Discrete Input/Output Maps and their Relation to Proper Orthogonal Decomposition

Manuel Baumann, Jan Heiland, and Michael Schmidt

Abstract Current control design techniques require system models of moderate size to be applicable. The generation of such models is challenging for complex systems which are typically described by partial differential equations (PDEs), and model-order reduction or low-order-modeling techniques have been developed for this purpose. Many of them heavily rely on the state space models and their discretizations. However, in control applications, a sufficient accuracy of the models with respect to their input/output (I/O) behavior is typically more relevant than the accurate representation of the system states. Therefore, a discretization framework has been developed and is discussed here, which heavily focuses on the I/O map of the original PDE system and its direct discretization in the form of an I/O matrix and with error bounds measuring the relevant I/O error. We also discuss an SVD-based dimension reduction for the matrix representation of an I/O map and how it can be interpreted in terms of the Proper Orthogonal Decomposition (POD) method which gives rise to a more general POD approach in time capturing. We present numerical examples for both, reduced I/O maps and generalized POD.

Manuel Baumann
Delft Institute of Applied Mathematics, Faculty EWI, Mekelweg 4, 2628 CD Delft, The Netherlands, e-mail: m.m.baumann@tudelft.nl

Jan Heiland
Max-Planck-Institute in Magdeburg, Sandtorstr. 1, 39106 Magdeburg, Germany, e-mail: heiland@mpi-magdeburg.mpg.de

Michael Schmidt
University of Applied Sciences Offenburg, Badstraße 24, 77652 Offenburg, Germany, e-mail: schmidt@hs-offenburg.de

1 Introduction

To come up with a real-time controller for a system of partial differential equations, the synthesis of a surrogate model of moderate size that still inherits the important system dynamics is a necessary step.

A classical approach to obtain such surrogate models is to start with a finite but very high-dimensional state space model stemming from a spatial discretization of the system's original PDE state space model. In a next step, this model is reduced by so-called model-order reduction techniques (e.g. moment matching, balanced truncation) to a state space model of moderate size, for which current control design methods become feasible [2, 5, 15]. Often, a linearization step has to be carried out at some stage in addition. Another approach is low-order modeling POD, where the original state space model is analyzed to identify few and particularly relevant state space trajectories and which allows an approximate description of the state trajectory of the full system as a linear combination of the few ones [6, 14, 19, 20].

The classical methods mentioned above have in common that they heavily rely and focus on a state space representation. However, for control applications, only the I/O behavior of the system is of interest. Furthermore, state space representations of control systems can have simple I/O behaviors. Control engineers frequently use this insight when using black-box models from system identification methods applied to measured input/output sequences.

The high relevance of the I/O behavior has motivated the development of an alternative approach to generate surrogate models for control applications based on a direct discretization of the I/O behavior of the original infinite-dimensional system. A theoretical framework for the discretization of I/O maps of many important classes of linear infinite-dimensional systems has been established over the last years. Nonlinear systems can be treated in many cases after a linearization step [22, Ch. 3]. The result is an approximate representation of the I/O behavior via a matrix, and the approximation error is always considered with respect to the introduced error in the I/O behavior, cf. Fig. 1.

Our article is organized as follows: In Section 2, we will recall the essentials of the theoretical framework for the direct discretization of I/O maps established in [22] including error estimates. In Section 3, we will discuss the interesting feature of performing a special form of singular value decompositions on the I/O matrix representation, allowing to further reduce the size of the representation, but also to identify relevant locations for actuators and sensors. We will show that the well-known POD method can be interpreted as a special case of the direct I/O map discretization and a subsequent reduction. This observation gives rise to a generalized POD approach by extending the time discretization from snapshots to a wider class of time discretizations. Finally, in Section 4, we present numerical examples of applications of the direct discretization method, its application in optimal control, and of the generalized POD approach.

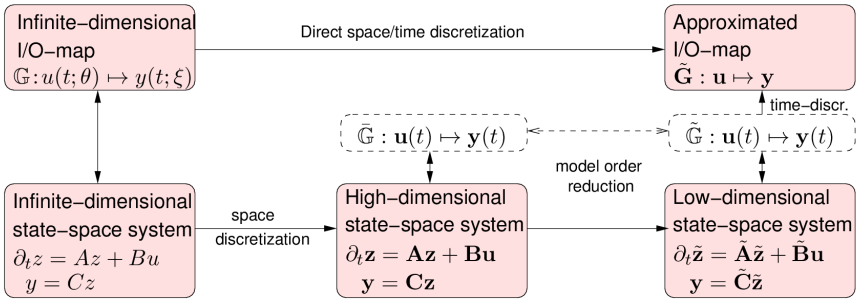

Fig. 1 Scheme of the classical model reduction approach and the direct discretization of the I/O map.

2 Direct Discretizations of I/O Maps

2.1 I/O Maps of Linear Systems

We follow [22] in style and notation and consider linear time-invariant (LTI) systems of first order:

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

Here, for time $t \in [0,T]$, the state $z(t)$ takes values in a Hilbert space Z. We assume that A is a densely defined unbounded operator $A\colon Z \supset D(A) \to Z$ generating a C^0-semigroup $(S(t))_{t\geq 0}$ on Z.

For Hilbert spaces U and Y, we assume the controls to be functions $u \in \mathcal{U} := L^2\big(0,T;U\big)$ and the observations $y \in \mathcal{Y} \subset L^2\big(0,T;Y\big)$.

If $B\colon U \to Z$ and $C\colon Z \to Y$ are bounded, then a bounded linear I/O map $\mathbb{G} \in \mathscr{L}(\mathcal{U},\mathcal{Y})$ can be associated with (1) by applying C to the unique mild solution defined as

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

see, e.g., [21, Ch. 4].

For later reference, we note that we can represent $\mathbb{G}$ as a convolution with a kernel function $K \in L^2(-T,T;\mathscr{L}(U,Y))$ via

$$(\mathbb{G}u)(t) = \int_0^T K(t-s)u(s)\,ds, \quad t \in [0,T], \tag{2}$$

with

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

This framework is suited for various differential equations (1), for instance heat equations, wave equations, transport equations, as well as linearizations of Navier-Stokes equations, cf. [22].

2.2 Discretization of I/O Maps in Two Steps

For the discretization of the I/O map

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

of the abstract system (1), we consider two steps:

1. *Approximation of signals.* For finite-dimensional subspaces $\bar{\mathcal{U}} \subset \mathcal{U}$ and $\bar{\mathcal{Y}} \subset \mathcal{Y}$ with orthogonal bases $\{u_1, \dots, u_{\bar{p}}\} \subset \bar{\mathcal{U}}$ and $\{y_1, \dots, y_{\bar{q}}\} \subset \bar{\mathcal{Y}}$ and corresponding orthogonal projections $\mathbb{P}_{\bar{\mathcal{U}}}$ and $\mathbb{P}_{\bar{\mathcal{Y}}}$, we find that the approximation
$$\mathbb{G}_S = \mathbb{P}_{\bar{\mathcal{Y}}} \mathbb{G} \mathbb{P}_{\bar{\mathcal{U}}}$$
is a finite-dimensional linear map which can be expressed as a matrix $\mathbf{G} \in \mathbb{R}^{\bar{q} \times \bar{p}}$ with elements $\mathbf{G}_{ij} = (y_i, \mathbb{G}u_j)_{\mathcal{Y}}$.
2. *Approximation of system dynamics.* The components $\mathbf{G}_{ij} = (y_i, \mathbb{G}u_j)_{\mathcal{Y}}$ can be obtained by computing the response of the model successively for inputs $u_1, \dots, u_{\bar{p}}$ and by testing it against all $y_1, \dots, y_{\bar{q}}$. For time-invariant state space systems and for bases with a space-time-like tensor structure
$$u_{(j,l)}(t) = \phi_j(t)\mu_l, \quad y_{(i,k)}(t) = \psi_i(t)\nu_k,$$
this task reduces to determining the observations $(\nu_k, Cz_l(t))_Y$ from the states $z_l(t) = S(t)B\mu_l$, where μ_l, $l = 1, \dots, p$ and ν_k, $k = 1, \cdots, q$ form bases of finite-dimensional subspaces of U and Y, where ϕ_j, $j = 1, \cdots, r$ and ψ_i, $i = 1, \cdots, s$ are bases of the time dimensions, and where $S(t)$ is the system's evolution semigroup.

Because the system's response is typically evaluated numerically, one has to consider an approximation $\mathbb{G}_{DS}$ of $\mathbb{G}_S$. The resulting total error ϵ_{DS} can be decomposed into the *signal* approximation error ϵ_S and the *dynamical* approximation error ϵ_D, i.e.

$$\underbrace{\|\mathbb{G} - \mathbb{G}_{DS}\|}_{=:\epsilon_{DS}} \leq \underbrace{\|\mathbb{G} - \mathbb{G}_S\|}_{=:\epsilon_S} + \underbrace{\|\mathbb{G}_S - \mathbb{G}_{DS}\|}_{=:\epsilon_D}, \tag{3}$$

in appropriate norms. In Thm. 1, it is shown that one can adjust $\bar{\mathcal{U}}$ and $\bar{\mathcal{Y}}$ and the accuracy of the numerical computations such that the errors are balanced.

