P}_{\mathcal{Y},h_2,\tau_2}\mathbb{G}_D\mathbb{P}_{\mathcal{U},h_1,\tau_1}$ by showing via (7)-(15) that the matrix representation of $\mathbb{P}_{\mathcal{Y},h_2,\tau_2}\mathbb{G}_D\mathbb{P}_{\mathcal{U},h_1,\tau_1}$ coincides with (17). Then $\|K_m(t)\|_{\mathscr{L}(U,Y)} = \|\mathbf{K}(t)\|_{\mathbb{R}^{q\times p};w}$ and $\|\tilde{K}(t)\|_{\mathscr{L}(U,Y)} = \|\tilde{\mathbf{K}}(t)\|_{\mathbb{R}^{q\times p};w}$ for all $t \in [0, T]$. Lemma 1 yields

$$\|\mathbb{G}_m - \mathbb{G}_D\|_{\mathscr{L}(\mathcal{U},\mathcal{Y})} \leq \sqrt{T}\|K_m - \tilde{K}\|_{L^2(0,T;\mathscr{L}(U,Y))} = \sqrt{T}\|\mathbf{K} - \tilde{\mathbf{K}}\|_{L^2(0,T;\mathbb{R}_w^{q\times p})}.$$

Defining $\mathbf{E}(t) = \mathbf{K}(t) - \tilde{\mathbf{K}}(t)$, for $\mathbf{u} \in \mathbb{R}^p$ with $\|\mathbf{u}\|_{\mathbb{R}^p} = 1$ and $t \in [0, T]$, by using the equivalence vector norms in $\mathbb{R}^p$ we have that

$$\|\mathbf{E}(t)\mathbf{u}\|_{\mathbb{R}^q} \leq \sum_{l=1}^p |\mathbf{u}_l|\|\mathbf{E}_{:,l}(t)\|_{\mathbb{R}^q} \leq \sqrt{p}\left(\sum_{l=1}^p \|\mathbf{E}_{:,l}(t)\|_{\mathbb{R}^q}^2\right)^{1/2},$$

and hence

$$\|\mathbf{E}\|^2_{L^2(0,T;\mathbb{R}^{q\times p})} \leq p\sum_{l=1}^{p}\int_0^T \|\mathbf{E}_{:,l}(t)\|^2_{\mathbb{R}^q}\,dt \leq p^2 \max_{l=1,\dots,p}\int_0^T \|\mathbf{E}_{:,l}(t)\|^2_{\mathbb{R}^q}\,dt,$$

which concludes the proof. □

Remark 11. *Calculating the columns of* $\mathbf{K}$ *directly and estimating* ϵ_D *via* (19)*, the quotient of the eigenvalues of the mass matrices* $\mathbf{M}_{U,h_1}$ *and* $\mathbf{M}_{Y,h_2}$ *has to be compensated by the approximation accuracy of* $\mathbf{K}_{:,l}$*. This may be problematic if hierarchical basis functions are chosen, since the quotient grows unboundedly with decreasing* h_1 *and* h_2*. One may circumvent this problem by calculating* $\mathbf{K}$ *with respect to different bases. Approximating the columns of* $\mathbf{K}^w(t) = \mathbf{M}_{\mathcal{Y}}^{1/2}\mathbf{K}(t)\mathbf{M}_{\mathcal{U}}^{-1/2}$ *via an adapted problem* (16)*, we have* $\epsilon_D \leq p\sqrt{T}\max_{1\leq l\leq p}\|\mathbf{K}^w_{:,l} - \tilde{\mathbf{K}}^w_{:,l}\|_{L^2(0,T;\mathbb{R}^q)}$*. Note that the necessary back transformations have to be carried out with sufficient accuracy.*

4.3 Error estimation for the homogeneous PDE

In order to approximate the system dynamics, the homogeneous PDE (16) has to be solved via a fully-discrete numerical scheme for p different initial values. A *first* goal in error control is to choose the time and space grids (and possibly other discretization parameters) such that

$$\|\mathbf{K}_{:,l} - \tilde{\mathbf{K}}_{:,l}\|_{L^2(0,T;\mathbb{R}^q)} < \texttt{tol} \quad \text{resp.} \quad \|\mathbf{K}^w_{:,l} - \tilde{\mathbf{K}}^w_{:,l}\|_{L^2(0,T;\mathbb{R}^q)} < \texttt{tol} \tag{21}$$

is *guaranteed* for a given $\texttt{tol} > 0$ by means of reliable error estimators. A *second* goal is to achieve this accuracy in a *cost-economic* way. A special difficulty in solving (16) numerically is the handling of initial values b_l, which, in general, only belong to Z but not necessarily to $D(A)$. Considering the example heat equation, this means that the space and time derivatives of the exact solution $z_l \in C^1((0,T], H^2(\Omega)\cap H^1_0(\Omega))$ may become very large for small t, but decay quickly for $t > 0$. In fact, in general we only have the analytic bound

$$\|\partial_t z(t)\|_{L^2(\Omega)} = \|\triangle z(t)\|_{L^2(\Omega)} \leq \frac{c}{t}\|z^0\|_{L^2(\Omega)} \quad \text{for all } t \in (0,T],$$

with some constant $c > 0$ independent of z_0 and T, cf. [27, p. 148]. Adaptive space and time discretizations on the basis of a posteriori error estimates are the method of choice to deal with these difficulties [28]. Discontinuous Galerkin time discretizations in combination with standard Galerkin space

discretizations provide an appropriate framework to derive corresponding (a priori and a posteriori) error estimates, also for the case of adaptively refined grids which are in general no longer quasi-uniform [29, 27, 30]. We distinguish two types of error estimates.

