Technische Universität Berlin

Institut für Mathematik

Systematic discretization of input-output maps of linear infinite dimensional systems

Michael Schmidt

Preprint 2006-06

Preprint-Reihe des Instituts für Mathematik
Technische Universität Berlin

Report 2006-06 **June 2006**

Many model reduction techniques take a semi-discretization of the original PDE model as starting point and aim then at an accurate approximation of its input/output map. In this contribution, we discuss the direct discretization of the i/o map of the *original* infinite-dimensional system. First, the input and output signals are discretized in space and time, second, the system dynamics is approximated in form of the underlying evolution operator, leading to an approximated i/o map with matrix representation. The discretization framework, corresponding error estimates, a SVD-based system reduction method and a numerical application in an optimization problem are presented for a general class of linear time-invariant systems and illustrated for a heat control system.

2000 Mathematics Subject Classification.
Primary: 93C20, 35B37; Secondary: 49J20, 93C05

Keywords: input-output map, discretization, infinite-dimensional control system, model reduction, optimization

Systematic discretization of input-output maps of linear infinite dimensional systems

June 26, 2006

Michael Schmidt
Institut für Mathematik, MA 4-5
Technische Universität Berlin
10623 Berlin, Germany

Many model reduction techniques take a semi-discretization of the original PDE model as starting point and aim then at an accurate approximation of its input/output map. In this contribution, we discuss the direct discretization of the i/o map of the *original* infinite-dimensional system. First, the input and output signals are discretized in space and time, second, the system dynamics is approximated in form of the underlying evolution operator, leading to an approximated i/o map with matrix representation. The discretization framework, corresponding error estimates, a SVD-based system reduction method and a numerical application in an optimization problem are presented for a general class of linear time-invariant systems and illustrated for a heat control system.

Contents

1. Introduction

The control of complex physical systems is a big challenge in many engineering applications as well as in mathematical research. Frequently, these control systems are modeled by infinite-dimensional state space systems on the basis of (instationary and nonlinear) partial differential equations (PDEs). On the one hand, space-discretizations resolving most of the state information typically lead to very large semi-discrete systems, on the other hand, classical design techniques for real-time controllers require linear models of very moderate size.

Numerous approaches to bridge this gap are proposed in the literature. For instance, low-dimensional modeling and model reduction techniques aim to extract the system information that is essential for the specific control purpose, which in many applications proves to be the system's *input/output (i/o) map*. Robust feedback control is then a good choice to take the inevitable modeling and approximation errors into account.

In many engineering applications the empirical black-box modeling, i.e. a system identification on the basis of measured or simulated i/o-sequences, has proven to be successful, even for complex control tasks like the control of the recirculation bubble behind a backward facing step, see e.g. [7]. Physical insight may be used to derive low-dimensional state space models capturing essential state dynamics, information which necessarily gets lost in black-box modeling, see e.g. [22, 37, 40] for the control of flows. As a consequence, these methods usually lack analytical estimates for the model accuracy.

For some mathematical model reduction techniques such estimates exist, for others they are an open research topic. Most mathematical methods take the space-discretized PDE as starting point, some of them even a time-invariant linearization. One can roughly distinguish between methods based on singular-value decompositions (SVDs), like balanced truncation (BT) or proper orthogonal decomposition (POD), and methods based on Krylov subspace iterations, like moment matching (MM). Many of these techniques like MM, BT and recent variants of POD focus on an accurate approximation of the *i/o-map*. For a survey on model reduction see e.g. [2], [3] and the references therein. For error estimates in the case of BT and POD see e.g. [23], [36]. The preceding PDE space discretizations, however, rarely take an efficient approximation of the i/o map into account: On the one hand, the state space discretization typically aims at a reduction of the global state space error and is thus still oriented at a *state simulation problem*. On the other hand, the discretization of distributed inputs and outputs in space (leading to a multiple-input-multiple-output system) is rarely rigorously chosen with respect to the resulting i/o error.

In this paper we investigate a new approach. Motivated by the evident efficiency of system-identification in many engineering applications, we focus directly on the i/o map of the *original* infinite-dimensional system, in the following denoted by

$$\mathbb{G} : u = u(t, \theta) \mapsto y = y(t, \xi),$$

where u are input signals and y are output signals, which may vary in time t and space $\theta \in \Theta$ respectively $\xi \in \Xi$. Here Θ and Ξ denote the domains of the spatially distributed inputs and outputs.

We suggest a framework for its direct discretization and approximation with rigorous error control for the case of *instationary linear* systems.

The framework and its prospects are briefly presented in Section 2 for general linear bounded i/o maps. In Section 3 we consider linear state space systems fitting into this setting and give the heat equation as concrete example which will be reconsidered throughout the text. The framework mainly consists of two steps. First, the input and output signal spaces are discretized in space and time, leading to a matrix representation $\mathbf{G}$ of $\mathbb{G}$, see Section 4. Second, the matrix has to be numerically approximated. Assuming that $\mathbb{G}$ arose from a PDE state space model, this can be done by calculating the approximated outputs corresponding to those states which are excited by the approximated inputs, see Section 5.

In Section 6 the numerical solution of the underlying PDE and appropriate error estimations are discussed, before a total error estimation is given in Section 7. In Section 8 some numerical convergence results are presented, as well as a matrix reduction

method which is based on the singular value decomposition and which is especially suited for systems with signals which are distributed in space and time. Furthermore we present how the matrix representation can be used in an optimization process. The paper concludes with some final remarks and an outlook.

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 $v : [0,T] \to \mathbb{R}$ 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, for details see e.g. [17]. By $\mathscr{C}([0,T];X)$ and $\mathscr{C}^\alpha([0,T];X)$ we denote the space of functions $v : [0,T] \to X$ which are continuous respectively α-times continuously differentiable. Frequently, such abstract function spaces will be denoted by calligraphic capital letters like $\mathcal{U}$ and $\mathcal{Y}$, whereas the underlying function space like $L^2(\Omega)$ will be written in standard capital letters like U and Y.

Operators acting between such abstract function spaces as $\mathcal{U}$ and $\mathcal{Y}$ are denoted for instance as $\mathbb{G}$, $\mathbb{P}$ or $\mathbb{I}$, with a slight abuse of a typeset which is often reserved for sets, like $\mathbb{N}$ for the set of natural numbers, $\mathbb{N}_0$ for the natural numbers including 0, $\mathbb{R}$ for the real numbers and $\mathbb{R}^{\alpha,\beta}$ for the set of real $\alpha \times \beta$ matrices. 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 $\mathbf{G}$.

2. Framework

We consider bounded linear i/o maps

$$\mathbb{G} : \mathcal{U} \to \mathcal{Y} \qquad u \mapsto y,$$

with input signals u and output signals y from respective Hilbert spaces $\mathcal{U}$ and $\mathcal{Y}$.

1. step: Approximation of signals. 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 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 $\mathbf{G} \in \mathbb{R}^{\bar{q} \times \bar{p}}$, for instance with elements

$$\mathbf{G}_{ij} = (y_i, \mathbb{G} u_j)_{\mathcal{Y}}$$

if orthonormal bases are chosen in $\bar{\mathcal{U}}$ and $\bar{\mathcal{Y}}$, see Section 4.
2. step: Approximation of observed and excited system dynamics. In order to calculate $\mathbf{G}$ numerically, we have to numerically approximate the observations in $\bar{\mathcal{Y}}$ of states excited by controls in $\bar{\mathcal{U}}$. If $\mathbb{G}$ corresponds to a known linear PDE state space model, then $\mathbf{G}$ can be approximated by integrating this model numerically for inputs u_j, $j = 1, \ldots, \bar{p}$, thus leading to an approximation $\tilde{\mathbf{G}}$ of $\mathbf{G}$ and an corresponding approximation $\mathbb{G}_{DS}$ of $\mathbb{G}_S$. In some cases, the numerical burden can be significantly reduced by choosing basis functions with a space-time tensor structure, like

$$u_i(t;\theta) = \phi_{j(i)}(t)\mu_{l(i)}(\theta), \quad y_j(t;\xi) = \psi_{i(j)}(t)\nu_{k(j)}(\xi),$$

see Section 5. In this case, the approximation of $\mathbb{G}$ reduces to the approximation of observations $(\nu_{k(j)}, z_{l(i)}(t))_Y$, where $z_{l(j)}(t)$ are the system's impulse responses corresponding to initial values $\mu_{l(i)}$, i.e. basically the system's evolution operator is approximated by numerically solving the corresponding underlying homogeneous PDEs. For such a choice of basis functions, $\mathbf{G}$ becomes a fourth-order tensor, but we will represent $\mathbf{G}$ as a block-structured matrix for ease of notation.
Some prospects of this framework.

(i) *Error estimation.* The total error ϵ_{DS} can be estimated by the *signal* approximation error ϵ_S and the *dynamical* approximation error ϵ_D, i.e.

$$\underbrace{\|\mathbb{G} - \mathbb{G}_{DS}\|}_{=:\epsilon_{DS}} \leq \underbrace{\|\mathbb{G} - \mathbb{G}_S\|}_{=:\epsilon_S} + \underbrace{\|\mathbb{G}_S - \mathbb{G}_{DS}\|}_{=:\epsilon_D}, \tag{1}$$

where the norms still have to be specified. A natural goal is to choose $\bar{\mathcal{U}}$ and $\bar{\mathcal{Y}}$ in the first step and the accuracy tolerances for the numerical solutions of the underlying PDEs in the second step such that ϵ_S and ϵ_D balance and that $\epsilon_S + \epsilon_D < tol$ for a given tolerance tol.

(ii) *Successive signal approximation refinement.* The error ϵ_S can be successively reduced by adding further basis functions $u_{\bar{p}+1}, u_{\bar{p}+2}, \ldots$ and $y_{\bar{q}+1}, y_{\bar{q}+2}, \ldots$ and thus enlarging $\bar{\mathcal{U}}$ and $\bar{\mathcal{Y}}$. This corresponds to adding respective columns and rows to the matrix $\mathbf{G}$. Such a procedure is numerically feasible, for instance by choosing hierarchical bases in $\bar{\mathcal{U}}$ and $\bar{\mathcal{Y}}$, like wavelets or hierarchical finite elements. However, one must ensure that ϵ_D is sufficiently small in order to reduce ϵ_{DS} significantly.

(iii) *Matrix reduction via SVD.* The matrix representation of $\mathbb{G}_{DS}$ allows for low rank approximations with error estimates on the basis of singular value decompositions (SVDs). The corresponding singular vectors represent the most relevant input and output signals.

(iv) *Actuators and sensors for distributed inputs and outputs.* Thinking of practical applications, input signals $u(t;\theta)$ and output signals $y(t;\xi)$ are often generated and measured by actuators and sensors with limited spatial and temporal resolutions, such that 'realizable' input and output signals naturally belong to finite

dimensional subspaces $\bar{\mathcal{U}}$ and $\bar{\mathcal{Y}}$, respectively. Error estimates of the form (1) and the extraction of relevant input and output signals on the basis of SVDs may thus provide useful information for efficient sensor and actuator design, see Section 8. Note that classical approaches (where the control system is first discretized in space and then model reduction is applied) rarely take the error due to input and output space-discretizations into account.

(v) *Control Design.* 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.

3. I/o maps of ∞-dimensional LTI state space systems

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

$$\partial_t z(t) = Az(t) + Bu(t), \quad t \in (0,T], \tag{2a}$$

$$z(0) = z^0, \tag{2b}$$

$$y(t) = Cz(t), \quad t \in [0,T]. \tag{2c}$$

Here for every time $t \in [0,T]$, the state $z(t)$ is supposed to belong to a Hilbert space Z like $Z = L^2(\Omega)$ or $Z = H^1(\Omega) \times L^2(\Omega)$, where $\Omega \subset \mathbb{R}^{d_\Omega}$ is a Lipschitz-domain of space dimension d_Ω. A is a densely defined unbounded operator $A : Z \supset D(A) \to Z$, generating a $\mathscr{C}^0$-semigroup $(S(t))_{t\geq 0}$ on Z. The control operator B belongs to $\mathscr{L}(U,Z)$ and the observation operator C to $\mathscr{L}(Z,Y)$, with $U = L^2(\Theta)$ and $Y = L^2(\Xi)$ and Lipschitz domains $\Theta \subset \mathbb{R}^{d_1}$ and $\Xi \subset \mathbb{R}^{d_2}$, $d_1, d_2 \in \mathbb{N}$.

We 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 to (2), for details see e.g. [41, Ch. 4]. It is well-known that for initial values $z_0 \in D(A)$ and controls $u \in \mathscr{C}^1([0,T];Z)$, a unique *classical solution* $z \in \mathscr{C}([0,T];Z) \cap \mathscr{C}^1((0,T);Z)$ of (2) exists. For $z_0 \in Z$ and $u \in \mathcal{U}$, the well-defined function

$$z(t) = S(t)z_0 + \int_0^t S(t-s)Bu(s)\,ds, \quad t \in [0,T], \tag{3}$$

is called a *mild solution* of (2). A mild solution of (2) is unique, belongs to $\mathscr{C}([0,T];Z)$ and is the uniform limit of classical solutions [41]. Hence, the output signal $y(t) = Cz(t)$ is well-defined and belongs to $\mathcal{Y} \cap \mathscr{C}([0,T];Y)$. In particular, the ouput 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 (2) by $u \mapsto y(u)$. It is possible to represent $\mathbb{G}$ as a convolution with the kernel function $K \in L^2(-T,T;\mathscr{L}(U,Y))$,

$$K(t) = \begin{cases} CS(t)B, & t \geq 0 \\ 0, & t < 0 \end{cases}. \tag{4}$$

Lemma 1. *The i/o-map $\mathbb{G}$ of (2) has the representation*

$$(\mathbb{G}u)(t) = \int_0^T K(t-s)u(s)\,ds, \quad t \in [0,T], \tag{5}$$

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

$$\|\mathbb{G}\|_{\mathscr{L}(\mathcal{U},\mathcal{Y})} \le \sqrt{T}\|K\|_{L^2(0,T;\mathscr{L}(U,Y))}. \tag{6}$$

Proof. Since C is bounded, the representation of $y = Cz$ based on (3) can be reformulated as in (5), see e.g. [17] for the theory of Bochner integrals. For general $K \in L^2(-T,T;\mathscr{L}(U,Y))$, a generalized Hölder's inequality yields 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

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

and by integrating over $[0,T]$ we obtain (6). □

Example 1. As prototype for a parabolic system, we consider the heat equation with homogeneous Dirichlet boundary conditions. In this case, $Z = L^2(\Omega)$ and the operator A in (2) coincides with the Laplace operator

$$A = \triangle : D(A) = H^2(\Omega) \cap H_0^1(\Omega) \subset Z \to Z. \tag{7}$$

Since A is the infinitesimal generator of an analytic $\mathscr{C}^0$-semigroup of contractions $(S(t))_{t\ge 0}$, the mild solution z of (2) exhibits the following stability and regularity properties, see e.g. [41, Ch. 7] and [21].

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

$$\|z\|_{H^1(0,T;L^2(\Omega))} + \|z\|_{L^2(0,T;H^2(\Omega))} \le c\|u\|_{\mathcal{U}}. \tag{8}$$

(ii) If $z_0 \in D(A)$ and $u \equiv 0$, then $z \in \mathscr{C}^1([0,T];D(A))$.

(iii) If $z_0 \in Z$ and $u \equiv 0$, then $z \in \mathscr{C}^1((0,T];D(A))$ and

$$\|z(t)\|_{L^2(\Omega)} \le \|z^0\|_{L^2(\Omega)} \quad \text{for all } t \in [0,T]. \tag{9a}$$

$$\|z(t)\|_{H^2(\Omega)} \le c\|\triangle z(t)\|_{L^2(\Omega)} \quad \text{for all } t \in [0,T], \tag{9b}$$

$$\|\partial_t z(t)\|_{L^2(\Omega)} = \|\triangle z(t)\|_{L^2(\Omega)} \le \frac{c'}{t}\|z^0\|_{L^2(\Omega)} \quad \text{for all } t \in (0,T], \tag{9c}$$

where c and c' are positive constants independent of T and z^0.

A heat equation example with concrete choices of Ω, B and C will be introduced in Section 8.

We note that if the observation preserves the inherent state regularity in the sense that

$$C_{|H^2(\Omega)} \in \mathscr{L}(H^2(\Omega),H^2(\Xi)), \tag{10}$$

then $\mathbb{G} \in \mathscr{L}(\mathcal{U}, \mathcal{Y}_s)$ and also $\mathbb{G}_{|\mathcal{U}_s} \in \mathscr{L}(\mathcal{U}_s, \mathcal{Y}_s)$ with input and output signal spaces of higher regularity in space and time

$$\mathcal{U}_s = H^1(0,T;L^2(\Theta)) \cap L^2(0,T;H^2(\Theta)), \quad \mathcal{Y}_s = H^1(0,T;L^2(\Xi)) \cap L^2(0,T;H^2(\Xi))$$

with norms $\|\cdot\|_{\mathcal{U}_s} = \|\cdot\|_{H^1(0,T;L^2(\Theta))} + \|\cdot\|_{L^2(0,T;H^2(\Theta))}$ and $\|\cdot\|_{\mathcal{Y}_s} = \|\cdot\|_{H^1(0,T;L^2(\Xi))} + \|\cdot\|_{L^2(0,T;H^2(\Xi))}$. In fact, for $u \in \mathcal{U}$, we have $\|u\|_{\mathcal{U}} \le \|u\|_{\mathcal{U}_S}$, and for $u \in \mathcal{U}$, we have

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

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 (8).

Remark 1. We consider *bounded* control operators B and observation operators C, which means, by definition, that $B \in \mathscr{L}(U,Z)$ and $C \in \mathscr{L}(Z,Y)$. Bounded B and C allow to describe distributed controls and observations, whereas boundary controls and observations as well as point-wise controls and observations are modeled via *unbounded* operators

$$B \in \mathscr{L}(U, Z_{-1}), \qquad C \in \mathscr{L}(Z_1, Y),$$

with spaces $Z_1 = D(A) \subset Z$ and $Z_{-1} = (D(A^*))' \supset Z$ equipped with appropriate norms. For so-called *admissible* (unbounded) control and observation operators, the state is still a continuous $L^2(0,T;Z)$-function with representation (3) and an observation $C_\lambda z(t)$ can still be well-defined via an expanded observation operator C_λ, see e.g. the survey articles [30] and [46] and the references therein. However, the kernel (4) is no longer well-defined and thus $\mathbb{G}$ can no longer be represented as (5).

Time-varying linear systems, like linearizations of nonlinear systems along trajectories, contain time-dependent operators $B(t)$, $A(t)$ and $C(t)$. They lead in the case of distributed controls and observations and under appropriate assumptions to representations of the form (5) but with kernel functions $K(t,s) = C(t)S(t,s)B(s) \in L^2([0,T]^2;L^2(U,Y))$ with evolution operators $S(t,s)$. For more details on time-varying systems, also with unbounded control and observation, see e.g. [29].

4. Signal approximation

In this section we consider general linear i/o-maps $\mathbb{G} \in \mathscr{L}(\mathcal{U}, \mathcal{Y})$ with $\mathcal{U} = L^2(0,T;U)$ and $\mathcal{Y} = L^2(0,T;Y)$, where $U = L^2(\Theta)$ and $Y = L^2(\Xi)$, i.e. the results also hold for time-varying systems with unbounded control and observation.

4.1. Space and time discretizations

In order to discretize input signals $u \in \mathcal{U}$ and output signals $y \in \mathcal{Y}$ in space, we choose two families $\{U_{h_1}\}_{h_1>0}$ and $\{Y_{h_2}\}_{h_2>0}$ of subspaces

$$U_{h_1} \subset U, \qquad Y_{h_2} \subset Y$$

of finite dimensions $p(h_1) = dim(U_{h_1})$ and $q(h_2) = dim(Y_{h_2})$, and denote the corresponding orthogonal projections onto these subspaces by $P_{U,h_1} \in \mathscr{L}(U)$ and $P_{Y,h_2} \in$

$\mathscr{L}(Y)$. Correspondingly, we choose for the discretization in time two families $\{\mathcal{R}_{\tau_1}\}_{\tau_1>0}$ and $\{\mathcal{S}_{\tau_2}\}_{\tau_2>0}$ of subspaces

$$\mathcal{R}_{\tau_1} \subset L^2(0,T), \qquad \mathcal{S}_{\tau_2} \subset L^2(0,T)$$

of finite dimensions $r(\tau_1) = dim(\mathcal{R}_{\tau_1})$ and $s(\tau_2) = dim(\mathcal{S}_{\tau_2})$, and denote the corresponding orthogonal projections onto these subspaces by $P_{\mathcal{R},\tau_1} \in \mathscr{L}(L^2(0,T))$ and $P_{\mathcal{S},\tau_2} \in \mathscr{L}(L^2(0,T))$. The projections P_{U,h_1} and $P_{\mathcal{R},\tau_1}$ induce in a natural way operators $\mathbb{P}_{U,h_1}$, $\mathbb{P}_{\mathcal{R},\tau_1} \in \mathscr{L}(\mathcal{U})$ via

$$(\mathbb{P}_{U,h_1}u)(t;\theta) = (P_{U,h_1}u(t;\cdot))(\theta), \quad (\mathbb{P}_{\mathcal{R},\tau_1}u)(t;\theta) = (P_{\mathcal{R},\tau_1}u(\cdot;\theta))(t) \tag{11}$$

for almost every $t \in [0,T]$ and $\theta \in \Theta$. The operators $\mathbb{P}_{Y,h_2}, \mathbb{P}_{\mathcal{R},\tau_1} \in \mathscr{L}(\mathcal{Y})$ arise similarly. Most of the time, we will omit the discretization parameters in $p = p(h_1)$, $q = q(h_2)$, $r = r(\tau_1)$ and $s = s(\tau_2)$.