In what follows, we consider discrete I/O maps and their numerical approximation. We have summarized the symbols used for the definition of the I/O maps on the different levels of approximation in Table 1.

∞-dimensional system	discrete I/O spaces	numerical approximation
$\mathbb{G}\colon \mathcal{U} \to \mathcal{Y}$	$\mathbb{G}_S\colon \bar{\mathcal{U}} \to \bar{\mathcal{Y}}$	$\mathbb{G}_{DS}\colon \bar{\mathcal{U}} \to \bar{\mathcal{Y}}$
	$\mathbb{G}_S\colon \mathcal{U}_{h_1,\tau_1} \to \mathcal{Y}_{h_2,\tau_2}$	$\mathbb{G}_{DS}\colon \mathcal{U}_{h_1,\tau_1} \to \mathcal{Y}_{h_2,\tau_2}$
	$\mathbf{G}\colon \mathbb{R}^{pr} \to \mathbb{R}^{qs}$	$\tilde{\mathbf{G}}\colon \mathbb{R}^{pr} \to \mathbb{R}^{qs}$
	$\mathbf{H} = \mathbf{M}_{\bar{\mathcal{Y}}}\mathbf{G}\colon \mathbb{R}^{pr} \to \mathbb{R}^{qs}$	$\tilde{\mathbf{H}}\colon \mathbb{R}^{pr} \to \mathbb{R}^{qs}$
		$\hat{\mathbf{G}}\colon \mathbb{R}^{\hat{p}\hat{r}} \to \mathbb{R}^{\hat{q}\hat{s}}$

Table 1 The considered I/O maps, their discretization, and their numerical approximation. The first row contains the operators for a general system and for a space-time like tensor structure of the discrete signal spaces $\bar{\mathcal{U}} = \mathcal{U}_{h_1,\tau_1} = \mathcal{R}_{\tau_1} \cdot U_{h_1}$ and $\bar{\mathcal{Y}} = \mathcal{Y}_{h_2,\tau_2} = \mathcal{S}_{\tau_2} \cdot Y_{h_2}$. The second line lists their matrix representations and the third line contains the low-dimensional approximation of $\tilde{\mathbf{G}}$.

2.3 Discretization of Signals

2.3.1 Space-time Discretization and Matrix Representation

Following the notation used in [17, 22], we recall the definitions and notions for direct discretization of I/O maps. In order to discretize the input signals $u \in \mathcal{U}$ and $y \in \mathcal{Y}$ in space and time, we choose four families $\{U_{h_1}\}_{h_1>0}$, $\{Y_{h_2}\}_{h_2>0}$, $\{\mathcal{R}_{\tau_1}\}_{\tau_1>0}$, and $\{\mathcal{S}_{\tau_2}\}_{\tau_2>0}$ of subspaces $U_{h_1} \subset U$, $Y_{h_2} \subset Y$, $\mathcal{R}_{\tau_1} \subset L^2(0,T)$ and $\mathcal{S}_{\tau_2} \subset L^2(0,T)$ of dimensions $p(h_1) = \dim(U_{h_1})$, $q(h_2) = \dim(Y_{h_2})$, $r(\tau_1) = \dim(\mathcal{R}_{\tau_1})$, and $s(\tau_2) = \dim(\mathcal{S}_{\tau_2})$ and define

$$\mathcal{U}_{h_1,\tau_1} := \mathcal{R}_{\tau_1} \cdot U_{h_1} \quad \text{and} \quad \mathcal{Y}_{h_2,\tau_2} = \mathcal{S}_{\tau_2} \cdot Y_{h_2}.$$

We denote the orthogonal projections onto these subspaces by $\mathbb{P}_{\mathcal{U},h_1,\tau_1} \in \mathscr{L}(\mathcal{U})$ and $\mathbb{P}_{\mathcal{Y},h_2,\tau_2} \in \mathscr{L}(\mathcal{Y})$. To approximate $\mathbb{G}$, we define

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

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

$\mathbb{R}^{q\times q}$, $\mathbf{M}_{\mathcal{R},\tau_1} \in \mathbb{R}^{r\times r}$ and $\mathbf{M}_{\mathcal{S},\tau_2} \in \mathbb{R}^{s\times s}$, for instance via

$$[\mathbf{M}_{U,h_1}]_{l_1 l_2} = (\mu_{l_1}, \mu_{l_2})_U, \qquad l_1, l_2 = 1, \dots, p.$$

These mass matrices induce 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 $\|\cdot\|_{p;w}$. We write signals $u \in \mathcal{U}_{h_1,\tau_1}$ and $y \in \mathcal{Y}_{h_2,\tau_2}$ as

$$u(t) = \sum_{l=1}^{p}\sum_{j=1}^{r} \mathbf{u}_j^l \phi_j(t)\mu_l, \qquad y(t) = \sum_{k=1}^{q}\sum_{j=1}^{s} \mathbf{y}_i^k \psi_i(t)\nu_k, \tag{4}$$

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

We obtain a matrix representation $\mathbf{G}$ of $\mathbb{G}_S$ by setting

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

where some dependencies on h_1, τ_1, h_2, τ_2 have been omitted. With the norms $\|\cdot\|_{pr;w}$ and $\|\cdot\|_{qs;w}$ in the product spaces with the mass matrices

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

the isomorphisms $\kappa_{\mathcal{U},h_1,\tau_1} \in \mathscr{L}(\mathcal{U}_{h_1,\tau_1}, \mathbb{R}^{pr}_w)$ and $\kappa_{\mathcal{Y},h_2,\tau_2} \in \mathscr{L}(\mathcal{Y}_{h_2,\tau_2}, \mathbb{R}^{qs}_w)$ associating functions with coefficient vectors are unitary mappings and we can define the discrete $\mathscr{L}(\mathcal{U},\mathcal{Y})$-norm as

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

For later reference, we define $\mathbf{H} = \mathbf{H}(h_1,\tau_1,h_2,\tau_2) := \mathbf{M}_{\mathcal{Y},h_2,\tau_2}\mathbf{G} \in \mathbb{R}^{qs\times pr}$, which is a matrix of $q \times p$ blocks $\mathbf{H}^{kl} \in \mathbb{R}^{s\times r}$ with block elements

$$\mathbf{H}^{kl}_{ij} = [\mathbf{M}_{\mathcal{Y}}\kappa_{\mathcal{Y}}\mathbb{P}_{\mathcal{Y}}\mathbb{G}(\mu_l\phi_j)]^l_j = (\nu_k\psi_i, \mathbb{G}(\mu_l\phi_j))_{\mathcal{Y}}. \tag{5}$$

We have the following convergence result:

Lemma 1 (Lem. 3.2, [22]). *For all $(h_1,\tau_1,h_2,\tau_2) \in \mathbb{R}^4_+$, we have*

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

If the subspaces $\{\mathcal{U}_{h_1,\tau_1}\}_{h_1,\tau_1>0}$ and $\{\mathcal{Y}_{h_2,\tau_2}\}_{h_2,\tau_2>0}$ are nested, i.e.

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

then $\|\mathbf{G}(h_1,\tau_1,h_2,\tau_2)\|_{qs\times pr;w}$ is monotonically increasing and bounded, i.e. convergent, for $(h_1,\tau_1,h_2,\tau_2) \searrow 0$.

2.3.2 An Example for Signal Discretizations

Let $U = Y = L^2(0,1)$ and let U_{h_1} and Y_{h_2} be spanned by continuous piecewise linear functions and let $\mathcal{R}_{\tau_1}$ and $\mathcal{S}_{\tau_2}$ be spanned by piecewise constant functions. For equidistant grids one can easily construct nested bases as illustrated in Fig. 2.

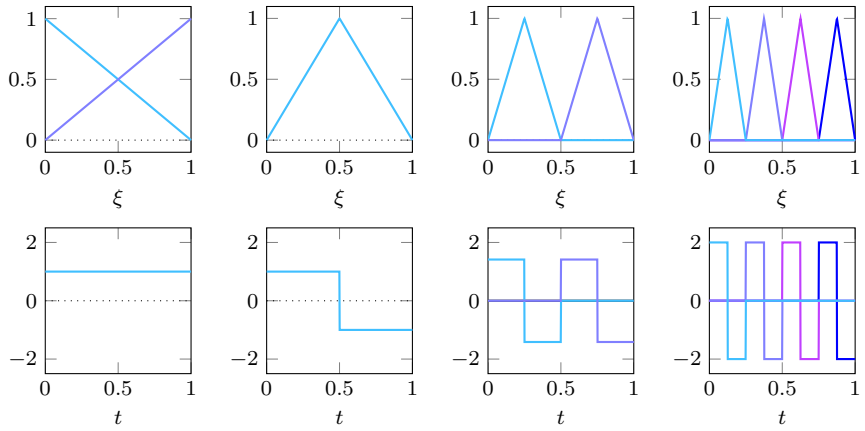

Fig. 2 Four levels of bases for nested $L^2(0,1)$-subspaces: Hierarchical basis for piecewise linear functions (above) and Haar wavelets for piecewise constant functions (below).