Global state error estimates measure the error $(z_l - z_{l,\mathtt{tol}})$ in some global norm. For parabolic problems, a priori and a posteriori estimates for the error in $L^\infty(0,T;L^2(\Omega))$ and $L^\infty(0,T;L^\infty(\Omega))$ can be found in [29]. Such results permit to guarantee (21) in view of

$$\|\mathbf{K}_{:,l} - \tilde{\mathbf{K}}_{:,l}\|_{L^2(0,T;\mathbb{R}^q)} \leq \|C\|_{\mathscr{L}(Z,Y)} \left(\sum_{i=1}^{q} \|\nu_i\|_Y^2 \right)^{1/2} \|z - z_{\mathtt{tol}}^{(l)}\|_{L^2(0,T;Z)}. \quad (22)$$

Goal-oriented error estimates can be used to measure the error $\|\mathbf{K}_{:,l} - \tilde{\mathbf{K}}_{:,l}\|_{L^2(0,T;\mathbb{R}^q)}$ directly. This may be advantageous, since (22) may be very conservative: the error in the *observations* $\mathbf{K}_{:,l}$ can be small even if some norm of the *state* error is large. The core of these error estimation techniques is an exact error representation formula, which can be evaluated if one knows the residual and the solution of an auxiliary dual PDE. This leads to the *dual-weighted residuals* (DWR) approach, see e.g. [31, 32, 33, 34, 35, 36, 37, 38, 39] and the references therein.

The previous discussion justifies the following assumption.

Assumption 12. *Given a tolerance* $\mathtt{tol} > 0$*, we can ensure (by using appropriate error estimators and mesh refinements) that the solutions* z_l *of* (16) *and the solutions* $z_{l,\mathtt{tol}}$ *calculated by means of an appropriate fully-discrete numerical scheme satisfy*

$$\|\mathbf{K}_{:,l} - \tilde{\mathbf{K}}_{:,l}\|_{L^2(0,T;\mathbb{R}^q)} < \mathtt{tol}, \qquad l = 1, \dots, p. \quad (23)$$

5 Total Error Estimates

We present estimates for the total error in the approximation of $\mathbb{G}$. Using general-purpose ansatz spaces $\mathcal{U}_{h_1,\tau_1}$ and $\mathcal{Y}_{h_2,\tau_2}$ for the signal approximation, we only obtain error results in a weaker $\mathscr{L}(\mathcal{U}_s, \mathcal{Y})$-norm.

Theorem 13. *Consider the I/O map* $\mathbb{G} \in \mathscr{L}(\mathcal{U}, \mathcal{Y})$ *of the infinite-dimensional linear time-invariant system* (3) *and assume that*

(i) $\mathbb{G}_{|\mathcal{U}_s} \in \mathscr{L}(\mathcal{U}_s, \mathcal{Y}_s)$ *with spaces of higher regularity in space and time*

$$\mathcal{U}_s = H^{\alpha_1,\beta_1}((0,T)\times\Theta), \quad \mathcal{Y}_s = H^{\alpha_2,\beta_2}((0,T)\times\Xi), \quad \alpha_1, \beta_1, \alpha_2, \beta_2 \in \mathbb{N}.$$

(ii) *The families of subspaces* $\{\mathcal{U}_{h_1,\tau_1}\}_{h_1,\tau_1}$ *and* $\{\mathcal{Y}_{h_2,\tau_2}\}_{h_2,\tau_2}$ *satisfy*

$$\begin{aligned}\|u - \mathbb{P}_{\mathcal{U},h_1,\tau_1} u\|_{\mathcal{U}} &\leq (c_{\mathcal{R}}\tau_1^{\alpha_1} + c_U h_1^{\beta_1})\|u\|_{\mathcal{U}_s}, \qquad u \in \mathcal{U}_s,\\ \|y - \mathbb{P}_{\mathcal{Y},h_2,\tau_2} y\|_{\mathcal{Y}} &\leq (c_{\mathcal{S}}\tau_2^{\alpha_2} + c_Y h_2^{\beta_2})\|y\|_{\mathcal{Y}_s}, \qquad y \in \mathcal{Y}_s,\end{aligned}$$

with positive constants $c_{\mathcal{R}}$, $c_{\mathcal{S}}$, c_U *and* c_Y.

(iii) *The error in solving for the state dynamics can be made arbitrarily small, i.e., Assumption 12 holds.*

Let $\delta > 0$ *be given. Then one can choose subspaces* $\mathcal{U}_{h_1^*,\tau_1^*}$ *and* $\mathcal{Y}_{h_2^*,\tau_2^*}$ *such that*

$$\tau_1^* < \left(\frac{\delta}{8c_{\mathcal{R}}\|\mathbb{G}\|_{\mathscr{L}(\mathcal{U},\mathcal{Y})}}\right)^{1/\alpha_1}, \quad h_1^* < \left(\frac{\delta}{8c_U\|\mathbb{G}\|_{\mathscr{L}(\mathcal{U},\mathcal{Y})}}\right)^{1/\beta_1}, \tag{24a}$$

$$\tau_2^* < \left(\frac{\delta}{8c_{\mathcal{S}}\|\mathbb{G}\|_{\mathscr{L}(\mathcal{U}_s,\mathcal{Y}_s)}}\right)^{1/\alpha_2}, \quad h_2^* < \left(\frac{\delta}{8c_Y\|\mathbb{G}\|_{\mathscr{L}(\mathcal{U}_s,\mathcal{Y}_s)}}\right)^{1/\beta_2}, \tag{24b}$$

and one can solve the PDEs (16) *numerically for* $l = 1, \ldots, p(h_1)$ *such that one of the following conditions holds.*

(i) $$\|\mathbf{K}^w_{:,l} - \tilde{\mathbf{K}}^w_{:,l}\|_{L^2(0,T;\mathbb{R}^q)} < \frac{\delta}{2\sqrt{T}p(h_1^*)}, \tag{25a}$$

(ii) $$\|\mathbf{K}_{:,l} - \tilde{\mathbf{K}}_{:,l}\|_{L^2(0,T;\mathbb{R}^q)} < \frac{\delta}{2\sqrt{T}p(h_1^*)}\sqrt{\frac{\lambda_{\min}(\mathbf{M}_{U,h_1^*})}{\lambda_{\max}(\mathbf{M}_{Y,h_2^*})}}, \tag{25b}$$