Lemma 2. *$\mathbb{P}_{U,h_1}$ and $\mathbb{P}_{\mathcal{R},\tau_1}$ are commuting orthogonal projections and*

$$\mathbb{P}_{\mathcal{U},h_1,\tau_1} := \mathbb{P}_{U,h_1}\mathbb{P}_{\mathcal{R},\tau_1} = \mathbb{P}_{\mathcal{R},\tau_1}\mathbb{P}_{U,h_1}$$

is the orthogonal $\mathscr{L}(\mathcal{U})$-projection onto the subspace

$$\mathcal{U}_{h_1,\tau_1} = \{u \in \mathcal{U} : u(t;\cdot) \in U_{h_1},\, u(\cdot;\theta) \in \mathcal{R}_{\tau_1} \quad \forall a.e.\, t \in [0,T], \theta \in \Theta\}.$$

In a similar way, $\mathbb{P}_{\mathcal{Y},h_2,\tau_2} = \mathbb{P}_{Y,h_2}\mathbb{P}_{\mathcal{S},\tau_2} = \mathbb{P}_{\mathcal{S},\tau_2}\mathbb{P}_{Y,h_2}$ is the orthogonal $\mathscr{L}(\mathcal{Y})$-projection onto the subspace

$$\mathcal{Y}_{h_2,\tau_2} = \{y \in \mathcal{Y} : y(t;\cdot) \in Y_{h_2},\, y(\cdot;\xi) \in \mathcal{S}_{\tau_2} \quad \forall a.e.\, t \in [0,T], \xi \in \Xi\}.$$

Proof. We introduce an orthonormal basis $\{\mu_i\}_{i=1,\dots,p}$ of U_{h_1} and an orthonormal basis $\{\phi_i\}_{i=1,\dots,r}$ of $\mathcal{R}_{\tau_1}$. Then, for $u \in \mathcal{U}$,

$$\tilde{u}(t;\theta) = (\mathbb{P}_{\mathcal{U},h_1}u)(t;\theta) = \sum_{i=1}^{p} \tilde{u}_i(t)\mu_i(\theta), \qquad \bar{u}(t;\theta) = (\mathbb{P}_{\mathcal{R},\tau_1}u)(t;\theta) = \sum_{i=1}^{r} \bar{u}_i(\theta)\phi_i(t),$$

with $\tilde{u}_i(t) = (u(t;\cdot),\mu_i(\cdot))_U$ and $\bar{u}_i(\theta) = (u(\cdot;\theta),\phi_i(\cdot))_{L^2(0,T)}$. Obviously, $\mathbb{P}_{\mathcal{U},h_1}$ and $\mathbb{P}_{\mathcal{R},\tau_1}$ are orthogonal projections. $\mathbb{P}_{\mathcal{U},h_1}$ and $\mathbb{P}_{\mathcal{R},\tau_1}$ commute, since

$$\mathbb{P}_{\mathcal{R},\tau_1}\tilde{u} = \sum_{i=1}^{r}\sum_{k=1}^{p} \underbrace{(\tilde{u}_k,\phi_i)_{L^2(0,T)}}_{=:(\tilde{u}_k)_i}\mu_k\phi_i, \qquad \mathbb{P}_{\mathcal{U},h_1}\bar{u} = \sum_{i=1}^{r}\sum_{k=1}^{p} \underbrace{(\bar{u}_i,\mu_k)_U}_{=:(\bar{u}_i)_k}\mu_k\phi_i,$$

and by Fubini's theorem,

$$(\tilde{u}_k)_i = \int_0^T (u(t;\cdot),\mu_k(\cdot))_U\phi_i(t)\,dt = \int_\Theta (u(\cdot;\theta),\phi_i(\cdot))_{L^2(0,T)}\mu_k(\theta)\,d\theta = (\bar{u}_i)_k.$$

Hence, $\mathbb{P}^2_{\mathcal{U},h_1,\tau_1} = \mathbb{P}_{\mathcal{U},h_1,\tau_1}$ and $\mathbb{P}^*_{\mathcal{U},h_1,\tau_1} = \mathbb{P}_{\mathcal{U},h_1,\tau_1}$. □

We introduce an approximation multi-parameter $\mathbf{h} = (h_1,\tau_1,h_2,\tau_2) \in \mathbb{R}^4_+$ and define the approximation $\mathbb{G}_S = \mathbb{G}_S(\mathbf{h})$ of $\mathbb{G}$ via

$$\mathbb{G}_S(\mathbf{h}) = \mathbb{P}_{\mathcal{Y},h_2,\tau_2}\mathbb{G}\mathbb{P}_{\mathcal{U},h_1,\tau_1}. \tag{12}$$

4.2. Matrix representation of $\mathbb{G}_S$

We introduce for $h_1, h_2, \tau_1, \tau_2 > 0$ families of bases

$$\begin{aligned} &\{\mu_1, \dots, \mu_p\} \text{ of } U_{h_1}, \qquad &&\{\nu_1, \dots, \nu_q\} \text{ of } Y_{h_2}, \\ &\{\phi_1, \dots, \phi_r\} \text{ of } \mathcal{R}_{\tau_1}, \qquad &&\{\psi_1, \dots, \psi_s\} \text{ of } \mathcal{S}_{\tau_2}. \end{aligned}$$

Here $\mu_j = \mu_j^{(h_1)}$, $\nu_j = \nu_j^{(h_2)}$, $\phi_j = \phi_j^{(\tau_1)}$ and $\psi_j = \psi_j^{(\tau_2)}$, but we omit the superscripts for ease of notation. We define the mass matrix $\mathbf{M}_{U,h_1} \in \mathbb{R}^{p,p}$ by

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

and define the mass matrices $\mathbf{M}_{Y,h_2} \in \mathbb{R}^{q,q}$, $\mathbf{M}_{\mathcal{R},\tau_1} \in \mathbb{R}^{r,r}$ and $\mathbf{M}_{\mathcal{S},\tau_2} \in \mathbb{R}^{s,s}$ correspondingly. The mass-matrices induce, for instance via

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

weighted scalar products and corresponding norms in the respective spaces, which we indicate by a subscript w, like $\mathbb{R}^p_w$ with $(\cdot,\cdot)_{p;w}$ and $\|\cdot\|_{p;w}$, in contrast to the canonical spaces like $\mathbb{R}^p$ with $(\cdot,\cdot)_p$ and $\|\cdot\|_p$. Denoting by $\lambda_{\min}(\mathbf{M}_{U,h_1})$ and $\lambda_{\max}(\mathbf{M}_{U,h_1})$ the smallest and the largest eigenvalue of $\mathbf{M}_{U,h_1}$, respectively, we have

$$\sqrt{\lambda_{\min}(\mathbf{M}_{U,h_1})}\|\alpha\|_p \le \|\alpha\|_{p;w} \le \sqrt{\lambda_{\max}(\mathbf{M}_{U,h_1})}\|\alpha\|_p.$$

We define the *unitary* coordinate isomorphisms $\kappa_{U,h_1} : U_{h_1} \to \mathbb{R}^p_w$, $\kappa_{Y,h_2} : Y_{h_2} \to \mathbb{R}^q_w$, $\kappa_{\mathcal{R},\tau_1} : \mathcal{R}_{\tau_1} \to \mathbb{R}^r_w$ and $\kappa_{\mathcal{S},\tau_2} : \mathcal{S}_{\tau_2} \to \mathbb{R}^s_w$, for instance via

$$\kappa_{U,h_1} : U_{h_1} \to \mathbb{R}^p_w, \qquad u = \sum_{j=1}^p \alpha_j \mu_j \mapsto [\alpha_j]_j = \mathbf{M}^{-1}_{U,h_1}[(u, \mu_j)_U]_j.$$

We observe that

$$\{\mu_i^{(h_1)} \phi_j^{(\tau_1)}\}_{j=1,\dots,r}^{i=1,\dots,p} \subset \mathcal{U} \quad \text{and} \quad \{\nu_i^{(h_2)} \psi_j^{(\tau_2)}\}_{j=1,\dots,s}^{i=1,\dots,q} \subset \mathcal{Y}$$

are bases of $\mathcal{U}_{h_1,\tau_1}$ and $\mathcal{Y}_{h_2,\tau_2}$, respectively. We then define the block-structured mass matrices

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

which induce weighted scalar products $(\cdot,\cdot)_{pr;w}$ in $\mathbb{R}^{pr}_w$ and $(\cdot,\cdot)_{qs;w}$ in $\mathbb{R}^{qs}_w$, and corresponding norms $\|\cdot\|_{pr;w}$ and $\|\cdot\|_{qs;w}$. We again define *unitary* coordinate isomorphisms $\kappa_{\mathcal{U},h_1,\tau_1} : \mathcal{U}_{h_1,\tau_1} \to \mathbb{R}^{pr}_w$ and $\kappa_{\mathcal{Y},h_2,\tau_2} : \mathcal{Y}_{h_2,\tau_2} \to \mathbb{R}^{qs}_w$ by

$$\begin{aligned} u(t;\theta) &= \sum_{k=1}^p \sum_{i=1}^r \mathbf{u}_i^k \phi_i(t) \mu_k(\theta) \mapsto [\mathbf{u}]_i^k = \mathbf{M}^{-1}_{\mathcal{U},h_1,\tau_1}[(u, \phi_i \mu_k)_{\mathcal{U}}]_i^k, \\ y(t;\xi) &= \sum_{l=1}^q \sum_{j=1}^s \mathbf{y}_j^l \psi_j(t) \nu_k(\xi) \mapsto [\mathbf{y}]_j^l = \mathbf{M}^{-1}_{\mathcal{Y},h_2,\tau_2}[(y, \psi_j \nu_l)_{\mathcal{Y}}]_j^l. \end{aligned}$$

Here $[\mathbf{u}]_i^k$ denotes a block structured vector with p blocks $\mathbf{u}^k \in \mathbb{R}^r$, other block structured vectors are defined similarly.

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

$$\mathbf{G} = \mathbf{G}(\mathbf{h}) = \kappa_{\mathcal{Y}} \mathbb{P}_{\mathcal{Y}} \mathbb{G} \mathbb{P}_{\mathcal{U}} \kappa_{\mathcal{U}}^{-1} \in \mathbb{R}^{qs,pr}, \tag{13}$$

where the dependencies on h_1, τ_1, h_2, τ_2 have been partially omitted. Defining the block-structured matrix

$$\mathbf{H}(\mathbf{h}) = \mathbf{M}_{\mathcal{Y},h_2,\tau_2} \mathbf{G}(\mathbf{h}) = [\mathbf{H}_{ij}^{kl}]_{ij}^{kl}$$

with blocks $\mathbf{H}^{kl} = [\mathbf{H}_{ij}^{kl}]_{ij} \in \mathbb{R}^{s,r}$ for $k = 1, \dots, q$ and $l = 1, \dots, p$, we obtain the element-wise representation

$$\mathbf{H}_{ij}^{kl} = [\mathbf{M}_{\mathcal{Y}} \kappa_{\mathcal{Y}} \mathbb{P}_{\mathcal{Y}} \mathbb{G}(\mu_l \phi_j)]_i^k = (\nu_k \psi_i, \mathbb{G}(\mu_l \phi_j))_{\mathcal{Y}}, \tag{14}$$

and introduce the norm

$$\|\mathbf{G}(\mathbf{h})\|_{\mathbf{h}} := \sup_{\mathbf{u} \in \mathbb{R}^{pr}} \frac{\|\mathbf{G}\mathbf{u}\|_{qs;w}}{\|\mathbf{u}\|_{pr;w}} = \|\mathbf{M}_{\mathcal{Y},h_2,\tau_2}^{1/2} \mathbf{G}(\mathbf{h}) \mathbf{M}_{\mathcal{U},h_1,\tau_1}^{-1/2}\|_{qs,pr}. \tag{15}$$

Lemma 3. *For all $\mathbf{h} = (h_1, \tau_1, h_2, \tau_2) \in \mathbb{R}_+^4$, we have*

$$\|\mathbf{G}(\mathbf{h})\|_{\mathbf{h}} = \|\mathbb{G}_S(\mathbf{h})\|_{\mathscr{L}(\mathcal{U},\mathcal{Y})} \le \|\mathbb{G}\|_{\mathscr{L}(\mathcal{U},\mathcal{Y})}. \tag{16}$$

Assume, in addition, that 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

$$\mathcal{U}_{h_1,\tau_1} \subset \mathcal{U}_{h_1',\tau_1'} \qquad \text{for } h_1' < h_1 \quad \text{and} \quad \tau_1' < \tau_1, \tag{17a}$$

$$\mathcal{Y}_{h_2,\tau_2} \subset \mathcal{Y}_{h_2',\tau_2'} \qquad \text{for } h_2' < h_2 \quad \text{and} \quad \tau_2' < \tau_2. \tag{17b}$$

In this case, $\|\mathbf{G}(\mathbf{h})\|_{\mathbf{h}}$ monotonically increases if $\mathbf{h} \in \mathbb{R}_+^4$ decreases with respect to the partial ordering

$$(h_1', \tau_1', h_2', \tau_2') \le (h_1, \tau_1, h_2, \tau_2) :\Leftrightarrow h_1' \le h_1, \quad \tau_1' \le \tau_1, \quad h_2' \le h_2, \quad \tau_2' \le \tau_2,$$

and $\|\mathbf{G}(\mathbf{h})\|_{\mathbf{h}}$ is convergent for $\mathbf{h} \searrow 0$.

Proof. First, we calculate

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

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

$$\begin{aligned}
\|\mathbb{G}_S u\|_{\mathcal{Y}} = \|\kappa_{\mathcal{Y},h_2,\tau_2}^{-1} \mathbf{G} \kappa_{\mathcal{U},h_1,\tau_1} \mathbb{P}_{\mathcal{U},h_1,\tau_1} u\|_{\mathcal{Y}} = \|\mathbf{G}\mathbf{u}\|_{qs;w} \le \|\mathbf{G}\|_{\mathbf{h}} \|\mathbf{u}\|_{pr;w} = \|\mathbf{G}\|_{\mathbf{h}} \|u\|_{\mathcal{U}}, \\
\|\mathbf{G}\mathbf{u}\|_{qs;w} \le \|\kappa_{\mathcal{Y},h_2,\tau_2} \mathbb{G}_S \kappa_{\mathcal{U},h_1,\tau_1}^{-1} \mathbf{u}\|_{qs;w} = \|\mathbb{G}_S u\|_{\mathcal{Y}} \le \|\mathbb{G}_S\|_{\mathscr{L}(\mathcal{U},\mathcal{Y}} \|\mathbf{u}\|_{pr;w}
\end{aligned}$$

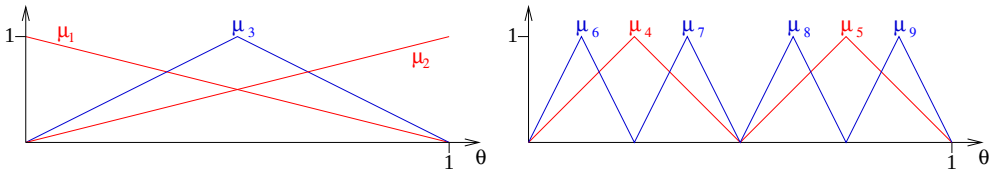

Figure 1: First nine basis functions of a hierarchical basis for $L^2(0,1)$-subspaces of piecewise linear functions.

and, thus, $\|\mathbb{G}_S\|_{\mathscr{L}(\mathcal{U},\mathcal{Y})} = \|\mathbf{G}\|_{\mathbf{h}}$. Finally, we assume that $\mathbf{h}' < \mathbf{h}$ and that (17) holds. Since $\|\mathbb{P}_{\mathcal{Y},h_2,\tau_2} y\|_{\mathcal{Y}} \leq \|\mathbb{P}_{\mathcal{Y},h_2',\tau_2'} y\|_{\mathcal{Y}}$ for all $y \in \mathcal{Y}$, we have

$$\|\mathbb{G}_S(\mathbf{h})\|_{\mathbf{h}} = \sup_{u \in \mathcal{U}_{h_1,\tau_1}} \frac{\|\mathbb{P}_{\mathcal{Y},h_2,\tau_2} \mathbb{G} u\|_{\mathcal{Y}}}{\|u\|_{\mathcal{U}}} \leq \sup_{u \in \mathcal{U}_{h_1',\tau_1'}} \frac{\|\mathbb{P}_{\mathcal{Y},h_2',\tau_2'} \mathbb{G} u\|_{\mathcal{Y}}}{\|u\|_{\mathcal{U}}} = \|\mathbb{G}_S(\mathbf{h}')\|_{\mathbf{h}'}$$

Since $\|\mathbb{G}_S(\mathbf{h})\|_{\mathbf{h}}$ is bounded above by $\|\mathbb{G}\|_{\mathscr{L}(\mathcal{U},\mathcal{Y})}$, the convergence is assured. □

4.3. An example for signal discretizations

As an example, consider the case $U = Y = L^2(0,1)$, and 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. One can easily assure nested spaced as in (17), for instance by taking hierarchical finite element bases for U_{h_1} and Y_{h_2} and Haar-wavelet bases for $\mathcal{R}_{\tau_1}$ and $\mathcal{S}_{\tau_2}$.

To get the (space) discretization of U, define for $p \in \mathbb{N}$, $p \geq 2$ and $h_1(p) = 1/(p-1)$ the family of equidistant partitions $\{\mathscr{T}_{h_1}\}_{h_1}$ of $(0,1]$,

$$\mathscr{T}_{h_1} = \{I_n\}, \qquad \text{with } I_n = ((n-1)h_1, nh_1], \quad n = 1, \ldots, p-1.$$

The space U_{h_1} of continuous piecewise linear functions with respect to $\mathscr{T}_{h_1}$ is spanned by the nodal basis

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

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 $\hat{\mu}_l$, as in Fig. 1, see e.g. [48], [49].

Then,

$$U_{h_1} = \operatorname{span}\{\hat{\mu}_1, \ldots, \hat{\mu}_{p(h_1)}\}, \quad \text{for all } h_1 \in \{2^{-n}\}_{n \in \mathbb{N}_0}$$

with basis functions $\hat{\mu}_k$ independent of h_1.

Remark 2. As we will see later, the eigenvalues of the mass matrices $\mathbf{M}_{U,h_1}$ and $\mathbf{M}_{Y,h_2}$ occur in some versions of the error estimates. Choosing, for instance, a *nodal* basis in U_{h_1}, the eigenvalues of the corresponding mass matrix $\mathbf{M}_{U,h_1}$ satisfy,

$$c_\lambda h_1 \leq \lambda_{\min}(\mathbf{M}_{U,h_1}) \leq \lambda_{\max}(\mathbf{M}_{U,h_1}) \leq C_\lambda h_1, \tag{18}$$

with two positive constants c_λ, C_λ independent of h_1, see e.g. [45]. Choosing a *hierarchical* basis, the eigenvalues of the mass matrix satisfy

$$c'_\lambda h_1 \leq \lambda_{\min}(\mathbf{M}_{U,h_1}) \leq \lambda_{\max}(\mathbf{M}_{U,h_1}) \leq C'_\lambda. \tag{19}$$

We note that (19) is only based on numerical experiments, see Appendix A.1, where also numerical estimates for c_λ, C_λ, c'_λ and C'_λ are calculated.

For the (time) discretization of $L^2(0,T)$, define for $r \in \mathbb{N}$ and $\tau_1 = 1/r$ the family of equidistant partitions $\{\Gamma_{\tau_1}\}_{\tau_1}$ of $(0,T]$.

$$\Gamma_{h_1} = \{I_n\}, \qquad \text{with } I_n = ((n-1)\tau_1, n\tau_1], \quad n = 1, \dots, r.$$

The space $\mathcal{R}_{\tau_1}$ of piecewise constant functions with respect to Γ_{τ_1} can be spanned by the nodal basis

$$\{\phi_1^{(\tau_1)}(t), \dots, \phi_r^{(\tau_1)}(t)\} \quad \text{with } \phi_j^{(\tau_1)}(t) = \chi_{I_n}(t), \qquad j = 1, \dots, r$$

and mass matrix $\mathbf{M}_{Y,\tau_1} = \tau_1 Id_r$. The spaces are nested by requiring $\tau_1 \in \{2^{-n}\}_{n\in\mathbb{N}_0}$. One obtains a hierarchical basis for $\mathcal{R}_{\tau_1}$ by choosing ϕ_i as Haar-wavelets $\hat{\phi}_i$, i.e.

$$\hat{\phi}_i(t) = \frac{1}{\sqrt{T}} \begin{cases} \chi_{[0,T]}(t) & , i = 1, \\ 2^{n/2}\hat{\phi}(2^n t/T - m) & , i = 2^{n+1} + m + 1, \end{cases}$$

with

$$\hat{\phi}(t) = \begin{cases} 1 & , 0 < t \leq 1/2, \\ -1 & , 1/ < t \leq 1, \\ 0 & , \text{otherwise.} \end{cases}$$

In this case, $\mathbf{M}_{\mathcal{R},\tau_1} = Id_r$ and

$$\mathcal{R}_{\tau_1} = \operatorname{span}\{\hat{\phi}_1, \dots, \hat{\phi}_{r(\tau_1)}\}, \quad \text{for all } \tau_1 \in \{2^{-n}\}_{n\in\mathbb{N}_0}.$$

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's inequality shows that there exist positive constants c_U and $c_\mathcal{R}$, independent of h_1, τ_1 and T, such that

$$\|u - P_{U_{h_1}} u\|_{L^2(0,1)} \leq \|u - \mathbb{I}_{U_{h_1}} u\|_{L^2(0,1)} \leq c_U h_1^2 \|u''\|_{L^2(0,1)} \quad \text{for } u \in H^2(0,1), \tag{20a}$$

$$\|v - P_{\mathcal{R}_{\tau_1}} v\|_{L^2(0,T)} \leq \|v - \mathbb{I}_{\mathcal{R}_{\tau_1}} v\|_{L^2(0,T)} \leq c_\mathcal{R} \tau_1 \|\dot{v}\|_{L^2(0,T)} \quad \text{for } v \in H^1(0,T), \tag{20b}$$

see e.g. [12], [50]. Here $\mathbb{I}_{U_{h_1}} u$ is the piecewise linear interpolant of u with respect to $\mathscr{T}_{h_1}$ and $\mathbb{I}_{\mathcal{R}_{\tau_1}} v$ is the piecewise constant interpolant of v with respect to Γ_{τ_1}. By Fubini's theorem, it follows that the induced projection $\mathbb{P}_{\mathcal{U},h_1,\tau_1} = \mathbb{P}_{U,h_1}\mathbb{P}_{\mathcal{R},\tau_1}$ onto

$$\mathcal{U}_{h_1,\tau_1} = \{u \in \mathcal{U},\, u_{|I_n} \equiv u^{(n)},\, u^{(n)} \in U_{h_1}, \quad n = 1, \dots, N\}$$

satisfies

$$\|u - \mathbb{P}_{\mathcal{U},h_1,\tau_1} u\|_\mathcal{U} \leq (c_U h_1^2 + c_\mathcal{R}\tau_1)\|u\|_{\mathcal{U}_s} \quad \text{for all } u \in \mathcal{U}_s, \tag{21}$$

with

$$\mathcal{U}_s = H^1(0,T;L^2(0,1)) \cap L^2(0,T;H^2(0,1)),$$

and $\|u\|_{\mathcal{U}_s} = \|u\|_{H^1(0,T;L^2(0,1))} + \|u\|_{L^2(0,T;H^2(0,1))}$. We define Y_{h_2},$\mathcal{R}_{\tau_2}$ and $\mathcal{Y}_{h_2,\tau_2}$ accordingly and a corresponding estimate as (21) holds for the projection $\mathbb{P}_{\mathcal{Y},h_2,\tau_2} y$ of elements $y \in \mathcal{Y}_s = \mathcal{U}_s$. Note that higher approximation orders can be achieved, if ansatz functions of higher polynomial degree are used and if the input and output signals exhibit corresponding higher regularity in space and time, see e.g. [12], [50].