If we denote the orthogonal projections onto U_{h_1} and $\mathcal{R}_{\tau_1}$ by P_{U,h_1} and $P_{\mathcal{R},\tau_1}$, respectively, one can show that there exist $c_U = 1/2$ and $c_{\mathcal{R}} = 1/\sqrt{2}$, independent of h_1, τ_1 and T, such that

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

see e.g. [7]. From *Fubini*'s theorem, one can infer that a similar estimate holds for time space tensorized bases, i.e., the projection $\mathbb{P}_{\mathcal{U},h_1,\tau_1}$ satisfies

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

The same is true for the similarly defined projection $\mathbb{P}_{\mathcal{Y},h_2,\tau_2}$.

Similar estimates also hold for higher spatial dimensions as they are classical results from the interpolation theory in Sobolev spaces, see e.g. [7]. For spaces of higher regularity, using basis functions of higher polynomial degree, higher approximation orders are also possible.

2.3.3 Signal Approximation Error

As shown in [22, Lem. 3.3], the error in the signal approximation can be decomposed via $\epsilon_s := \|\mathbb{G} - \mathbb{G}_S\|_{\mathscr{L}(\mathcal{U},\mathcal{Y})} = \epsilon_{s,inp} + \epsilon_{s,outp}$ with

$$\epsilon_{s,inp} := \sup_{u \in \ker \mathbb{P}_{\mathcal{U},h_1,\tau_1}} \frac{\|\mathbb{G}u\|_{\mathcal{Y}}}{\|u\|_{\mathcal{U}}}, \quad \epsilon_{s,outp} := \max_{u \in \mathcal{U}_{h_1,\tau_1}} \frac{\|(I - \mathbb{P}_{\mathcal{Y},h_2,\tau_2})\mathbb{G}u)\|_{\mathcal{Y}}}{\|u\|_{\mathcal{U}}}.$$

As laid out in the following remarks, a good approximation in $\|\cdot\|_{\mathscr{L}(\mathcal{U},\mathcal{Y})}$ can only be achieved, if the subspaces $\mathcal{U}_{h_1,\tau_1}$ and $\mathcal{Y}_{h_2,\tau_2}$ are chosen *specifically* for $\mathbb{G}$. In short, the response of the chosen inputs needs to be well captured by the chosen output discretization. On the other hand, the discretization of the input space has to catch the major output modes of the system.

Remark 1. The usual requirement for families of approximating subspaces $\mathcal{U}_{h_1,\tau_1}$ and $\mathcal{Y}_{h_2,\tau_2}$ that they become dense in their superspaces is sufficient for pointwise convergence of $\|(\mathbb{G} - \mathbb{G}_S)u\|_{\mathcal{Y}} \to 0$ for every $u \in \mathcal{U}$, i.e. for convergence in the *strong operator topology*, but not for *uniform* convergence $\|\mathbb{G} - \mathbb{G}_S\|_{\mathscr{L}(\mathcal{U},\mathcal{Y})} \to 0$, cf. [22, Rem. 3.11].

Remark 2. For compact $\mathbb{G} \in \mathscr{L}(\mathcal{U},\mathcal{Y})$, 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 $\mathbb{G}u = \sum_{k=1}^{\infty} \sigma_k (u, \hat{u}_k)_{\mathcal{U}} \hat{y}_k$ for all $u \in \mathcal{U}$, see e.g. [26]. Thus, if we choose $\mathcal{U}_{h_1,\tau_1}$ 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 an efficient approximation $\mathbb{G}_S$ of $\mathbb{G}$ with $\|\mathbb{G} - \mathbb{G}_S\|_{\mathscr{L}(\mathcal{U},\mathcal{Y})} \leq \sigma_{r+1}$. However, we must point out that the I/O map of a linear system is not compact unless it is the zero map. This can be deducted from the common fact that the I/O map of a causal linear system is a *Toeplitz* operator [23] which is compact only if it is the zero operator [10, Rem. 7.15].

2.4 System Dynamics Approximation

We discuss the efficient computation of the discrete I/O map $\mathbf{G} = \mathbf{M}_{\mathcal{Y}}^{-1}\mathbf{H}$ of a state space system, via the approximation of the associated convolution kernel $K \in L^2(0,T;\mathscr{L}(U,Y))$.

2.4.1 Kernel Function Approximation

If we insert (2) in (5), after a change of variables, we obtain

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

where $\mathbf{W}\colon [0,T] \to \mathbb{R}^{s\times r}$ and $\mathbf{K}\colon [0,T] \to \mathbb{R}^{q\times p}$,

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

so that

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

Remark 3. For piecewise polynomial ansatz functions $\psi_i(t)$ and $\phi_j(t)$, $\mathbf{W}(t)$ can be calculated exactly. For particular choices, $\mathbf{W}(t) \in \mathbb{R}^{r\times r}$ is a lower triangular Toeplitz matrix for all $t \in [0,T]$. Hence, the matrices $\mathbf{H}_{ij} = \int_0^T \mathbf{W}_{ij}(t)\mathbf{K}(t)\,dt \in \mathbb{R}^{q\times p}$ satisfy $\mathbf{H}_{ij} = \mathbf{H}_{i-j}$ for $1 \le i,j \le r$ and $\mathbf{H}_{ij} = 0$ for $1 \le i < j \le r$. This implies that $\mathbf{H}_{ij}$ are *Markov parameters* of a discrete-time linear time-invariant causal *MIMO* system [22, Ch. 3.4.3].

For state space systems (1), the matrix-valued function $\mathbf{K}$ reads

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

where $c_k^* = C^*\nu_k \in Z$ and $b_l = B\mu_l$ for $k = 1,\dots,q$ and $l = 1,\dots,p$. Accordingly, one can obtain $\mathbf{K}$ by solving p homogeneous systems

$$\dot{z}_l(t) = Az_l(t), \qquad t \in (0,T], \tag{8a}$$
$$z_l(0) = b_l, \tag{8b}$$

recalling that (8) has the mild solution $z_l(t) = S(t)b_l \in C([0,T];L^2(\Omega))$. Typically, $z_l(t)$ is numerically approximated by $z_{l,\mathtt{tol}}(t)$, which gives only an approximation $\tilde{\mathbf{K}}$ of $\mathbf{K}$ and, thus, an approximation $\tilde{\mathbf{H}}$ of $\mathbf{H}$. Here, the subscript `tol` refers to a parameter that controls the error $z_l - z_{l,\mathtt{tol}}$ as we will specify later. Then, the approximation $\mathbb{G}_{DS}$ of $\mathbb{G}_S$, depending on h_1, h_2, τ_1, τ_2 and `tol`, is given by

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

Note that it might be preferable to consider an adjoint system and approximate the kernel functions via an adjoint system, cf. [22, Rem. 3.14].

2.4.2 System Dynamics Error

The approximation error in the system dynamics is related to the approximation error of the states via (8) as follows:

Proposition 1 (Thm. 3.6, [22]). *The system dynamics error ϵ_D satisfies*

$$\|\mathbb{G}_S - \mathbb{G}_{DS}\|_{\mathscr{L}(\mathcal{U},\mathcal{Y})} \leq \sqrt{T}\|\mathbf{K} - \tilde{\mathbf{K}}\|_{L^2(0,T;\mathbb{R}_w^{q\times p})}$$
$$\leq p\sqrt{T}\sqrt{\frac{\lambda_{\max}(\mathbf{M}_{Y,h_2})}{\lambda_{\min}(\mathbf{M}_{U,h_1})}} \max_{1\leq l\leq p} \|\mathbf{K}_{:,l} - \tilde{\mathbf{K}}_{:,l}\|_{L^2(0,T;\mathbb{R}^q)}.$$

Here $\mathbf{K}_{:,l}$ and $\tilde{\mathbf{K}}_{:,l}$ denote the l'th column of $\mathbf{K}(t)$ and $\tilde{\mathbf{K}}(t)$ that are defined via the exact and the approximate solutions of (8)*, 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 $\mathbf{M}_{U,h_1}$. $\mathbb{R}_w^{q\times p}$ denotes the space of real $q \times p$-matrices equipped with the weighted matrix norm $\|\mathbf{M}\|_{q\times p;w} = \sup_{\mathbf{u}\neq 0} \|\mathbf{M}\mathbf{u}\|_{q;w}/\|\mathbf{u}\|_{p;w}$.*

2.4.3 Error in the State Approximation

As laid out in Section 2.4.1, the approximation of the system dynamics is realized via numerically solving homogeneous PDEs (8) for p different initial values.