(iii) $$\|z_l - z_{l,\mathtt{tol}}\|_{L^2(0,T;Z)} < \frac{\delta}{2\sqrt{T}p(h_1^*)}\sqrt{\frac{\lambda_{\min}(\mathbf{M}_{U,h_1^*})}{\lambda_{\max}(\mathbf{M}_{Y,h_2^*})}}\|C\|^{-1}_{\mathscr{L}(Z,Y)}\left(\sum_{i=1}^{q(h_2^*)}\|\nu_i\|_Y^2\right)^{-1/2}. \tag{25c}$$

In this case,

$$\|\mathbb{G} - \mathbb{G}_{DS}\|_{\mathscr{L}(\mathcal{U}_s,\mathcal{Y})} < \delta.$$

Moreover, the signal error $\epsilon'_S := \|\mathbb{G} - \mathbb{G}_S\|_{\mathscr{L}(\mathcal{U}_s,\mathcal{Y})}$ *and the system dynamics error* $\epsilon_D := \|\mathbb{G}_S - \mathbb{G}_{DS}\|_{\mathscr{L}(\mathcal{U},\mathcal{Y})}$ *are balanced in the sense that* $\epsilon'_S, \epsilon_D < \delta/2$.

Proof. For $u \in \mathcal{U}_s$, we have

$$\begin{aligned}
\|\mathbb{G}u - \mathbb{G}_S u\|_{\mathcal{Y}} \leq & \|\mathbb{G}u - \mathbb{P}_{\mathcal{Y},h_2,\tau_2}\mathbb{G}u\|_{\mathcal{Y}} + \|\mathbb{P}_{\mathcal{Y},h_2,\tau_2}\mathbb{G}u - \mathbb{P}_{\mathcal{Y},h_2,\tau_2}\mathbb{G}\mathbb{P}_{\mathcal{U},h_1,\tau_1}u\|_{\mathcal{Y}}, \\
\leq & (c_{\mathcal{S}}\tau_2^{\alpha_2} + c_Y h_2^{\beta_2})\|\mathbb{G}u\|_{\mathcal{Y}_s} \\
& + (c_{\mathcal{R}}\tau_1^{\alpha_1} + c_U h_1^{\beta_1})\|\mathbb{P}_{\mathcal{Y}}\|_{\mathscr{L}(\mathcal{Y})}\|\mathbb{G}\|_{\mathscr{L}(\mathcal{U},\mathcal{Y})}\|u\|_{\mathcal{U}_s}, \\
\leq & \Big\{ (c_{\mathcal{S}}\tau_2^{\alpha_2} + c_Y h_2^{\beta_2})\|\mathbb{G}\|_{\mathscr{L}(\mathcal{U}_{rs},\mathcal{Y}_s)} \\
& + (c_{\mathcal{R}}\tau_1^{\alpha_1} + c_U h_1^{\beta_1})\|\mathbb{G}\|_{\mathscr{L}(\mathcal{U},\mathcal{Y})} \Big\} \|u\|_{\mathcal{U}_s},
\end{aligned}$$

and thus (24) ensures that $\epsilon_S' = \|\mathbb{G} - \mathbb{G}_S\|_{\mathscr{L}(\mathcal{U}_s,\mathcal{Y})} < \delta/2$. Proposition 10 in combination with (25) and in view of (22) ensures that $\epsilon_D = \|\mathbb{G}_S - \mathbb{G}_{DS}\|_{\mathscr{L}(\mathcal{U},\mathcal{Y})} < \delta/2$, which concludes the proof. □

6 Applications and numerical results

6.1 Test problems

As test cases, we consider two heat equations on different domains $\Omega \subset \mathbb{R}^2$ as depicted in Fig. 4. In both cases the control and observation operators are defined on rectangular subsets of Ω $\Omega_c = (a_{c,1}, a_{c,2}) \times (b_{c,1}, b_{c,2})$ and $\Omega_m = (a_{m,1}, a_{m,2}) \times (b_{m,1}, b_{m,2})$, where the control is active and the observation takes place, respectively.

Test case 6.1. *Setting $U = Y = L^2(0,1)$, we define $C \in \mathscr{L}(L^2(\Omega), Y)$ and $B \in \mathscr{L}(U, L^2(\Omega))$ by*

$$(Bu)(x_1,x_2) = \begin{cases} u(\theta_1(x_1))\omega_c(x_2), & (x_1,x_2) \in \Omega_c \\ 0, & (x_1,x_2) \notin \Omega_c \end{cases}$$

and

$$(Cz)(\xi) = \int_{a_{m,1}}^{b_{m,1}} \frac{z(x_1, \theta_2(\xi))}{b_{m,1} - a_{m,1}}\, dx_1,$$

where $\omega_c \in L^2(a_{c,2}, b_{c,2})$ is a weight function and $\theta_1 : [a_{c,1}, b_{c,1}] \to [0,1]$ and $\theta_2 : [0,1] \to [a_{m,1}, b_{m,1}]$ are affine-linear transformations.

Note that C preserves an inherent spatial state regularity, i.e., $C_{|H^2(\Omega)} \in \mathscr{L}(H^2(\Omega), H^2(0,1))$.