Remark 3. In the simple 1D setting the proof of (20) directly provides the numerical values $c_U = c_Y = 1/2$ and $c_{\mathcal{R}} = c_{\mathcal{S}} = 1/\sqrt{2}$. For more complex domains Θ and Ξ the interpolation constants have to be estimated numerically, see e.g. [18].

4.4. Signal approximation error estimates

We first consider the signal error ϵ_D with respect to the $\mathscr{L}(\mathcal{U},\mathcal{Y})$-norm.

Lemma 4. *The signal approximation error in the norm $\|\cdot\|_{\mathscr{L}(\mathcal{U},\mathcal{Y})}$ is given by*

$$\epsilon_s := \|\mathbb{G} - \mathbb{G}_S\|_{\mathscr{L}(\mathcal{U},\mathcal{Y})} = \max_{u \in \mathcal{U}_{h_1,\tau_1}} \frac{\|(I - \mathbb{P}_{\mathcal{Y},h_2,\tau_2})\mathbb{G}u)\|_{\mathcal{Y}}}{\|u\|_{\mathcal{U}}} + \sup_{u \in \mathcal{U}'_{h_1,\tau_1}} \frac{\|\mathbb{G}u\|_{\mathcal{Y}}}{\|u\|_{\mathcal{U}}}, \tag{22}$$

where $\mathcal{U}'_{h_1,\tau_1} = ker \mathbb{P}_{\mathcal{U},h_1,\tau_1}$.

Proof. We decompose $\mathbb{G} - \mathbb{G}_S$ as

$$\mathbb{G} - \mathbb{G}_S = \mathbb{G}(\mathbb{I} - \mathbb{P}_{\mathcal{U},h_1,\tau_1}) + (\mathbb{I} - \mathbb{P}_{\mathcal{Y},h_2,\tau_2})\mathbb{G}\mathbb{P}_{\mathcal{U},h_1,\tau_1}, \tag{23}$$

and observe that

$$\|(\mathbb{I} - \mathbb{P}_{\mathcal{Y},h_2,\tau_2})\mathbb{G}\mathbb{P}_{\mathcal{U},h_1,\tau_1}\|_{\mathscr{L}(\mathcal{U}_r,\mathcal{Y})} = \max_{u \in \mathcal{U}_{h_1,\tau_1}} \frac{\|(I - \mathbb{P}_{\mathcal{Y},h_2,\tau_2})\mathbb{G}u)\|_{\mathcal{Y}}}{\|u\|_{\mathcal{U}}}, \tag{24}$$

$$\|\mathbb{G}(\mathbb{I} - \mathbb{P}_{\mathcal{U},h_1,\tau_1})\|_{\mathscr{L}(\mathcal{Y})} = \sup_{u \in \mathcal{U}'_{r;h_1,\tau_1}} \frac{\|\mathbb{G}u\|_{\mathcal{Y}}}{\|u\|_{\mathcal{U}}}. \tag{25}$$

Applying the triangle inequality to (23), and considering a sequence $u_j = u^* + u'_j$, where u^* is the maximizer of (24) and $(u'_j)_j \subset \mathcal{U}_{h_1,\tau_1}$ is a supremal sequence in (25), shows (22). □

We note that the usual approximation conditions

$$\forall u \in \mathcal{U}: \quad dist(u,\mathcal{U}_{h_1,\tau_1}) \to 0 \quad \text{if } h_1, \tau_1 \to 0,$$

$$\forall y \in \mathcal{Y}: \quad dist(y,\mathcal{Y}_{h_2,\tau_2}) \to 0 \quad \text{if } h_2, \tau_2 \to 0,$$

imply $\|(\mathbb{G} - \mathbb{G}_S)u\|_{\mathcal{Y}} \to 0$ for every $u \in \mathcal{U}$, but are *not* sufficient to guarantee a uniform convergence $\|\mathbb{G} - \mathbb{G}_S\|_{\mathscr{L}(\mathcal{U},\mathcal{Y})} \to 0$. Considering, for instance, the identity operator $\mathbb{G} = Id \in \mathscr{L}(\mathcal{U},\mathcal{Y})$ in the case $\mathcal{U} = \mathcal{Y}$, the second term in (22) equals one for every

finite-dimensional $\mathcal{U}_{h_1,\tau_1}$. Similar effects can be expected for operators of the form $\mathbb{G} = \mathbb{G}_1 + \alpha\mathbb{P}_{\bar{\mathcal{U}}}$, where $\mathbb{G}_1 \in \mathscr{L}(\mathcal{U}, \mathcal{Y} = \mathcal{U})$ and $\alpha\mathbb{P}_{\bar{\mathcal{U}}}$ is a feedthrough operator for all input signals from a infinite-dimensional subspace $\bar{\mathcal{U}}$.

Hence, we can only hope for a good approximation in $\|\cdot\|_{\mathscr{L}(\mathcal{U},\mathcal{Y})}$ if the subspaces $\mathcal{U}_{h_1,\tau_1}$ and $\mathcal{Y}_{h_2,\tau_2}$ can be chosen *specifically* for $\mathbb{G}$ such that output signals from input signals $u \in \mathcal{U}_{h_1,\tau_1}$ are well approximated in $\mathcal{Y}_{h_2,\tau_2}$ and that neglected input signal components in $\mathcal{U}'_{h_1,\tau_1}$ only lead to small output signals. For instance, if $\mathbb{G} \in \mathscr{L}(\mathcal{U}, \mathcal{Y})$ is a compact operator, then there exist orthonormal systems $\{\hat{u}_1, \hat{u}_2 \dots\}$ of $\mathcal{U}$ and $\{\hat{y}_1, \hat{y}_2, \dots\}$ of $\mathcal{Y}$ and nonnegative numbers $\sigma_1 \geq \sigma_2 \geq \dots$ with $\sigma_k \to 0$ such that

$$\mathbb{G}u = \sum_{k=1}^{\infty} \sigma_k (u, \hat{u}_k)_{\mathcal{U}} \hat{y}_k, \quad \text{for all } u \in \mathcal{U},$$

see e.g. [47]. Choosing $\mathcal{U}_{h_1,\tau_1}$ and $\mathcal{Y}_{h_2,\tau_2}$ as the span of $\hat{u}_1, \dots, \hat{u}_r$ and $\hat{y}_1, \dots, \hat{y}_s$, respectively, with $s = r$ and $r \in \mathbb{N}$, we obtain a very efficient approximation $\mathbb{G}_S$ of $\mathbb{G}$ with $\|\mathbb{G} - \mathbb{G}_S\|_{\mathscr{L}(\mathcal{U},\mathcal{Y})} \leq \sigma_{r+1}$.

Next we consider the case where less specific information about $\mathbb{G}$ is available and we only know that

$$\mathbb{G}_{|\mathcal{U}_s} \in \mathscr{L}(\mathcal{U}_s, \mathcal{Y}_s), \tag{26}$$

with spaces of higher regularity in space and time

$$\mathcal{U}_s = H^{\alpha_1}(0, T; L^2(\Theta)) \cap L^2(0, T; H^{\beta_1}(\Theta)), \tag{27a}$$

$$\mathcal{Y}_s = H^{\alpha_2}(0, T; L^2(\Xi)) \cap L^2(0, T; H^{\beta_2}(\Xi)), \tag{27b}$$

where α_1, α_2, β_1, $\beta_2 \in \mathbb{N}$. Such a situation holds, for instance, for the heat equation, cf. (10). Corresponding norms can be defined by

$$\|u\|_{\mathcal{U}_s} = \|u\|_{H^{\alpha_1}(0,T;L^2(\Theta))} + \|u\|_{L^2(0,T;H^{\beta_1}(\Theta))},$$
$$\|y\|_{\mathcal{Y}_s} = \|y\|_{H^{\alpha_2}(0,T;L^2(\Xi))} + \|y\|_{L^2(0,T;H^{\beta_2}(\Xi))}.$$

Choosing U_{h_1}, Y_{h_2}, $\mathcal{R}_{\tau_1}$ and $\mathcal{S}_{\tau_2}$ for instance as spaces of piecewise polynomial functions of appropriate degrees one can achieve estimates

$$\|u - \mathbb{P}_{\mathcal{U},h_1,\tau_1} u\|_{\mathcal{U}} \leq (c_{\mathcal{R}} \tau_1^{\alpha_1} + c_U h_1^{\beta_1}) \|u\|_{\mathcal{U}_s}, \qquad u \in \mathcal{U}_s, \tag{28a}$$

$$\|y - \mathbb{P}_{\mathcal{Y},h_2,\tau_2} y\|_{\mathcal{Y}} \leq (c_{\mathcal{S}} \tau_2^{\alpha_2} + c_Y h_2^{\beta_2}) \|y\|_{\mathcal{Y}_s}, \qquad y \in \mathcal{Y}_s, \tag{28b}$$

with positive interpolation constants $c_{\mathcal{R}}$, $c_{\mathcal{S}}$, c_U and c_Y, cf. (20) and refer e.g. to [12] for interpolation theory in Sobolev spaces in the case of more general settings.

Proposition 1. *Assume that $\mathbb{G}$ satisfies (26) with some α_1, α_2, β_1, $\beta_2 \in \mathbb{N}$. Assume further that U_{h_1}, Y_{h_2}, $\mathcal{R}_{\tau_1}$ and $\mathcal{S}_{\tau_2}$ are chosen such that (28) holds with positive interpolation constants c_U, c_Y, $c_{\mathcal{R}}$ and $c_{\mathcal{S}}$. In this case we have*

$$\|\mathbb{G} - \mathbb{G}_S\|_{\mathscr{L}(\mathcal{U}_s,\mathcal{Y})} \leq c'_{\mathcal{R}} \tau_1^{\alpha_1} + c'_U h_1^{\beta_1} + c'_{\mathcal{S}} \tau_2^{\alpha_2} + c'_Y h_2^{\beta_2}, \tag{29}$$

with $c'_U = \|\mathbb{G}\|_{\mathscr{L}(\mathcal{U},\mathcal{Y})} c_U$, $c'_Y = \|\mathbb{G}\|_{\mathscr{L}(\mathcal{U}_s,\mathcal{Y}_s)} c_Y$, $c'_{\mathcal{R}} = \|\mathbb{G}\|_{\mathscr{L}(\mathcal{U},\mathcal{Y})} c_{\mathcal{R}}$ and $c'_{\mathcal{S}} = \|\mathbb{G}\|_{\mathscr{L}(\mathcal{U}_s,\mathcal{Y}_s)} c_{\mathcal{S}}$, being independent of h_1, h_2, τ_1, τ_2 and T.

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

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

□

For the heat equation in Example 1 we obtain the following corollary.

Corollary 1. *Consider the heat equation in Example 1 with $U = Y = L^2(0,1)$ and input signals restricted to*

$$\mathcal{U}_s = \mathcal{Y}_s = L^2(0,T;H^2(0,1)) \cap H^1(0,T;L^2(0,1)) \tag{30}$$

and assume that

$$C_{|H^2(\Omega)} \in \mathscr{L}(H^2(\Omega), H^2(0,1)).$$

If we choose the subspaces $\mathcal{U}_{h_1,\tau_1}$ and $\mathcal{Y}_{h_2,\tau_2}$ as spaces of functions which are piecewise constant in time and continuous piecewise linear in space with respect to regular grids of grid sizes τ_1, τ_2, h_1 and h_2, respectively (cf. Section 4.3), then the corresponding approximation $\mathbb{G}_S$ satisfies

$$\|\mathbb{G} - \mathbb{G}_S\|_{\mathscr{L}(\mathcal{U}_s,\mathcal{Y})} \leq c'_{\mathcal{R}}\tau_1 + c'_U h_1^2 + c'_{\mathcal{S}}\tau_2 + c'_Y h_2^2 \tag{31}$$

with $c'_U = \|\mathbb{G}\|_{\mathscr{L}(\mathcal{U},\mathcal{Y})}/2$, $c'_Y = \|\mathbb{G}\|_{\mathscr{L}(\mathcal{U}_s,\mathcal{Y}_s)}/2$, $c'_{\mathcal{R}} = \|\mathbb{G}\|_{\mathscr{L}(\mathcal{U},\mathcal{Y})}/\sqrt{2}$ and $c'_{\mathcal{S}} = \|\mathbb{G}\|_{\mathscr{L}(\mathcal{U}_s,\mathcal{Y}_s)}/\sqrt{2}$, being independent of h_1, h_2, τ_1, τ_2 and T.

Proof. In view of (10) and the following remarks, condition (26) is satisfied with $\alpha_1 = \alpha_2 = 1$ and $\beta_1 = \beta_2 = 2$. (28) holds in view of (21) and Remark 3 with $c_U = c_Y = 1/2$ and $c_{\mathcal{R}} = c_{\mathcal{S}} = 1/\sqrt{2}$, and hence Proposition 1 applies. □

Remark 4. Considering *unconstrained* optimization problems for parabolic equations with Tychonov regularization, it is known that the optimal control u_* is smooth in the sense of (30), see e.g. [44], such that the consideration of restrictions of $\mathcal{U}$ to $\mathcal{U}_s$ is reasonable if the approximation $\mathbb{G}_{DS}$ is later used to approximate u_*.

Remark 5. Property (26) can be expected for systems which either have smooth states (like the heat equation) or which at least preserve the regularity of the input signals (like the wave equation).

5. System dynamics approximation

This section deals with the efficient approximation of $\mathbb{G}_S$ respectively of its matrix representation $\mathbf{G} = \mathbf{M}_{\mathcal{Y}}^{-1}\mathbf{H}$ defined in (14). In general, $\mathbf{G}$ can be calculated by testing the output corresponding to *each* input basis function of $\mathcal{U}_{h_1,\tau_1}$ against *all* output

basis functions of $\mathcal{Y}_{h_2,\tau_2}$. If we consider time-invariant systems with bounded control and observation, then the i/o-maps $\mathbb{G}$ can be represented as a convolution with a kernel function $K(t)$ as in (5). In this case, the approximation of $\mathbf{G}$ reduces to the approximation of $K(t) \in \mathscr{L}(\mathcal{U}, \mathcal{Y})$ by $\tilde{K}(t) \in \mathscr{L}(U_{h_1}, Y_{h_2})$ such that only the input and output basis functions with respect to the spatial distribution have to be considered and the numerical burden is significantly reduced.

5.1. Kernel function approximation

Inserting (5) in (14), by a change of variables we obtain

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

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

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

and thus

$$\mathbf{H} = \mathbf{M}_{\mathcal{Y}}\mathbf{G} = \int_0^T \mathbf{K}(t) \otimes \mathbf{W}(t) \, dt. \tag{33}$$

We note that $\mathbf{W}(t)$ can be exactly calculated if piecewise polynomial ansatz functions $\psi_i(t)$ and $\phi_j(t)$ are chosen.

For systems of the form (2), 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

$$c_k^* = C^*\nu_k \in Z, \quad k = 1, \dots, q, \qquad b_l = B\mu_l, \quad l = 1, \dots, p.$$

Hence, $\mathbf{K}$ can be calculated by solving p homogeneous systems

$$\dot{z}_l(t) = Az_l(t), \qquad t \in (0,T], \tag{34a}$$

$$z_l(0) = b_l, \tag{34b}$$

since (34) has the mild solution $z_l(t) = S(t)b_l \in \mathscr{C}([0,T]; L^2(\Omega))$. We obtain an approximation $\tilde{\mathbf{H}}$ of $\mathbf{H}$ by replacing $z_l(t)$ by numerical approximations $z_{l,tol}(t)$, i.e.

$$\tilde{\mathbf{H}} = \int_0^T \tilde{\mathbf{K}}(t) \otimes \mathbf{W}(t) \, dt, \tag{35}$$

with

$$\tilde{\mathbf{K}}_{kl}(t) = (\nu_k, Cz_{l,tol}(t))_Y = (c_k^*, z_{l,tol}(t))_Z.$$

Here the subscript *tol* indicates that the error $z_l - z_{l,tol}$ is assumed to satisfy some tolerance criterion which will be specified later.

Note that in this way the matrix function $\mathbf{K}$ is approximated *column-wise*. The corresponding approximation $\mathbb{G}_{DS}$ of $\mathbb{G}_S$ is given by

$$\mathbb{G}_{DS} = \kappa_{\mathcal{Y}}^{-1}\tilde{\mathbf{G}}\kappa_{\mathcal{U}}\mathbb{P}_{\mathcal{U}}, \quad \text{with } \tilde{\mathbf{G}} = \mathbf{M}_{\mathcal{Y}}^{-1}\tilde{\mathbf{H}} \tag{36}$$

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

Remark 6. As we will see later, it may be preferable to calculate $\mathbf{K}$ with respect to different bases. Setting

$$\mathbf{K}^w(t) = \mathbf{M}_{\mathcal{Y}}^{1/2}\mathbf{K}(t)\mathbf{M}_{\mathcal{U}}^{-1/2},$$

we have $\mathbf{K}_{kl}^w(t) = (\nu_k^w, CS(t)B\mu_l^w)_Y$ with

$$\mu_l^w = \sum_{j=1}^{p}[\mathbf{M}_{\mathcal{U}}^{-1/2}]_{lj}\mu_j, \qquad \nu_k^w = \sum_{i=1}^{q}[\mathbf{M}_{\mathcal{Y}}^{1/2}]_{ki}\nu_i,$$

and thus (34) has to be solved with $b_l^w = B\mu_l^w$ instead of b_l.

Remark 7. The calculation of $\tilde{\mathbf{H}}$ can be parallelized in an obvious way by calculating the p solutions $z_{l,tol}$ in parallel and we note that no state trajectories have to be stored.

Remark 8. The kernel may also be calculated *row-wise* via

$$\tilde{\mathbf{K}}_{kl}(t) = (B^* z_{k,tol}(t), \mu_l)_U = (z_{k,tol}(t), b_l)_Z,$$

where now the $z_{k,tol}$ are numerical solutions of the adjoint autonomous systems

$$\dot{z}_k(t) = A^* z_k(t), \qquad t \in (0,T], \tag{37a}$$
$$z_k(0) = c_k^*. \tag{37b}$$

This approach may be preferable if $q < p$ or if one wants to improve the output approximation by adding basis functions $\nu_{q+1}, \nu_{q+2}, \ldots$ see Section 8.

Remark 9. In general, the matrix $\mathbf{H}$ is not sparse. Using hierarchical linear finite elements for the signal space discretization and Haar wavelets for the time discretization, a matrix $\mathbf{H} \in \mathbb{R}^{qs,pr}$ with $p = q = 33$ and $r = s = 32$ has about 64% non-zero elements, whereas for $p = q = 65$ and $r = s = 64$ the proportion of non-zero elements is 56%. Hence, the memory requirements become significant if a high resolution of the signals in space and time is required, and the question if $\mathbf{H}$ can be represented in a data-sparse way becomes essential. In the following make a first step in this direction.

We aim to approximate the entries $\mathbf{K}_{kl}(t)$ of the matrix-valued kernel function $\mathbf{K}(t)$ by a function of the type

$$\check{\mathbf{K}}_{kl}(t-s) = \sum_{m}^{M}\sum_{n=1}^{N}\alpha_{mn}L_m(t)L_n(s), \quad s,t \in [0,T], \tag{38}$$

with $\alpha_{mn} \in \mathbb{R}$. For instance, one can realize such an approximation via interpolation, cf. [24]. We introduce a grid on $[0, t]$, i.e.

$$0 = t_1 < t_1 < \cdots < t_N = T \tag{39}$$

and let L_j be the corresponding Lagrange interpolation functions, i.e.

$$L_j(t) = \prod_{\substack{1 \le i \le N \\ i \neq j}} \frac{t - t_i}{t_j - t_i}, \quad j = 1, \ldots, N.$$

Then $\check{\mathbf{K}}_{kl}(\tau)$, defined by

$$\check{\mathbf{K}}_{kl}(t - s) = \sum_{m,n=1}^{N} \mathbf{K}_{kl}(t_m - t_n) L_m(t) L_n(s),$$

interpolates $\mathbf{K}_{kl}(\tau)$ in all $\tau = (t_m - t_n)$ with $m, n = 0, \ldots, N$. Hence,

$$\mathbf{H}_{ij}^{kl} \approx \check{\mathbf{H}}_{ij}^{kl} = \sum_{m,n=1}^{N} \Psi_{im} \check{\mathbf{K}}_{mn}^{kl} \Phi_{jn}$$

with matrices $\Psi \in \mathbb{R}^{s,N}$, $\Phi \in \mathbb{R}^{r,N}$ and $\check{\mathbf{K}}^{kl} \in \mathbb{R}^{N,N}$ having elements

$$\Psi_{im} = \int_0^T \psi_i(t) L_m(t)\, dt, \quad \check{\mathbf{K}}_{mn}^{kl} = \mathbf{K}_{kl}(t_m - t_n), \quad \Phi_{jn} = \int_0^T \phi_j(t) L_n(t)\, dt.$$

Thus, $\check{\mathbf{H}}^{kl} = \Psi \check{\mathbf{K}}^{kl} \Phi^T$ and, combining the matrices $\check{\mathbf{K}}^{kl}$ to a blockstructured matrix $\check{K} \in \mathbb{R}^{qN,pN}$ we have

$$\check{\mathbf{H}} = (\mathbf{I}_q \otimes \Psi)(\check{\mathbf{K}})(\mathbf{I}_p \otimes \Phi^T).$$

If, for instance, a regular time grid (39) is chosen, then the blocks $\check{\mathbf{K}}^{kl}$ are lower triangular Toeplitz matrices, and only $q \cdot p \cdot N$ elements have to be stored. Note that such a regular time grid may not be optimal to capture the initial dynamics, cf. Fig. 2 and recall the problems of interpolation with polynomials of higher degree. Investigations on the use of other decompositions of the type (38), using e.g. spline functions or Taylor expansions, are hence an important future task.

5.2. Dynamics approximation error

The following proposition relates the system dynamics error ϵ_D to the errors made in solving the PDE (34) for $l = 1, \ldots, p$.