In view of the error estimates, we will focus on the approximation properties of the chosen discretization, i.e. whether

$$\|\mathbf{K}_{:,l} - \tilde{\mathbf{K}}_{:,l}\|_{L^2(0,T;\mathbb{R}^q)} < \texttt{tol} \tag{9}$$

is *guaranteed* for a given $\texttt{tol} > 0$ rather than on efficiency.

For the concrete system (8), the handling of the initial values b_l is an issue, as they in general only belong to Z but not necessarily to $D(A)$. For the heat equation, this typically leads to large but quickly decaying gradients in the exact solution $z_l \in C^1((0,T], H^2(\Omega)\cap H_0^1(\Omega))$.

This is reflected in the analytical bound for the general case

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

with some constant $c > 0$ independent of z^0 and T, cf. [18, p. 148].

To capture large gradients in the simulation, adaptive space and time discretizations are the methods of choice [11]. As for the guaranteed accuracy of the state approximation, the combination of discontinuous Galerkin time discretizations with standard Galerkin space discretizations comes with (a priori and a posteriori) error estimators, that also work for adapted meshes [12, 18, 24]. We distinguish two types of error estimates.

1. *Global state error estimates* that measure the global error $(z_l - z_{l,\texttt{tol}})$, see [12] for a priori and a posteriori estimates for parabolic problems that guarantee (9) by ensuring that

$$\|\mathbf{K}_{:,l} - \tilde{\mathbf{K}}_{:,l}\|^2_{L^2(0,T;\mathbb{R}^q)} \leq \|C\|^2_{\mathscr{L}(Z,Y)} \sum_{i=1}^{q} \|\nu_i\|_Y^2 \|z - z^{(l)}_{\texttt{tol}}\|^2_{L^2(0,T;Z)}. \tag{10}$$

2. *Goal-oriented error estimates* that measure the error $\|\mathbf{K}_{:,l}-\tilde{\mathbf{K}}_{:,l}\|_{L^2(0,T;\mathbb{R}^q)}$ directly. This is advantageous when the error in the *observations* $\mathbf{K}_{:,l}$ is small although the error in the states is large, see, e.g., [1].

Thus, for typical applications, via a suitable choice of the approximation schemes, we can fulfill the following assumption:

Assumption 1. For a given tolerance $\mathtt{tol}$, the approximations $z_{l,\mathtt{tol}}$ to the solutions z_l of (8) can be computed such that

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

2.5 Total Error Estimates

We show how the previously introduced error estimates sum up to an estimate of the total error in the approximation of $\mathbb{G}$.

Theorem 1 (Thm. 3.7, [22]). *Consider the I/O map $\mathbb{G}\in\mathscr{L}(\mathcal{U},\mathcal{Y})$ of the infinite-dimensional linear time-invariant system (2) and assume that $\mathbb{G}_{|\mathcal{U}_s}\in\mathscr{L}(\mathcal{U}_s,\mathcal{Y}_s)$ with spaces $\mathcal{U}_s\subset U$ and $\mathcal{Y}_s\subset Y$ such that for α_1, β_1, α_2, $\beta_2\in\mathbb{N}$, 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*

$$\|u-\mathbb{P}_{\mathcal{U},h_1,\tau_1}u\|_{\mathcal{U}} \le (c_{\mathcal{R}}\tau_1^{\alpha_1}+c_U h_1^{\beta_1})\|u\|_{\mathcal{U}_s}, \qquad u\in\mathcal{U}_s, \tag{11a}$$

$$\|y-\mathbb{P}_{\mathcal{Y},h_2,\tau_2}y\|_{\mathcal{Y}} \le (c_{\mathcal{S}}\tau_2^{\alpha_2}+c_Y h_2^{\beta_2})\|y\|_{\mathcal{Y}_s}, \qquad y\in\mathcal{Y}_s, \tag{11b}$$

with positive constants $c_{\mathcal{R}}$, $c_{\mathcal{S}}$, c_U and c_Y. And assume that the error in solving for the state dynamics can be made arbitrarily small, i.e. Assumption 1 holds.

Let $\delta>0$ be given. Assume, the chosen subspaces $\mathcal{U}_{h_1^*,\tau_1^*}$ and $\mathcal{Y}_{h_2^*,\tau_2^*}$ fulfill

$$\tau_1^* < \left(\frac{\delta}{8c_{\mathcal{R}}\|\mathbb{G}\|_{\mathscr{L}(\mathcal{U},\mathcal{Y})}}\right)^{1/\alpha_1}, \quad h_1^* < \left(\frac{\delta}{8c_U\|\mathbb{G}\|_{\mathscr{L}(\mathcal{U},\mathcal{Y})}}\right)^{1/\beta_1}, \tag{12a}$$

$$\tau_2^* < \left(\frac{\delta}{8c_{\mathcal{S}}\|\mathbb{G}\|_{\mathscr{L}(\mathcal{U}_s,\mathcal{Y}_s)}}\right)^{1/\alpha_2}, \quad h_2^* < \left(\frac{\delta}{8c_Y\|\mathbb{G}\|_{\mathscr{L}(\mathcal{U}_s,\mathcal{Y}_s)}}\right)^{1/\beta_2}, \tag{12b}$$

and that one can solve numerically the homogeneous systems (8) for $k=1,\dots,p(h_1)$ such that

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

Then,

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

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. The proof of [22, Thm. 3.7] is readily extended to the considered situtation noting that [22, Thm. 3.5] is valid for any $\mathcal{U}_s$ and $\mathcal{Y}_s$ for which $\mathbb{G}_{|\mathcal{U}_s} \in \mathscr{L}(\mathcal{U}_s, \mathcal{Y}_s)$ and (11) holds.

Remark 4. Considering subspaces $\mathcal{U}_s$ and $\mathcal{Y}_s$ allows for possibly better error estimates and does not exclude the case that $\mathcal{U}_s = \mathcal{U}$ or $\mathcal{Y}_s = \mathcal{Y}$. In particular, for modelling distributed control and observation, one can choose $Z = L^2(\Omega)$ for a spatial domain Ω and defines $U = L^2(\Theta)$ and $Y = L^2(\Sigma)$ as function spaces over domains of control $\Theta \subset \Omega$ and observation $\Sigma \subset \Omega$. Then $\mathcal{U}_s$ and $\mathcal{Y}_s$ can be chosen as subspaces of higher regularity like

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

for which (11) is obtained using standard approximation schemes [7].

3 Higher Order SVD for I/O Maps and POD

3.1 Dimension Reduction of the I/O Map via SVDs

An accurate resolution of the signal spaces by general basis functions may lead to large dimensions of the discrete I/O map $\tilde{\mathbf{G}}$. We show how to employ a *Tucker decomposition* or *higher order singular value decomposition* (HOSVD) [8] to reduce the degrees of freedom in input and output space while preserving accuracy and the tensor structure.

We consider the discrete spaces as introduced in Section 2.3.1 with their dimensions p, q, r, s and their indexing via j, i, l, k as in (4).

For $\tilde{\mathbf{G}} \in \mathbb{R}^{qs\times pr}$ considered as a fourth-order tensor $\tilde{\mathbf{G}} \in \mathbb{R}^{s\times r\times q\times p}$ with $\tilde{\mathbf{G}}_{ijkl} = \tilde{\mathbf{G}}^{kl}_{ij}$ there exists a HOSVD

$$\tilde{\mathbf{G}} = \mathbf{S} \times_1 \mathbf{U}^{(\psi)} \times_2 \mathbf{U}^{(\phi)} \times_3 \mathbf{U}^{(\nu)} \times_4 \mathbf{U}^{(\mu)}, \tag{13}$$

with the *core tensor* $\mathbf{S} \in \mathbb{R}^{s\times r\times q\times p}$ satisfying some orthogonality properties and unitary matrices $\mathbf{U}^{(\psi)} \in \mathbb{R}^{s\times s}$, $\mathbf{U}^{(\phi)} \in \mathbb{R}^{r\times r}$ and $\mathbf{U}^{(\nu)} \in \mathbb{R}^{q\times q}$, $\mathbf{U}^{(\mu)} \in \mathbb{R}^{p\times p}$. Here, $\times_1, \ldots, \times_4$ denote tensor-matrix multiplications. We define a *matrix unfolding* $\tilde{\mathbf{G}}^{(\psi)} \in \mathbb{R}^{s\times rqp}$ of the tensor $\tilde{\mathbf{G}}$ via

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

which is putting all elements belonging to $\psi_1, \psi_2, \ldots, \psi_s$ into one respective row. Similarly, we define the unfoldings $\tilde{\mathbf{G}}^{(\phi)} \in \mathbb{R}^{r\times qps}$, $\tilde{\mathbf{G}}^{(\nu)} \in \mathbb{R}^{q\times psr}$ and $\tilde{\mathbf{G}}^{(\mu)} \in \mathbb{R}^{p\times srq}$. Then we can calculate $\mathbf{U}^{(\psi)}$, $\mathbf{U}^{(\phi)}$, $\mathbf{U}^{(\nu)}$ and $\mathbf{U}^{(\mu)}$ in (13) by

means of four SVDs like

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

with $\Sigma^{(\psi)}$ diagonal with entries $\sigma_1^{(\psi)} \geq \sigma_2^{(\psi)} \geq \dots \sigma_s^{(\psi)} \geq 0$ and $\mathbf{V}^{(\psi)}$ columnwise orthonormal. The $\sigma_i^{(\psi)}$ are the *n-mode singular values* of the tensor $\tilde{\mathbf{G}}$.