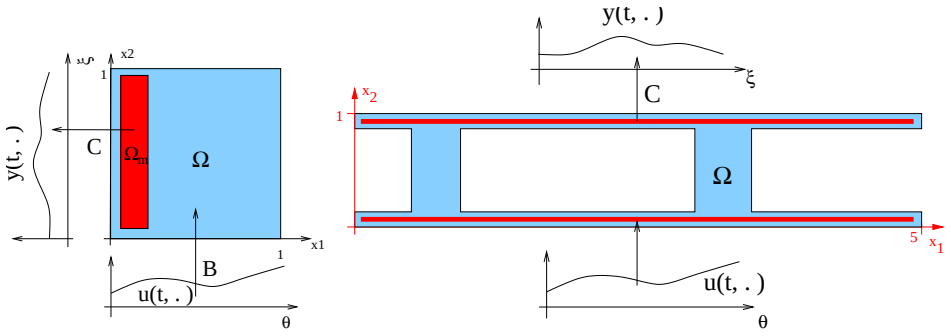
For the state equation we consider a heat equation with homogeneous Dirichlet boundary conditions on $(0,T] \times \Omega$ *with* $T = 1$ *and* $\Omega = (0,1)^2$. *We choose* $\Omega_c = \Omega$, $\Omega_m = (0.1, 0.2) \times (0.1, 0.9)$ *and* $\omega_c(x_2) = \sin(\pi x_2)$. *In this case, the output obtained by inputs of the special form* $u(t;\theta) = \sin(\omega_T \pi t)\sin(m\pi\theta)$ *with* $\omega_T, m \in \mathbb{N}$ *can be explicitly computed in terms of the eigenfunctions of the Laplace operator.*

Test case 6.2. *As second test case, we consider two infinitely long plates of width* 5 *and height* 0.2, *which are connected by two rectangular bars as shown in the cross section in Fig. 4. We assume that the plates are surrounded by an insulating material and that we can heat the bottom plate and measure the temperature distribution in the upper plate.*

The input operator is chosen as in Test case 6.1 for the output operator, we just switch the variables in the definition of C.

As state equation we consider a heat equation with homogeneous Neumann boundary conditions on $(0,T] \times \Omega$ *with* $T = 1$ *and* Ω *as in Fig. 4, and choose* $\Omega_c = (0.05, 4.95) \times (0.05, 0.15)$, $\Omega_m = (0.05, 4.95) \times (0.85, 0.95)$ *and* $\omega_c(x_2) = \sin(\pi(x_2 - 0.05)/0.1)$.

The matrix approximations $\tilde{\mathbf{G}}$ of the I/O-maps $\mathbb{G}$ corresponding to the test cases have been calculated by means of a heat equation solver, which is based on the C++ FEM software library DEAL.II[40]. It realizes a discontinuous Galerkin scheme with adaptive space and time grids and applies goal-oriented DWR-based error control to ensure (21).



(a) Test case 6.1 (b) Test case 6.2

Figure 4: Test cases heat equation: (a) with homogeneous Dirichlet boundary conditions, (b) with homogeneous Neumann boundary conditions.

6.2 Tests of convergence

The following numerical convergence tests have all been carried out with approximations $\mathbb{G}_{DS}(h_1, \tau_1, h_2, \tau_2, \texttt{tol})$ of the I/O-map $\mathbb{G}$ corresponding to Test case 6.1. Hierarchical linear finite elements in U_{h_1} and Y_{h_2} and Haar wavelets in $\mathcal{R}_{\tau_1}$ and $\mathcal{S}_{\tau_2}$ have been chosen. The tolerance `tol` refers to the estimate (23).

Convergence of single outputs. Considering Test case 6.1 with inputs $u(t;\theta) = \sin(\omega_T \pi t)\sin(m\pi\theta)$, and exactly known outputs $y = \mathbb{G}u$, we investigate the relative error $\|y - \tilde{y}\|_{\mathcal{Y}}/\|u\|_{\mathcal{U}_s}$, with $\tilde{y} = \mathbb{G}_{DS}(h_1, \tau_1, h_2, \tau_2, \texttt{tol})u$, for varying discretization parameters h_1, τ_1, h_2, τ_2 and `tol`. Choosing, e.g., $m = 5$ and $\omega_T = 10$, we observe a quadratic convergence in $h_1 = h_2$ (cf. Fig. 6.2-a) and a linear convergence in $\tau_1 = \tau_2$ (cf. Fig. 6.2-b) in correspondence to Thm. 13. However, the error does not converge to zero but to a positive plateau value, which is due to the system dynamics error and which becomes smaller for lower tolerances `tol`. For input signals with $m > 5$ and $\omega_T > 10$ the convergence order can only be observed for smaller discretization parameters h_1, h_2, τ_1 and τ_2.

Convergence of the norm $\|\mathbb{G}_S(h_1, \tau_1, h_2, \tau_2)\|_{\mathscr{L}(\mathcal{U},\mathcal{Y})}$ for nested subspaces. Successively improving the signal approximation by adding additional basis functions, the norm $\|\mathbb{G}_S(h_1, \tau_1, h_2, \tau_2)\|_{\mathscr{L}(\mathcal{U},\mathcal{Y})}$ converges, cf. Lemma 5. We approximate $\|\mathbb{G}_S\|_{\mathscr{L}(\mathcal{U},\mathcal{Y})}$ by $\|\mathbb{G}_{DS}\|_{\mathscr{L}(\mathcal{U},\mathcal{Y})}$, where $\mathbb{G}_{DS}$ has been calculated with $\texttt{tol} = 4.0e-5$. In Fig. 6.2-c, the approximations $\|\mathbb{G}_S(h_1, \tau_1, h_2, \tau_2)\|_{\mathscr{L}(\mathcal{U},\mathcal{Y})} = \|\mathbb{G}_S(\frac{1}{p-1}, \frac{1}{r}, \frac{1}{q-1}, \frac{1}{s})\|_{\mathscr{L}(\mathcal{U},\mathcal{Y})}$ are plotted for increasing subspace dimensions $p = q = r + 1 = s + 1 = 2, 3, \ldots, 65$.

6.3 Matrix reduction on the basis of SVDs

In order to resolve the input and output signal spaces accurately by means of general purpose basis functions, a large number of basis functions is needed in general. In order to reduce the large size of the resulting I/O-matrices $\tilde{\mathbf{G}}$, we apply a reduction method known as *Tucker decomposition* or *higher order singular value decomposition* (HOSVD) [41]. It is based on singular value decompositions (SVDs) and preserves the space-time tensor structure of the input and output signal bases.