Proposition 2. *The system dynamics error* $\epsilon_D = \|\mathbb{G}_S - \mathbb{G}_{DS}\|_{\mathscr{L}(\mathcal{U},\mathcal{Y})}$ *satisfies*

$$\epsilon_D \le \sqrt{T} \|\mathbf{K} - \tilde{\mathbf{K}}\|_{L^2(0,T;\mathbb{R}_w^{q,p})} = \sqrt{T} \|\mathbf{K}^w - \tilde{\mathbf{K}}^w\|_{L^2(0,T;\mathbb{R}^{q,p})}$$

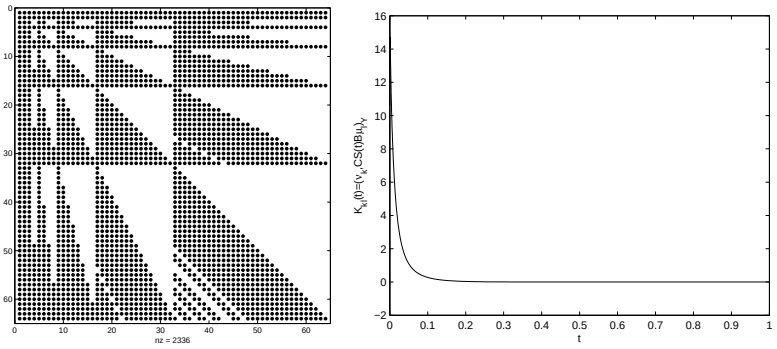

Figure 2: Left:One block $\mathbf{H}^{kl} \in \mathbb{R}^{s,r}$ of a matrix $\mathbf{H} \in \mathbb{R}^{qs,pr}$ with $p = q = 65$ and $r = s = 64$. For the signal discretization hierarchical linear finite elements in space and Haar-wavelets in time have been used. Right: A typical time evolution of an entry $\mathbf{K}_{kl}(t)$ for the example heat equation.

and

$$\textit{(i)} \quad \epsilon_D \leq p\sqrt{T} \max_{1\leq l\leq p} \|\mathbf{K}^w_{:,l} - \tilde{\mathbf{K}}^w_{:,l}\|_{L^2(0,T;\mathbb{R}^q)}, \tag{40a}$$

$$\textit{(ii)} \quad \epsilon_D \leq p\sqrt{T}\sqrt{\frac{\lambda_{\max}(\mathbf{M}_{Y,h_2})}{\lambda_{\min}(\mathbf{M}_{U,h_1})}} \max_{1\leq l\leq p} \|\mathbf{K}_{:,l} - \tilde{\mathbf{K}}_{:,l}\|_{L^2(0,T;\mathbb{R}^q)}. \tag{40b}$$

Here $\mathbf{K}_{:,l}$, $\tilde{\mathbf{K}}_{:,l}$, $\mathbf{K}^w_{:,l}$ and $\tilde{\mathbf{K}}^w_{:,l}$ denote the l'th column of $\mathbf{K}(t)$, $\tilde{\mathbf{K}}(t)$, $\mathbf{K}^w(t)$ and $\tilde{\mathbf{K}}^w(t)$, 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 $\mathbf{M}_{U,h_1}$. $\mathbb{R}^{q,p}$ denotes the space of real $q \times p$-matrices equipped with the canonical spectral norm, whereas $\mathbb{R}^{q,p}_w$ is the same space equipped with the weighted matrix norm $\|\mathbf{M}\|_{q,p;w} = \sup_{\|\mathbf{u}\|_{p;w}=1} \|\mathbf{M}\mathbf{u}\|_{q;w}$.

Proof. $\mathbf{K}$ is the matrix function representation of the space-projected kernel function

$$K_m : [-T, T] \to \mathscr{L}(U, Y), \qquad K_m(t) = P_{Y,h_2} K(t) P_{U,h_1},$$

with corresponding i/o-map $\mathbb{G}_m = \mathbb{G}_m(h_1, h_2)$,

$$(\mathbb{G}_m u)(t) = \int_0^T K_m(t-s)u(s)\,ds, \qquad t \in [0, T]. \tag{41}$$

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

$$\tilde{K} : [-T, T] \to \mathscr{L}(U, Y), \qquad \tilde{K}(t) = \kappa^{-1}_{Y,h_2} \tilde{\mathbf{K}}(t) \kappa_{U,h_1} P_{U,h_1},$$

with corresponding i/o-map

$$(\mathbb{G}_D u)(t) = \int_0^T \tilde{K}(t-s)u(s)\,ds, \qquad t \in [0,T]. \tag{42}$$

We observe that $\mathbb{G}_{DS}$ as defined in (36) 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 according to (13)-(33), 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 (35). We note that $\|K_m(t)\|_{\mathscr{L}(U,Y)} = \|\mathbf{K}(t)\|_{q,p;w}$ and $\|\tilde{K}(t)\|_{\mathscr{L}(U,Y)} = \|\tilde{\mathbf{K}}(t)\|_{q,p;w}$ for all $t \in [0,T]$. Lemma 1 yields

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

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 of the $1-$norm and 2-vector norms in $\mathbb{R}^p$ we have that

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

and hence

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

which concludes the proof. □

Calculating the columns of $\mathbf{K}$ directly and estimating ϵ_D via (40a), the term depending on 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 $\mathbf{K}_{:,l}$. In view of (19) this may be problematic if hierarchical bases functions are chosen. In this case it may be reasonable to approximate the columns of $\mathbf{K}$ by calculating the columns of $\mathbf{K}^w$ and to estimate ϵ_D via (40b). A bad conditioning of the mass matrices must then be taken into account to ensure that the corresponding basis transformations are carried out with sufficient accuracy.

6. Numerical solution of the homogeneous PDEs

In this section we discuss the numerical solution of (34) and how

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

can be ensured for a given $tol > 0$.

6.1. Weak formulation and numerical schemes

We assume that the operator $\hat{A} = -A$ is self-adjoint and positive definite, and for ease of presentation, we assume that A has a compact inverse. We can define $\hat{A}^{1/2} : D(\hat{A}^{1/2}) \to Z$ as the self-adjoint operator

$$\hat{A}z = \sum_{i=1}^{\infty} \lambda_i^{1/2}(z, e_i)e_i, \quad z \in D(\hat{A}^{1/2}) = \{v \in Z : \sum_{i=1}^{\infty} \lambda_i (e_i, v)_Z^2 < \infty\}$$

where the λ_i and the e_i are the eigenvalues and eigenfunctions of $\hat{A}$, see e.g. [43, Ch.12]. Then we can extend the bilinear form $\hat{a} : D(\hat{A}) \times D(\hat{A}) \to \mathbb{R}$, $\hat{a}(v, w) = (\hat{A}v, w)_Z$ in a natural way to

$$\hat{a} : D(\hat{A}^{1/2}) \times D(\hat{A}^{1/2}) \to \mathbb{R}, \qquad \hat{a}(v, w) = (\hat{A}^{1/2}v, \hat{A}^{1/2}w)_Z.$$

Defining $V = D(\hat{A}^{1/2})$, the weak formulation of (34) is now the following. Find $z_l \in W(0, T) = \{v \in L^2(0, T; V), \quad \dot{v} \in L^2(0, T; V')\}$ such that

$$\langle \partial_t z_l, v\rangle_{V',V} + \hat{a}(z_l, v) = 0 \quad \text{for all } v \in V \text{ and a.e. } t \in (0, T] \tag{43}$$

and $z_l(0) = b_l \in Z$. The initial condition makes sense, since $W(0, T)$ is continuously embedded in $C([0, T])$, see e.g. [21, Thm. 5.9.3]. If $\hat{a}$ is bounded and satisfies a Garding inequality, then a unique *weak* solution $z_l \in W(0, T)$ of (43) exists, see e.g. [17]

Remark 10. For the heat equation in Example 1, i.e. $A = \triangle$ with $D(A) = H^2(\Omega) \cap H_0^1(\Omega)$, we have $V = H_0^1(\Omega)$ and $\hat{a}(v, w) = (\nabla v, \nabla w)_L^2(\Omega)$, and thus a unique weak solution of (43) exists.

Remark 11. Choosing a finite dimensional subspace $V_{h_3} \subset V$, the standard Galerkin discretization of (43) consists in finding a solution $z_{l,h_3} \in L^2(0, T; V_{h_3})$ such that

$$\langle \partial_t z_{l,h_3}, v\rangle_{V',V} + \hat{a}(z_{l,h_3}, v) = 0 \quad \text{for all } v \in V_{h_3} \text{ and a.e. } t \in (0, T] \tag{44}$$

and $z_h(0) = b_{l,h_3}$ with some approximation $b_{l,h_3} \in V_{h_3}$ of $b_l \in Z$. We can write (44) equivalently as an implicit ODE system

$$\mathbf{M}\dot{\mathbf{z}}^{(l)}(t) + \mathbf{A}\mathbf{z}^{(l)}(t) = 0, \quad t \in (0, T], \qquad \mathbf{z}^{(l)}(0) = \mathbf{b}^{(l)} \tag{45}$$

with mass matrix $\mathbf{M} = [(\varphi_j, \varphi_i)_{L^2(\Omega)}]_{ij} \in \mathbb{R}^{\hat{n},\hat{n}}$ and stiffness matrix $\mathbf{A} = [a(\varphi_j, \varphi_i)]_{ij} \in \mathbb{R}^{\hat{n},\hat{n}}$, where $\mathbf{M}$ and $\mathbf{A}$ are positive definite. Here $\{\varphi_1, \ldots, \varphi_{\hat{n}}\}$ are shape functions forming a basis of V_{h_3} such that $b_{l,h_3} = \sum_{i=1}^{\hat{n}} \mathbf{b}_i^{(l)}\varphi_i$ and $z_{l,h_3}(t) = \sum_{i=1}^{\hat{n}} \mathbf{z}_i^{(l)}(t)\varphi_i$. For this, standard stiff ODE discretization schemes can be applied in order to obtain a fully discrete approximation scheme, see e.g. [25].

We now consider discontinuous Galerkin time discretization schemes of (43), which lead in combination with classical space disretizations to fully discrete schemes. They

have a variational interpretation being especially suited for the error analysis and they offer much flexibility with respect to adaptive mesh refinement and time stepping.

Consider partitions of $(0,T]$ into subintervals $I_n = (t_{n-1}, t_n]$ of length $k_n = |I_n|$ and maximal length $\tau_3 = \max k_n$, where $0 = t_0 < t_1 < \cdots < t_N = T$. We define piecewise polynomial spaces

$$\mathcal{V}_{h_3,\tau_3} = \{v : [0,T] \to V : v_{|I_n} = \sum_{j=0}^{q_\mathcal{V}} v_j t^j, \quad v_j \in V_{h_{3,n}}, \quad n = 1, \ldots, N\},$$

with $q_\mathcal{V} \in \mathbb{N}_0$ and $h_3 = \max h_{3,n}$. Here $V_{h_{3,n}}$ are finite-dimensional subspaces of V which may differ from time step to time step and $h_{3,n}$ denote the corresponding discretization parameters. The fully discrete scheme for solving (43) may be formulated as follows, see e.g. [19, 31]. Find $z_{l,tol} \in \mathcal{V}_{h_3,\tau_3}$ such that

$$B(z_{l,tol}, v) = (b_l, v_0^+), \quad \text{for all } v \in \mathcal{V}_{h_3,\tau_3}, \tag{46}$$

where B is a bilinear form on $\mathcal{V}_{h_3,\tau_3}$ given by

$$B(w,v) = \sum_{n=1}^{N} \int_{I_n} \{(\partial_t w, v)_Z + \hat{a}(w,v)\}\, dt + \sum_{n=2}^{N} ([w]_{n-1}, v_{n-1}^+)_Z + (w_0^+, v_0^+)_Z,$$

with jump terms and one-sided limits

$$[v]_n = v_n^+ - v_n^-, \quad v_n^+ = \lim_{t \to t_n^+} v(t), \quad v_n^- = \lim_{t \to t_n^-} v(t).$$

For results concerning the unique solvability of (46), see e.g. [43, p. 183].

Example 2. For instance, one may choose the spaces $V_{h_{3,n}}$ from a family of spaces of continuous piecewise linear functions as follows. Assume, for simplicity, that $\Omega \subset \mathbb{R}^2$ is a bounded convex polygonal domain. For $h > 0$, let $\mathscr{T}_h = \{\bar{\omega}\}$ be a triangulation of Ω with $h = \max_{\bar{\omega} \in \mathscr{T}_h} h_{\bar{\omega}}$, where $h_{\bar{\omega}}$ denotes the diameter of a triangle $\bar{\omega}$. We then choose the spaces $V_{h_{3,n}}$ from the family $\{V_h\}_{h>0}$ of spaces

$$V_h = \{v \in \mathscr{C}^0(\bar{\Omega}) \cap H_0^1(\Omega) : v_{|\bar{\omega}} \text{ linear } \forall \bar{\omega} \in \mathscr{T}_h\}. \tag{47}$$

We say that the family $\{\mathscr{T}_h\}_h$ satisfies the *minimal angle condition*, if there exists $\beta > 0$ independent of the triangulations $\mathscr{T}_h \in \{\mathscr{T}_h\}_h$ such that

$$\frac{\rho_{\bar{\omega}}}{h_{\bar{\omega}}} \geq \beta, \qquad \forall \bar{\omega} \in \mathscr{T}_h, \tag{48}$$

where $\rho_{\bar{\omega}}$ denotes the diameter of a circle inscribed in $\bar{\omega}$.

We say that the family $\{\mathscr{T}_h\}_h$ is *quasi-uniform* if there is a constant c independent of $\mathscr{T}_h \in \{\mathscr{T}_h\}_h$ such that

$$h_{\bar{\omega}} \leq ch, \qquad \forall \bar{\omega} \in \mathscr{T}_h.$$

Remark 12. For $q_{\mathcal{V}} = 0$ and fixed space discretizations $V_{h_3} = V_{h_{3,1}} = \ldots, V_{h_{3,N}}$, (46) coincides with the backward Euler scheme applied to (45). If the solutions $z_{1,tol}, \ldots, z_{p,tol}$ are calculated on the same time grid, (35) becomes $\tilde{\mathbf{H}} = \sum_{n=1}^{N} \tilde{\mathbf{K}}(t_n) \otimes \int_{I_n} \mathbf{W}(t)\, dt$. The integrals $\int_{I_n} \mathbf{W}(t)\, dt$ can be exactly calculated for many choices of ϕ_i and ψ_j. Other choices of $q_{\mathcal{V}}$ also allow for an exact time integration in (35) if only polynomial ansatz functions are implied. Note that the Crank-Nicolson scheme can also be formulated in terms of (46), see e.g. [43, p. 185].

6.2. Global state error estimates

The errors $\|\mathbf{K}_{:,l} - \tilde{\mathbf{K}}_{:,l}\|_{L^2(0,T;\mathbb{R}^q)}$ can be controlled via the global state error in solving (34) by

$$\|\mathbf{K}_{:,l} - \tilde{\mathbf{K}}_{:,l}\|_{L^2(0,T;\mathbb{R}^q)} \leq \|C\|_{\mathscr{L}(Z,Y)} \left(\sum_{i=1}^{q} \|\nu_i\|_Y^2 \right)^{1/2} \|z - z_{tol}^{(l)}\|_{L^2(0,t;Z)}. \tag{49}$$

Considering for example the numerical solution of the homogeneous heat equation by means of a discontinuous Galerkin scheme (46) with $q_{\mathcal{V}} = 0$, we quote the following result.

Lemma 5 ([31])**.** *Assume, for simplicity, that $\Omega \subset \mathbb{R}^2$ is a bounded convex polygonal domain and that the spaces $V_{h_{3,n}}$ are chosen as subspaces of continuous piecewise linear functions as in* (47)*. Assume further, that the corresponding triangulations are quasi-uniform and satisfy a minimal angle condition* (48) *, and let*

$$V_{h_{3,n}} \subseteq V_{h_{3,n-1}}, \qquad k_{n-1} \leq \gamma k_n, \qquad n = 2, \ldots, N,$$

with a constant γ independent of the time grid. Then, there exists a constant C, depending only on γ and the minimum angle constant β, such that the solution z_l of (43) *and the solution $z_{l,tol}$ of* (46) *satisfy*

$$\begin{aligned} &\|z_l - z_{l,tol}\|_{L^2(0,T;L^2(\Omega))} \leq \sqrt{T} \|z_l - z_{l,tol}\|_{L^\infty(0,T;L^2(\Omega))} \\ \leq & C\sqrt{T} \max_{n=1,\ldots,N} \left\{ (1 + \log(t_n/k_n))^{1/2} \left(k_n \max_{t \in I_n} \|\partial_t z_l(t)\|_{L^2(\Omega)} + h_{3,n}^2 \max_{t \in I_n} \|z_l(t)\|_{H^2(\Omega)} \right) \right\} \end{aligned} \tag{50}$$

More general results, in particular for *nonquasi-uniform* triangulations, can be found e.g. in [19, 20].

Remark 13. If $b_l \in D(A) = H^2(\Omega) \cap H_0^1(\Omega)$, then we have $z_l \in C^1([0,T], D(A))$ and the estimate (50) can be used to choose global discretization parameters $h_3 = \max h_{3,n}$ and $\tau_3 = \max \tau_{3,n}$ in order to ensure $\|z_l - z_{l,tol}\|_{L^2(0,T;L^2(\Omega))} < tol$ for a given tolerance tol. However, in the generic situation we only have the regularity $b_l \in L^2(\Omega)$ and $z_l \in C^1((0,T], D(A))$, and the norms of the time and space derivatives of z_l may become very large for small t, cf. (9c). We call, in agreement with [31], an initial phase where this happens an *initial transient*. Since $D(A)$ is dense in $L^2(\Omega)$, one

can always find global discretization parameters h_3 and τ_3 small enough to ensure $\|z_l - z_{l,tol}\|_{L^2(0,T;L^2(\Omega))} < tol$ also for nonsmooth initial values. However, it is not economic to use such fine mesh sizes (in space and time) for larger $t > 0$ where the solution has much more regularity, such that adaptive meshing on the basis of a posteriori error estimates may be very advantageous. For different possibilities of corresponding a posteriori error controls, see e.g. [1] or [18] and the references therein.

Remark 14. Any semi-discretization can resolve non-smooth initial transients only up to a certain accuracy. For illustration purposes, we quote the following error estimate from [31, p. 151] for the system (44) in the case of the heat equation in Example 1. Assuming that V_{h_3} consists of continuous piecewise linear functions with respect to a quasi-uniform triangulation of Ω, there exists a positive constants c such that the solution z_l of (43) and the solution $z_{l,h}$ of (44) satisfy

$$\|z_l - z_{l,h}\|_{L^\infty(0,T;L^2(\Omega))} \leq C(1 + |log(T/h_3^2)|) \max_{t\in(0,T)} h_3^2 \|z_l(t)\|_{H^2(\Omega)},$$

but in view of (9) this estimate is useless for general initial values $b_l \in L^2(\Omega)$. This should be kept in mind when approximating PDE control systems by means of semi-discretizations in space and considering later impulse responses of controls with jumps in time. In our approach, the problem of resolving initial transients with sufficient accuracy can be directly taken into account by using a posteriori error control in solving (43). For nonsmooth data results for semi-discretizations of the heat equation, cf. e.g. [43, Ch. 3] and the references therein.

Absorbing the "almost bounded" logarithmic quantity and C in (50) into a new constant $\tilde{C}$, one may for instance aim to choose the local time steps k_n and the local mesh sizes $h_{3,n}$ such that

$$k_n \max_{t\in I_n} \|\partial_t z_l(t)\|_{L^2(\Omega)} \leq \frac{tol}{2\tilde{C}\sqrt{T}} \quad \text{and} \quad h_{3,n}^2 \max_{t\in I_n} \|\partial_t z_l(t)\|_{L^2(\Omega)} \leq \frac{tol}{2\tilde{C}'\sqrt{T}},$$

where we used that $\|z_l(t)\|_{H^2(\Omega)} \leq c\|\partial_t z_l(t)\|_{L^2(\Omega)}$, cf. (9). Approximating e.g.

$$k_n \max_{t\in I_n} \|\partial_t z_l(t)\|_{L^2(\Omega)} \approx \|z_{l,tol}(t_n) - z_{l,tol}(t_{n-1})\|_{L^2(\Omega)},$$

we can choose k_n and $h_{3,n}$ on the basis of numerically computable quantities, see e.g. [31].

The preceding discussion justifies the following assumption.

Assumption 1. *Given a tolerance $tol > 0$, we can ensure (by using appropriate error estimators and mesh refinements) that the solution z_l of* (43) *and the solution $z_{l,tol}$ of* (46) *satisfy*

$$\|z_l - z_{l,tol}\|_{L^2(0,T;Z)} < tol, \qquad l = 1, \ldots, p.$$

6.3. Goal-oriented error estimation

As we have observed in the last subsection, the accurate *and* economic calculation of non-smooth initial transients requires a posteriori error control and an adaptive choice of mesh sizes in space and time. Moreover, the estimate (49) 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. Therefore, we briefly present the concept of *goal-oriented* error estimators, which estimate the error with respect to linear functionals of the state z like

$$J_l(z) = (z, \sigma_l)_{L^2(0,T;L^2(\Omega))}, \tag{51}$$

with a weight function $\sigma_l \in L^2(0,T;L^2(\Omega))$. This technique is based on the so-called *dual-weighted residuals* (DWR) approach. Duality arguments in a priori estimations have been employed for the first time in [4] for the post-processing of 'quantities of physical interest' in elliptic problems. Substantial contributions and generalizations have since then been made in [1], [6], [8], [9], [10], [28], [32], [33] and the references therein.

We first show how our problem fits into the formulation (51). Assume for the moment that the error $\mathbf{E}(t) = \mathbf{K}(t) - \tilde{\mathbf{K}}(t)$ is known. Define

$$\sigma_l(t) = \sum_{k=1}^{q} \mathbf{E}_{kl}(t)\nu_k^*, \quad l = 1, \dots, p.$$

Then

$$\begin{aligned} J_l(z_l - z_{l,tol}) &= \int_0^T (z_l(t) - z_{l,tol}(t), \sigma_l(t))\, dt \\ &= \int_0^T (z_l(t) - z_{l,tol}(t), \sum_{k=1}^{q} (z_l(t) - z_{l,tol}(t), \nu_k^*)_{L^2(\Omega)} \nu_k^*)_{L^2(\Omega)}\, dt \\ &= \int_0^T \sum_{k=1}^{q} \|(z_l(t) - z_{l,tol}(t), \nu_k^*)\|^2_{L^2(\Omega)} = \|\mathbf{K}_{:,l} - \tilde{\mathbf{K}}_{:,l}\|^2_{L^2(0,T;\mathbb{R}^q)}. \end{aligned}$$

In practice we have to work with estimates $\hat{\sigma}_l(t) = \sum_{k=1}^{q} \tilde{\mathbf{E}}_{kl}(t)\nu_k^*$. and approximated functionals $\hat{J}_l(z) = (z, \hat{\sigma}_l)_{L^2(0,T;L^2(\Omega))}$.