From (13), we derive an approximation $\hat{\mathbf{G}} \in \mathbb{R}^{s \times r \times q \times p}$ of $\tilde{\mathbf{G}}$ by discarding the smallest n-mode singular values $\{\sigma_{\hat{s}+1}^{(\psi)}, \dots, \sigma_s^{(\psi)}\}$, $\{\sigma_{\hat{r}+1}^{(\phi)}, \dots, \sigma_r^{(\psi)}\}$, $\{\sigma_{\hat{q}+1}^{(\nu)}, \dots, \sigma_q^{(\nu)}\}$ and $\{\sigma_{\hat{p}+1}^{(\mu)}, \dots, \sigma_p^{(\mu)}\}$, i.e. by setting the corresponding parts of $\mathbf{S}$ to zero. Then we have

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

see [8]. Unlike the matrix case, this approximation needs not be optimal in a least square sense, see e.g. [9] for best norm approximations.

Finally, expressing the signals in terms of the corresponding leading singular vectors contained in $\mathbf{U}^{(\psi)}$, $\mathbf{U}^{(\phi)}$, $\mathbf{U}^{(\nu)}$, and $\mathbf{U}^{(\mu)}$, we obtain a low-dimensional representation of $\hat{\mathbf{G}}$ in the smaller space $\mathbb{R}^{\hat{q}\hat{r} \times \hat{p}\hat{s}}$.

Apart from being used for defining a reduced model, the major HOSVD modes can be interpreted as the most relevant input and output signals - an insight that can be exploited for sensor and actuator design and placement.

3.2 I/O Maps and the Classical POD Method

To illustrate the relation of POD and HOSVD-reduced I/O maps, we consider the finite-dimensional LTI system

$$\dot{v} = Av + f$$

with the output $y = v$ and the sole (one-dimensional) input $Bu = f$.

Consider the grid $\mathcal{T} = \{t_i\}_{i=1}^s$ of time instances

$$0 = t_1 < t_2 < \dots < t_s = T$$

and let $\bar{\mathcal{Y}}$ be the span of the nodal vectors of $v(t)$ taken at $t \in \mathcal{T}$. Then, the corresponding $\tilde{\mathbf{G}}$ for a single input dimension is a $s \times 1 \times q \times 1$ tensor that can be unfolded into the matrix

$$\tilde{\mathbf{G}}^{(\nu)} = \mathbf{X} = \begin{bmatrix} v_1(t_1) & \dots & v_1(t_s) \\ \vdots & \ddots & \vdots \\ v_q(t_1) & \dots & v_q(t_s) \end{bmatrix} \in \mathbb{R}^{q \times s}. \tag{14}$$

As laid out in the discussion of the truncated HOSVD in Section 3.1, a reduced basis for the space dimension of the state space can be obtained via a truncated SVD of $\tilde{\mathbf{G}}^{(\nu)}$. We observe that this reduced basis would be the well-known reduced basis used in POD, see [25] for an introduction.

3.3 A Generalized Approach in Time Capturing for POD

In POD, one uses samplings from the time evolution to compress the spatial state dimension. Instead of considering the matrix of snapshots (14) from discrete time instances, we propose considering the measurement matrix

$$\mathbf{X}_{gen} = \begin{bmatrix} (v_1,\psi_1)_{\mathcal{S}} & \dots & (v_1,\psi_s)_{\mathcal{S}} \\ \vdots & \ddots & \vdots \\ (v_q,\psi_1)_{\mathcal{S}} & \dots & (v_q,\psi_s)_{\mathcal{S}} \end{bmatrix} \in \mathbb{R}^{q\times s}, \tag{15}$$

of the same dimension that is obtained by testing the spatial components of the state v against the basis functions of a discrete $\mathcal{S}_{\tau_2} = \operatorname{span}\{\psi_1,\cdots,\psi_s\} \subset L^2(0,T)$. Note that for smooth trajectories, the standard snapshot matrix is obtained from (15) by testing against *delta distributions* located at $t \in \mathcal{T}$.

The L^2 orthogonal projection of the state vector $\mathbf{v}$ onto the space spanned by the measurements is given as

$$\hat{\mathbf{v}}(t) = \mathbf{X}_{gen}\mathbf{M}_{\mathcal{S}}^{-1}\boldsymbol{\psi}(t),$$

where $\boldsymbol{\psi} := [\psi_1,...,\psi_s]^{\mathsf{T}}$ and where $\mathbf{M}_{\mathcal{S}}$ is the mass matrix of $\mathcal{S}_{\tau_2}$, i.e. $[\mathbf{M}_{\mathcal{S}}]_{i_1 i_2} = (\psi_{i_1},\psi_{i_2})_{\mathcal{S}}$. As for the classical POD, cf. [25, Thm. 1.12], where $\mathbf{M}_{\mathcal{S}} = I$, one can show that the generalized POD modes are the eigenvalues corresponding to the largest eigenvalues of the operator

$$\begin{aligned}\mathbf{R} &= \int_0^T \hat{\mathbf{v}}\hat{\mathbf{v}}^{\mathsf{T}}\,dt = \int_0^T \mathbf{X}_{gen}\mathbf{M}_{\mathcal{S}}^{-1}\boldsymbol{\psi}(t)\boldsymbol{\psi}(t)^{\mathsf{T}}\mathbf{M}_{\mathcal{S}}^{-1}\mathbf{X}_{gen}^{\mathsf{T}}\,dt \\ &= \mathbf{X}_{gen}\mathbf{M}_{\mathcal{S}}^{-1}\underbrace{\int_0^T \boldsymbol{\psi}(t)\boldsymbol{\psi}(t)^{\mathsf{T}}\,dt}_{=\mathbf{M}_{\mathcal{S}}}\mathbf{M}_{\mathcal{S}}^{-1}\mathbf{X}_{gen}^{\mathsf{T}} = \mathbf{X}_{gen}\mathbf{M}_{\mathcal{S}}^{-1}\mathbf{X}_{gen}^{\mathsf{T}}.\end{aligned}$$

Therefore, in an implementation, the generalized POD modes can be obtained via an SVD of

$$\mathbf{X}_{gen}\mathbf{M}_{\mathcal{S}}^{-1/2}. \tag{16}$$

4 Numerical Examples and Applications

By means of four numerical examples we present numerical convergence results, an application of the discretized I/O map for the solution of optimization problems including the use of SVD, as well as numerical tests comparing the I/O maps motivated POD variant to the standard approach.

4.1 Convergence of the I/O Map for a Heat Equation

We consider a heat equation with homogeneous Dirichlet boundary conditions, which for a domain Ω with a C^2-boundary and for $Z = L^2(\Omega)$ becomes a system of type (1) with the Laplace operator

$$A = \triangle : D(A) = H^2(\Omega) \cap H_0^1(\Omega) \subset Z \to Z. \tag{17}$$

Since A is the infinitesimal generator of an analytic C^0-semigroup of contractions $(S(t))_{t\geq 0}$, the mild solution z of (1) exhibits the following stability and regularity properties, see e.g. [21, Ch. 7] and [13].