Considering $\tilde{\mathbf{G}} \in \mathbb{R}^{qs\times pr}$ as a fourth-order tensor $\tilde{\mathbf{G}} \in \mathbb{R}^{s\times r\times q\times p}$ with $\tilde{\mathbf{G}}_{ijkl} = \tilde{\mathbf{G}}^{kl}_{ij}$, it is shown in [41] that there exists a HOSVD

$$\tilde{\mathbf{G}} = \mathbf{S} \times_1 \mathbf{U}^{(\psi)} \times_2 \mathbf{U}^{(\phi)} \times_3 \mathbf{U}^{(\nu)} \times_4 \mathbf{U}^{(\mu)}. \tag{26}$$

Here $\mathbf{S} \in \mathbb{R}^{s\times r\times q\times p}$ is a so-called *core tensor*, satisfying some orthogonality

properties, $\mathbf{U}^{(\psi)} \in \mathbb{R}^{s\times s}$, $\mathbf{U}^{(\phi)} \in \mathbb{R}^{r\times r}$, $\mathbf{U}^{(\nu)} \in \mathbb{R}^{q\times q}$, $\mathbf{U}^{(\mu)} \in \mathbb{R}^{p\times p}$ are unitary matrices and $\times_1, \ldots, \times_4$ denote tensor-matrix multiplications. We define a so-called *matrix unfolding* $\tilde{\mathbf{G}}^{(\psi)} \in \mathbb{R}^{s\times rqp}$ of the tensor $\tilde{\mathbf{G}}$ by

$$\tilde{\mathbf{G}}^{(\psi)}_{im} = \mathbf{G}_{ijkl}, \quad m = (k-1)ps + (l-1)s + i,$$

i.e., we put all elements belonging to $\psi_1, \psi_2, \ldots, \psi_s$ into one respective row, and we define the unfoldings $\tilde{\mathbf{G}}^{(\phi)} \in \mathbb{R}^{r\times qps}$, $\tilde{\mathbf{G}}^{(\nu)} \in \mathbb{R}^{q\times psr}$ and $\tilde{\mathbf{G}}^{(\mu)} \in \mathbb{R}^{p\times srq}$ in a similar cyclic way. Then, $\mathbf{U}^{(\psi)}$, $\mathbf{U}^{(\phi)}$, $\mathbf{U}^{(\nu)}$ and $\mathbf{U}^{(\mu)}$ in (26) can be calculated by means of four SVDs of the respective form

$$\tilde{\mathbf{G}}^{(\psi)} = \mathbf{U}^{(\psi)}\Sigma^{(\psi)}(\mathbf{V}^{(\psi)})^T,$$

where $\Sigma^{(\psi)}$ is diagonal with entries $\sigma_1^{(\psi)} \geq \sigma_2^{(\psi)} \geq \ldots \sigma_s^{(\psi)} \geq 0$ and $\mathbf{V}^{(\psi)}$ is columnwise orthonormal. The $\sigma_i^{(\psi)}$ are so-called *n-mode singular values* (or in our case ψ-mode singular values) of the tensor $\tilde{\mathbf{G}}$ and correspond to the Frobenius norms of certain subtensors of the core tensor $\mathbf{S}$.

On the basis of (26) we can define an approximation $\hat{\mathbf{G}} \in \mathbb{R}^{s\times r\times q\times p}$ of $\tilde{\mathbf{G}}$ by discarding the smallest n-mode singular values $\{\sigma^{(\psi)}_{\hat{s}+1}, \ldots, \sigma^{(\psi)}_s\}$, $\{\sigma^{(\phi)}_{\hat{r}+1}, \ldots, \sigma^{(\psi)}_r\}$, $\{\sigma^{(\nu)}_{\hat{q}+1}, \ldots, \sigma^{(\nu)}_q\}$ and $\{\sigma^{(\mu)}_{\hat{p}+1}, \ldots, \sigma^{(\mu)}_p\}$, i.e., we set the corresponding parts of $\mathbf{S}$ to zero. Then we have [41]

$$\|\tilde{\mathbf{G}} - \hat{\mathbf{G}}\|_F^2 \leq \sum_{i=\hat{s}+1}^{s} \sigma_i^{(\psi)} + \sum_{j=\hat{r}+1}^{r} \sigma_j^{(\phi)} + \sum_{k=\hat{q}+1}^{q} \sigma_k^{(\nu)} + \sum_{l=\hat{p}+1}^{p} \sigma_l^{(\mu)}.$$

The truncation of $\hat{\mathbf{G}} \in \mathbb{R}^{qr\times ps}$ after a basis transformation corresponding to $\mathbf{U}^{(\psi)}$, $\mathbf{U}^{(\phi)}$, $\mathbf{U}^{(\nu)}$ and $\mathbf{U}^{(\mu)}$ yields a low-dimensional representation $\bar{\mathbf{G}} \in \mathbb{R}^{\hat{q}\hat{r}\times\hat{p}\hat{s}}$.

In Figure 6 the HOSVD has been applied to a matrix $\tilde{\mathbf{G}} \in \mathbb{R}^{qs\times pr}$ for the Test case 6.2 with $p = 17$, $q = 65$ and $r = s = 64$. The first row shows the respective n-mode singular values. Underneath the first and most relevant two transformed/new basis functions $\hat{\mu}_i$, $\hat{\nu}_i$, $\hat{\phi}_i$ and $\hat{\psi}_i$, are plotted. It is not surprising that the positions of the connections between the plates can be recovered as large values of the corresponding spatial input and output basis functions.