Next, we present an exact error representation formula for $J_l(z_k - z_{l,tol})$. We extend the bilinear form B to a space $\mathcal{V}$ with $\mathcal{V} \supset W(0,T)$ and observe that the exact solution $z_l \in W(0,T)$ of (43) also satisfies (46), and hence we have the so-called *Galerkin orthogonality* of the error and the ansatz space, i.e.

$$B(z_l - z_{l,tol}, v) = 0 \quad \forall v \in \mathcal{V}_{h_3,\tau_3}.$$

For each $l = 1, \dots, p$ we introduce the dual problem

$$B^*(w_l, v) = B(v, w_l) = J_l(v) \quad \forall v \in \mathcal{V} \tag{52}$$

and assume that a unique solution $w_l \in W(0,T)$ exists. Note that w_l can be interpreted as the Riesz-representor of J_l. Then

$$J_l(z_l - z_{l,tol}) = B(z_l - z_{l,tol}, w_l), \tag{53a}$$
$$= B(z_l - z_{l,tol}, w_l - w) \quad \forall w \in \mathcal{V}_{h_3,\tau_3}, \tag{53b}$$
$$= (b_l, (w_l - w)_0^+)_Z - B(z_{l,tol}, w_l - w) \quad \forall w \in \mathcal{V}_{h_3,\tau_3}, \tag{53c}$$

which can be interpreted as the *residual* of the approximated solution $z_{l,tol}$ tested or *weighted* with the *dual* solution w_l. The representation (53) can be localized by carrying out time and space integrations separately for each time step and spatial mesh element. For the example of the heat equation, (53) becomes

$$\begin{aligned} J_l(z_l - z_{l,tol}) = &- \sum_{n=1}^{N} \int_{I_n} \{(\partial_t z_{l,tol}, w_l - w)_{L^2(\Omega)} + (\nabla z_{l,tol}, \nabla(w_l - w))_{L^2(\Omega)}\}\, dt \\ &- \sum_{n=2}^{N} ([z_{l,tol}]_{n-1}, (w_l - w)_{n-1}^+)_{L^2(\Omega)} + (b_l - (z_{l,tol})_0^+, (w_l - w)_0^+)_{L^2(\Omega)} \\ = &\sum_{n=1}^{N} \sum_{\bar\omega \in \mathcal{T}_{h_n}} \eta_{n,\bar\omega}. \end{aligned} \tag{54}$$

Here $\eta_{n,\bar\omega} = \eta_{n,\bar\omega}^{(1)} + \eta_{n,\bar\omega}^{(2)}$ are local error indicators with

$$\begin{aligned} \eta_{n,\bar\omega}^{(1)} &= -\int_{I_n} \{(\partial_t z_{l,tol}, w_l - w)_{L^2(\bar\omega)} + (\nabla z_{l,tol}, \nabla(w_l - w))_{L^2(\bar\omega)}\} dt, \quad n = 1, \dots, N, \\ \eta_{n,\bar\omega}^{(2)} &= -([z_{l,tol}]_{n-1}, (w_l - w)_{n-1}^+)_{L^2(\bar\omega)}, \quad n = 2, \dots, N \\ \eta_{1,\bar\omega}^{(2)} &= (b_l - (z_{l,tol})_0^+, (w_l - w)_0^+)_{L^2(\bar\omega)}, \end{aligned}$$

and with mesh elements $\bar\omega$ from the respective mesh $\mathcal{T}_{h_n}$. Frequently, the terms $\eta_{n,\bar\omega}^{(1)}$ are defined in a slightly different manner, when the second term in $\eta_{n,\bar\omega}^{(1)}$ is integrated by parts over $\bar\omega$, and the resulting boundary terms of neighboring terms are combined to jump terms which are then equally allocated to the respective local error indicators [6] [26] . Often, w is chosen to be the interpolant of w_l in $\mathcal{V}_{h_3,\tau_3}$.

There are several ways to use (54) or similar representations for error control and adaptive meshing. Solving the dual problem numerically with sufficient accuracy, the indicators can directly be evaluated, cf. Section 8.2. Another possibility is to apply the Cauchy-Schwarz inequality to the local error indicators and take the resulting upper bounds as new local error indicators, see e.g. [1]. Aiming to extract information on the local contribution to the time discretization error and to the spatial discretization error, the local error indicators may be split by appropriate projections into respective space and time error indicators, see e.g. [26] and [39].

Often, the error due to a numerical approximation of the dual solutions w_l is neglected, since it vanishes asymptotically with higher order. But in order to obtain

reliable error results also for moderately accurate solutions of the dual problem, the global state error in the dual solution should be itself controlled by appropriate a posteriori estimations, see e. g. [1].

An example how the local error estimators $\eta_{n,\bar{\omega}}$ can be used in a mesh adaption strategy is presented in Section 8.2.

The previous discussion justifies the following assumption.

Assumption 2. *Given a tolerance $tol > 0$, we can ensure (by using appropriate error estimators and mesh refinements) that the solution z_l of (43) and the solution $z_{l,tol}$ of (46) satisfy*

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

We note that Assumption 1 implies Assumption 2, and that Assumption 2 implies that we can ensure, for every given $tol > 0$ to solve (43) such that $\|\mathbf{K}^w_{:,l} - \tilde{\mathbf{K}}^w_{:,l}\|_{L^2(0,T;\mathbb{R}^q)} < tol$.

7. Total error estimates

We will summarize previous error results for the total error in the approximation of $\mathbb{G}$ and of its adjoint $\mathbb{G}^*$.

Theorem 1. *Consider the i/o map $\mathbb{G} \in \mathscr{L}(\mathcal{U}, \mathcal{Y})$ of an infinite-dimensional LTI system as given in (5) 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

$$\mathcal{U}_s = H^{\alpha_1}(0,T;L^2(\Theta)) \cap L^2(0,T;H^{\beta_1}(\Theta)),$$
$$\mathcal{Y}_s = H^{\alpha_2}(0,T;L^2(\Xi)) \cap L^2(0,T;H^{\beta_2}(\Xi)),$$

where α_1, α_2, β_1, $\beta_2 \in \mathbb{N}$.

(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

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

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 2 holds.

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

$$\tau_1^* < \left(\frac{\delta}{8c_{\mathcal{R}}\|\mathbb{G}\|_{\mathscr{L}(\mathcal{U},\mathcal{Y})}}\right)^{1/\alpha_1}, \quad h_1^* < \left(\frac{\delta}{8c_U\|\mathbb{G}\|_{\mathscr{L}(\mathcal{U},\mathcal{Y})}}\right)^{1/\beta_1}, \tag{55a}$$

$$\tau_2^* < \left(\frac{\delta}{8c_{\mathcal{S}}\|\mathbb{G}\|_{\mathscr{L}(\mathcal{U}_s,\mathcal{Y}_s)}}\right)^{1/\alpha_2}, \quad h_2^* < \left(\frac{\delta}{8c_Y\|\mathbb{G}\|_{\mathscr{L}(\mathcal{U}_s,\mathcal{Y}_s)}}\right)^{1/\beta_2}, \tag{55b}$$

and one can solve the PDEs (43) *for* $l = 1, \dots, p(h_1)$ *such that one of the following conditions holds.*

$$\textit{(i)} \quad \|\mathbf{K}^w_{:,l} - \tilde{\mathbf{K}}^w_{:,l}\|_{L^2(0,T;\mathbb{R}^q)} < \frac{\delta}{2\sqrt{T}p(h_1^*)}, \tag{56a}$$

$$\textit{(ii)} \quad \|\mathbf{K}_{:,l} - \tilde{\mathbf{K}}_{:,l}\|_{L^2(0,T;\mathbb{R}^q)} < \frac{\delta}{2\sqrt{T}p(h_1^*)}\sqrt{\frac{\lambda_{\min}(\mathbf{M}_{U,h_1^*})}{\lambda_{\max}(\mathbf{M}_{Y,h_2^*})}}, \tag{56b}$$

$$\textit{(iii)} \quad \|z_l - z_{l,tol}\|_{L^2(0,T;Z)} < \frac{\delta}{2\sqrt{T}p(h_1^*)}\sqrt{\frac{\lambda_{\min}(\mathbf{M}_{U,h_1^*})}{\lambda_{\max}(\mathbf{M}_{Y,h_2^*})}}\|C\|^{-1}_{\mathscr{L}(Z,Y)}\left(\sum_{i=1}^{q(h_2^*)}\|\nu_i\|_Y^2\right)^{-1/2}. \tag{56c}$$

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. Since $\|u\|_{\mathcal{U}} \leq \|u\|_{\mathcal{U}_s}$ for all $u \in \mathcal{U}_s$, we have

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

Proposition 1 in combination with (55) ensures that

$$\epsilon_S := \|\mathbb{G} - \mathbb{G}_{DS}\|_{\mathscr{L}(\mathcal{U}_s,\mathcal{Y})} < \delta/2, \tag{57}$$

and Proposition 2 in combination with (56) and in view of (49) ensures that

$$\epsilon_D := \|\mathbb{G}_S - \mathbb{G}_{DS}\|_{\mathscr{L}(\mathcal{U},\mathcal{Y})} < \delta/2, \tag{58}$$

which concludes the proof. □

Remark 15. (i) The tolerances for the PDE calculations are coupled to the input signal approximation, if error estimators for (56a) are used. They are even coupled to the choice of basis functions for the input and output signals if (56b) or (56c) are used.

(ii) If the signal approximation error ϵ_S can be measured in the $\mathscr{L}(\mathcal{U},\mathcal{Y})$-norm, then one obtains also a total error estimation in the $\mathscr{L}(\mathcal{U},\mathcal{Y})$-norm.

(iii) If $B \in \mathscr{L}(H^{\beta_1}(\Theta), D(A))$, then one can obtain total error estimates containing the global space and time mesh sizes h_3 and τ_3 for the numerical solution of the PDEs (43) via (50).

Corollary 2. *Consider the heat control system in Example 1 with* $U = Y = L^2(0,1)$ *and let* $\delta > 0$ *be given. We assume that*

$$C_{|H^2(\Omega)} \in \mathscr{L}(H^2(\Omega), H^2(0,1)) \tag{59}$$

and that the input signals are restricted to $\mathcal{U}_s = \mathcal{Y}_s = H^1(0,T;L^2(0,1)) \cap L^2(0,T;H^2(0,1))$, *i.e.* $\mathbb{G}_{|\mathcal{U}_s} \in \mathscr{L}(\mathcal{U}_s, \mathcal{Y}_s)$. *We choose* U_{h_1} *and* Y_{h_2} *as spaces of continuous piecewise linear functions with respect to equidistant grids on* $[0,1]$, *and we choose* $\mathcal{R}_{\tau_1}$ *and* $\mathcal{S}_{\tau_2}$ *as spaces of piecewise constant functions with respect to equidistant grids on* $[0,T]$, *with dimensions satisfying*

$$p > 2\sqrt{\frac{\|\mathbb{G}\|_{\mathscr{L}(\mathcal{U},\mathcal{Y})}}{\delta}} + 1, \quad q > 2\sqrt{\frac{\|\mathbb{G}\|_{\mathscr{L}(\mathcal{U}_s,\mathcal{Y}_s)}}{\delta}} + 1, \tag{60a}$$

$$r > \frac{\sqrt{2}4\|\mathbb{G}\|_{\mathscr{L}(\mathcal{U},\mathcal{Y})}}{\delta}, \qquad s > \frac{\sqrt{2}4\|\mathbb{G}\|_{\mathscr{L}(\mathcal{U}_s,\mathcal{Y}_s)}}{\delta}. \tag{60b}$$

If the homogeneous heat equations (43) *are solved for* $l = 1, \dots, p$ *such that one of the conditions (i)-(iii) in* (56) *holds, then*

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

Proof. We recall that (59) implies conditions (i) and (ii) of Theorem 1 with $\alpha_1 = \alpha_2 = 1$ and $\beta_1 = \beta_2 = 2$, cf. Corollary 1. Remark 13 ensures condition (iii). For chosen ansatz functions on *regular* grids we have $h_1 = 1/(p-1)$, $h_2 = 1/(q-1)$, $\tau_1 = 1/r$ and $\tau_2 = 1/s$. Hence condition (55) becomes (60) in view of Remark 3. □

Remark 16. We consider the question how Corollary 2 can be used to choose the dimensions in (60) and the tolerances in (56) in practice.

If we choose *nodal* bases (i.e. the usual hat functions) in Y_{h_2}, then

$$\sum_{l=1}^{q} \|\nu_l\|^2_{L^2(0,1)} = 2/3$$

independent of $h_2 = 1/(q-1)$. If we choose *hierarchical* bases in Y_{h_2}, we have, with $q = 2^n + 1$ and $h_2 = 1/2^n$ for some $n \in \mathbb{N}$, that

$$\sum_{l=1}^{q} \|\nu_l\|^2_{L^2(0,1)} = (n+2)/3.$$

Note that $\lambda_{\max}(\mathbf{M}_{U,h_1^*})$ and $\lambda_{\max}(\mathbf{M}_{Y,h_2^*})$ can be either directly numerically calculated for fixed choices of p and q or they can be estimated via (18) and (19) and numerical approximations of c_λ, C_λ and c'_λ, C'_λ. It remains to estimate or calculate the operator norms of $\mathbb{G}$ and C. Recalling Lemma 3, $\|\mathbb{G}\|_{\mathscr{L}(\mathcal{U},\mathcal{Y})}$ can be numerically approximated by $\|\tilde{\mathbf{G}}\|_{\mathbf{h}}$ with some sufficiently accurate $\tilde{\mathbf{G}}$, though one should be aware that this may lead to choices of insufficiently large dimensions in view of $\|\mathbf{G}\|_{\mathbf{h}} \leq \|\mathbb{G}\|_{\mathscr{L}(\mathcal{U},\mathcal{Y})}$. The question, how good estimates of $\|\mathbb{G}\|_{\mathscr{L}(\mathcal{U}_s,\mathcal{Y}_s)}$ can be obtained, has still to be investigated, as well as the question, how sharp the error estimations are.

The next result shows that $(\mathbb{G}_{DS})^* \in \mathscr{L}(\mathcal{Y},\mathcal{U})$ automatically approximates the adjoint $\mathbb{G}^*$ with $\|\mathbb{G}^* - (\mathbb{G}_{DS})^*\|_{\mathscr{L}(\mathcal{Y}_s,\mathcal{U})} < c\,\delta$ with a $\mathbb{G}$-specific constant c, under the assumption that $\mathbb{G}^*_{|\mathcal{Y}_s} \in \mathscr{L}(\mathcal{Y}_s,\mathcal{U}_s)$. Note that $\mathbb{G}^* \in \mathscr{L}(\mathcal{Y},\mathcal{U})$ is given by

$$(\mathbb{G}^* y)(s) = \int_0^T K(s-t)^* y(t) dt.$$

Theorem 2. *The adjoint $(\mathbb{G}_{DS})^* \in \mathscr{L}(\mathcal{Y},\mathcal{U})$ of $\mathbb{G}_{DS} \in \mathscr{L}(\mathcal{U},\mathcal{Y})$ has the matrix representation*

$$\tilde{\mathbf{G}}^* := \mathbf{M}_{\mathcal{U}}^{-1}\tilde{\mathbf{G}}^T\mathbf{M}_{\mathcal{Y}} = \mathbf{M}_{\mathcal{U}}^{-1}\tilde{\mathbf{H}}^T \in \mathbb{R}^{pr,qs} \tag{61}$$

For a given $\delta > 0$, assume that all conditions in Theorem 1 hold, ensuring that $\|\mathbb{G} - \mathbb{G}_{DS}\|_{\mathscr{L}(\mathcal{U}_s,\mathcal{Y})} < \delta$. If, in addition, $\mathbb{G}^*_{|\mathcal{Y}_s} \in \mathscr{L}(\mathcal{Y}_s,\mathcal{U}_s)$, then

$$\|\mathbb{G}^* - (\mathbb{G}_{DS})^*\|_{\mathscr{L}(\mathcal{Y}_s,\mathcal{U})} < \delta(\frac{1}{2} + c_*) \tag{62}$$

with $c_* = \frac{1}{4}(\|\mathbb{G}^*\|_{\mathscr{L}(\mathcal{Y}_s,\mathcal{U}_s)}/\|\mathbb{G}\|_{\mathscr{L}(\mathcal{Y},\mathcal{U})} + \|\mathbb{G}^*\|_{\mathscr{L}(\mathcal{Y},\mathcal{U})}/\|\mathbb{G}\|_{\mathscr{L}(\mathcal{Y}_s,\mathcal{U}_s)})$.

Proof. We first observe that $\tilde{\mathbf{G}}^*$ is the adjoint of $\tilde{\mathbf{G}} : \mathbb{R}^{pr}_w \to \mathbb{R}^{qs}_w$, since

$$(\tilde{\mathbf{G}}\mathbf{u},\mathbf{y})_{qs;w} = \mathbf{u}^T\tilde{\mathbf{G}}^T\mathbf{M}_{\mathcal{Y}}\mathbf{y} = (\mathbf{u},\mathbf{M}_{\mathcal{U}}^{-1}\tilde{\mathbf{G}}^T\mathbf{M}_{\mathcal{Y}}\mathbf{y})_{pr;w}.$$

For $u \in \mathcal{U}$ and $y \in \mathcal{Y}$ and omitting the dependencies on h_1, h_2, τ_1, τ_2 and *tol*, we have

$$\begin{aligned}(\mathbb{G}_{DS}u, y)_{\mathcal{Y}} &= (\mathbb{P}_{\mathcal{Y}}\kappa_{\mathcal{Y}}^{-1}\tilde{\mathbf{G}}\kappa_{\mathcal{U}}\mathbb{P}_{\mathcal{U}}u, y)_{\mathcal{Y}} = (\tilde{\mathbf{G}}\kappa_{\mathcal{U}}\mathbb{P}_{\mathcal{U}}u, \kappa_{\mathcal{Y}}\mathbb{P}_{\mathcal{Y}}y)_{qs;w}\\ &= (\kappa_{\mathcal{U}}\mathbb{P}_{\mathcal{U}}u, \tilde{\mathbf{G}}^*\kappa_{\mathcal{Y}}\mathbb{P}_{\mathcal{Y}}y)_{qs;w} = (u, \mathbb{P}_{\mathcal{U}}\kappa_{\mathcal{U}}^{-1}\tilde{\mathbf{G}}^*\kappa_{\mathcal{Y}}\mathbb{P}_{\mathcal{Y}}y)_{\mathcal{U}}\\ &= (u, (\mathbb{G}_{DS})^* y)_{\mathcal{U}},\end{aligned}$$

where we have used that $\mathbb{P}_{\mathcal{U}} = \mathbb{P}^*_{\mathcal{U}}$, $\mathbb{P}_{\mathcal{Y}} = \mathbb{P}^*_{\mathcal{Y}}$, $\kappa^*_{\mathcal{U}} = \kappa^{-1}_{\mathcal{U}}$ and $\kappa^*_{\mathcal{Y}} = \kappa^{-1}_{\mathcal{Y}}$. To show (62), we estimate

$$\|\mathbb{G}^* - (\mathbb{G}_{DS})^*\|_{\mathscr{L}(\mathcal{Y}_s,\mathcal{U})} \le \|\mathbb{G}^* - (\mathbb{G}_S)^*\|_{\mathscr{L}(\mathcal{Y}_s,\mathcal{U})} + \|\mathbb{G}^*_S - (\mathbb{G}_{DS})^*\|_{\mathscr{L}(\mathcal{Y},\mathcal{U})},$$

where $(\mathbb{G}_S)^* = \mathbb{P}_{\mathcal{U}}\mathbb{G}^*\mathbb{P}_{\mathcal{Y}}$ is the adjoint of $\mathbb{G}_S \in \mathscr{L}(\mathcal{U},\mathcal{Y})$. In analogy to Proposition 1, one can show

$$\epsilon^*_S := \|\mathbb{G}^* - (\mathbb{G}_S)^*\|_{\mathscr{L}(\mathcal{Y}_s,\mathcal{U})} \le c''_{\mathcal{R}}\tau_1^{\alpha_1} + c''_U h_1^{\beta_1} + c''_{\mathcal{S}}\tau_2^{\alpha_2} + c''_Y h_2^{\beta_2}$$

with $c''_U = \|\mathbb{G}^*\|_{\mathscr{L}(\mathcal{Y}_S,\mathcal{U}_S)}c_U$, $c''_Y = \|\mathbb{G}^*\|_{\mathscr{L}(\mathcal{Y},\mathcal{U})}c_Y$, $c''_{\mathcal{R}} = \|\mathbb{G}^*\|_{\mathscr{L}(\mathcal{Y}_s,\mathcal{U}_s)}c_{\mathcal{R}}$ and $c''_{\mathcal{S}} = \|\mathbb{G}^*\|_{\mathscr{L}(\mathcal{Y},\mathcal{U})}c_{\mathcal{S}}$. Hence, (55) implies

$$\epsilon^*_S \le \frac{\delta}{8}\left(\frac{c''_{\mathcal{U}}}{c_{\mathcal{U}}} + \frac{c''_{\mathcal{R}}}{c_{\mathcal{R}}} + \frac{c''_{\mathcal{Y}}}{c_{\mathcal{Y}}} + \frac{c''_{\mathcal{S}}}{c_{\mathcal{S}}}\right) = \frac{\delta}{4}\left(\frac{\|\mathbb{G}^*\|_{\mathscr{L}(\mathcal{Y}_s,\mathcal{U}_s)}}{\|\mathbb{G}\|_{\mathscr{L}(\mathcal{Y},\mathcal{U})}} + \frac{\|\mathbb{G}^*\|_{\mathscr{L}(\mathcal{Y},\mathcal{U})}}{\|\mathbb{G}\|_{\mathscr{L}(\mathcal{Y}_s,\mathcal{U}_s)}}\right).$$

In order to estimate $\epsilon^*_D := \|\mathbb{G}^*_S - (\mathbb{G}_{DS})^*\|_{\mathscr{L}(\mathcal{Y},\mathcal{U})}$, we recall the definition of $\mathbb{G}_m$ in (41) and of $\mathbb{G}_D$ in (42) and obtain

$$\begin{aligned}\epsilon^*_D &\le \|\mathbb{P}_{\mathcal{U}}\|_{\mathscr{L}(\mathcal{U})}\|(\mathbb{G}_m)^* - (\mathbb{G}_D)^*\|_{\mathscr{L}(\mathcal{Y},\mathcal{U})}\|\mathbb{P}_{\mathcal{Y}}\|_{\mathscr{L}(\mathcal{Y})}\\ &\le \sqrt{T}(\int_0^T \|K_m(t)^* - \tilde{K}(t)^*\|^2_{\mathscr{L}(Y,U)}\,dt)^{1/2}.\end{aligned} \tag{63}$$

We observe that

$$K_m(t) = P_U \kappa_U^{-1} \mathbf{K}(t)^* \kappa_Y P_Y \quad \text{and} \quad \tilde{K}(t)^* P_U \kappa_U^{-1} \tilde{\mathbf{K}} \kappa_Y P_Y \tag{64}$$

with $\mathbf{K}(t)^* = \mathbf{M}_U^{-1} \mathbf{K}(t)^T \mathbf{M}_Y$ and $\tilde{\mathbf{K}}(t)^* = \mathbf{M}_U^{-1} \tilde{\mathbf{K}}(t)^T \mathbf{M}_Y$, and that

$$\|\mathbf{K}(t)^* - \tilde{\mathbf{K}}(t)^*\|_{\mathbb{R}_w^{p,q}} = \|\mathbf{K}(t) - \tilde{\mathbf{K}}(t)\|_{\mathbb{R}_w^{q,p}} \quad \text{for } t \in [0,T]. \tag{65}$$

Since each of the conditions in (56) ensures $\|\mathbf{K} - \tilde{\mathbf{K}}\|_{\mathbb{R}_w^{q,p}} < \delta/2\sqrt{T}$, we have by means of (63) - (65) that $\epsilon_D^* < \delta/2$, which concludes the proof. □

Remark 17. It remains to investigate the accuracy of the respective approximation of the (possibly regularized) pseudo-inverses $(\mathbb{G}^*\mathbb{G} + \alpha\mathbb{I})^{-1}\mathbb{G}^*$ with $\alpha \geq 0$, which play an important role e.g. in optimal control problems.