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

$$\|z\|_{H^{1,2}((0,T)\times\Omega)} \leq c\|u\|_{\mathcal{U}}. \tag{18}$$

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

For the numerical tests, we will consider domains $\Omega \subset \mathbb{R}^2$ and define control and observation operators as follows: For given points a_c, b_c, a_m, $b_m \in \bar{\Omega}$, let $\Omega_c = (a_{c,1}, a_{c,2}) \times (b_{c,1}, b_{c,2})$ and $\Omega_m = (a_{m,1}, a_{m,2}) \times (b_{m,1}, b_{m,2})$ be rectangular subsets of Ω where the control is active and the observation takes place, respectively. Let $U = Y = L^2(0,1)$ and define $C \in \mathscr{L}(L^2(\Omega), Y)$ and $B \in \mathscr{L}(U, L^2(\Omega))$ via

$$(Cz)(\xi) = \int_{a_{m,1}}^{b_{m,1}} \frac{z(x_1, x_2(\xi))}{b_{m,1} - a_{m,1}}\, dx_1, \tag{19}$$

$$(Bu)(x_1,x_2) = \begin{cases} u(\theta(x_1))\omega_c(x_2), & (x_1,x_2) \in \Omega_c \\ 0, & (x_1,x_2) \notin \Omega_c \end{cases}, \tag{20}$$

where $\omega_c \in L^2(a_{c,2}, b_{c,2})$ is a weight function and $\theta : [a_{c,1}, b_{c,1}] \to [0,1]$ and $x_1 : [0,1] \to [a_{m,1}, b_{m,1}]$ are affine-linear transformations.

Since $C_{|H^2(\Omega)} \in \mathscr{L}(H^2(\Omega), H^2(0,1))$, we have $\mathbb{G} \in \mathscr{L}(\mathcal{U}, \mathcal{Y}_s)$ as well as

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

Also, for $u \in \mathcal{U}_s$, we have $\|u\|_{\mathcal{U}} \leq \|u\|_{\mathcal{U}_s}$, and for $u \in \mathcal{U}$, we have $\|\mathbb{G}u\|_{\mathcal{Y}_s} \leq c'\|z\|_{H^{1,2}((0,T)\times\Omega)} \leq cc'\|u\|_{\mathcal{U}}$, where c is the constant used in the stability estimate (18) and $c' = \max\{\|C\|_{\mathscr{L}(L^2(\Omega),L^2(\Xi))}, \|C\|_{\mathscr{L}(H^2(\Omega),H^2(\Xi))}\}$.

For the concrete test case, we set $T = 1$ and $\Omega = (0,1)^2$ and choose $\Omega_c = \Omega$, $\Omega_m = (0.1, 0.2) \times (0.1, 0.9)$, and $\omega_c(x_2) = \sin(\pi x_2)$, see Fig. 3. For this choice, for inputs of the form $u(t;\theta) = \sin(\omega_T \pi t)\sin(m\pi\theta)$ with $\omega_T, m \in \mathbb{N}$, there is an analytic expression for the outputs in terms of the eigenfunctions of the Laplace operator (17).

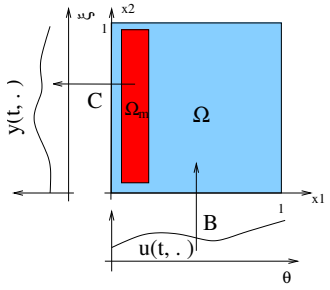

Fig. 3 Setup for the heat equation with Dirichlet boundary conditions and known solutions.

For the finite-dimensional approximation of the I/O map $\mathbb{G}$, we choose hierarchical linear finite elements in U_{h_1} and Y_{h_2} and Haar wavelets in $\mathcal{R}_{\tau_1}$ and $\mathcal{S}_{\tau_2}$ and compute $\mathbb{G}_{DS}(h_1, \tau_1, h_2, \tau_2, \texttt{tol})$. The tolerance `tol` refers to the accuracy of the numerical approximation of the system dynamics, cf. Ass. 1.

Convergence of single outputs. For the present setup, with inputs $u(t;\theta) = \sin(\omega_T\pi t)\sin(m\pi\theta)$, and known outputs $y = \mathbb{G}u$, we investigate the relative error $\|y - \tilde{y}\|_{\mathcal{Y}}/\|u\|_{\mathcal{U}_s}$, with $\tilde{y} = \mathbb{G}_{DS}(h_1, \tau_1, h_2, \tau_2, \texttt{tol})u$, for varying discretization parameters h_1, τ_1, h_2, τ_2 and `tol`. For $m = 5$ and $\omega_T = 10$, we observe a quadratic convergence with respect to decreasing $h_1 = h_2$ (cf. Fig. 4(a)) and a linear convergence for $\tau_1 = \tau_2$ (cf. Fig. 4(b)). It is notable, however, that due to the system dynamics error, the error converges to a positive plateau value depending on the tolerance `tol`.

Convergence of the norm $\|\mathbb{G}_S(h_1, \tau_1, h_2, \tau_2)\|_{\mathscr{L}(\mathcal{U},\mathcal{Y})}$ for nested subspaces. Successively improving the signal approximation by adding additional basis functions, the norm $\|\mathbb{G}_S(h_1, \tau_1, h_2, \tau_2)\|_{\mathscr{L}(\mathcal{U},\mathcal{Y})}$ converges, cf. Lem. 1. We approximate $\|\mathbb{G}_S\|_{\mathscr{L}(\mathcal{U},\mathcal{Y})}$ by $\|\mathbb{G}_{DS}\|_{\mathscr{L}(\mathcal{U},\mathcal{Y})}$, where $\mathbb{G}_{DS}$ has been calculated with $\texttt{tol} = 4.0e-5$. In Fig. 4(c), the approximations $\|\mathbb{G}_S(h_1, \tau_1, h_2, \tau_2)\|_{\mathscr{L}(\mathcal{U},\mathcal{Y})} = \|\mathbb{G}_S(\frac{1}{p-1}, \frac{1}{r}, \frac{1}{q-1}, \frac{1}{s})\|_{\mathscr{L}(\mathcal{U},\mathcal{Y})}$ are plotted for increasing subspace dimensions $p = q = r + 1 = s + 1 = 2, 3, \ldots, 65$.

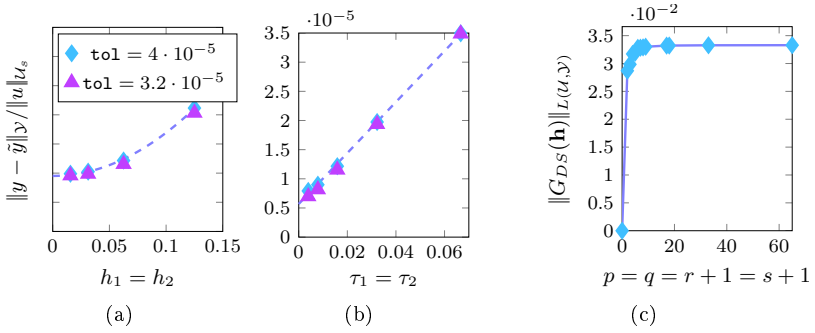

Fig. 4 The relative output errors for input $u(t;\theta) = \sin(10\pi t)\sin(5\pi\theta)$ for varying $h_1 = h_2$ and fixed $\tau_1 = \tau_2 = 1/64$ (a) and for varying $\tau_1 = \tau_2$ and fixed $h_1 = h_2 = 1/17$ (b). The dashed lines are the linear and the quadratic fit. (c): the norm of the discrete I/O map $\|\mathbb{G}_{DS}(\mathbf{h})\|_{\mathscr{L}(\mathcal{U},\mathcal{Y})}$ versus approximation space dimensions $p = q = r + 1 = s + 1$ for a fixed tolerance $\texttt{tol} = 4.0e-5$.

4.2 I/O Maps and Higher Order SVD for Optimal Control

We demonstrate the use of discrete I/O maps in optimization problems like

$$J(u,y) = \frac{1}{2}\|y - y_D\|_{\mathcal{Y}}^2 + \alpha\|u\|_{\mathcal{U}}^2 \to \min, \quad \text{s. t. } y = \mathbb{G}u,\ u \in \mathcal{U}_{ad}, \tag{21}$$

where $\mathcal{U}_{ad} \subset \mathcal{U}$ is the subset of admissible controls, $y_D \in \mathcal{Y}$ is a target output signal, and $\alpha > 0$ is a regularization parameter. With $\mathbf{y}_D = \kappa_{\mathcal{Y},h_2,\tau_2}\mathbb{P}_{\mathcal{Y},h_2,\tau_2}y_D$ and $\bar{U}_{ad} = \{\mathbf{u} \in \mathbb{R}^{pr} : \mathbf{u} = \kappa_{\mathcal{U},h_1,\tau_1}\mathbb{P}_{\mathcal{U},h_1,\tau_1}u,\ u \in \mathcal{U}_{ad}\}$, we define the discrete approximation to (21) as,

$$\bar{J}_{\mathbf{h}}(\mathbf{u},\mathbf{y}) = \frac{1}{2}\|\mathbf{y} - \mathbf{y}_D\|_{qs;w}^2 + \alpha\|\mathbf{u}\|_{pr;w}^2 \to \min \quad \text{s. t. } \mathbf{y} = \tilde{\mathbf{G}}\mathbf{u},\ \mathbf{u} \in \bar{U}_{ad} \tag{22}$$