Remark 14. *The application of a HOSVD is useful in two ways.* First, *it delivers a low-dimensional matrix-representation of the system, which is small enough to be used for real-time feedback control design.* Second, *it allows to identify relevant input and output signals, which can be exploited in actuator and sensor design, i.e., to decide where actuators and sensors have to be placed and which resolution in time and space they should have.*

6.4 Application in optimization problems

We investigate the use of the I/O-map approximation in optimization problems

$$\min J(u,y) \qquad \text{subject to } y = \mathbb{G}u, \quad u \in \mathcal{U}_{ad}. \tag{27}$$

Here $\mathcal{U}_{ad} \subset \mathcal{U}$ is the subset of admissible controls, $J : \mathcal{U} \times \mathcal{Y} \to \mathbb{R}$ is a quadratic cost functional $J(u,y) = \frac{1}{2}\|y - y_D\|^2_{\mathcal{Y}} + \alpha\|u\|^2_{\mathcal{U}}$, $y_D \in \mathcal{Y}$ is a desired output signal, and $\alpha > 0$ is a regularization parameter. We define the discretized cost functional

$$\bar{J}_{\mathbf{h}} : \mathbb{R}^{pr} \times \mathbb{R}^{qs} \to \mathbb{R}, \qquad \bar{J}_{\mathbf{h}}(\mathbf{u},\mathbf{y}) = \frac{1}{2}\|\mathbf{y} - \mathbf{y}_D\|^2_{qs;w} + \alpha\|\mathbf{u}\|^2_{pr;w},$$

with $\mathbf{y}_D = \kappa_{\mathcal{Y},h_2,\tau_2}\mathbb{P}_{\mathcal{Y},h_2,\tau_2}y_D$, and instead of (27) we solve

$$\min \bar{J}_{\mathbf{h}}(\mathbf{u},\mathbf{y}) \qquad \text{subject to } \mathbf{y} = \tilde{\mathbf{G}}\mathbf{u}, \quad \mathbf{u} \in \bar{U}_{ad}, \tag{28}$$

with $\bar{U}_{ad} = \{\mathbf{u} \in \mathbb{R}^{pr} : \mathbf{u} = \kappa_{\mathcal{U},h_1,\tau_1}\mathbb{P}_{\mathcal{U},h_1,\tau_1}u,\ u \in \mathcal{U}_{ad}\}$. Considering optimization problems without control constraints, i.e., $\mathcal{U}_{ad} = \mathcal{U}$ and $\bar{U}_{ad} = \mathbb{R}^{pr}$, the solution $\bar{\mathbf{u}}$ of (28) is characterized by

$$(\tilde{\mathbf{G}}^T\mathbf{M}_{\mathcal{Y}}\tilde{\mathbf{G}} + \alpha\mathbf{M}_{\mathcal{U}})\bar{\mathbf{u}} = \tilde{\mathbf{G}}^T\mathbf{M}_{\mathcal{Y}}\mathbf{y}_D. \tag{29}$$

As concrete example, we consider Test case 6.2 and choose $y_D = \mathbb{G}u_0$ to be the output for an input $u_0 \equiv 1$ which is equal to 1 on all of $[0,T] \times (0,1)$. We then try to find an *optimal* input u_* that minimizes the cost functional (27).

First we solve (29) with an approximated I/O map $\tilde{\mathbf{G}} \in \mathbb{R}^{17\cdot64\times65\cdot64}$ and $\alpha = 10^{-4}$, yielding an approximation $\bar{u} \approx u_*$.

The solution takes 0.33 seconds on a normal desktop PC. The u-norm is reduced by 27.9% and the relative deviation of $\mathbb{G}\bar{u}$ from y_D is 9.4%. In Fig. 7 the same calculations have been carried out with $\hat{\mathbf{G}} \in \mathbb{R}^{3\cdot5\times3\cdot5}$, where $\hat{\mathbf{G}}$ arises from a HOSVD-based matrix reduction of $\tilde{\mathbf{G}} \in \mathbb{R}^{17\cdot64\times65\cdot64}$, where all but the 3 most relevant spatial and the 5 most relevant temporal input and output basis functions have been truncated. Using this approximation, the norm of u is reduced by 27.4%, whereas the relative deviation of $\mathbb{G}\bar{u}$ from y_D is 9.5%. The cost functional has been reduced by 44.5%, and the calculation of $\bar{u}$ took less than 0.0004 seconds. The outputs resulting from u_0 and $\bar{u}$ have been calculated in simulations independent from the calculation of the I/O-matrix.

7 Final remarks and outlook

We have presented a systematic framework for the discretization of I/O-maps of linear infinite-dimensional control systems with *spatially distributed* inputs and outputs. Global error estimates have been provided, which allow to choose the involved discretization parameters in such a way that a desired overall accuracy is achieved and that the signal and the system dynamics approximation errors are balanced. Moreover, the error results are capable to take into account many practical and technical restrictions in sensor and actuator design, like limited spatial and temporal resolutions or the use of piecewise constant controls and observations due to digital devices.

The numerical costs of the approach are primarily governed by the numerical calculation of p underlying homogeneous PDEs, where p is the number of input basis functions in space. This, however can be done beforehand, and in parallel, provided there is enough storage available that allows to store these solutions. This, however, can become problematic when the spatial resolution of the input signal space has to be very accurate. In this case, code-optimization, e.g. due to parallelization and appropriate updating of mass and stiffness matrices from prior calculations, promises to have a large potential for speed-up.

The SVD-based dimension reduction for the matrix representation can be considered as an alternative model reduction approach, and the resulting reduced I/O-models proved to be successful in first numerical optimization applications. Moreover, the SVD-based reduction may be able to provide useful insight for efficient actuator and sensor design by filtering out relevant input and output signals.

References

[1] A. C. Antoulas, *Approximation of large-scale dynamical systems* (Society for Industrial and Applied Mathematics (SIAM), Philadelphia, 2005).

[2] P. Benner, V. Mehrmann and D. Sorensen (eds.), *Dimension Reduction of Large-Scale Systems* (Springer, 2005).

[3] D. M. Luchtenburg, B. Gunter, B. R. Noack, R. King and G. Tadmor, *J. Fluid Mech.* **623**, 283 (2009).