8. Applications and numerical results

8.1. A test problem

As test problem, we consider the heat equation with homogeneous Dirichlet boundary conditions in Example 1 with input and output operators of the following form, cf. Fig 3. Let $\Omega \subset \mathbb{R}^2$ be a set with sufficiently smooth boundary, and let

$$\begin{aligned} \Omega_c &= \{(x_1,x_2) \in \Omega : \quad a_{c,1} < x_1 \leq b_{c,1}, \quad a_{c,2} < x_2 \leq b_{c,2}\}, \\ \Omega_m &= \{(x_1,x_2) \in \Omega : \quad a_{m,1} < x_1 \leq b_{m,1}, \quad a_{m,2} < x_2 \leq b_{m,2}\}, \end{aligned}$$

with 4 appropriate points a_c, b_c, a_m, $b_m \in \bar{\Omega}$, be rectangular subdomains where the control is active and the observation takes place, respectively. Setting $U = Y = L^2(0,1)$, we define $B \in \mathscr{L}(U, L^2(\Omega))$ and $C \in \mathscr{L}(L^2(\Omega), Y)$ by

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

$$(Cz)(\xi) = \frac{1}{b_{m,1} - a_{m,1}} \int_{a_{m,1}}^{b_{m,1}} z(x_1, x_2(\xi))\, dx_1$$

where $\omega_c \in L^2(a_{c,2}, b_{c,2})$ is a weight function and

$$\begin{aligned} \theta &: [a_{c,1}, b_{c,1}] \to [0,1], \qquad \theta(x_1) = (x_1 - ac, 1)/(b_{c,1} - a_{c,1}), \\ x_2 &: [0,1] \to [a_{m,1}, b_{m,1}], \quad x_2(\xi) = a_{m,2} + \xi(b_{m,2} - a_{m,2}). \end{aligned}$$

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

As special case, let $T = 1$ and

$$\Omega = [0,1]^2, \quad \Omega_c = \Omega, \quad \Omega_m = [0.1, 0.2] \times [0.1, 0.9], \quad \omega_c(x_2) = \sin(\pi x_2). \tag{66}$$

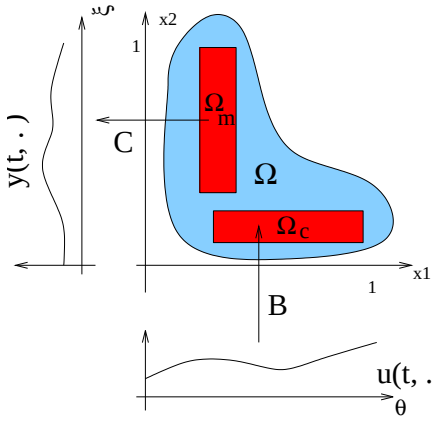

Figure 3: Test case 1 - Heat equation with homogeneous Dirichlet boundary conditions: The input (heating) is active on Ω_c, its intensity can be varied in time t and in the x_1 direction. The observation is the distribution of the temperature at time t in the strip Ω_m in the x_2-direction, averaged in x_1 direction.

For inputs of the special form

$$u(t;\theta) = \sin(\omega_T \pi t)\sin(m\pi\theta), \quad \omega_T, m \in \mathbb{N} \tag{67}$$

we have

$$(Bu)(t;x_1,x_2) = \sin(\omega_T \pi t)\sin(m\pi x_1)\sin(\pi x_2).$$

Since the Laplace operator $\triangle : H^2([0,1]^2) \cap H_0^1([0,1]^2) \to L^2([0,1]^2)$ has the eigenfunctions $v_{m,n}(x_1,x_2) = 2\sin(m\pi x_1)\sin(n\pi x_2)$ to the eigenvalue $\lambda_{m,n} = -2(m^2+n^2)\pi^2$, the corresponding state and, hence, the output can be explicitely calculated and we obtain

$$y(t,\xi) = \frac{\left(\omega_T\,\pi\,e^{\lambda_{m,1}\,t} - \omega_T\,\pi\,\cos\left(\omega_T\,\pi t\right) - \lambda_{m,1}\sin\left(\omega_T\,\pi t\right)\right)}{\lambda_{m,1}^2 + {\omega_T}^2\pi^2} \cdot \ldots$$
$$\frac{\cos(0.1m\pi) - \cos(0.2m\pi)}{0.1m\pi}\sin(\pi(0.1+0.8\xi))$$

8.2. Numerical PDE solution in practice

In order to calculate the matrix approximation $\tilde{\mathbf{G}}$ of the i/o-map $\mathbb{G}$ of the test problem (66), the underlying homogeneous heat equations (34) are solved numerically. Two different implementations have been carried out and investigated.

The first implementation realizes the standard approach of first carrying out a FEM space discretization and then applying a stiff ODE solver, cf. Remark 11. The corresponding solver code is written in MATLAB [38] and uses the FEM software FEMLAB [13]. The spatial meshes are controlled by means of FEMLAB's mesh parameter '`hauto`',

which automatically adjusts the global mesh size and a number of related mesh parameters [13]. FEMLAB works with triangular meshes and provides several types of shape functions. In the presented calculations quadratic Lagrange elements have been used. Local mesh refinements in Ω are carried out *a priori* in the support of $b_l = B\mu_l$ and $\nu_k^* = C^*\nu_k$ and in the neighborhoods of the supports in order to capture steep gradients which can be expected there. However, no error control with respect to the space discretization error is applied so far. By default, FEMLAB solves the resulting stiff ODEs by means of MATLAB's *ode15s*, a variable order solver based on numerical differentiation formulas and optionally using backward differentiation formulas, which has an adaptive step size control with respect to the user-provided tolerance parameters '`atol`' and '`rtol`' [13]. A resulting typical step size distribution can be seen in Fig. 4.

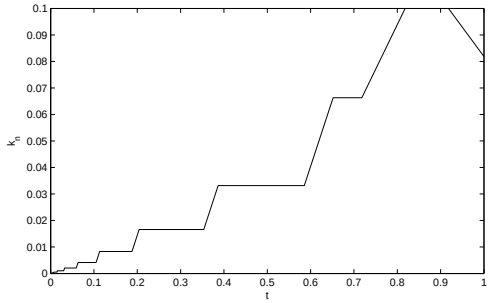

Figure 4: Typical time step size distributions. FEMLAB's stiff ODE solver (227 time steps).

The second implementation, which is based on the C++ FEM software package DEAL.II [5], realizes a discontinuous Galerkin scheme (46) and applies goal-oriented DWR-based error control, see [34] for details. We choose the ansatz spaces $\mathcal{V}_{h_3,\tau_3}$ as functions which are piecewise constant in time and piecewise polynomial in space, leading to a backward-Euler like scheme. However, the spatial grids can vary from time step to time step. The spatial grids are based on quadrilaterals and hexahedra and several shape functions can be chosen. We work with linear Lagrangian ansatz functions for the primal problem and quadratic ansatz functions for the dual problem.

We apply the DWR-based error estimation for $J_l(z_l - z_{l,tol}) = \|\mathbf{K}_{:,l} - \tilde{\mathbf{K}}_{:,l}\|_{L^(0,T;\mathbb{R}^q)}$ in the following way, cf. Section 6.3. The dual problems (52) are solved numerically, i.e. for $l = 1, \ldots, p$ we find a numerical approximation $\tilde{w}_l \in \mathcal{W}_{h_3^*,\tau_3^*}$ of the exact solution $w_l \in W(0,T)$ satisfying

$$B(v, \tilde{w}_l) = \hat{J}_l(v) \qquad \text{for all } v \in \mathcal{W}_{h_3^*,\tau_3^*}. \tag{68}$$

Here, $\mathcal{W}_{h_3^*,\tau_3^*}$ is an appropriately chosen finite-dimensional space, and $\hat{J}_l(v) = (v, \hat{\sigma}_l)_{L^2(0,T;L^2(\Omega))}$ is the approximation of the exact error functional $J_l(v) = (v, \sigma_l)_{L^2(0,T;L^2(\Omega))}$ resulting from estimating the error $\mathbf{E}(t) = \mathbf{K}(t) - \tilde{\mathbf{K}}(t)$ by $\hat{\mathbf{E}}(t)$. The error $\hat{\mathbf{E}}$ can, for instance,

be estimated by interpolating a calculated solution $z_{l,tol} \in \mathcal{V}_{h_3,\tau_3}$ in a larger ansatz space $\mathcal{V}_{h_3',\tau_3'} \supset \mathcal{V}_{h_3,\tau_3}$ and taking the difference as an estimation for the state error. Note that the dual problems (52) correspond, in the case of the heat equation and the chosen type of error functionals, to heat equations backwards in time, cf. [26],

$$\begin{aligned} -\partial_t w_l(t) &= \triangle w_l(t) + \sigma_l(t), \quad t \in [0,T), \\ w_l(T) &= 0. \end{aligned}$$

The approximate dual solutions $\tilde{w}_l$ are then used to directly evaluate the error representation formula (53), i.e. for all $w \in \mathcal{V}_{h_3,\tau_3}$,

$$\begin{aligned} J_l(z_l - z_{l,tol}) \approx \tilde{J}_l(z_l - z_{l,tol}) &:= B(z_l - z_{l,tol}, \tilde{w}_l) \\ &= (b_l, (\tilde{w}_l - w)_0^+)_{L^2(\Omega)} + B(z_{l,tol}, \tilde{w}_l - w) \\ &= \sum_{n=1}^{N} \sum_{\bar{\omega} \in \mathscr{T}_{h_n}} \tilde{\eta}_{m,\bar{\omega}} \end{aligned}$$

A solution $z_{l,tol}$ is accepted if for some fixed $\alpha \in [0,1)$ and given tolerance $\mathtt{tol}$,

$$\alpha\,\mathtt{tol} \le \tilde{J}_l(z_l - z_{l,tol}) < \mathtt{tol}. \tag{69}$$

Otherwise the local error indicators $\tilde{\eta}_{m,\bar{\omega}}$ are used in the following way to adapt the time grid and the meshes $\mathscr{T}_{h_n}$. An $\alpha > 0$ makes sense if one is interested in a reduced number of degrees of freedom in order to achieve a given accuracy.

Defining $\tilde{\eta}_n = \sum_{\bar{\omega} \in \mathscr{T}_{h_n}} \tilde{\eta}_{n,\bar{\omega}}$ and scaling $\tilde{\eta}_{n,\bar{\omega}}$ and $\tilde{\eta}_n$ to $\bar{\eta}_{n,\bar{\omega}}$ and $\bar{\eta}_n$ such that

$$\tilde{\eta}_{n,\bar{\omega}} = \frac{k_n}{T}\frac{|\bar{\omega}|}{|\Omega|}\bar{\eta}_{n\bar{\omega}}, \qquad \tilde{\eta}_n = \frac{k_n}{T}\bar{\eta}_n,$$

we observe that (69) is ensured if one of the following conditions holds

$$\begin{aligned} &\text{(i)} \quad \alpha\mathtt{tol} \le \bar{\eta}_n < \mathtt{tol}, \quad \text{for all } n = 1,\dots,N, \\ &\text{(ii)} \quad \alpha\mathtt{tol} \le \bar{\eta}_{n,\bar{\omega}} < \mathtt{tol}, \quad \text{for all } n = 1,\dots,N, \quad \bar{\omega} \in \mathscr{T}_{h_n}. \end{aligned}$$

A simple mesh adaption strategy which aims at achieving (69) is the following. Refine a time step k_n if $\bar{\eta}_n > \mathtt{tol}$, and refine a mesh element $\bar{\omega} \in \mathscr{T}_{h_n}$ if $\bar{\eta}_{n,\bar{\omega}} > \mathtt{tol}$. Reduce the number of time steps by merging neighboring time steps k_n which all satisfy $\bar{\eta}_n < \alpha\mathtt{tol}$ and reduce the number of mesh elements by merging neighboring elements $\bar{\omega}$ with $\bar{\eta}_{n,\bar{\omega}} < \alpha\mathtt{tol}$.

The underlying idea of balancing the error contribution over the space time domain is the basis of many mesh adaption strategies. Note that it is not clear that an iterative scheme based on this refinement strategy does actually produce a solution satisfying (69), in particular for α near 1, and nothing is said about the optimality of the space and time grids in the sense of a minimal error with respect to the degrees of freedom. Moreover, for the practical mesh and grid handling it is preferable to limit the number of hanging nodes, i.e. a refinement strategy should not only focus on single mesh elements and time steps but also take the neighboring elements into account. Many other refinement strategies are discussed in the literature, cf. [18].

Remark 18. Numerical experiments show that $\tilde{J}_l(z_l - z_{l,tol})$ only approximates $J_l(z_l - z_{l,tol})$ *quantitatively* well if the dual solution w_l is approximated with rather high accuracy [34]. This is comprehensible in view of the following considerations. Neglecting rounding errors and assuming that $J_l = \hat{J}_l$, we focus on the estimation error introduced by approximating $\tilde{w}_l \approx w_l$. We note that the primal state error $e_l := z_l - z_{l,tol}$ and the dual state error $e_l^* := w_l - \tilde{w}_l$ satisfy the following Galerkin orthogonality conditions,

$$B(e_l, v) = 0 \quad \text{for } v \in \mathcal{V}_{h_3,\tau_3}, \qquad \text{and} \qquad B(w, e_l^*) = 0 \quad \text{for } w \in \mathcal{W}_{h_3^*,\tau_3^*}.$$

A direct consequence is that the ansatz space $\mathcal{W}_{h_3^*,\tau_3^*}$ for $\tilde{w}_l$ must not coincide with $\mathcal{V}_{h_3,\tau_3}$, since then $\tilde{J}_l(e_l) = B(e_l, \tilde{w}_l) = 0$. Denoting the B-orthogonal projection onto $\mathcal{W}_{h_3^*,\tau_3^*}$ by $\mathbb{P}_{\mathcal{W}}$, i.e.

$$B((\mathbb{I} - \mathbb{P}_{\mathcal{W}})v, w) = 0 \quad \text{for all } v \in \mathcal{V},\ w \in \mathcal{W}_{h_3^*,\tau_3^*},$$

we observe

$$\tilde{J}_l(e_l) = B(e_l, \tilde{w}_l) = B(\mathbb{P}_{\mathcal{W}} e_l, \tilde{w}_l) = B(\mathbb{P}_{\mathcal{W}} e_l, \tilde{w}_l) + B(\mathbb{P}_{\mathcal{W}} e_l, e_l^*), \tag{70a}$$

$$= B(\mathbb{P}_{\mathcal{W}} e_l, w_l) = J_l(\mathbb{P}_{\mathcal{W}} e_l). \tag{70b}$$

This means, the practical implementation of the DWR approach catches only the contribution of $\mathbb{P}_{\mathcal{W}} e_l$ to the exact error $J_l(e_l)$, and that $J_l(e_l) - \tilde{J}_l(e_l) = B((\mathbb{I} - \mathbb{P}_{\mathcal{W}})e_l, e_l^*)$ is neglected.

In practical applications one aims to start with coarse (primal and dual) meshes and to refine subsequently only where it is necessary, and to spend less numerical effort on the error estimation than on the original problem. First numerical tests indicate that the approximated local error indicators $\tilde{\eta}_{n,\bar{\omega}}$ approximate the exact terms $\eta_{n,\bar{\omega}}$ *qualitatively* well even for coarse primal and dual meshes, such that it seems reasonable to use them for adaptive space and time mesh refinement anyway, systematic investigations of this topic are ongoing [34].

We conclude this subsection with a remark on the numerical costs and with a short discussion of the two implementations.

Remark 19. The numerical solution of the (possibly many) homogeneous PDEs represents the key cost factor in the proposed approximation framework, such that code-optimization is an important issue. For instance, much calculation time can be saved if the assembling of the mass and stiffness matrices does not have to be carried out for each initial value $b_l = B\mu_l$, $l = 1, \ldots, p$ again. However, the support of the b_l differ and local refinements in the neighborhood of their supports are reasonable, such that different spatial meshes may be useful. An intelligent updating of meshes and the corresponding mass and stiffness matrices may therefore speed up the calculation of the i/o-approximation considerably.

The FEMLAB implementation stands out for its high flexibility and easy adaptivity to new problem formulations, i.e. geometries, boundary conditions and the partial differential equations can easily be modified. However, a posteriori error control for

the time *and* space discretization error has still to be implemented in order to obtain reliable approximations and thus to be useful for the proposed i/o-map approximation framework. Though FEMLAB provides tools for comfortable spatial mesh refinement, tools for the rigorous error control of time-depending problems are missing and have to be developped by the user, and spatial meshes which vary in time will be very difficult to implement.

The DEAL.II-based code has the advantage of providing reliable results with errors within a given tolerance range and producing problem-optimized space and time grids by means of DWR-based error control. However, in the current implementation this is achieved by a very accurate solution of the dual PDEs and thus with considerable extra costs.

This inconvenience is acceptable in the present academical context of investigating the influence of the accuracy in the PDE solutions on the overall accuracy of the i/o map approximation. In practice, the effort of calculating a very accurate dual solution is too high and would better be invested directly in a more accurate solution of the original problems. An optimization and cost-reduction for the error estimation procedure is thus necessary, probably involving appropriate a posteriori estimations for the global error in the dual solution, see e.g. [1]. We note that the extra costs due to the numerical solution of the dual problem may be outweighed by the reduction of computational costs due to optimized meshes and by the advantage of reliable results more easily in the case of nonlinear problems, see e.g. [27].

Remark 20. Since the DWR-based error estimation is still under development, the i/o approximations used in the presented numerical experiments have predominantly been calculated with the FEMLAB implementation and therefore miss quantitative information about their accuracy with respect to the dynamics approximation. I/o approximations with rather accurate error information on the basis of the DEAL.II-based implementation are ongoing work.

8.3. Tests of convergence

Convergence of single outputs

We consider the test problem (66), apply inputs of the form (67) with several choices for m and ω_T, and compare $y = \mathbb{G}u$ with $\tilde{y} = \mathbb{G}_{DS}u$ for different approximations $\mathbb{G}_{DS}(\mathbf{h})$.

In Fig. 5, the input $u(t;\theta) = \sin(3\pi t)\sin(3\pi\theta)$ has been chosen, and the relative error $\|\mathbb{G}u - \mathbb{G}_{DS}u\|_{\mathcal{Y}}/\|u\|_{\mathcal{U}_s}$ is plotted for the following approximations $\mathbb{G}_{DS}$. The black lines (labeled in the legend as G_1) correspond to approximations $\mathbb{G}_{DS}$ which have been calculated with the FEMLAB-based code with moderately high tolerance tol_{PDE} (i.e. FEMLAB parameters `hauto` $= 5$, `atol` $= 1e-3$ and `rtol` $= 1e-2$). The red lines (labeled in the legend as G_2) correspond to approximations $\mathbb{G}_{DS}$ which have been calculated with the same code but a lower tolerance tol_{PDE} (i.e. `hauto` $= 3$, `atol` $= 1e-6$ and `rtol` $= 1e-5$). The signal discretizations have been carried out with hierarchical linear finite elements in space and Haar wavelets in time.

In the upper row, the number of *space* ansatz functions μ_i and ν_i is varied with $p = q = 5, 9, 17, 33, 65$ whereas the number of time basis functions ϕ_i and ψ_i is fixed to $r = s = 64$. According to Cor. 1, a quadratic convergence in $h_1 = h_2$ is expected and can also be approximately observed: The dotted lines in the upper left image show best (in the least square sense) quadratic fits $f(h_1) = a + ch_1^2$ and the corresponding estimated limit a. The corresponding values of the two constants $a, c \in \mathbb{R}$ are given in Table 1. The error does not seem to converge to zero but to a value a, which can be mainly explained by the still present dynamic approximation error. Note that a corresponding to the better dynamical approximation G_2 is lower than a corresponding to G_1. The right upper image shows the convergence *to the estimated limit* a in a double-logarithmic scale.

In the lower row, the number of *time* ansatz functions ϕ_i and ψ_i is varied with $r = s = 4, 8, 16, 32, 64$ whereas the number of space basis functions μ_i and ν_i is fixed to $p = q = 65$. Here we expect and also observe a linear convergence, the dotted lines show now linear best fits $f(h_1) = a + ch_1$ and the corresponding estimated limits a. The lower right image shows again the convergence *to the estimated limit* a in a double-logarithmic scale.

Fig. 6 shows similar numerical results for a more oscillatory input $u(t;\theta) = \sin(10\pi t)\sin(10\pi\theta)$. Though the order of magnitude of the corresponding relative errors is comparable to the results of the less oscillatory input, a quadratic, respectively, linear convergence can no longer be observed. It is remarkable that the approximations $\mathbb{G}_{DS}$ calculated with different tolerances tol_{PDE} lead to the same errors with respect to the space discretization (cf. 6-upper row), and that the errors corresponding to $\mathbb{G}_{DS}$ with more accurate dynamics approximation become larger than the errors corresponding with less accurate dynamics for small $\tau_1 = \tau_2$. Since we do not have convincing explications for these observations, further investigations are necessary.

Fig. 6 shows the output errors over the space-time domain for the input $u(t;\theta) = \sin(10\pi t)\sin(10\pi\theta)$, again for different approximations $\mathbb{G}_{DS}$. This time, only the numbers p, q, r and s of basis functions for the signal discretizations are varied. The dynamics has been approximated with the FEMLAB-based code with moderately high tolerances (`hauto` $= 5$, `atol` $= 1e-3$ and `rtol` $= 1e-2$).