For an optimization problem without control constraints, i.e. $\mathcal{U}_{ad} = \mathcal{U}$ and $\bar{U}_{ad} = \mathbb{R}^{pr}$, the solution $\bar{\mathbf{u}}$ of (22) is readily defined by

$$(\tilde{\mathbf{G}}^{\mathsf{T}}\mathbf{M}_{\mathcal{Y}}\tilde{\mathbf{G}} + \alpha\mathbf{M}_{\mathcal{U}})\bar{\mathbf{u}} = \tilde{\mathbf{G}}^{\mathsf{T}}\mathbf{M}_{\mathcal{Y}}\mathbf{y}_D. \tag{23}$$

As the test case for optimal control, we consider a model for the heat conduction in two infinitely long plates of width 5 and height 0.2 which are connected by two rectangular bars and which are surrounded by an insulating material. That is, for $t \in (0,1]$ we solve a heat equation with homogeneous Neumann boundary conditions on a domain Ω as in Fig. 5. We assume, that we can heat the bottom plate and measure the temperature distribution in the upper plate, i.e. to model the control and observation

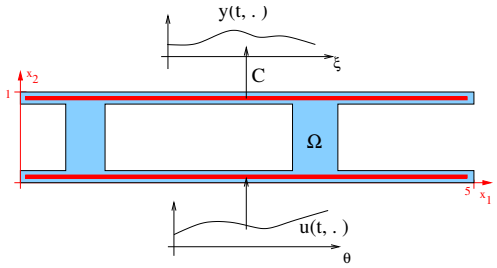

Fig. 5 Setup of the test case for the heat equation with homogeneous Neumann boundary conditions.

we set $\Omega_c = (0.05, 4.95) \times (0.05, 0.15)$, $\Omega_m = (0.05, 4.95) \times (0.85, 0.95)$ and $\omega_c(x_2) = \sin(\pi(x_2 - 0.05)/0.1)$ and define B and C similarly to (19).

As the target output, we choose $y_D = \mathbb{G}u_0$, i.e. the output for the input $u_0 \equiv 1$. Then, the solution of the optimal control problem (21) will give an optimized input u_* that leads to a similar output as u_0 but at lower costs.

As the reference for a subsequent HOSVD based reduction, we solve (23) with an approximated I/O map $\tilde{\mathbf{G}} \in \mathbb{R}^{17\cdot64\times65\cdot64}$ and $\alpha = 10^{-4}$ for an approximation $\bar{u}$ to u_*. This took 0.33 seconds on a desktop PC. The norm of the input was reduced by 27.9% with a relative deviation of $\mathbb{G}\bar{u}$ from y_D of 9.4%. Having employed a HOSVD to reduce the I/O map to $\hat{\mathbf{G}} \in \mathbb{R}^{3\cdot5\times3\cdot5}$, i.e. having truncated all but the 3 most relevant spatial and the 5 most relevant temporal input and output modes, the calculation of $\bar{u}$ took less than 0.0004 seconds. The optimal control on the base of $\hat{\mathbf{G}}$ came with a norm reduction of 27.4% if compared to u_0 while admitting a deviation from the target of 9.5%. See Fig. 6 for an illustration of the optimization results.

The insights provided by an HOSVD are illustrated in Fig. 7, namely the distribution of the n-mode singular values indicating how many degrees of freedom are needed to capture the dynamics and the most relevant input and output modes, cf. Section 3.1. From the modes one can draw conclusions on effective actuations. In the given example, the locations of the connecting bars are clearly visible in the spatial input modes μ_1 and μ_2.

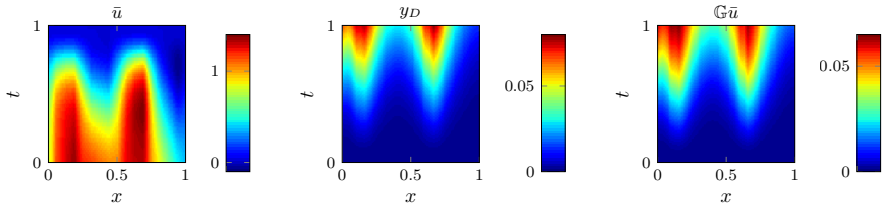

Fig. 6 Application of the SVD-reduced approximated I/O map $\hat{\mathbf{G}} \in \mathbb{R}^{3\cdot5\times3\cdot5}$ in an optimization problem. From left to right: the optimal control $\bar{u}$, the target output $y_D = \mathbb{G}u_0$, and the output $\mathbb{G}\bar{u}$.

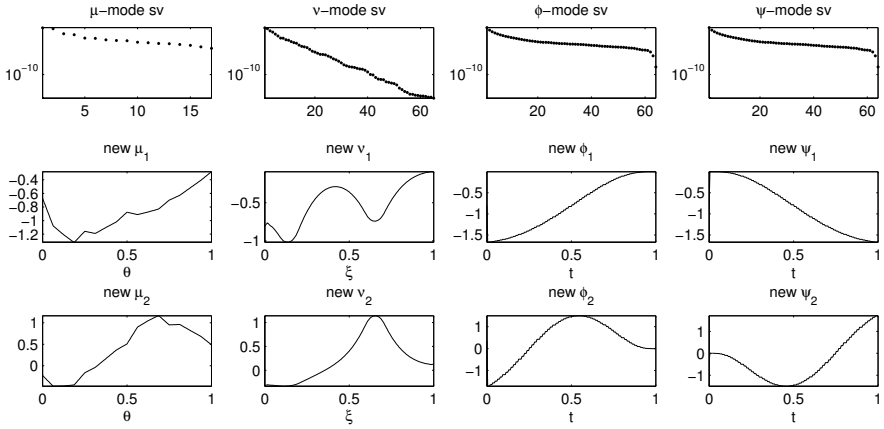

Fig. 7 Results of an HOSVD applied to the I/O map of the test case illustrated in Fig. 5. In the first row the n-mode singular values in a semilogarithmic scale are plotted. In the 2nd and 3rd row the respective most relevant modes are shown.

4.3 Driven Cavity & Generalized POD

We consider a driven cavity flow at *Reynolds number* $Re = 2000$ in the unit square. For the spatial discretization we use *Taylor-Hood* elements on a uniform triangulation of the domain by 50^2 triangles, see Fig. 8 for an illustration of the flow and the discretization. We linearize the equations about the velocity solution α of the associated steady state Stokes problem, and consider the time evolution of the velocity $v(t) \in \mathbb{R}^{N_v}$ and the pressure $p(t) \in \mathbb{R}^{N_p}$, N_v, $N_p \in \mathbb{N}$, modeled by

$$M\dot{v}(t) = -A(\alpha, Re)v(t) + J^{\mathsf{T}}p(t) + f(t) \tag{24a}$$

$$0 = Jv(t) \tag{24b}$$

$$v(0) = \alpha \tag{24c}$$

in the time interval $(0, 5]$. Here, $M \in \mathbb{R}^{N_v, N_v}$ is the velocity mass matrix, $J \in \mathbb{R}^{N_p, N_v}$ and its transpose J^{T} are discretizations of the divergence and the gradient operator, $A \in \mathbb{R}^{N_v, N_v}$ models the diffusion and the linearized convection depending on Re and v_0, and $f \in \mathbb{R}^{N_v}$ is a source terms that arises from the linearization and the incorporation of the boundary conditions. See [16] for details of the modeling.

We compare the I/O map motivated variant of Section 3.3 of POD to the standard approach that we refer to as POD. To the generalization we refer as gmPOD. To the columns of the measurement matrices (14) or (15) we will refer to as *measurements* which is more general than the standard term *snapshot*. By s, we denote the number of measurements, i.e. the number of snapshots or the dimension of the test space, respectively. By $\hat{k}$, we denote

the number of modes, i.e. leading singular vectors of the measurement matrix, that are taken for the reduced order model (often called *POD-dimension*).

Say $U_{\hat{k}}$ is the matrix of the chosen $\hat{k}$ modes obtained via a truncated SVD of (14) for POD or (16) for gmPOD. Then, a reduced model of (24) is given as

$$\hat{M}\dot{\hat{v}}(t) = \hat{A}\hat{v}(t) + \hat{f}(t, U_{\hat{k}}\hat{v}(t)), \tag{25a}$$
$$\hat{v}(0) = \hat{\alpha}, \tag{25b}$$

where $\hat{M} := U_{\hat{k}}^{\mathsf{T}} M U_{\hat{k}}$, $\hat{A} := U_{\hat{k}}^{\mathsf{T}} A U_{\hat{k}}$, and $\hat{f} = U_{\hat{k}}^{\mathsf{T}} f$, with M, A, and f being the mass matrix, coefficient matrix, and a nonlinearity or source term. The solution $\hat{v}$ is related to the actual solution v via the ansatz $v = U_{\hat{k}}\hat{v}$. Accordingly, the initial value is typically chosen as $\hat{\alpha} = U_{\hat{k}}^{\mathsf{T}} v(0)$. We consider only the velocity component for the measurements for (24a). Since the measurements fulfill $JX = 0$, we have that $U_{\hat{k}}^{\mathsf{T}} J^{\mathsf{T}} = 0$.