[4] B. R. Noack, M. Schlegel, B. Ahlborn, G. Mutschke, M. Morzynski, P. Comte and G. Tadmor, *J. Non-Equilib. Thermodyn.* **33**, 103 (2008).

[5] M. Pastoor, L. Henning, B. R. Noack, R. King and G. Tadmor, *J. Fluid Mech.* **608**, 161 (2008).

[6] G. Berkooz, P. Holmes and J. L. Lumley, The proper orthogonal decomposition in the analysis of turbulent flows, in *Annual review of fluid mechanics, Vol. 25*, (Annual Reviews Inc., Palo Alto, 1993) pp. 539–575.

[7] V. Mehrmann and T. Stykel, Balanced truncation model reduction for large-scale systems in descriptor form, in *Dimension Reduction of Large-Scale Systems*, eds. P. Benner, V. Mehrmann and D. Sorensen (Springer, Heidelberg, 2005).

[8] R. Becker, M. Garwon, C. Gutknecht, G. Bärwolff and R. King, *J. of Process Control* **15**, 691 (2005).

[9] L. Henning, D. Kuzmin, V. Mehrmann, M. Schmidt, A. Sokolov and S. Turek, Flow control on the basis of a FEATFLOW-MATLAB coupling, in *Active Flow Control. Papers contributed to the Conference "Active Flow Control 2006", Berlin, Germany, September 27 to 29, 2006*, ed. R. King (Springer, Berlin, 2006).

[10] S. Gugercin and A. C. Antoulas, *Internat. J. Control* **77**, 748 (2004).

[11] R. W. Freund, *Model Reduction Methods Based on Krylov Subspaces*, tech. rep., Bell Laboratories, Lucent Technologies (2001).

[12] C. W. Rowley, *Internat. J. Bifur. Chaos Appl. Sci. Engrg.* **15**, 997 (2005).

[13] E. Emmrich, *Gewöhnliche und Operator-Differentialgleichungen* (Vieweg, Wiesbaden, 2004).

[14] J.-L. Lions and E. Magenes, *Non-homogeneous boundary value problems and applications. Vol. II* (Springer, New York, 1972).

[15] A. Pazy, *Semigroups of linear operators and applications to partial differential equations* (Springer, New York, 1983).

[16] L. C. Evans, *Partial differential equations*, Graduate Studies in Mathematics, Vol. 19 (American Mathematical Society, Providence, 1998).

[17] A. Lunardi, *Analytic semigroups and optimal regularity in parabolic problems* (Birkhäuser, Basel, 1995).

[18] H. Sohr, *The Navier-Stokes equations* (Birkhäuser, Basel, 2001).

[19] E. Emmrich and V. Mehrmann, *Analysis of a class of operator differential-algebraic equations arising in fluid mechanics. Part 1. The finite dimensional case*, tech. rep. (2010).

[20] J. Heiland, V. Mehrmann and M. Schmidt, A new discretization framework for input/output maps and its application to flow control, in *Active Flow Control. Papers contributed to the Conference "Active Flow Control II 2010", Berlin, Germany, May 26 to 28, 2010*, ed. R. King (Springer, Berlin, 2006).

[21] H. Yserentant, Hierarchical bases in the numerical solution of parabolic problems, in *Large scale scientific computing (Oberwolfach, 1985)*, (Birkhäuser, Boston, MA, 1987) pp. 22–36.

[22] H. Yserentant, Hierarchical bases, in *ICIAM 91 (Washington, DC, 1991)*, (SIAM pp. 256–276.

[23] A. Cohen, *Numerical analysis of wavelet methods*, Studies in Mathematics and its Applications, Vol. 32 (North-Holland Publishing Co., Amsterdam, 2003).

[24] P. G. Ciarlet, *The finite element method for elliptic problems*, Classics in Applied Mathematics, Vol. 40 (SIAM, Philadelphia, PA, 2002).

[25] E. Zeidler, *Nonlinear functional analysis and its applications. II/A* (Springer, New York, 1990).

[26] W. Hackbusch, B. N. Khoromskij and E. E. Tyrtyshnikov, *J. Numer. Math.* **13**, 119 (2005).

[27] C. Johnson, *Numerical solution of partial differential equations by the finite element method* (Cambridge University Press, Cambridge, 1987).

[28] K. Eriksson, D. Estep, P. Hansbo and C. Johnson, Introduction to adaptive methods for differential equations, in *Acta numerica, 1995*, Acta Numer. (Cambridge University Press, Cambridge) pp. 105–158.

[29] K. Eriksson and C. Johnson, *SIAM J. Numer. Anal.* **32**, 706 (1995).

[30] V. Thomée, *Galerkin finite element methods for parabolic problems*, Springer Series in Computational Mathematics, Vol. 25 (Springer, Berlin, 1997).

[31] I. Babuska and A. Miller, *Int. J. Numer. Methods Eng.* **20**, 2311 (1984).

[32] M. Ainsworth and J. T. Oden, *A posteriori error estimation in finite element analysis* (Wiley-Interscience, New York, 2000).

[33] W. Bangerth and R. Rannacher, *Adaptive finite element methods for differential equations* (Birkhäuser, Basel, 2003).

[34] R. Becker, V. Heuveline and R. Rannacher, *Int. J. Numer. Methods Fluids* **40**, 105 (2002).

[35] R. Becker and R. Rannacher, *East-West J. Numer. Math.* **4**, 237 (1996).

[36] R. Becker and R. Rannacher, *Acta Numer.* **10**, 1 (2001).

[37] V. Heuveline and R. Rannacher, *J. Numer. Math.* **11**, 95 (2003).

[38] C. Johnson and R. Rannacher, On errror control in cfd, in *Numerical methods for the Navier-Stokes equations. Proceedings of the international workshop held, Heidelberg, Germany, October 25-28, 1993. Notes Numer. Fluid Mech. 47, 133-144* , ed. K.-F. Hebeker (Vieweg, Braunschweig, 1994)

[39] C. Johnson, R. Rannacher and M. Boman, *SIAM J. Numer. Anal.* **32**, 1058 (1995).