Another important future task is the investigation, how the preceeding error results correspond to the total error estimation in Thm. 1 and Cor. 2, and how such numerical results can be used in order to estimate unknown constants in these estimates.

Successive signal approximation and convergence of the norm $\|\mathbb{G}_S(\mathbf{h})\|_{\mathscr{L}(\mathcal{U},\mathcal{Y})}$

We investigate numerically the convergence of $\|\mathbb{G}_S(\mathbf{h})\|_{\mathscr{L}(\mathcal{U},\mathcal{Y})}$ for growing nested subspaces U_{h_1}, Y_{h_2}, $\mathcal{R}_{\tau_1}$ and $\mathcal{S}_{\tau_2}$, compare Lemma 3. As an example, we consider the approximation $\mathbb{G}_S(\mathbf{h})$ of the i/o-map $\mathbb{G}$ corresponding to the test problem (66) and choose hierarchical linear finite elements in U_{h_1} and Y_{h_2} and Haar wavelets in $\mathcal{R}_{\tau_1}$ and $\mathcal{S}_{\tau_2}$. We approximate $\|\mathbb{G}_S(\mathbf{h})\|_{\mathscr{L}(\mathcal{U},\mathcal{Y})} \approx \|\mathbb{G}_{DS}(\mathbf{h})\|_{\mathscr{L}(\mathcal{U},\mathcal{Y})}$, where $\mathbb{G}_{DS}$ has been calculated with the FEMLAB-based solver with tolerances `hauto`$= 5$, `atol`$= e-2$, `rtol`$= e-3$ and local support refinements. Figure 8 shows the approximated $\|\mathbb{G}_S(\mathbf{h})\|_{\mathscr{L}(\mathcal{U},\mathcal{Y})}$ for

Fig.	m	ω_T	varied param.	α	tol_{PDE}	a	c
5(top)	3	3	$h_1 = h_2$	2	low (G_1)	3.57e-5	3.49e-3
					high (G_2)	1.71e-5	3.43e-3
5(down)	3	3	$\tau_1 = \tau_2$	1	low (G_1)	2.96e-5	9.49e-4
					high (G_2)	2.14e-5	9.97e-4
6(top)	10	10	$h_1 = h_2$	2	low (G_1)	(-5.64e-5)	(1.21-e3)
					high (G_2)	(-5.34e-5)	(1.24e-3)
6(down)	10	10	$\tau_1 = \tau_2$	1	low (G_1)	(4.72e-6)	(7.07e-5)
					high (G_2)	(8.18e-6)	(3.45e-5)

Table 1: Coefficients a and c of best fitting curves $f(h) = a+ch^\alpha$ with varied parameter h and expected convergence order α.

subspace dimensions

$$p = q = r + 1 = s + 1 = 2, 3, \ldots, 65.$$

In the first row of Figure 9 we only vary one subspace dimension while keeping the others fixed to a maximal $p_{\max} = q_{\max} = 65$ or $r_{\max} = s_{\max} = 64$, respectively. In the second row first the numbers $r = s$ of time bases functions are increased synchronously while fixing $p = q = 65$, second the numbers of space basis functions are increased while fixing $r = s = 64$. In the third image, the number of input basis functions $p = r + 1$ is increased synchronously, while fixing $q = s + 1 = 65$, and in the fourth, the number of output basis functions $q = s+1$ is increased synchronously, while fixing $p = r + 1 = 65$.

It remains to investigate the dependence of this norm convergence of the dynamical error ϵ_D, i.e. for different $\mathbb{G}_{DS} \approx \mathbb{G}_S$. Moreover, it would be interesting to investigate further test problems. Aiming to test also the convergence for higher subspace dimensions $p, q > 65$ and $r, s > 64$, we note that the norm calculations get increasingly costly due to the involved singular value decompositions.

Convergence in the operator norm

Next we investigate the behavior of $\|\mathbb{G}_{DS}(h_1, h_2, \tau_1, \tau_2, \texttt{tol})\|_{\mathscr{L}(\mathcal{U},\mathcal{Y})}$ for varying discretization parameters. Again, we consider approximations $\mathbb{G}_{DS}$ of the i/o-map of the test problem (66) with one fixed `tol`, calculated with the FEMLAB code. In Figure 10,

$$\|\mathbb{G}_{DS}(h_1, h_2, \tau_1, \tau_2, \texttt{tol}) - \mathbb{G}_{DS}(\frac{h_1}{2}, \frac{h_2}{2}, \frac{\tau_1}{2}, \frac{\tau_2}{2}, \texttt{tol})\|_{\mathscr{L}(\mathcal{U},\mathcal{Y})} \tag{71}$$

is plotted over $h_1 = h_2 = \tau_1 = \tau_2$. The norms are calculated via the matrix norm (15) and appropriate embeddings of $\tilde{\mathbf{G}}(h_1, h_2, \tau_1, \tau_2, \texttt{tol}) \in \mathbb{R}_w^{pq,rs}$ into the larger space $\mathbb{R}_w^{\bar{p}\bar{q},\bar{r}\bar{s}}$ with $\tilde{\mathbf{G}}(\frac{h_1}{2}, \frac{h_2}{2}, \frac{\tau_1}{2}, \frac{\tau_2}{2}, \texttt{tol}) \in \mathbb{R}_w^{\bar{p}\bar{q},\bar{r}\bar{s}}$.

In Figure 11,

$$\|\mathbb{G}_{DS}(h_1, h_2, \tau_1, \tau_2, \texttt{tol}) - \mathbb{G}_{DS}(\frac{h_1}{2}, \frac{h_2}{2}, \tau_1, \tau_2, \texttt{tol})\|_{\mathscr{L}(\mathcal{U},\mathcal{Y})} \tag{72}$$

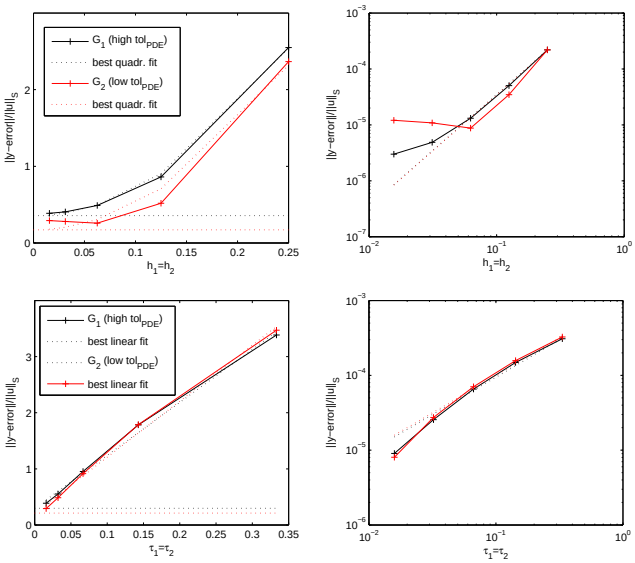

Figure 5: Relative output errors for input (67) with $m = 3$ and $\omega_T = 3$. Upper row: error for $h_1 = h_2 \to 0$ and $\tau_1 = \tau_2 = 1/64$. Lower row: error for $\tau_1 = \tau_2 \to 0$ and $h_1 = h_2 = 1/64$. (Note: the values in the images on the left have to be scaled by the factor 10^{-4}.)

is plotted over $h_1 = h_2$ for fixed $\tau_1 = \tau_2 = 1/64$ and in Figure 12,

$$\|\mathbb{G}_{DS}(h_1, h_2, \tau_1, \tau_2, \texttt{tol}) - \mathbb{G}_{DS}(h_1, h_2, \frac{\tau_1}{2}, \frac{\tau_2}{2}, \texttt{tol})\|_{\mathscr{L}(\mathcal{U},\mathcal{Y})} \tag{73}$$

is plotted over $\tau_1 = \tau_2$ for fixed $h_1 = h_2 = 1/64$.

Unfortunately, it is not possible to observe a convergence rate of at least one as suggested by Corollary 2. A possible reason is that the dynamical approximation error ϵ_D is too large and destroys the order. It therefore remains to investigate the convergence with approximations $\mathbb{G}_{DS}$ with lower error ϵ_D. It would be preferable to investigate directly the $\mathscr{L}(\mathcal{U}_s, \mathcal{Y})$-norm, but we do not have a convenient discrete counterpart of this norm for the chosen signal basis functions.

8.4. Matrix reduction on the basis of SVDs

In order to resolve the input signal space and the output signal space by means of general purpose basis functions like finite elements or Haar wavelets with sufficient accuracy, a large number of basis functions is needed in general, thus leading to i/o-matrices $\tilde{\mathbf{G}}$ of relative large dimensions. We propose a matrix reduction which is based on singular value decompositions (SVDs) but which preserves the 'space-time'

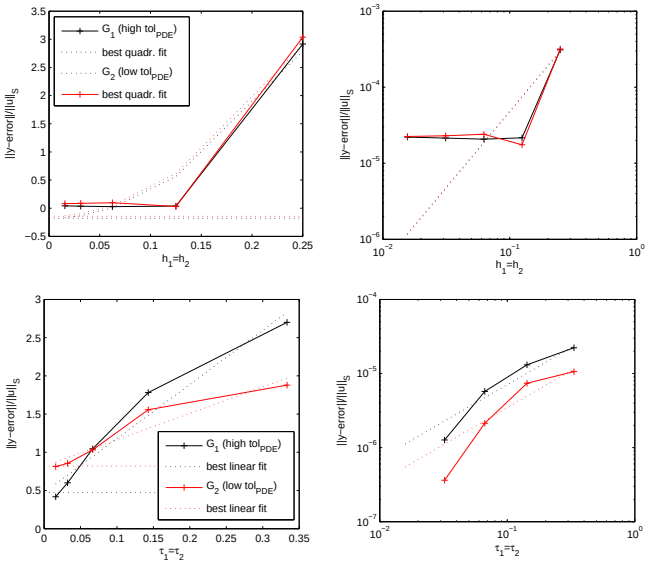

Figure 6: Relative output errors for input (67) with $m = 10$ and $\omega_T = 10$. Left: error for $h_1 = h_2 \to 0$ and $\tau_1 = \tau_2 = 1/64$. Right: error for $\tau_1 = \tau_2 \to 0$ and $h_1 = h_2 = 1/64$. (Note: the values in the images on the left have to be scaled by the factor 10^{-4}.)

tensor structure of the input and output signal bases [42]. We note that similar reductions have been developed independently before, and can be found in literature e.g. as 'higher order singular value decomposition' (HOSVD), see e.g. [14], [15] and [16] and the references therein. These tensor decompositions are mostly motivated by applications from higher order statistics, e.g. in psychology, but also in computer facial recognition, image processing, telecommunications and other fields [35], but are not aware that they have been applied for the model reduction of linear control systems.

Such a system reduction can be useful in the following ways.

- To obtain a low-dimensional matrix-representation of the system, which is small enough to be used for real-time feedback control design.
- To identify relevant input and output signals. They may be instructive for the actuator and sensor design, i.e. they might help to answer where actuators and sensors have to be placed and which resolution in time and space they should have.

Roughly, the matrix reduction for $\tilde{\mathbf{G}}$ works in the following way.

1. Put all elements belonging to $\mu_1, \mu_2, \ldots, \mu_p$ into one respective column and carry out SVD on the corresponding $(qsr) \times p$ matrix to determine the relevant linear combinations of μ_i's.

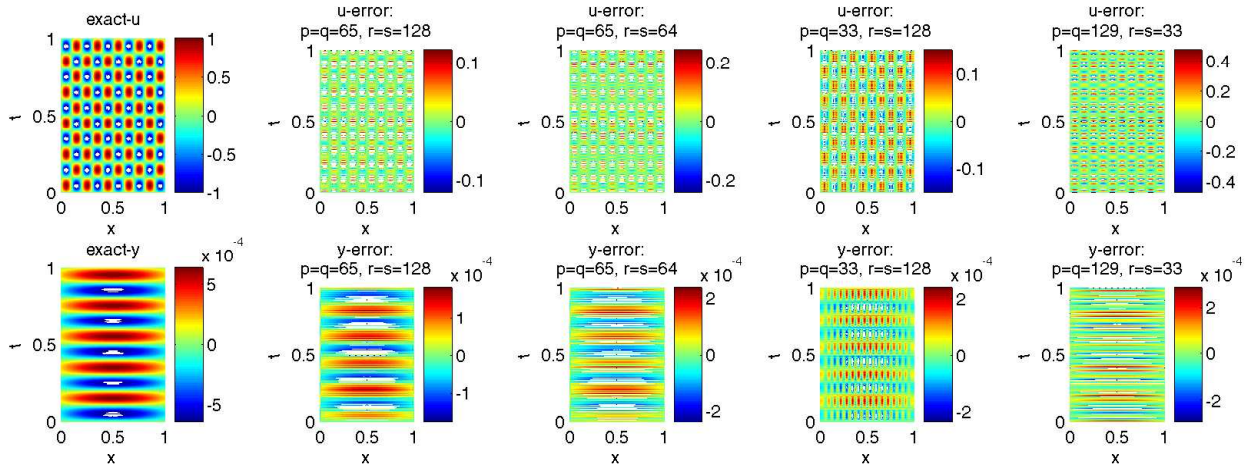

Figure 7: Output errors over the space-time domain for input with $m = 10$, $\omega_T = 10$

2. Put all elements belonging to $\nu_1, \nu_2, \ldots, \nu_q$ into one respective row and carry out SVD on the corresponding $q \times (psr)$ matrix to determine the relevant linear combinations of ν_i's.

3. Put all elements belonging to $\phi_1, \phi_2, \ldots, \phi_r$ into one respective column and carry out SVD on the corresponding $(qsp) \times r$ matrix determine the relevant linear combinations of ϕ_i's.

4. Put all elements belonging to $\psi_1, \nu_2, \ldots, \nu_s$ into one respective row and carry out SVD on the corresponding $s \times (psq)$ matrix to determine the relevant linear combinations of ψ_i's.

5. Transform $\tilde{\mathbf{G}}$ with respect to the new basis functions

6. Truncate columns and rows corresponding to new time basis functions of low relevance and truncate block rows and block colums corresponding to space time basis functions of low relevance.

For a detailed and formalized description see [42] and primariliy [14], [15] and [16] for truncation error results. The relevance of basis functions can be evaluated by means of the decay of the singular values of the respective SVDs. A more significant criterion is to consider so-called *n-mode singular values* which are related to a so-called *core tensor*, which is the higher-order tensor analogon to the matrix Σ in a SVD $A = U\Sigma V'$ of a matrix A, cf. [14].

In Figure 13 the tensor SVD has been applied to a matrix $\tilde{\mathbf{G}} \in \mathbb{R}^{qs,pr}$ approximating the i/o-map $\mathbb{G}$ of the test problem, with $p = q = 65$ hierarchical linear finite element basis functions in space and $r = s = 64$ Haar wavelet basis functions in time. The first row shows the singular values of the SVDs corresponding to steps (1)-(4) in the above

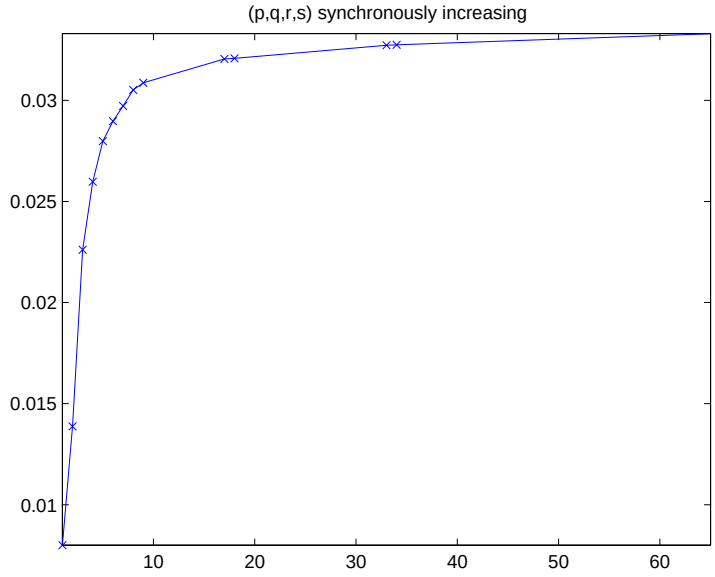

Figure 8: Norm $\|\mathbb{G}_{DS}(\mathbf{h})\|_{\mathscr{L}(\mathcal{U},\mathcal{Y})}$ for synchronously increasing approximation space dimensions $p = q = r + 1 = s + 1$ and one fixed tolerance `tol`.

procedure in a semilogarithmic scale. Underneath the first and most relevant three transformed/new basis functions $\hat{\mu}_i$, $\hat{\nu}_i$, $\hat{\phi}_i$ and $\hat{\psi}_i$, are plotted. It is not surprising that μ_1 is large in the neighborhood of $[0.1, 0.2]$, since the observation takes place in $[0.1, 0.2] \times [0.1, 0.9]$. We do not have an explanation yet, why the transformed basis functions $\hat{\nu}_i$ do not show the symmetry of the underlying problem. The most relevant input and output time basis functions seem to be the sinusoidal functions and the corresponding singular values decay slowly. The direct coupling of inputs and outputs via overlapping input and output domains may be a reason why also higher frequencies keep to be important, though the heat equation is known for damping such higher oscillations.

It remains to investigate other test problems and to apply truncation error results in order to obtain global error estimates for $\hat{\mathbb{G}}_{DS} - \mathbb{G}$, where $\hat{\mathbb{G}}_{DS}$ is the tensor SVD reduced approximation of $\mathbb{G}$. An application of a reduced approximation $\hat{\mathbb{G}}_{DS}$ in an optimization problem follows in the next subsection.

8.5. Unconstrained optimization

We investigate the use of the i/o-map approximation in optimization problems of the following form.

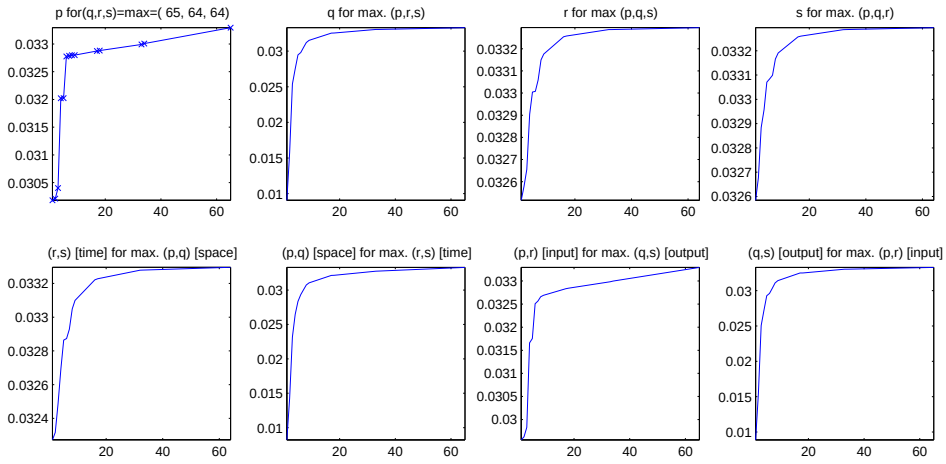

Figure 9: Norm $\|\mathbb{G}_{DS}(\mathbf{h})\|_{\mathscr{L}(\mathcal{U},\mathcal{Y})}$: one fixed tolerance `tol`, varying the dimensions of one subspace (upper row) or two subspaces (lower row) while keeping the other dimensions fixed.

Let J be a quadratic cost functionals of the *tracking type*,

$$J : \mathcal{U} \times \mathcal{Y} \to \mathbb{R}, \qquad J(u,y) = \frac{1}{2}\|y - y_D\|_{\mathcal{Y}}^2 + \alpha\|u\|_{\mathcal{U}}^2. \tag{74}$$

Here $y_D \in \mathcal{Y}$ is a desired output signal that one wants the system to achieve and $\alpha > 0$ is a regularization parameter. The optimization problem consists in the minimization

$$\min J(u,y) \qquad \text{subject to } y = \mathbb{G}u, \quad u \in \mathcal{U}_{ad} \tag{75}$$

where $\mathcal{U}_{ad} \subset \mathcal{U}$ is the subset of admissible controls.

We define the discretize cost functional

$$\bar{J}_{\mathbf{h}} : \mathbb{R}^{pr} \times \mathbb{R}^{qs} \to \mathbb{R}, \qquad \bar{J}_{\mathbf{h}}(\mathbf{u},\mathbf{y}) = \frac{1}{2}\|\mathbf{y} - \mathbf{y}_D\|_{qs;w}^2 + \alpha\|\mathbf{u}\|_{pr;w}^2, \tag{76}$$

with $\mathbf{y}_D = \kappa_{\mathcal{Y},h_2,\tau_2}\mathbb{P}_{\mathcal{Y},h_2,\tau_2}y_D$, and instead of (75) we solve

$$\min \bar{J}_{\mathbf{h}}(\mathbf{u},\mathbf{y}) \qquad \text{subject to } \mathbf{y} = \tilde{\mathbf{G}}\mathbf{u}, \quad \mathbf{u} \in \bar{U}_{ad} \tag{77}$$

with $\bar{U}_{ad} = \{\mathbf{u} \in \mathbb{R}^{pr} : \mathbf{u} = \kappa_{\mathcal{U},h_1,\tau_1}\mathbb{P}_{\mathcal{U},h_1,\tau_1}u,\ u \in \mathcal{U}_{ad}\}$.

We note that the continuous formulation of (76) and (77) is given by

$$J_{\mathbf{h}} : \mathcal{U} \times \mathcal{Y} \to \mathbb{R}, \qquad J_{\mathbf{h}}(u,y) = \frac{1}{2}\|\mathbb{P}_{\mathcal{Y},h_2,\tau_2}(y - y_D)\|_{\mathcal{Y}}^2 + \alpha\|\mathbb{P}_{\mathcal{U},h_1,\tau_1}u\|_{\mathcal{U}}^2, \tag{78}$$

$$\min J_h(u,y) \qquad \text{subject to } y = \mathbb{G}_{DS}u, \quad u \in \mathbb{P}_{\mathcal{U},h_1,\tau_1}\mathcal{U}_{ad}. \tag{79}$$

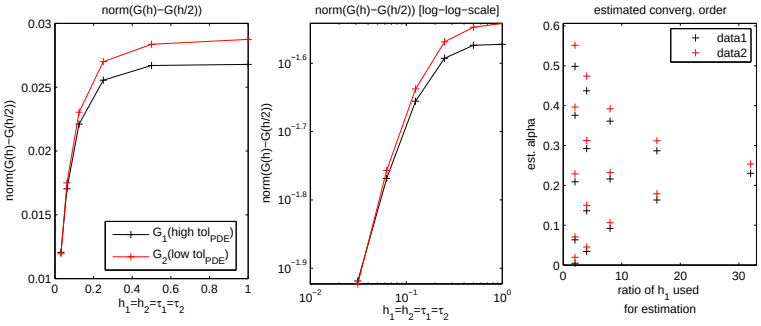

Figure 10: The figure shows (71) over $h_1 = h_2 = \tau_1 = \tau_2$ in a decimal and in a double logarithmic scale, and estimations of the numerical convergence order α.