As the basis for the (generalized) time capturing, we use the hierarchical basis of piecewise linear functions illustrated in Fig. 2.

The time integration for the full order system was implemented using the implicit trapezoidal rule on a uniform time grid. The corresponding reduced order systems were numerically integrated using *Scipy*'s (for the current test case) and *Matlab*'s (for the following test case) built-in ODE solvers `integrate.odeint` and `ODE45`. The source code of the POD and gmPOD tests is available from the author's *Github* account [4].

We consider the error $e_{s,\hat{k}} \approx \left(\int_0^T \|v(t) - U_{\hat{k}}\hat{v}(t)\|_{L^2(\Omega)}^2 dt\right)^{1/2}$ where v is the solution of the full order system and $\hat{v}$ solves the reduced order system constructed by means of s measurements and $\hat{k}$ POD modes. Here, T is the endpoint of the considered time interval and Ω is the spatial domain of the considered test case. The error is evaluated numerically using the piecewise trapezoidal rule in time and the finite element space norm on the chosen discretization of the full order model.

The conducted tests were designed to evaluate the evolution of the error $e_{s,\hat{k}}$ with respect to the number of measurements s and the dimension of the reduced model $\hat{k}$. Both parameters are limiting factors: the amount of needed memory increases linearly with s and the computational complexity to compute the POD modes via an SVD increases like $\hat{k}$ times the square of the spatial dimension. Also, the effort for the numerical solution of the reduced systems scales with the number of POD modes necessary for the required accuracy.

For the test case with the linearized flow equations (24), we find that gmPOD significantly outperforms POD. Throughout the investigated range of measurements and number of chosen POD modes, the approximation error after the gmPOD reduction is much smaller than after a POD reduction of the same dimension, see Fig. 10(b).

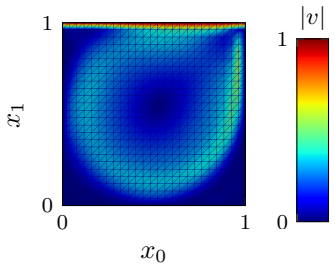

Fig. 8 Illustration of the driven cavity at the steady state for $Re = 2000$ and the triangulation of the domain.

4.4 Nonlinear Burgers' Equation & Generalized POD

As the last test case, we consider a system stemming from a spatial discretization of the nonlinear Burgers' equation,

$$\partial_t z(t,x) + \partial_x\big(\frac{1}{2}z(t,x)^2 - \nu\partial_x z(t,x)\big) = 0, \tag{26}$$

with the spatial coordinate $x \in (0,1)$, the time variable $t \in (0,1]$, the viscosity parameter $\nu = 0.01$ completed by zero Dirichlet boundary conditions and a step function as initial conditions as illustrated in Fig. 9(a). See [3] and [4] for details on the implementation.

We compute the gmPOD-reduced model and compare it to the standard POD reduction as explained in Section 4.3. Again, the generalized POD approach gmPOD outperforms the classical variant in terms of accuracy versus number of snapshots and POD modes, see Fig. 10(a) for the error plot and Fig. 9(b-c) for an illustration of the approximation error.

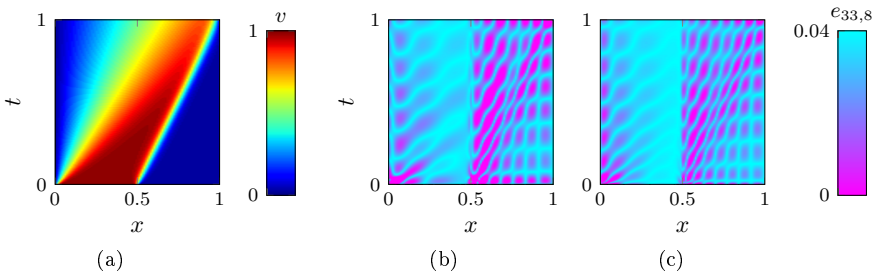

Fig. 9 Full order approximation of Burgers' equation (a) and the error made by POD (b) and gmPOD (c) using $s = 33$ snapshots and a reduced model of dimension $\hat{k} = 8$.

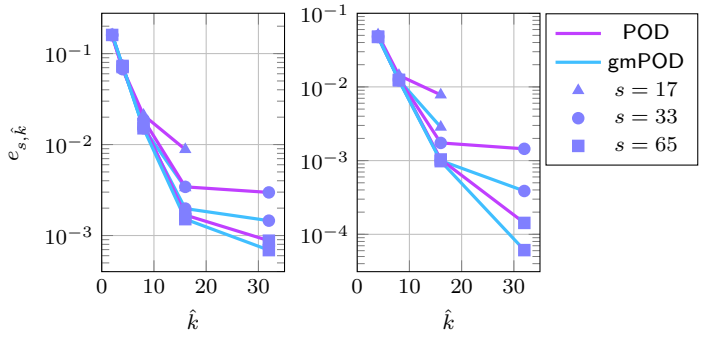

Fig. 10 The time and space approximation error $e_{s,\hat{k}}$: (a) for the semi-discrete Burgers' equation (Section 4.4) and (b) for the linearized flow equations (Section 4.3) for varying numbers s of measurements and varying numbers $\hat{k}$ of POD modes used for the reduced model.

5 Final Remarks and Outlook

The presented framework is suitable to provide direct discretizations of I/O maps of linear infinite-dimensional control systems with *spatially distributed* inputs and outputs. This allows for tailoring numerical approximations to the I/O behavior of systems which is particularly important in control setups. The provided methodology, thus, comes with error estimates in the relevant $\mathcal{L}(\mathcal{U}, \mathcal{Y})$ norm. We have illustrated the approximation properties in a numerical example.

The discrete I/O map can be expressed as a tensor and reduced by higher order SVDs to the most relevant input and output modes. Apart from the reduction, a HOSVD also gives insights for controller and sensor design.

We have shown that the idea of sampling trajectories in time to compress spatial information is related to POD reduced order modeling. The freedom to choose the test functions for the sampling can be used to define POD reduced models on the basis of measurements that are more general than snapshots. The newly introduced generalized POD variant has shown to significantly outperform classical POD reduction for a linear and a nonlinear test case.

The proposed variant of POD seems promising in further aspects that are subject to future work. Unlike in the standard approach, the sampling is done in a subspace of the state space which can be exploited for improved error estimates. Secondly, the nonlocalized measurements have smoothing properties that might be of advantage for noisy measurements. Finally, the freedom in bases choice allows for problem specific measurement functions which may reduce the number of necessary snapshots.

References

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

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

[3] M. Baumann. Nonlinear Model Order Reduction using POD/DEIM for Optimal Control of Burgers' Equation. Master's thesis, Delft University of Technology, The Netherlands, 2013.

[4] M. Baumann and J. Heiland. genpod - matlab and python implementation with test cases. `https://github.com/ManuelMBaumann/genpod.git`, September 2014.

[5] P. Benner, V. Mehrmann, and D. Sorensen (Editors). *Dimension Reduction of Large-Scale Systems*, volume 45 of *LNSCE.* Springer, Heidelberg, 2005.

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

[7] 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.

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

[9] 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.

[10] R. G. Douglas. *Banach Algebra Techniques in Operator Theory.* Academic Press, New York, NY, 1972.

[11] 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 University Press, Cambridge, 1995.

[12] 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.

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

[14] 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.

[15] 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.

[16] J. Heiland and V. Mehrmann. Distributed control of linearized Navier-Stokes equations via discretized input/output maps. *Z. Angew. Math. Mech.*, 92(4):257–274, 2012.

[17] J. Heiland, V. Mehrmann, and M. Schmidt. A new discretization framework for input/output maps and its application to flow control. In R. King, editor, *Active Flow Control. Papers contributed to the Conference "Active Flow Control II 2010", Berlin, Germany, May 26 to 28, 2010*, pages 375–372. Springer, Berlin, 2010.

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

[19] 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.

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

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

[22] M. Schmidt. *Systematic Discretization of Input/Output Maps and other Contributions to the Control of Distributed Parameter Systems*. PhD thesis, TU Berlin, Fakultät Mathematik, Berlin, Germany, 2007.

[23] O. J. Staffans. *Well-posed Linear Systems*. Cambridge University Press, 2005.

[24] V. Thomée. *Galerkin finite element methods for parabolic problems*. Springer, Berlin, 1997.

[25] S. Volkwein. *Model reduction using proper orthogonal decomposition*. Lecture Notes, Institute of Mathematics and Scientific Computing, University of Graz, Austria, 2011.

[26] D. Werner. *Funktionalanalysis*. Springer, Berlin, 2000.