[40] W. Bangerth, R. Hartmann and G. Kanschat, `deal.II` *Differential Equations Analysis Library, Technical Reference.* http://www.dealii.org/, 5.2 edn.(September, 2005).

[41] L. De Lathauwer, B. De Moor and J. Vandewalle, *SIAM J. Matrix Anal. Appl.* **21**, 1253 (2000).

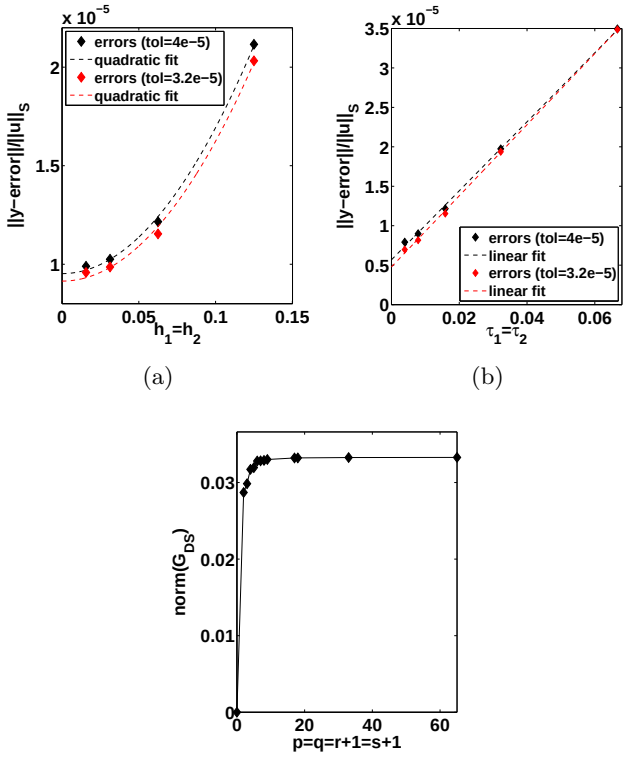


Figure 5: (a) Relative output errors for input $u(t;\theta) = \sin(10\pi t)\sin(5\pi\theta)$, varying $h_1 = h_2$ and fixed $\tau_1 = \tau_2 = 1/64$. (b) Relative output errors for input $u(t;\theta) = \sin(10\pi t)\sin(5\pi\theta)$, varying $\tau_1 = \tau_2$ and fixed $h_1 = h_2 = 1/17$. (c) Norm $\|\mathbb{G}_{DS}(\mathbf{h})\|_{\mathscr{L}(\mathcal{U},\mathcal{Y})}$ for synchronously increasing approximation space dimensions $p = q = r + 1 = s + 1$ and fixed tolerance $\texttt{tol} = 4.0e - 5$.

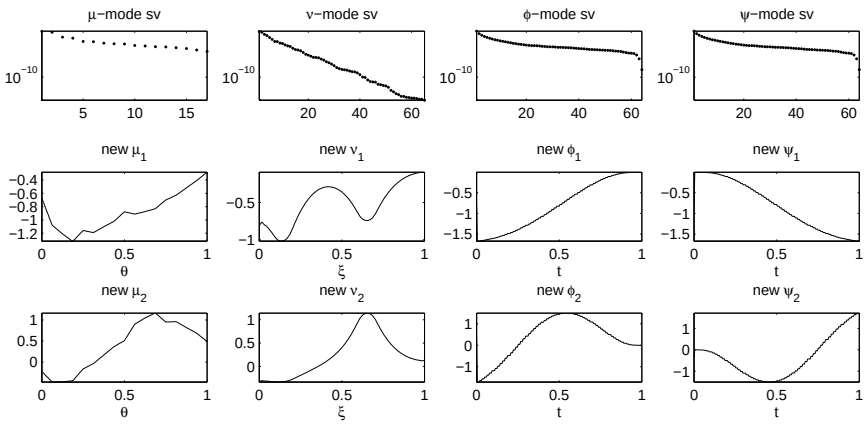


Figure 6: HOSVD applied to the I/O map of Test case 6.2. First row: n-mode singular values in semilogarithmic scales. 2nd and 3rd row: Respective two most relevant basis functions.

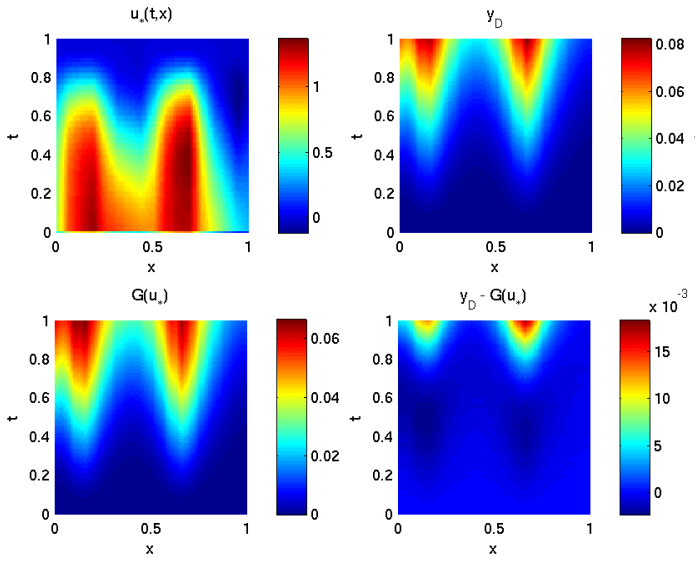


Figure 7: Application of the SVD-reduced approximated I/O map $\hat{\mathbf{G}} \in \mathbb{R}^{3\cdot5\times3\cdot5}$ in an optimization problem. From top left to bottom right: optimized control $\bar{u}$, original output $y_D = \mathbb{G}u_0$, optimized output $\mathbb{G}\bar{u}$ and their difference.