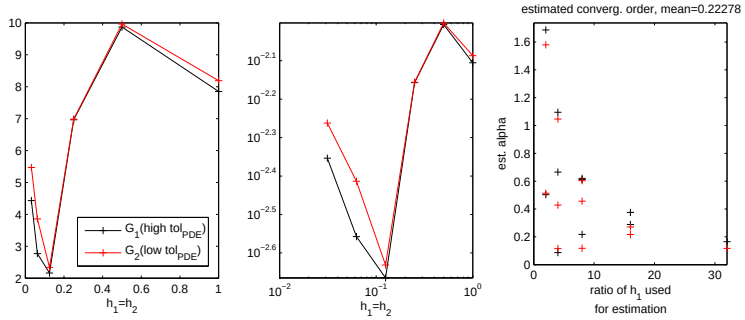

Figure 11: The figure shows (72) over $h_1 = h_2$ in a decimal and in a double logarithmic scale, and estimations of the numerical convergence order α.

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 (77) is characterized by

$$(\tilde{\mathbf{G}}^T \mathbf{M}_{\mathcal{Y}} \tilde{\mathbf{G}} + \alpha \mathbf{M}_{\mathcal{U}}) \bar{\mathbf{u}} = \tilde{\mathbf{G}}^T \mathbf{M}_{\mathcal{Y}} \mathbf{y}_D. \tag{80}$$

As concrete example, we consider the test problem (66) 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 *optimized* input u_* of less energy, such that $\mathbb{G}u_* \approx y_D$, or more exactly, u_* that minimizes the cost functional (75).

Fig. 14 shows the numerical results for $\alpha = 10^{-6}$. Here an i/o map approximation $\tilde{\mathbf{G}} \in \mathbb{R}^{65\cdot 64 \times 65\cdot 64}$ with $p = q = 65$ hierarchical linear basis functions in space and $r = s = 64$ Haar wavelets in time is used, the calculation of $\bar{\mathbf{u}}^* = \kappa_{\mathcal{U}} \bar{u}^*$ via (80) on a normal desktop PC takes 12.9 seconds. The u-norm is reduced by 22.7% and the relative deviation of $\mathbb{G}\bar{u}_*$ from y_D is only 2.4%. The outputs resulting from u_0 and u_* have been calculated in simulations independent from the calculation of the i/o-matrix.

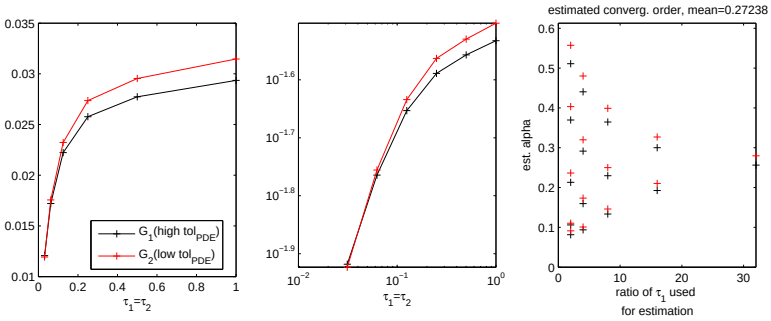

Figure 12: The figure shows (73) over $\tau_1 = \tau_2$ in a decimal and in a double logarithmic scale, and estimations of the numerical convergence order α.

In Fig. 15 the same calculations have been carried out with $\hat{\mathbf{G}} \in \mathbb{R}^{5\cdot5,5\cdot5}$, where $\hat{\mathbf{G}}$ arises from a tensor SVD-based matrix reduction of $\tilde{\mathbf{G}} \in \mathbb{R}^{65\cdot64,65\cdot64}$, where all but the five most relevant input and output basis functions in space and time have been truncated, cf. Fig. 13. Using this approximation, the norm of u is reduced by 21.3%, whereas the relative deviation of $\mathbb{G}u_*$ from y_D is 13%. The cost functional has been reduced by 38.0%, and the calculation of u_* took 0.0005 seconds on a normal desktop PC.

It remains to investigate, if analytical results concerning the performance of the approximated optimal solution can be developed. Moreover, it would be interesting to investigate, how $\mathbb{G}_{DS}$ can be directly used in feedback control, e.g. in the framework of model predictive control [11].

9. Final remarks and outlook

We have presented a systematic framework for the discretization of the 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 many practical and technical restrictions in sensor and actuator design like limited spatial and temporal resolutions or the use of piece-wise constant controls and observations due to digital devices directly into account.

For the moment only the example of a heat control system has been considered in detail. It remains to investigate other linear control systems, e.g. based on the wave equation or Stokes equations, and for which the general results also apply. The transfer of results to time-varying or nonlinear systems or pointwise and boundary controls and observations will be less straight forward and presents an interesting research topic.

The numerical costs of the approach are primarily governed by the numerical cal-

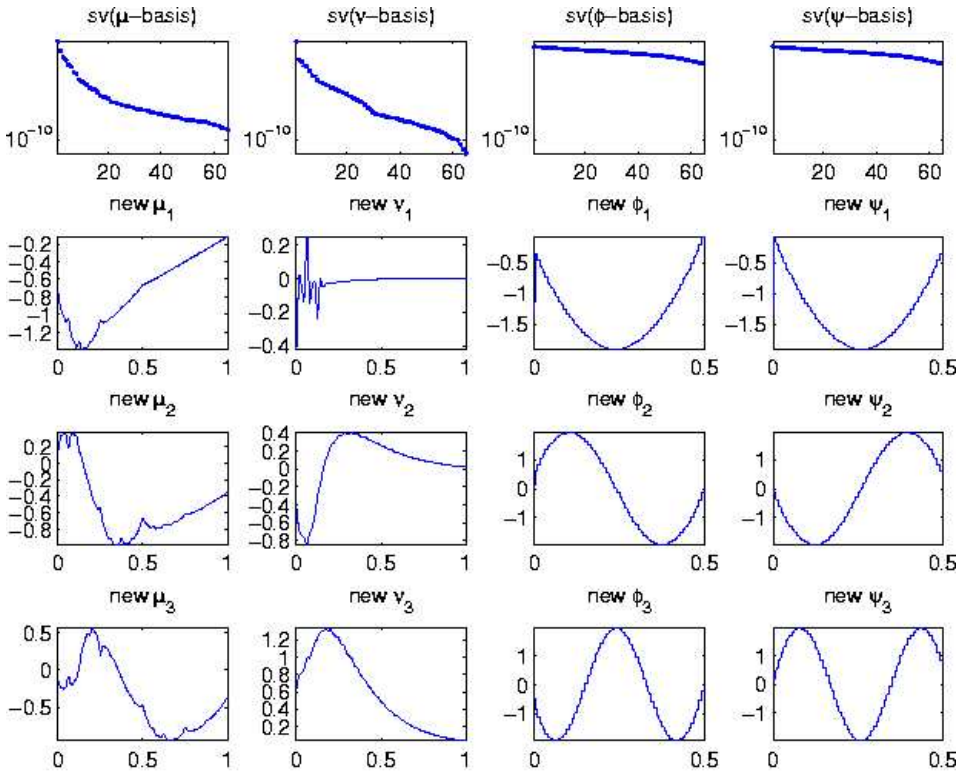

Figure 13: Tensor SVD-based reduction for the approximation of the i/o map of the testproblem (66). First row: Singular values of the 4 respective SVD in semilogarithmic scale. 2nd to 4th row: Respective three most relevant basis functions.

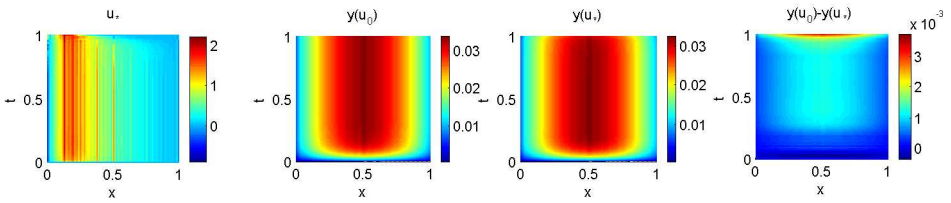

Figure 14: Application of an approximated i/o map $\tilde{\mathbf{G}} \in \mathbb{R}^{65\cdot 64,65\cdot 64}$ in an optimization problem. From left to right: optimized control u_*, original output $y_D = \mathbb{G}u_0$, optimized output $\mathbb{G}u_*$ and their difference.

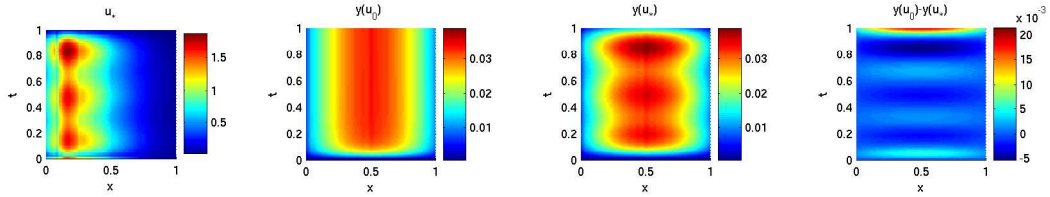

Figure 15: Application of the SVD-reduced approximated i/o map $\hat{\mathbf{G}} \in \mathbb{R}^{5\cdot 5,5\cdot 5}$ in an optimization problem. From left to right: optimized control u_*, original output $y_D = \mathbb{G}u_0$, optimized output $\mathbb{G}u_*$ and their difference.

culation of p underlying homogeneous PDEs, where p is the number of input basis functions in space, which can become problematic when the spatial resolution of the input signal space has to be 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 which has not yet been investigated.

A bottle-neck of the proposed framework may be the large dimensions and memory requirements of the approximations $\tilde{\mathbf{G}}$ due to the full space-time discretization of the i/o-map, in the case where high accuracy of the i/o approximation is required. Therefore it would be worthwile to investigate if data-sparse matrix formats like hierarchical matrices and possibly also related efficient arithmetics can be employed.

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 prove to be useful 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.

Acknowledgement

The author was supported by the *Deutsche Forschungsgemeinschaft* (SFB557). The author would like to thank Christian Kamm for his valuable assistance in the devel-

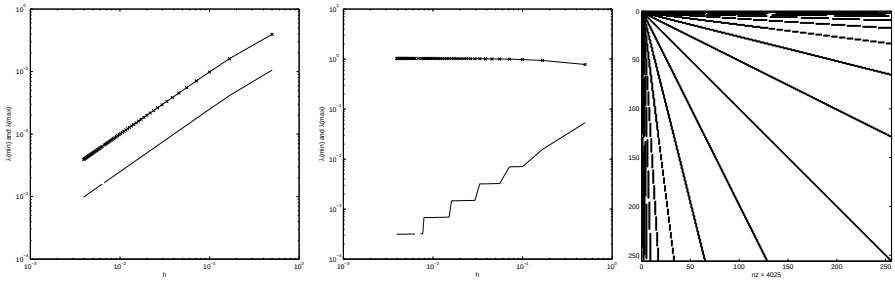

Figure 16: From left to right: maximal and minimal eigenvalues of mass matrices for nodal basis, maximal and minimal eigenvalues of mass matrices for hierarchical basis, nonzero elements in a mass matrix for a hierarchical basis

opment of the DEAL.II-based code for the numerical solution of PDEs.

A. Appendix

A.1. Eigenvalues of mass matrices

In Fig. 16 we can see the maximal and minimal eigenvalues of the mass matrices corresponding to nodal bases and to hierarchical bases, as well as the pattern of nonzero elements of a hierarchical mass matrix. The eigenvalue plots are in agreement with (18) and (19). In order to estimate the constants c_λ, C_λ, c'_λ and C'_λ in (18) and (19) numerically, the eigenvalues of the corresponding mass matrices $\mathbf{M}_{U,h_1}$ have been calculated for $h_1 = 1/(p-1)$, where $p = 3, \ldots, 257$, with MATLAB. For the mass matrices with respect to nodal bases we approximate

$$c_\lambda \approx \min_{3\le p\le 257} \lambda_{\min}(\mathbf{M}_{U,h_1(p)})/h_1 = 0.211325,$$
$$C_\lambda \approx \max_{3\le p\le 257} \lambda_{\min}(\mathbf{M}_{U,h_1(p)})/h_1 = 0.999975$$

and note that $\lambda_{\min}(\mathbf{M}_{U,h_1(p)})/h_1 \to 0.25$ for $h_1 \to 0$ numerically. For the mass matrices with respect to hierarchical bases we approximate

$$c'_\lambda \approx \min_{3\le p\le 257} \lambda_{\min}(\mathbf{M}_{U,h_1(p)})/h_1 = 0.041949,$$
$$C'_\lambda \approx \max_{3\le p\le 257} \lambda_{\min}(\mathbf{M}_{U,h_1(p)}) = 1.030179.$$

References

[1] M. Ainsworth and J. T. Oden. *A posteriori error estimation in finite element analysis*. Pure and Applied Mathematics (New York). Wiley-Interscience [John Wiley & Sons], New York, 2000.

[2] A. C. Antoulas. *Approximation of large-scale dynamical systems*, volume 6 of *Advances in Design and Control*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2005.

[3] A. C. Antoulas, D. C. Sorensen, and S. Gugercin. A survey of model reduction methods for large-scale systems. In *Structured matrices in mathematics, computer science, and engineering, I (Boulder, CO, 1999)*, volume 280 of *Contemp. Math.*, pages 193–219. Amer. Math. Soc., Providence, RI, 2001.

[4] I. Babuska and A. Miller. The post-processing approach in the finite element method. III. A posteriori error estimates and adaptive mesh selection. *Int. J. Numer. Methods Eng.*, 20:2311–2324, 1984.

[5] W. Bangerth, R. Hartmann, and G. Kanschat. `deal.II` *Differential Equations Analysis Library, Technical Reference*. http://www.dealii.org/, 5.2 edition, sept 2005. First edition 1999.

[6] W. Bangerth and R. Rannacher. *Adaptive finite element methods for differential equations*. Lectures in Mathematics ETH Zürich. Birkhäuser Verlag, Basel, 2003.

[7] R. Becker, M. Garwon, C. Gutknecht, G. Bärwolff, and R. King. Robust control of separated shear flows in simulation and experiment. *Journal of Process Control*, 15:691–700, 2005.

[8] R. Becker, V. Heuveline, and R. Rannacher. An optimal control approach to adaptivity in computational fluid mechanics. *Int. J. Numer. Methods Fluids*, 40(1-2):105–120, 2002.

[9] R. Becker and R. Rannacher. A feed-back approach to error control in finite element methods: basic analysis and examples. *East-West J. Numer. Math.*, 4(4):237–264, 1996.

[10] R. Becker and R. Rannacher. An optimal control approach to a posteriori error estimation in finite element methods. *Acta Numer.*, 10:1–102, 2001.

[11] E. F. Camacho and F. Allgöwer, editors. *Robust model predictive control*. John Wiley & Sons Ltd., Chichester, 2000. Internat. J. Robust Nonlinear Control 10 (2000), no. 13.

[12] P. G. Ciarlet. *The finite element method for elliptic problems*, volume 40 of *Classics in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2002. Reprint of the 1978 original [North-Holland, Amsterdam; MR0520174 (58 #25001)].

[13] Femlab 3.0 command reference. COMSOL AB, Tegnérgatan 23, 111 40 Stockholm, 2004.

[14] L. De Lathauwer, B. De Moor, and J. Vandewalle. A multilinear singular value decomposition. *SIAM J. Matrix Anal. Appl.*, 21(4):1253–1278, 2000.

[15] L. De Lathauwer, B. De Moor, and J. Vandewalle. On the best rank-1 and rank-$(R_1, R_2, \cdots, R_N)$ approximation of higher-order tensors. *SIAM J. Matrix Anal. Appl.*, 21(4):1324–1342, 2000.

[16] L. De Lathauwer and J. Vandewalle. Dimensionality reduction in higher-order signal processing and rank-$(R_1, R_2, \ldots, R_N)$ reduction in multilinear algebra. *Linear Algebra Appl.*, 391:31–55, 2004.

[17] E. Emmrich. *Gewöhnliche und Operator-Differentialgleichungen*. Vieweg Verlag, Wiesbaden, 2004.

[18] K. Eriksson, D. Estep, P. Hansbo, and C. Johnson. Introduction to adaptive methods for differential equations. In *Acta numerica, 1995*, Acta Numer., pages 105–158. Cambridge Univ. Press, Cambridge, 1995.

[19] K. Eriksson and C. Johnson. Adaptive finite element methods for parabolic problems. II. Optimal error estimates in $L_\infty L_2$ and $L_\infty L_\infty$. *SIAM J. Numer. Anal.*, 32(3):706–740, 1995.

[20] K. Eriksson and C. Johnson. Adaptive finite element methods for parabolic problems. IV. Nonlinear problems. *SIAM J. Numer. Anal.*, 32(6):1729–1749, 1995.

[21] L. C. Evans. *Partial differential equations*, volume 19 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 1998.

[22] J. Gerhard, M. Pastoor, R. King, B.R. Noack, A. Dillmann, M. Morzynski, and G. Tadmor. Model-based control of vortex shedding using low-dimensional galerkin models. *AIAA-Paper 2003-4262*, 2003.

[23] S. Gugercin and A. C. Antoulas. A survey of model reduction by balanced truncation and some new results. *Internat. J. Control*, 77(8):748–766, 2004.

[24] W. Hackbusch, B. N. Khoromskij, and E. E. Tyrtyshnikov. Hierarchical Kronecker tensor-product approximations. *J. Numer. Math.*, 13(2):119–156, 2005.

[25] E. Hairer and G. Wanner. *Solving ordinary differential equations. II*, volume 14 of *Springer Series in Computational Mathematics*. Springer-Verlag, Berlin, second edition, 1996. Stiff and differential-algebraic problems.

[26] R. Hartmann. A posteriori Fehlerschätzung und adaptive Schrittweiten- und Ortsgittersteuerung bei Galerkin-Verfahren für die Wärmeleitungsgleichung. Diplomarbeit, Institut für Angewandte Mathematik, Universität Heidelberg, 1998.

[27] R. Hartmann and P. Houston. Adaptive discontinuous Galerkin finite element methods for nonlinear hyperbolic conservation laws. *SIAM J. Sci. Comput.*, 24(3):979–1004, 2002.

[28] V. Heuveline and R. Rannacher. Duality-based adaptivity in the *hp*-finite element method. *J. Numer. Math.*, 11(2):95–113, 2003.

[29] B. Jacob. *Time-Varying Infinite Dimensional State-Space Systems*. PhD thesis, Universität Bremen, May 1995.

[30] B. Jacob and J. R. Partington. Admissibility of control and observation operators for semigroups: a survey. In *Current trends in operator theory and its applications*, volume 149 of *Oper. Theory Adv. Appl.*, pages 199–221. Birkhäuser, Basel, 2004.

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

[32] C. Johnson and R. Rannacher. On errror control in CFD. In *Hebeker, Friedrich-Karl (ed.) et al., 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* . Vieweg, Braunschweig, 1994.

[33] C. Johnson, R. Rannacher, and M. Boman. Numerics and hydrodynamic stability: toward error control in computational fluid dynamics. *SIAM J. Numer. Anal.*, 32(4):1058–1079, 1995.

[34] C. Kamm and M. Schmidt. Dual-weighted residual error estimation for the heat equation in practice. Technical report, TU Berlin, Institute of Mathematics, in preparation, 2006.

[35] M. E. Kilmer and C. D. Moravitz Martin. Decomposing a tensor. *SIAM News*, 37(9), 2004.

[36] K. Kunisch and S. Volkwein. Galerkin proper orthogonal decomposition methods for parabolic problems. *Numer. Math.*, 90(1):117–148, 2001.

[37] O. Lehmann, D.M. Luchtenburg, B.R. Noack, R. King, M. Morzynski, and G. Tadmor. Wake stabilization using pod galerkin models with interpolated modes. In *Proceedings of the 44th IEEE Conference on Decision and Control and European Conference ECC, Invited Paper 1618*, 2005.

[38] The MathWorks, Inc., Cochituate Place, 24 Prime Park Way, Natick, Mass, 01760. Matlab Version 7.0.4.352 (R14), 2005.

[39] S. Noelle and C. Steiner. Space-time splitting of adjoint error representations for hyperbolic conservation laws. *Submitted to Numerical Methods for Partial Differential Equations*.

[40] M. Pastoor, R. King, B.R. Noack, A. Dillmann, and G. Tadmor. Model-based coherent-structure control of turbulent shear ows using low-dimensional vortex models. *AIAA-Paper 2003-4261*, 2003.

[41] A. Pazy. *Semigroups of linear operators and applications to partial differential equations*, volume 44 of *Applied Mathematical Sciences*. Springer-Verlag, New York, 1983.

[42] M. Schmidt and C. Schröder. A structure preserving SVD-based reduction of input/output maps. Technical report, TU Berlin, Institute of Mathematics, in preparation, 2006.

[43] V. Thomée. *Galerkin finite element methods for parabolic problems*, volume 25 of *Springer Series in Computational Mathematics*. Springer-Verlag, Berlin, 1997.

[44] F. Tröltzsch. *Optimale Steuerung partieller Differentialgleichungen*. Vieweg, 2005.

[45] A. J. Wathen. Realistic eigenvalue bounds for the Galerkin mass matrix. *IMA J. Numer. Anal.*, 7(4):449–457, 1987.

[46] G. Weiss, O. J. Staffans, and M. Tucsnak. Well-posed linear systems—a survey with emphasis on conservative systems. *Int. J. Appl. Math. Comput. Sci.*, 11(1):7–33, 2001.

[47] D. Werner. *Funktionalanalysis*. Springer-Verlag, Berlin, extended edition, 2000.

[48] H. Yserentant. Hierarchical bases in the numerical solution of parabolic problems. In *Large scale scientific computing (Oberwolfach, 1985)*, volume 7 of *Progr. Sci. Comput.*, pages 22–36. Birkhäuser Boston, Boston, MA, 1987.

[49] H. Yserentant. Hierarchical bases. In *ICIAM 91 (Washington, DC, 1991)*, pages 256–276. SIAM, Philadelphia, PA, 1992.

[50] E. Zeidler. *Nonlinear functional analysis and its applications. II/A*. Springer-Verlag, New York, 1990. Linear monotone operators, Translated from the German by the author and Leo F. Boron.